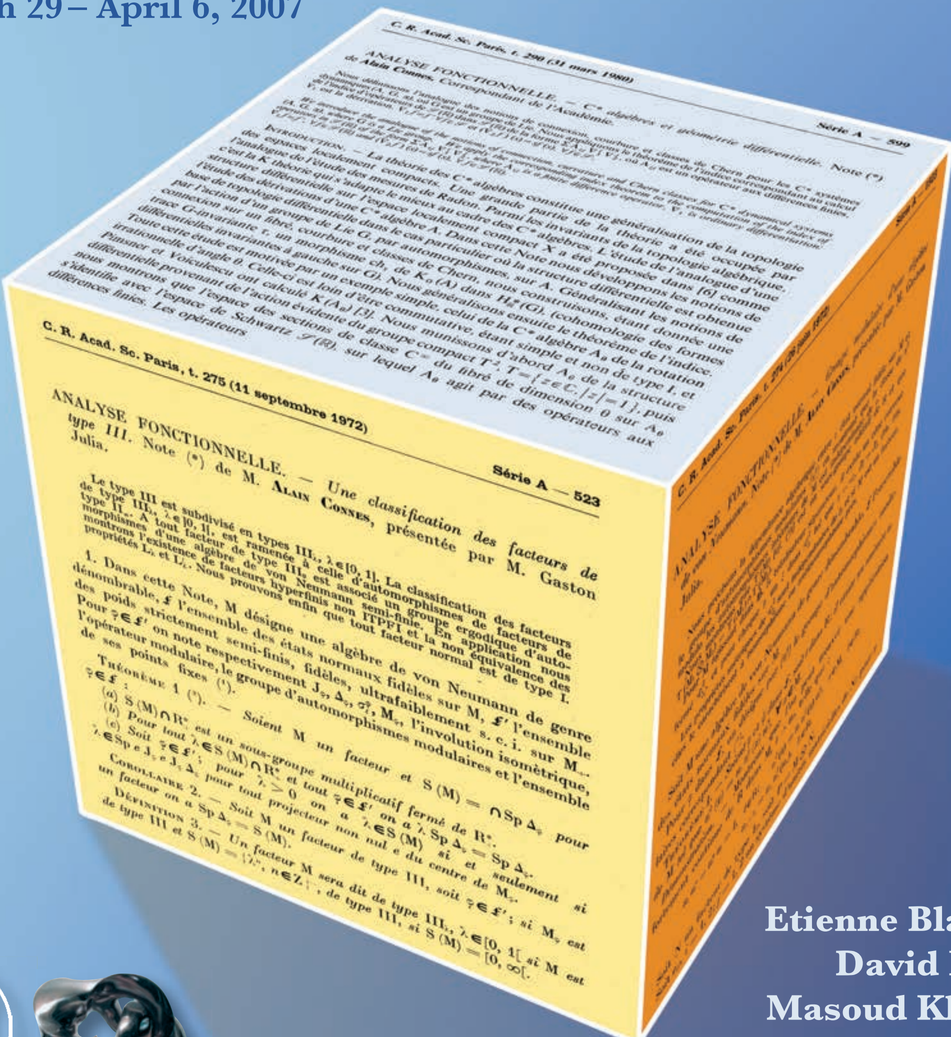
Clay Mathematics Proceedings

Volume 11

Quanta of Maths

Conference in honor of Alain Connes
Non Commutative Geometry
Institut Henri Poincaré, Institut des Hautes Études Scientifiques,
Institut de Mathématiques de Jussieu, Paris, France
March 29 – April 6, 2007



Etienne Blanchard
David Ellwood
Masoud Khalkhali
Matilde Marcolli
Henri Moscovici
Sorin Popa
Editors

American Mathematical Society
Clay Mathematics Institute

Quanta of Maths

Clay Mathematics Proceedings
Volume 11

Quanta of Maths

Conference in honor of Alain Connes
Non Commutative Geometry
Institut Henri Poincaré,
Institut des Hautes Études Scientifiques,
Institut de Mathématiques de Jussieu,
Paris, France
March 29– April 6, 2007

Etienne Blanchard
David Ellwood
Masoud Khalkhali
Matilde Marcolli
Henri Moscovici
Sorin Popa
Editors

American Mathematical Society
Clay Mathematics Institute

Photographs courtesy of Jean-François Dars.

2000 *Mathematics Subject Classification.* Primary 58B34.

Library of Congress Cataloging-in-Publication Data

Quanta of maths : non commutative geometry conference in honor of Alain Connes, March 29 to April 6, 2007, Institut Henri Poincaré, Institut des hautes études scientifiques, Institut de mathématiques de Jussieu, Paris, France / Etienne Blanchard . . . [et al.], editors.

p. cm. — (Clay mathematics proceedings ; v. 11)

Includes bibliographical references and index.

ISBN 978-0-8218-5203-3 (alk. paper)

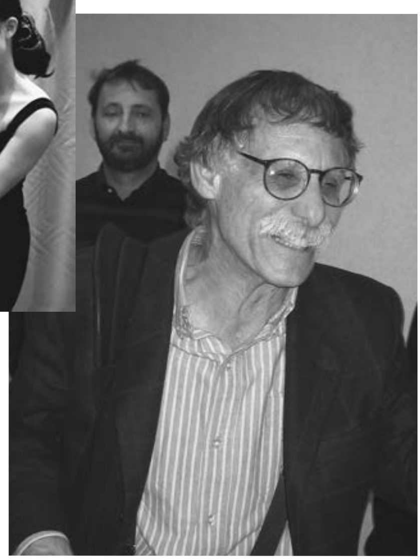
1. Noncommutative differential geometry—Congresses. I. Connes, Alain. II. Blanchard, Etienne, 1968– III. Institut Henri Poincaré. IV. Institut des hautes études scientifiques (Paris, France) V. Institut de mathématiques de Jussieu.

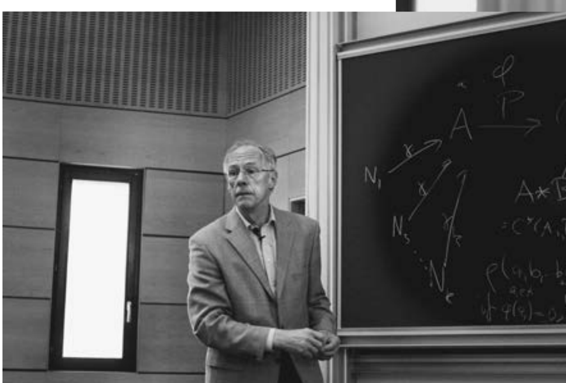
QA614.Q36 2011
512′.55—dc22

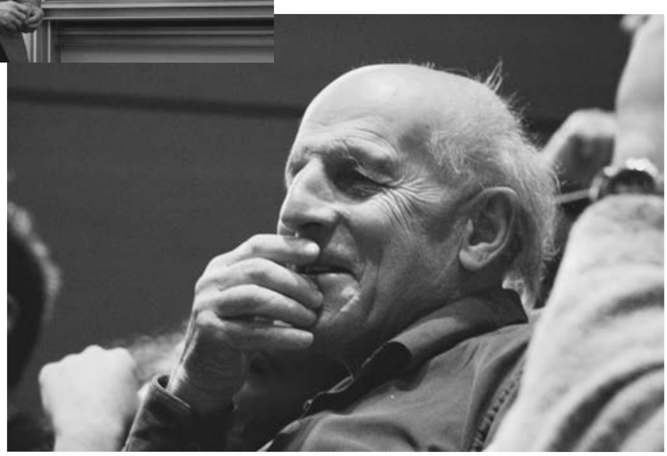
2010034787

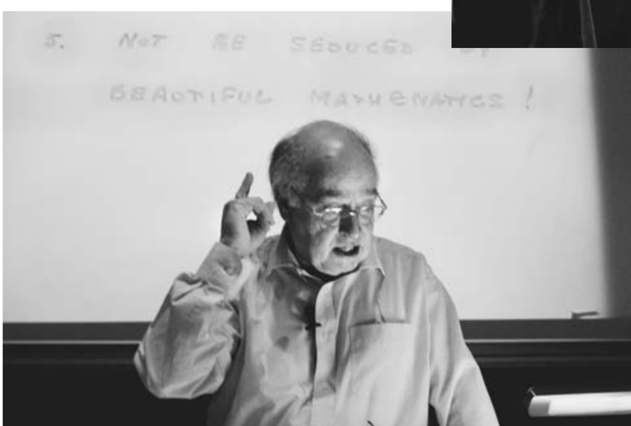
10 9 8 7 6 5 4 3 2 1 15 14 13 12 11 10

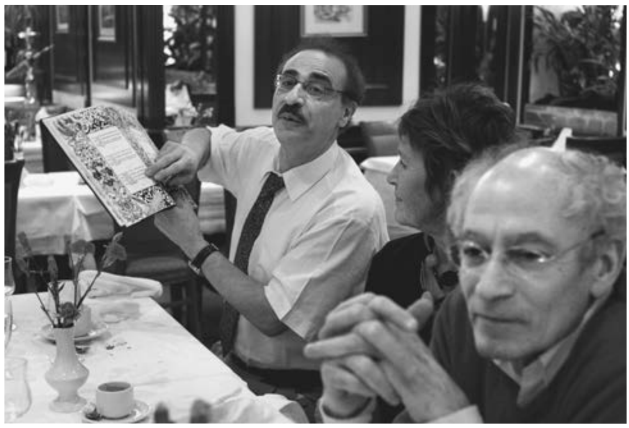


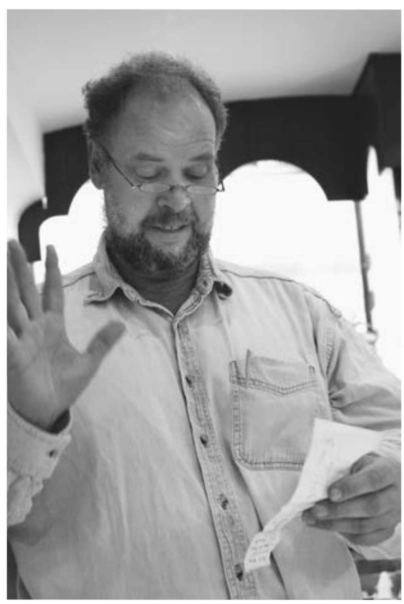








5. NOT BE SEDUCED
BEAUTIFUL MATHEMATICS !





Contents

Preface

The conference celebrating the 60th birthday of Alain Connes was hosted jointly by IHÉS (March 29–30, 2007) and Institut Henri Poincaré (April 2–7, 2007). The talks were a vibrant testimony of the health and vitality of the subject of noncommutative geometry, as reflected by its many interactions with other fields, including operator algebras, physics, analysis, topology, and number theory, which is the most recently opened frontier. The conference was a marvelous tribute to the breadth and depth of Alain Connes' contributions to mathematics.

The conference started with a talk by **Michael Atiyah**, entitled

Radical thoughts on the foundations of physics

Atiyah observed that all physical models since Newton, quantum mechanics included, have started from the basic premise that one can predict the future from full knowledge of the present. He suggested an alternative to this paradigm: besides the present, we may also need some knowledge of the past in order to predict the future. That is, perhaps the universe has memory, and somehow the short term memory of the past should play a role in predictions. This means that, instead of an ordinary differential equation or PDE, one should use a "delayed differential equation".

The next talk of the morning session was delivered by **Yuri Manin**, with the title

Cohomomorphisms and operads

and based on his joint paper with D. Borisov. Their work gives a unified axiomatic treatment of generalized operads as functors on categories of abstract labeled graphs, thus providing an approach to symmetry and moduli objects in non-commutative geometries.

In the afternoon, the first talk was by **Katia Consani**, and was entitled

Vanishing cycles: an adelic analogue

She drew an analogy between the geometry of the adèle class space and the theory of singularities, by which the complement of the idèle class group in the adèle class space is compared to the singular fiber of an algebraic degeneration.

It was followed by a talk given by **Matilde Marcolli**, entitled

How noncommutative geometry looks at number theory

She gave a broad survey of the interactions between noncommutative geometry and number theory, starting from the phase transitions with spontaneous symmetry breaking associated to $\mathbb{Q}$-lattices. She also described the new notion of endomotive, which together with cyclic cohomology gives the conceptual meaning in noncommutative geometry of the spectral realization of zeros of L-functions.

The first talk on Friday was given by **Erling Størmer**:

Survey of entropy for operator algebras

Erling gave vivid recollections of his collaboration with Alain in the seventies, and of the evolution since that time of the notion of entropy for automorphisms of factors, originally introduced by Connes and Størmer in order to classify shifts of the II_1 hyperfinite factor.

The second talk was by **Alain Connes**, and was entitled

Noncommutative geometry and physics

Alain described the physics part of his book with Matilde Marcolli, whose content is about equally divided between number theory and physics. He showed in particular how the work on renormalization and motivic Galois theory fits with the understanding, obtained in joint work with A. Chamseddine and M. Marcolli, of the extremely complex Lagrangian of gravity coupled with matter as unveiling the fine texture of space-time using the spectral action principle. At the end of the talk he explained the link between the two parts of the book based on the analogy between the electroweak phase transition in the standard model and the phase transitions which play a crucial role in the quantum statistical mechanics models involved in the approach to RH. In particular he proposed to extend the symmetry breaking to the full gravitational sector so that geometry appears only at low temperature as an emerging phenomenon.

The last talk on Friday was by

Don Zagier, and had the title

Quantum modular forms

Don explained a new type of modular objects, which live on the boundary of the usual domain of modular forms. He gave a number of examples of such quantum modular forms, coming from number theory, combinatorics (q-series) and quantum invariants of 3-manifolds and knots.

Friday ended with a delightful piano concert (Chopin) and poetry session, with pianist Lydie Solomon at the piano, and the poetess Nicole Barriere reciting 20th century poetry and one of her poems.

After a weekend break, the conference moved on Monday to the Institut Henri Poincaré (IHP), in the heart of Paris.

The first talk on Monday was given by

Dirk Kreimer, with the title

Diffeomorphism invariance, locality and the residue:
a physicist's harvest of a friend's work

He explained the role of locality and of the "residue" in the conceptual understanding of the Feynman diagram computations of quantum field theory. The residue also plays a basic role in noncommutative geometry, by filtering out the unimportant details and giving the proper meaning to the notion of locality in that framework.

Ali Chamseddine delivered the talk entitled

The little key to uncover the hidden noncommutative structure of space-time

Ali showed how, using the spectral action principle and spectral triples in noncommutative geometry, one can determine the finite noncommutative space whose product by the continuum describes the fine structure of space-time corresponding to the Standard Model coupled to gravity.

Gianni Landi gave the next talk, with the title

Quantum Groups and Quantum Spaces are Noncommutative Geometries

Gianni described his recent results with his collaborators, showing that quantum spheres and quantum groups such as $\mathrm{SU}_q(2)$ admit natural Dirac operators yielding spectral geometries with finite summability and fulfilling all regularity conditions of spectral triples.

Michel Dubois-Violette gave the talk

Moduli spaces for regular algebras

He discussed the noncommutative generalizations of polynomial algebras, which can be used in various noncommutative settings, noncommutative differential geometry, noncommutative algebraic geometry, etc. as well as in the applications in physics.

Masoud Khalkhali's talk was entitled

Hopf cyclic cohomology and noncommutative geometry:
some new thoughts and a tribute to Alain

He explained the role played by index theory and transverse geometry of foliations in Connes' discovery of cyclic cohomology in 1980-1981, and of Hopf cyclic cohomology in the late 1990's by Connes and Moscovici. He then gave a speculative survey on how Hopf cyclic theory can be extended to deal with more general types of symmetries, like those defined by quasi Hopf algebras and Hopf algebras in braided monoidal categories.

On Tuesday the theme of the conference moved to von Neumann algebras and was a tribute to Alain Connes' immense legacy in the subject. The speakers were as follows.

Anthony Wassermann, who talked about

Non-commutative geometry and conformal field theory

He gave a survey of the deep interaction between conformal field theory and operator algebras and in particular of the role played by the composition of correspondences between von Neumann algebras to describe the fusion rules turning the positive energy representations into a tensor category.

Vaughan Jones gave a talk entitled

Operations on planar algebras and subfactors

He explained the development of the theory of planar algebras which, among other applications, provided a very convenient axiomatization of the standard invariant of a subfactor and a useful new technique, similar in spirit to Conway's skein theory, for analyzing their structure.

Dietmar Bisch talked about

Free product of planar algebras and inclusions of subfactors

The standard invariant of a subfactor can be axiomatized in algebraic-combinatorial terms as a planar algebra. Jones and Bisch discovered a notion of "free product" of planar algebras, which gives rise to uncountably many new infinite depth subfactors. For instance, the Fuss-Catalan planar algebras can be viewed as a "free product" of two Temperley-Lieb planar algebras.

Sorin Popa's talk was entitled

Rigidity phenomena in von Neumann algebras of group actions

Sorin described a wealth of new startling results, which he obtained in the recent years by combining rigidity techniques based on property T together with his notion of malleability derived from Haagerup's compact approximation property. His work solved the long-standing problem of the computation of the fundamental group of type II_1 factors.

Stefan Vaes talked about

Explicit computations of all bifinite Connes' correspondences for certain II_1 factors

He showed how to use the rigidity techniques introduced by S. Popa to give an explicit description of the ring of all bifinite correspondences for certain II_1 factors, a result which illustrates the remarkable power of the new techniques.

Dimitri Shlyakhtenko then gave the talk

Free entropy dimension, L^2 derivations and stochastic calculus

He showed that the L^2-derivations introduced in his joint work with Alain give rise to a free stochastic differential equation which has a stationary solution. This solution gives lower bounds on Voiculescu's microstates free entropy dimension for generators of a finite von Neumann algebra M. It follows that for a class of R^ω embeddable groups, their microstates free entropy is given by the expression $\beta_1^{(2)} - \beta_0^{(2)} + 1$, with $\beta_j^{(2)}$ being the Cheeger-Gromov L^2 Betti numbers.

Wednesday's talks focused on the Baum-Connes conjecture, topology and C*-algebras.

Vincent Lafforgue gave the first talk, with the title

Strengthening property (T)

He introduced the strengthened property (T) which is an obstacle to prove the Baum-Connes conjecture with arbitrary coefficients for $\mathrm{SL}_3(\mathbb{R})$ or $\mathrm{SL}_3(\mathbb{Q}_p)$, using only the bivariant theory KK_{ban}, and stability under holomorphic functional calculus.

Gennadi Kasparov's talk was entitled

A K-theoretic index formula for transversally elliptic operators

Gennadi explained the K-theoretic version of the index theorem for transversally elliptic operators in the sense of Atiyah, which gives a conceptual understanding of the computation of the index formula in K-theory using C^*-algebra cross products and the Kasparov product. The talk provided yet another illustration of the power of his bivariant theory.

Nigel Higson's talk was on

The Baum-Connes conjecture and the Mackey analogy

He explained the connection between C^*-algebra K-theory and Mackey's proposal to study representation theory for a semisimple group G by developing an analogy between G and an associated semidirect product group. He showed how to use Mackey's point of view to give a new proof of the complex semisimple case of the Connes-Kasparov conjecture.

Marc Rieffel delivered a talk entitled

A new look at 'Matrix algebras converge to the sphere'

He showed how the approximation of the sphere by matrix algebras can be precisely understood using the notion of metric on noncommutative spaces obtained as a natural generalization of the Dirac distance between states introduced by Connes for spectral triples.

Guoliang Yu talked about

Higher index theory of elliptic operators and noncommutative geometry

He gave a survey on recent development of higher index theory in the context of noncommutative geometry, its applications, and its fascinating connection to the geometry of groups and metric spaces. Higher index theory has important applications to problems in differential topology and differential geometry such as the Novikov Conjecture on homotopy invariance of higher signatures and the existence problem of Riemannian metrics with positive scalar curvature.

Paul Baum lecture had the title

Noncommutative algebraic geometry and the representation theory of p-adic groups

He described his joint work with R. Plymen and A.-M. Aubert. Motivated by the tools of noncommutative geometry such as cyclic homology applied in the context of representation theory they conjecture the existence of a simple geometric structure underlying questions of reducibility of parabolically induced representations of reductive p-adic groups.

The first talk on Thursday was given by

Joachim Cuntz

C^*-algebras associated with the ax+b-semigroup over N

He presented a C^*-algebra which is naturally associated to the $ax+b$-semigroup over $\mathbb{N}$. It is simple and purely infinite and can be obtained from the algebra considered by Bost and Connes by adding one unitary generator which corresponds to addition. Its stabilization can be described as a crossed product of the algebra of continuous functions, vanishing at infinity, on the space of finite adeles for $\mathbb{Q}$ by the natural action of the $ax + b$-semigroup over $\mathbb{Q}$.

Uffe Haggerup's talk was entitled

Connes' classification of injective factors seen from a new perspective

Uffe presented his joint work with Magdalena Musat on the classification of preduals of injective factors up to completely bounded (cb) isomorphisms. They show that preduals of semifinite factors are not cb-isomorphic to preduals of type III factors and they obtain a characterization of those hyperfinite factors M whose preduals are cb-isomorphic to the predual of the unique hyperfinite type III_1-factor.

Dan Voiculescu gave the talk entitled

Aspects of free analysis

He surveyed the analysis around the free difference quotient derivation, which is the natural derivation for variables with the highest degree of noncommutativity and explained how a highly noncommutative extension of the spectral theory of resolvents is emerging from his free probability theory. Dan also discussed the possibility of using this spectral theory in relativistic quantum physics.

Alain Connes's talk was entitled

Thermodynamics of endomotives and the zeros of zeta

He described the number theory part of his joint book with Matilde Marcolli, and the phase transitions with spontaneous symmetry breaking which arise for the higher dimensional analogues of the Bost-Connes system, such as the GL(2)-system of two-dimensional $\mathbb{Q}$-lattices.

The last lecture on Thursday was delivered by

Henri Moscovici, with the title

Spectral geometry of noncommutative spaces

A long time collaborator and friend of Alain, Henri gave a survey of their joint work inspired by index theory, highlighting the local index formula in noncommutative geometry and its application to the transverse geometry of foliations, which in turn led to their discovery of Hopf cyclic cohomology. He then focused on their recent work on twisted spectral triples of type III. Twisting notwithstanding, the Chern character of such a spectral triple still lands in the standard cyclic cohomology of the underlying algebra, which raises the challenge of expressing it by a local formula.

The conference on Thursday ended with a reception at Institut de Mathématiques de Jussieu and the dedication of a birthday gift to Alain, a telescope to celebrate his farsighted visionary work!

There were four talks in the last day of the conference on Friday, all by former students of Alain.

The first was by

Alain Valette, entitled

Proper isometric actions on Hilbert and Banach spaces

He began by recalling in a witty manner the way he was introduced to property T by Alain, during a fast car drive from Bures to Paris. He then explained a beautiful piece of geometric group theory, inspired by the Baum-Connes conjecture.

George Skandalis talked about

Holonomy groupoid and C-algebra of a foliation*

He showed that many aspects of the interaction between noncommutative geometry and foliation theory generalize to the set up of singular foliations.

At the end of Georges' talk the whole audience gave him and the other organizers a big ovation for the splendid organizing job.

Marc Rosso's talk had the title

Quantum groups and algebraic combinatorics

He showed that the quantum shuffle approach to quantum groups and the combinatorics of Lyndon words provide new character formulas for irreducible representations and a new way of constructing the canonical bases.

Jean-Benoît Bost gave the last talk of the conference, on

Characteristic values of evaluation maps and Diophantine geometry

He sketched his conceptual geometric approach to Diophantine approximation as a theory of characteristic numbers in the context of Arakelov geometry and some intriguing analogies with noncommutative geometry and elliptic theory on noncommutative spaces.

This volume collects articles contributed by speakers at the conference, with their coauthors in some cases, together with a few others, offered for the occasion by the following mathematicians: **Etienne Blanchard**, **Dan Burghelea**, **Pierre Cartier**, **Max Karoubi**, **Jean-Louis Loday** and **Maria Ronco**, **Alejandro Perez** and **Carlo Rovelli**, **James Simons** and **Dennis Sullivan**, and **Raimar Wulkenhaar**. Many of the papers in this volume cover new results, and do not necessarily reflect the topic and content of the talks given at the conference.

The editors would like to heartily thank all the contributors. Along with them, we also express our enormous respect and admiration for the honoree, and wish him many more years of fruitful and inspirational research.

Acknowledgements. The editors are very grateful to Vida Salahi of the Clay Mathematics Institute for handling the whole editorial process and formatting the articles. Our thanks are also due to Arthur Greenspoon for kindly copyediting and checking the entire manuscript.
The volume's cover was designed by Marie Claude Vergne, IHES, and Teresa Levy of the AMS. We would like to thank both of them.

Finally, we would like to acknowledge the generous support of the following organizations, without whom this event would not have been possible. (Région Ile-de-France, Clay Mathematics Institute, Centre National de la Recherche Scientifique, Institut de Mathématiques de Jussieu et le Projet Algèbres d Opérateurs, Institut Henri Poincaré, Fédération de Recherche Mathématiques de Paris Centre, Université Pierre et Marie Curie, Université Paris Diderot, GDR Géométrie Non Commutative, Institut des Hautes Études Scientifiques, Institut Universitaire de France, Collège de France)

Etienne Blanchard, David Ellwood, Masoud Khalkhali, Matilde Marcolli, Henri Moscovici, Sorin Popa

August 2010

Clay Mathematics Proceedings
Volume 11, 2010

A Geometric Description of Equivariant K-Homology for Proper Actions

Paul Baum, Nigel Higson, and Thomas Schick

Dedicated with admiration and affection to
Alain Connes on his 60th birthday

ABSTRACT. Let G be a discrete group and let X be a G-finite, proper G-CW-complex. We prove that Kasparov's equivariant K-homology groups $KK^G_*(C_0(X), \mathbb{C})$ are isomorphic to the geometric equivariant K-homology groups of X that are obtained by making the geometric K-homology theory of Baum and Douglas equivariant in the natural way. This reconciles the original and current formulations of the Baum-Connes conjecture for discrete groups.

1. Introduction

In the original formulation of the Baum-Connes conjecture [**BC00**], the topological K-theory of a discrete group G (the "left-hand side" of the conjecture) was defined geometrically in terms of proper G-manifolds. Later, in [**BCH94**], the definition was changed so as to involve Kasparov's equivariant KK-theory. The change was made to accommodate new examples beyond the realm of discrete groups, such as p-adic groups, for which the geometric definition was not convenient or adequate. But it left open the question of whether the original and revised definitions are equivalent for discrete groups (for connected or almost-connected Lie groups the equivalence is straightforward). This is the question that we shall address in this paper.

In a recent article [**BHS07**], we gave a complete proof that the (non-equivariant) geometric K-homology theory of Baum and Douglas [**BD82**] agrees with Kasparov's K-homology on finite CW-complexes. Here we shall show that our techniques extend to show that the original and revised definitions of topological K-theory for a discrete group agree—provided that those techniques are supplemented by a key result of Lück and Oliver about equivariant vector bundles over G-finite, proper G-CW-complexes [**LO01**].

Lück and Oliver prove that if X is a G-finite, proper G-CW-complex, then there is a rich supply of equivariant vector bundles on X. It follows that the Grothendieck group

---

2010 *Mathematics Subject Classification.* Primary 19K33; Secondary 19K35, 19L47, 58J22.

The first- and second-named authors were partially supported by US National Science Foundation grants.

$K_G^0(X)$ of complex G-vector bundles is the degree zero group of a $\mathbb{Z}/2\mathbb{Z}$-graded cohomology theory on X, and it is essentially this fact that we shall need to carry over the arguments of [**BHS07**] to the equivariant case.

We shall show that the Lück-Oliver theorem is equivalent to the assertion that the crossed product C^*-algebra $C^*(X, G)$ associated to the action of G on X has an approximate identity consisting of projections. As a result $K_G^0(X)$ is isomorphic to the K_0-group of the crossed product C^*-algebra. This has some further benefits for us—for example it makes it clear that each complex G-vector bundle on a G-compact proper G-manifold has a unique smooth structure, up to isomorphism.

Returning to the Baum-Connes conjecture, the assertion that the old and the revised versions are the same is a consequence of the assertion that the natural G-equivariant development of the Baum-Douglas K-homology theory, which we shall write as $K_*^G(X)$, is isomorphic to Kasparov's group $KK_*^G(C_0(X), \mathbb{C})$ for any G-finite, proper G-CW-complex X. There is a natural map

$$\mu\colon K_*^G(X) \longrightarrow KK_*^G(C_0(X), \mathbb{C})$$

which is defined using the index of Dirac operators, and we shall prove that it is an isomorphism. What makes this nontrivial is that the groups $K_*^G(X)$ do not obviously constitute a homology theory. We shall address this problem by introducing groups $k_*^G(X)$ that manifestly *do* constitute a homology theory and by constructing a commutative diagram

$$\begin{tikzcd} k_*^G(X) \arrow[r] \arrow[d, equal] & K_*^G(X) \arrow[d, "\mu"] \\ k_*^G(X) \arrow[r, "\mu"'] & KK_*^G(C_0(X), \mathbb{C}). \end{tikzcd}$$

We shall prove that the map from $k_*^G(X)$ to $KK_*^G(C_0(X), \mathbb{C})$ is an isomorphism when X is a G-finite, proper G-CW-complex and that the map from $k_*^G(X)$ to $K_*^G(X)$ is surjective. This proves that the map from $k_*^G(X)$ to $KK_*^G(C_0(X), \mathbb{C})$ is an isomorphism.

2. Proper Actions

Throughout the paper we shall work with a fixed a countable discrete group G. By a G-*space* we shall mean a topological space with an action of G by homeomorphisms. We shall be concerned in the first place with *proper* G-*CW-complexes*. These are G-spaces with filtrations

$$\emptyset = X^{-1} \subseteq X^0 \subseteq X^1 \subseteq \cdots \subseteq X$$

such that X^k is obtained from X^{k-1} by attaching equivariant cells of the form $D^k \times G/H$ along their boundaries, where H is any finite subgroup of G. See [**Lüc05**, Section 1] for more details.

The Baum-Connes conjecture as formulated in [**BCH94**] involves *universal proper* G-*spaces*. In the context of proper G-CW-complexes, these may be characterized as follows:

2.1. THEOREM ([**Lüc05**, Theorem 1.9]). *There is a proper* G-CW*-complex* $\underline{E}G$ *with the property that if* Y *is any proper* G-CW*-complex, then there is a* G*-equivariant continuous map from* Y *into* $\underline{E}G$*, and moreover this map is unique up to equivariant homotopy.*

Clearly the G-CW-complex $\underline{E}G$ is unique up to equivariant homotopy. The universal space used in the formulation of the Baum-Connes conjecture is defined a bit differently (see [**BCH94**, Section 1]), but by the results of [**Lüc05**, Section 2] the same conjecture results if the above version of $\underline{E}G$ is used. Compare also Theorem 2.3 below.

A proper G-CW-complex is said to be G-*finite* if only finitely many equivariant cells are used in its construction. These are G-compact proper G-spaces in the sense of the following definition.

2.2. DEFINITION. We shall say that a G-space X is a G-*compact, proper* G-*space* if

(a) X is locally compact and Hausdorff.
(b) The quotient space X/G is compact and Hausdorff in the quotient topology.
(c) Each point of X is contained in an equivariant neighborhood U that maps continuously and equivariantly onto some proper orbit space G/H (where H is a finite subgroup of G).

Apart from G-finite G-CW-complexes, we shall also be concerned with smooth manifolds (with smooth actions of G) that satisfy these conditions. We shall call them G-*compact proper* G-*manifolds*.

The following result is a consequence of [**Lüc05**, Theorem 3.7] (the final statement reflects a simple feature of the CW-topology on $\underline{E}G$).

2.3. THEOREM. *If* X *is any* G-*finite proper* G-*space, then there is a* G-*equivariant map from* X *to* $\underline{E}G$. *It is unique up to equivariant homotopy, and its image is contained within a* G-*finite subcomplex of* $\underline{E}G$.

3. Equivariant Geometric K-Homology

In this section we shall present the equivariant version of the geometric K-homology theory of Baum and Douglas [**BD82**]. The definition presents no difficulties, so we shall be brief. The reader is referred to [**BD82**] or [**BHS07**] for treatments of the non-equivariant theory.

We shall work with principal bundles, rather than with spinor bundles as in [**BD82**] or [**BHS07**]. To fix notation, recall the following rudimentary facts about Clifford algebras, Spin^c-groups and Spin^c-structures. Denote by $\mathrm{Cliff}(n)$ the $\mathbb{Z}/2\mathbb{Z}$-graded complex $*$-algebra generated by skew-adjoint degree-one elements $e_1, \ldots, e_n$ such that

$$e_i e_j + e_j e_i = -2\delta_{ij} I.$$

We shall consider $\mathbb{R}^n$ as embedded into $\mathrm{Cliff}(n)$ in such a way that the standard basis of $\mathbb{R}^n$ is carried to $e_1, \ldots, e_n$.

Denote by $\mathrm{Spin}^c(n)$ the group of all even-grading-degree unitary elements in $\mathrm{Cliff}(n)$ that map $\mathbb{R}^n$ into itself under the adjoint action. This is a compact Lie group. The image of the group homomorphism

$$\alpha\colon \mathrm{Spin}^c(n) \longrightarrow GL(n, \mathbb{R})$$

given by the adjoint action is $SO(n)$ and the kernel is the circle group $U(1)$ of all unitaries in $\mathrm{Cliff}(n)$ that are multiples of the identity element.

There is a natural complex conjugation operation on $\mathrm{Cliff}(n)$ (since the relations defining the Clifford algebra involve only real coefficients) and the map $u \mapsto u\bar{u}^*$ is a homomorphism from $\mathrm{Spin}^c(n)$ onto $U(1)$. The combined homomorphism

$$\mathrm{Spin}^c(n) \longrightarrow SO(n) \times U(1)$$

is a double covering.

Let M be a smooth, proper G-manifold and let V be a smooth, real G-vector bundle over M of rank n. A G-Spin^c-*structure* on V is a homotopy class of reductions of the principal frame bundle of V (viewed as a G-equivariant right principal $GL(n, \mathbb{R})$-bundle) to a G-equivariant principal $\mathrm{Spin}^c(n)$-bundle. In other words it is a homotopy class of commutative diagrams

$$\begin{tikzcd} Q \arrow[r, "\varphi"] \arrow[d] & P \arrow[d] \\ M \arrow[r, equal] & M \end{tikzcd}$$

of smooth G-manifolds, where P is the bundle of ordered bases for the fibers of V, Q is a G-equivariant principal $\mathrm{Spin}^c(n)$-bundle, and

$$\varphi(qu) = \varphi(q)\alpha(u)$$

for every $q \in Q$ and every $u \in \mathrm{Spin}^c(n)$. A G-Spin^c-structure on V determines a G-invariant orientation, and a specific choice of Q within its homotopy class determines a Euclidean structure.

3.1. EXAMPLE. Every G-equivariant complex vector bundle carries a natural Spin^c-structure because there is a (unique) group homomorphism $U(k) \to \mathrm{Spin}^c(2k)$ that lifts the map

$$U(k) \longrightarrow SO(2k) \times U(1)$$

given in the right-hand factor by the determinant.

A G-Spin^c-*vector bundle* is a smooth real G-vector bundle with a given G-Spin^c-structure. The direct sum of two G-Spin^c-vector bundles carries a natural G-Spin^c-structure. This is obtained from the diagram

$$\begin{tikzcd} \mathrm{Spin}^c(m) \times \mathrm{Spin}^c(n) \arrow[r] \arrow[d] & \mathrm{Spin}^c(m+n) \arrow[d] \\ GL(n,\mathbb{R}) \times GL(m,\mathbb{R}) \arrow[r] & GL(m+n,\mathbb{R}) \end{tikzcd}$$

that is in turn obtained from the inclusions of $\mathrm{Cliff}(m)$ and $\mathrm{Cliff}(n)$ into $\mathrm{Cliff}(m+n)$ given by the formulas $e_k \mapsto e_k$ and $e_k \mapsto e_{m+k}$, respectively. In addition, if V and $V \oplus W$ carry Spin^c-structures, then there is a unique Spin^c-structure on W whose direct sum, as above, with the Spin^c-structure on V is the given Spin^c-structure on the direct sum. This is the *two out of three principle* for Spin^c-structures.

If M is a smooth G-manifold, then a G-Spin^c-*structure* on M is a G-Spin^c-structure on its tangent bundle, and a G-Spin^c-*manifold* is a smooth G-manifold together with a given G-Spin^c-structure.

3.2. DEFINITION. Let X be any G-space. An *equivariant* K-*cycle* for X is a triple (M, E, f) consisting of:

(a) A G-compact, proper G-Spin^c-manifold M without boundary.[1]
(b) A smooth complex G-vector bundle E over M.
(c) A continuous and G-equivariant map $f\colon M \to X$.

The geometric equivariant K-homology groups $K_*^G(X)$ will be obtained by placing a certain equivalence relation on the class of all equivariant K-cycles. Before describing it, we give constructions at the level of cycles that will give the arithmetic structure of the groups $K_*^G(X)$.

[1]The manifold M need not be connected. Moreover different connected components of M may have different dimensions.

If (M, E, f) and (M', E', f') are two equivariant K-cycles for X, then their *disjoint union* is the equivariant K-cycle $(M \sqcup M', E \sqcup E', f \sqcup f')$. The operation of disjoint union will give addition.

Let V be a G-Spin^c-vector bundle with a G-Spin^c-structure $\varphi \colon Q \to P$. Fix an orientation-reversing isometry of $\mathbb{R}^n$. Since it preserves the inner product, τ induces an automorphism of Cliff(n), and hence of Spin$^c(n)$, that we shall also denote by τ. Consider the map $\varphi_\tau \colon Q_\tau \to P$, where:

(a) Q_τ is equal to Q as a G-manifold, but has the twisted action $q \cdot_\tau u = q \cdot \tau(u)$ of the group Spin$^c(n)$.
(b) $\alpha_\tau(q) = \alpha(u)\tau$.

It defines the *opposite* G-Spin^c-vector bundle $-V$. Applying this to manifolds, we define the *opposite* of an equivariant K-cycle (M, E, f) to be the equivariant K-cycle $(-M, E, f)$. This will give the operation of additive inverse in the geometric groups $K_*^G(X)$.

If M is a Spin^c-G-manifold, then its boundary ∂M inherits a Spin^c-G-structure. This is obtained from the pullback diagram

$$\begin{tikzcd}
\mathrm{Spin}^c(n{-}1) \arrow[r] \arrow[d] & \mathrm{Spin}^c(n) \arrow[d] \\
GL(n{-}1, \mathbb{R}) \arrow[r] & GL(n, \mathbb{R})
\end{tikzcd}$$

associated to the lower-right-corner embedding of $GL(n{-}1, \mathbb{R})$ into $GL(n, \mathbb{R})$ and the inclusion of Cliff$(n{-}1)$ into Cliff(n) that maps the generators e_k to e_{k+1}. Using the outward-pointing normal first convention, the bundle of frames for $T\partial M$ maps to the restriction to the boundary of the bundle of frames for TM. A pullback construction gives the required reduction to Spin$^c(n{-}1)$.

3.3. DEFINITION. An equivariant K-cycle for X is a *boundary* if there is a G-compact, proper G-Spin^c-manifold W with boundary, a smooth, Hermitian equivariant vector bundle E over W and a continuous equivariant map $f \colon W \to X$ such that the given cycle is isomorphic to $(\partial W, E|_{\partial W}, f|_{\partial W})$. Two equivariant K-cycles for X, (M_1, E_1, f_1) and (M_2, E_2, f_2) are *bordant* if the disjoint union of one with the opposite of the other is a boundary.

The most subtle aspect of the equivalence relation on equivariant K-cycles that defines geometric K-homology involves certain sphere bundles over Spin^c-manifolds. To describe it we begin by considering a single sphere.

View S^{n-1} as the boundary of the unit ball in $\mathbb{R}^n$. The frame bundle for $\mathbb{R}^n$ can of course be identified with $\mathbb{R}^n \times GL(n, \mathbb{R})$ since the columns of any invertible matrix constitute a frame for $\mathbb{R}^n$. We can therefore equip $\mathbb{R}^n$ with the *trivial* Spin^c-structure $\mathbb{R}^n \times \mathrm{Spin}^c(n)$.

According to the prescription given prior to Definition 3.3, the associated Spin^c-structure on the sphere S^{n-1} is given by the right principal Spin$^c(n{-}1)$-bundle Q whose fiber at $v \in S^{n-1}$ is the space of all elements $u \in \mathrm{Spin}^c(n)$ whose image in $SO(n)$ is a matrix with first column equal to v. Observe that Q is Spin$^c(n)$-equivariant for the left action of Spin$^c(n)$ on the sphere given by the projection to $SO(n)$.

Let us now assume that $n = 2k{+}1$. We are going to fix a certain Spin$^c(n)$-equivariant complex vector bundle F on S^{2k}. The key property of F is that the Spin$^c(n)$-equivariant index of the Dirac operator (discussed in the next section) coupled to F is equal to the rank-one trivial representation of Spin$^c(n)$. An explicit calculation, given in [**BHS07**], shows that the dual of the positive part of the spinor bundle for S^{n-1} has the required property. It follows easily from the Bott periodicity theorem that F is essentially unique (up to addition of trivial bundles, any two F are isomorphic). For what follows, any choice of F will do. The bundle F and the trivial line bundle together generate $K(S^{2k})$, and for that reason we shall call it the *Bott generator*.

Following these preliminaries, we can describe the "vector bundle modification" step in the equivalence relation defining geometric K-homology.

Let V be a G-Spin^c-vector bundle of rank $2k$ over a G-Spin^c-manifold M and denote by $\widehat{M}$ the sphere bundle[2] of the direct sum vector bundle $\mathbb{R} \oplus V$. The manifold $\widehat{M}$ may be described as the fiber bundle

$$\widehat{M} = Q \times_{\mathrm{Spin}^c(2k+1)} S^{2k},$$

where Q is the principal G-Spin$^c(2k+1)$-bundle associated to $\mathbb{R} \oplus V$. Its tangent bundle is isomorphic to the pullback of the tangent bundle of M, direct sum the fiberwise tangent bundle $Q \times_{\mathrm{Spin}^c(2k+1)} TS^{2k}$. Both carry natural G-Spin^c-structures, and so $\widehat{M}$ is a G-Spin^c-manifold.

Form the G-equivariant complex vector bundle

$$Q \times_{\mathrm{Spin}^c(2k+1)} F,$$

from the Bott generator discussed above. We shall use the same symbol F for this bundle over $\widehat{M}$.

3.4. DEFINITION. Let (M, E, f) be an equivariant K-cycle and let V be a rank $2k$ G-Spin^c-vector bundle over M. The *modification* of (M, E, f) associated to V is the equivariant K-cycle

$$(M, E, f)^{\wedge} = \big(\widehat{M}, F \otimes \pi^*(E), f \circ \pi\big),$$

where:

(a) $\widehat{M}$ is the total space of the sphere bundle of $\mathbb{R}\oplus V$, equipped with the G-Spin^c-structure described above;
(b) π is the projection from $\widehat{M}$ onto M; and
(c) F is the G-equivariant complex vector bundle on $\widehat{M}$ described above.

We are now ready to define the geometric equivariant K-homology groups.

3.5. DEFINITION. Denote by $K^G(X)$ the set of equivalence classes of equivariant K-cycles over X, for the equivalence relation generated by the following three elementary relations:

(a) If (M, E_1, f) and (M, E_2, f) are two equivariant K-cycles with the same proper, G-compact G-Spin^c-manifold M and same map $f\colon M \to X$, then

$$(M \sqcup M, E_1 \sqcup E_2, f \sqcup f) \sim (M, E_1 \oplus E_2, f).$$

(b) If (M_1, E_1, f_1) and (M_2, E_2, f_2) are bordant equivariant K-cycles then

$$(M_1, E_1, f_1) \sim (M_2, E_2, f_2).$$

(c) If (M, E, f) is an equivariant K-cycle, if V is an even-rank G-Spin^c-vector bundle over M, and if $(M, E, f)^{\wedge}$ is the modification of (M, E, f) associated to V, then

$$(M, E, f) \sim (M, E, f)^{\wedge}.$$

[2] Strictly speaking, to form the sphere bundle we need a metric on V and so a specific choice of principal bundle Q within its homotopy class. Of course, any two sphere bundles will be bordant.

The set $K^G(X)$ is an abelian group with addition given by disjoint union. Denote by $K^G_{ev}(X)$ and $K^G_{odd}(X)$ the subgroups of $K^G(X)$ composed of equivalence classes of equivariant K-cycles (M, E, f) for which every connected component of M is even-dimensional or odd-dimensional, respectively. Then $K^G(X) \cong K^G_{ev}(X) \oplus K^G_{odd}(X)$.

4. Equivariant Kasparov Theory

In this section we shall define a natural transformation from geometric equivariant K-homology to Kasparov's equivariant K-homology. Once again, this is a straightforward extension to the equivariant context of the Baum-Douglas theory that was reviewed in detail already in the paper [**BHS07**]. Therefore we shall be brief.

Fix a *second countable* G-compact proper G-space X, for example a G-finite proper G-CW-complex. The second countability assumption is made for consistency with Kasparov's theory, which applies to second countable locally compact spaces, or separable C^*-algebras.

We shall denote by $KK^G_n(C_0(X), \mathbb{C})$ the Kasparov group $KK^G\big(C_0(X), \mathrm{Cliff}(n)\big)$ (the action of G on $\mathrm{Cliff}(n)$ is trivial). See [**Kas88**, Section 2]. There are canonical isomorphisms

$$KK^G_n(C_0(X), \mathbb{C}) \cong KK^G_{n+2}(C_0(X), \mathbb{C})$$

coming from the periodicity of Clifford algebras. Compare [**BHS07**]. As a result we may form the 2-periodic groups $KK^G_{ev\,/\,odd}(C_0(X), \mathbb{C})$.

The natural transformation

$$\mu\colon K^G_{ev\,/\,odd}(X) \longrightarrow KK^G_{ev\,/\,odd}(C_0(X), \mathbb{C})$$

into Kasparov theory is defined by associating to an equivariant K-cycle (M, E, f) a Dirac operator, and then constructing from the Dirac operator a cycle for Kasparov's analytic K-homology group.

The vector space $\mathrm{Cliff}(n)$ carries a natural inner product in which the monomials $e_{i_1} \cdots e_{i_k}$ form an orthonormal basis. If M is a G-compact, proper G-Spin^c-manifol, and if Q is a lifting to $\mathrm{Spin}^c(n)$ of the frame bundle of M, then the $\mathbb{Z}/2\mathbb{Z}$-graded Hermitian vector bundle

$$S = Q \times_{\mathrm{Spin}(n)} \mathrm{Cliff}(n)$$

that is formed using the left multiplication action of $\mathrm{Spin}^c(n)$ on $\mathrm{Cliff}(n)$ carries a right action of the algebra $\mathrm{Cliff}(n)$ and a commuting left action of TM as odd-graded skew-adjoint endomorphisms such that $v^2 = -\|v\|^2 I$. This is called the action of TM on the *spinor bundle* S by *Clifford multiplication*.

4.1. REMARK. There are other versions of the spinor bundle that do not carry a right Clifford algebra action. The bundle used here has the advantage of allowing a uniform treatment of both even- and odd-rank bundles V. In addition the real case may be treated similarly (although we shall not consider it in this paper).

4.2. DEFINITION. Let M be a G-compact proper G-Spin^c-manifold. Fix an associated principal Spin^c-bundle over M, and let S be the spinor bundle, as above. Let E be a smooth, Hermitian G-vector bundle over M. We shall call an odd-graded, symmetric, order one linear partial differential operator D acting on the sections of $S \otimes E$ a *Dirac operator* if it commutes with the right Clifford algebra action on the spinor bundle and if

$$[D, f]u = \mathrm{grad}(f) \cdot u,$$

for every smooth function f on M and every section u of $S \otimes E$, where $\mathrm{grad}(f) \cdot u$ denotes Clifford multiplication on S by the gradient of f.

Dirac operators in this sense always exist, and basic PDE theory gives the following result:

4.3. PROPOSITION. *The Dirac operator* D, *considered as an unbounded operator on* $L^2(M, S \otimes E)$ *with domain the smooth compactly supported sections, is essentially self-adjoint. The bounded Hilbert space operator* $F = D(I + D^2)^{-1/2}$ *commutes, modulo compact operators, with multiplication operators from* $C_0(M)$. *Moreover the product of* $I - F^2$ *with any multiplication operator from* $C_0(M)$ *is a compact operator.* □

Now the Hilbert space $L^2(M, S \otimes E)$ carries a right action of Cliff(n) that commutes with D and the action of $C_0(M)$. It also carries a unique Cliff(n)-valued inner product $\langle\ ,\ \rangle_{\mathrm{Cliff}}$ such that

$$\langle s_1, s_2\rangle = \tau\big(\langle s_1, s_2\rangle_{\mathrm{Cliff}}\big)$$

where on the left is the L^2-inner product, and on the right is the state τ on Cliff(n) that maps all nontrivial monomials $e_{i_1} \cdots e_{i_p}$ to zero. Using it we place a Hilbert Cliff(n)-module structure on $L^2(M, S \otimes E)$.

Proposition 4.3 implies that the operator $F = D(I + D^2)^{-1/2}$, viewed as an operator on the Hilbert Cliff(n)-module $L^2(M, S \otimes E)$, yields a cycle for Kasparov's equivariant KK-group $KK^G(C_0(M), \mathrm{Cliff}(n))$ (see [**Kas88**, Definition 2.2]).

4.4. DEFINITION. We shall denote by $[M, E] \in KK^G_n(C_0(M), \mathbb{C})$ the KK-class of the operator $F = D(I + D^2)^{-1/2}$.

The first main theorem concerning the classes $[M, E]$ is as follows:

4.5. THEOREM. *The correspondence that associates to each equivariant* K*-cycle* (M, E, f) *the* KK*-class*

$$f_*[M, E] \in KK^G_{\mathrm{ev}\,/\,\mathrm{odd}}(C_0(X), \mathbb{C})$$

gives a well-defined homomorphism

$$\mu\colon K^G_{\mathrm{ev}\,/\,\mathrm{odd}}(X) \longrightarrow KK^G_{\mathrm{ev}\,/\,\mathrm{odd}}(C_0(X), \mathbb{C}).$$

The non-equivariant case of the theorem is proved in [**BHS07**]. The proof for the equivariant case is exactly the same and therefore will not be repeated.

Our aim in this paper is to prove the second main theorem concerning the classes $[M, E]$.

4.6. THEOREM. *If* X *is any proper,* G*-finite* G*-CW-complex, then the index map*

$$\mu\colon K^G_{\mathrm{ev}\,/\,\mathrm{odd}}(X) \longrightarrow KK^G_{\mathrm{ev}\,/\,\mathrm{odd}}(C_0(X), \mathbb{C})$$

is an isomorphism.

The non-equivariant version of the theorem is due to Baum and Douglas, and is proved in detail in [**BHS07**]. Although the proof of the equivariant result is the same in outline, new issues must also be resolved having to do with the properties of equivariant vector bundles on G-compact proper G-spaces. These we shall consider next.

5. Equivariant Vector Bundles

Throughout this section we shall use the term G-*bundle* as an abbreviation for G-*equivariant complex vector bundle.* We shall review the basic theory of G-bundles over G-compact proper G-spaces, mostly as worked out by Lück and Oliver in [**LO01**]. In the next section we shall recast their results in the language of C^*-algebra K-theory.

5.1. THEOREM. *Let X be a G-compact, proper G-space. There is a G-bundle E over X such that for every $x \in X$, the fiber E_x is a multiple of the regular representation of the isotropy group G_x.* □

PROOF. This is proved for G-finite proper G-CW-complexes in [**LO01**, Corollary 2.8]. That result extends to more general X by pulling back along the map supplied by Theorem 2.3. □

5.2. COROLLARY (Compare [**LO01**, Lemma 3.8]). *Let Z be a G-compact proper G-space and let X be a closed, G-invariant subset of Z. If F is any G-bundle on X, then there is a G-bundle E on Z such that F embeds as a summand of $E|_X$.*

PROOF. Fix a G-bundle E on Z, as in Theorem 5.1. There are G-invariant open subsets $U_1, \dots, U_n$ of X such that:

(a) The sets cover X.
(b) For each j there is a finite subgroup of $F_j \subseteq G$ and an equivariant map $\pi_j \colon U_j \to G/F_j$.
(c) $F|_{X\cap U_j}$ is isomorphic to a bundle pulled back along π_j.
(d) $E|_{U_j}$ is also isomorphic to a bundle pulled back along π_j.

Replacing E by a direct sum $E \oplus \cdots \oplus E$, if necessary, we find that $F|_{U_j}$ may be embedded as a summand of $E|_{U_j}$, for every j. Making a second replacement of E by an n-fold direct sum $E \oplus \cdots \oplus E$ and using a standard partition of unity argument, we may now embed F into E, as required. □

More generally, if $f \colon X \to Z$ is a map between G-compact proper G-spaces, and if F is a G-bundle on X, then the same argument shows that F is isomorphic to a summand of the pullback along f of some G-bundle on Z.

5.3. DEFINITION. If S is any set, then denote by $\mathbb{C}[S]$ the free vector space on the set S, equipped with the standard inner product in which the elements of S are orthonormal. If S is equipped with an action of G, then we shall consider $\mathbb{C}[S]$ to be equipped with the corresponding permutation action of G.

We are interested primarily in the case where $S = G$, which we shall view as equipped with the usual left translation action of G.

5.4. DEFINITION. A *standard G-bundle* on a G-compact proper G-space X is a G-invariant subset E of $X\times\mathbb{C}[G]$ with the property that for every compact subset $K \subseteq X$ there is a *finite* subset $S \subseteq G$ such that the intersection of E with $K\times\mathbb{C}[G]$ is a (nonequivariant) complex vector subbundle of $K \times \mathbb{C}[S]$.

5.5. REMARK. We require that the restriction of E to K, as above, be a topological vector subbundle of the finite-dimensional trivial bundle $K\times\mathbb{C}[S]$. This fixes the topology on E and determines a G-bundle structure.

It follows from a standard partition of unity argument that every G-bundle on X is isomorphic to a standard G-bundle. We are going to prove the following result, which gives the set of standard G-bundles a useful directed set structure.

5.6. THEOREM. *Any two standard G-bundles are subbundles of a common third. Moreover the union of all standard G-bundles is $X\times\mathbb{C}[G]$.*

5.7. REMARK. In Section 8 we shall modify Definition 5.4 very slightly by replacing G with a countable disjoint union $G_\infty = G \sqcup G \sqcup \cdots$ (thought of as a left G-set). Theorem 5.6 remains true, with the same proof.

Since the theorem is obvious if G is finite, we shall assume G is infinite until the proof of Theorem 5.6 is concluded.

5.8. LEMMA. *If* S *is a finite subset of* G, *if* K *is a compact subset of* X, *and if* E *is any* G*-bundle over* X, *then there is a standard* G*-bundle* E_1 *that is isomorphic to* E *and whose restriction to* K *is orthogonal to* $K \times \mathbb{C}[S]$.

PROOF. Let E_1 be any standard bundle that is isomorphic to the equivariant vector bundle E. If $g \in G$, then the set

$$E_1 \cdot g = \{ (x, e \cdot g) : e \in E_{1,x} \}$$

is also a standard G-bundle. Here, in forming the vectors $e \cdot g$ we are using the *right* translation action of G on itself and hence on $\mathbb{C}[G]$. The bundle $g \cdot E_1$ is isomorphic to E_1, and hence to E. If we enlarge S, if necessary, so that $E_1|_K \subseteq K \times \mathbb{C}[S]$, and if we choose $g \in G$ so that $S \cap Sg = \emptyset$, then $(E_1 \cdot g)|_K$ is orthogonal to $K \times \mathbb{C}[S]$, as required. □

5.9. LEMMA. *Let* E *be a* G*-bundle on* X *and let* E_2 *be a standard* G*-bundle. There is a standard* G*-bundle that is isomorphic to* E *and orthogonal to* E_2.

PROOF. Let K be a compact subset of X whose G-saturation is X, and let S be a finite subset of G such that $E_2|_K \subseteq K \times \mathbb{C}[S]$. Now apply the previous lemma. □

5.10. LEMMA. *Let* U *be a* G*-invariant open subset of* X, *and let* Y *and* Z *be* G*-invariant closed subsets of* X *such that*

$$Z \subseteq U \subseteq Y \subseteq X.$$

Let F *be a standard* G*-bundle over* Y. *There is a standard* G*-bundle* E_1 *over* X *such that* $F|_Z \subseteq E_1|_Z$. *Moreover, given a standard* G*-bundle* E_2 *over* X *such that* $E_2|_Y$ *is orthogonal to* $F|_Y$, *the standard* G*-bundle* E_1 *may be chosen to be orthogonal to* E_2.

PROOF. According to Theorem 5.1, there is a G-bundle E over X such that F embeds in $E|_Y$. Any complement of F in $E|_Y$ may be embedded as a standard G-bundle F' on Y that is orthogonal to F, and after replacing F with $F \oplus F'$ we may assume that in fact there is a G-bundle E on X such that $E|_Y \cong F$.

Fix such an isomorphism $\Phi\colon E|_Y \to F$. Next, there is an embedding Ψ of E as a standard G-bundle on X such that $\Phi[E]$ is orthogonal to E_2 and $\Phi[E]|_Y$ is orthogonal to F (by a slight elaboration of Lemma 5.9). If we choose a G-invariant scalar function φ on X such that $\varphi = 1$ on Z and $\varphi = 0$ outside of U, and if we set $\psi = 1 - \varphi$, then $E_1 = (\varphi\Phi + \psi\Psi)[E]$ has the required properties. □

5.11. LEMMA. *Let* K *be any compact subset of* X, *and let* S *be a finite subset of* G. *There is a standard* G*-bundle that contains* $K \times \mathbb{C}[S]$.

PROOF. The compact set K may be written as a finite union of compact sets

$$K = K_1 \cup \cdots \cup K_n$$

where each K_j is included in a G-invariant open set that maps equivariantly onto some proper coset space G/H_j, in such a way that K_j maps to the identity coset. We shall use induction on n.

Let E_2 be a standard G-bundle that contains the set

$$\big(K_1 \cup \cdots \cup K_{n-1}\big) \times \mathbb{C}[S].$$

There is a G-compact subset $Y \subseteq X$ that contains a G-invariant neighborhood U of K_n and over which there is a standard G-bundle L such that

$$K_n \times \mathbb{C}[S] \subseteq L|_{K_n} \quad \text{and} \quad E_2|_Y \subseteq L.$$

Indeed we may choose Y so that it maps equivariantly to G/H_n, and if Y_n is the inverse image of the identity coset, then we may form

$$L = \bigcup_{g \in G} gY_n \times \mathbb{C}[Sg_n],$$

where S_n is a sufficiently large finite and H_n-invariant subset of G.

Now apply the previous lemma to the standard G-bundle $F = L \ominus E_2|_Y$ (the orthogonal complement of $E|_Y$ in L) to obtain a standard G-bundle E_1 on X such that E_1 is orthogonal to E_2 and

$$L|_{K_n} \ominus E_2|_{K_n} \subseteq E_1|_{K_n}.$$

The standard G-bundle $E_1 \oplus E_2$ then contains $K \times \mathbb{C}[S]$, as required. $\square$

PROOF OF THEOREM 5.6. Since there is a standard G-bundle that contains any given $K \times \mathbb{C}[S]$, it is clear that the union of all standard G-bundles is $X \times \mathbb{C}[G]$. Let E_1 and E_2 be standard G-bundles on X. Choose a compact set K whose G-translates cover X and choose a finite set $S \subseteq G$ such that $E_1|_K, E_2|_K \subseteq K \times \mathbb{C}[S]$. If E is a standard G-bundle containing $K \times \mathbb{C}[S]$, then it contains E_1 and E_2. $\square$

6. C*-Algebras and Equivariant K-Theory

6.1. DEFINITION. If S is any set, then denote by $M[S]$ the $*$-algebra of complex matrices $[T_{s_1 s_2}]$ with rows and columns parametrized by the set S, all but finitely many of whose entries are zero.

We shall be interested in the case where $S = G$. In this case the group G acts on $M[G]$ by automorphisms via the formula $(g \cdot T)_{g_1, g_2} = T_{g^{-1} g_1, g^{-1} g_2}$.

6.2. DEFINITION. Let X be a G-compact proper G-space. Let us call a function $F\colon X \to M[G]$ *standard* if its matrix element functions

$$F_{g_1, g_2} \colon x \mapsto F(x)_{g_1, g_2}$$

are continuous and compactly supported, and if for every compact subset K of X all but finitely many of them vanish outside of K. We shall denote by $\mathcal{C}(X, G)$ the $*$-algebra of all standard, G-equivariant functions from X to $M[G]$.

Note that if P is a projection in the $*$-algebra $\mathcal{C}(X, G)$, then the range of P (that is, the bundle over X whose fiber over $x \in X$ is the range of the projection operator $P(x)$ in $\mathbb{C}[G]$) is a standard G-bundle in the sense of Section 5. In fact every standard G-bundle is obtained in this way, which explains our interest in $\mathcal{C}(X, G)$. In fact we are even more interested in the following C^*-algebra completion of $\mathcal{C}(X, G)$.

6.3. DEFINITION. Let X be a G-compact proper G-space. Denote by $C^*(X, G)$ the C^*-algebra of G-equivariant, continuous functions from X into the compact operators on $\ell^2(G)$.

6.4. REMARK. The C^*-algebra $C^*(X, G)$ is isomorphic to the crossed product C^*-algebra $C_0(X) \rtimes G$. If E^{g_1, g_2} denotes the matrix with 1 in entry (g_1, g_2) and zero in every other entry, then the formula

$$f\,[g] \mapsto \sum_h h(f) E^{h, hg}$$

gives an isomorphism from $C_0(X) \rtimes G$ to $C^*(X, G)$, and the formula

$$F \mapsto \sum_{g \in G} F_{e,g}\,[g]$$

gives its inverse. Since the action of G on X is proper, the maximal and reduced crossed products are equal. Indeed, there is a unique C^*-algebra completion of the $*$-algebra $\mathcal{C}(X, G)$.

6.5. LEMMA. *Assume that G is infinite and X is a G-compact proper G-space. The correspondence between projections in $\mathcal{C}(X, G)$ and their ranges induces bijections among the following sets:*

(a) *Equivalence classes of projections in $\mathcal{C}(X, G)$.*
(b) *Equivalence classes of projections in $C^*(X, G)$.*
(c) *Isomorphism classes of standard G-bundles on X.*
(d) *Isomorphism classes of Hermitian G-bundles on X.*

If X is a G-compact proper G-manifold, then there is in addition a bijection with

(e) *Isomorphism classes of smooth Hermitian G-bundles on X.*

PROOF. Recall that two projections P and Q in a $*$-algebra are equivalent if and only if there is an element U such that $U^*U = P$ and $UU^* = Q$. The inclusion of $\mathcal{C}(X, G)$ into $C^*(X, G)$ is a simple example of a *holomorphically closed subalgebra*, and as a result the inclusion induces a bijection between the sets in (a) and (b). Compare [**Bla98**, Sections 3 and 4]. The sets in (a) and (c) are in bijective correspondence virtually by definition. The sets in (c) and (d) are in bijection thanks to Lemma 5.8, which in particular shows that every G-bundle is isomorphic to a standard bundle, if G is infinite.

If X is a manifold, then the inclusion of the smooth functions in $\mathcal{C}(X, G)$ into $C^*(X, G)$ is also a holomorphically closed subalgebra, and this gives the final part of the lemma since equivalence classes projections in the algebra of smooth functions correspond to isomorphism classes of (the obvious concept of) smooth standard G-bundles. □

6.6. REMARK. If G is finite, then the lemma remains true if $\mathcal{C}(X, G)$ and $C^*(X, G)$ are replaced by direct limits of matrix algebras over themselves.

6.7. THEOREM. *The C^*-algebra $C^*(X, G)$ has an approximate identity consisting of projections.*

PROOF. We claim that for every finite set of elements $F_1, \ldots, F_n$ in $\mathcal{C}(X, G)$ there is a projection P in $\mathcal{C}(X, G)$ such that

$$F_j = PF_j = F_jP$$

for all $j = 1, \ldots, n$. Indeed, the orthogonal projection onto any standard G-bundle is a projection in $\mathcal{C}(X, G)$ (and conversely). So the claim follows from Theorem 5.6, and the theorem follows since $\mathcal{C}(X, G)$ is dense in $C^*(X, G)$. □

6.8. THEOREM. *Let X be a G-compact proper G-space. The bijections in Lemma 6.5 determine a natural isomorphism between the Grothendieck group of G-bundles on X and the K_0-group of the C^*-algebra $C^*(X, G)$.*

PROOF. If A is any C^*-algebra with an approximate unit consisting of projections, then the natural map from the Grothedieck group of projections in matrix algebras over A into $K_0(A)$ is an isomorphism. So the theorem follows from the previous result. □

6.9. DEFINITION. If X is a G-compact proper G-space, and $j \in \mathbb{Z}/2\mathbb{Z}$, then denote by $K^j_G(X)$ the K_j-group of the C^*-algebra $C^*(X,G)$. If Y is a G-invariant closed subset of X, then denote by $K^j_G(X,Y)$ the K_0-group of the ideal in $C^*(X,G)$ consisting of functions that vanish on Y.

By the above, $K^0_G(X)$ is the Grothendieck group of isomorphism classes of G-bundles on X.

The relative groups $K^j_G(X,Y)$ satisfy excision (of the strongest possible type, that $K^j_G(X,Y)$ depends only on $X \setminus Y$). Elementary K-theory provides functorial coboundary maps

$$\partial \colon K^j_G(Y) \longrightarrow K^{j+1}_G(X,Y),$$

and these give the groups $K^j_G(X,Y)$ the structure of a $\mathbb{Z}/2\mathbb{Z}$-graded cohomology theory on G-compact proper G-spaces, in the sense that that they fit into functorial long exact sequences

$$\cdots \longrightarrow K^j_G(X) \longrightarrow K^j_G(Y) \longrightarrow K^{j+1}_G(X,Y) \longrightarrow K^{j+1}_G(X) \longrightarrow \cdots$$

Although we have accessed this fact using C^*-algebra K-theory, this is also the main result of [**LO01**].

We conclude by reviewing the Gysin maps in equivariant K-theory that we shall need in the next section. Let E be a complex G-bundle over a G-compact proper G-manifold M. As we noted earlier, E carries a canonical Spin^c-structure. Form the sphere bundle $\widehat{M}$ of the real bundle $\mathbb{R} \oplus E$, as in Section 3. The manifold M is equivariantly embedded as a retract in $\widehat{M}$ using the section

$$M \ni m \mapsto (1,0) \in \mathbb{R} \oplus E_m,$$

and associated to the embedding is a short exact sequence of K-theory groups

$$0 \longrightarrow K^0_G(\widehat{M}, M) \longrightarrow K^0_G(\widehat{M}) \longrightarrow K^0_G(M) \longrightarrow 0.$$

Let F be the complex G-bundle over $\widehat{M}$ that we defined in Section 3, and denote by F_0 the complex G-bundle obtained by restricting F to M, then pulling back the restriction to $\widehat{M}$ using the projection from $\widehat{M}$ down to M. The difference $[F]-[F_0]$ defines an element of $K^0_G(\widehat{M}, M)$. The relative group is a module over $K^0_G(M)$, and multiplication against $[F]-[F_0]$ gives the *Thom homomorphism*

$$K^0_G(M) \longrightarrow K^0_G(\widehat{M}, M).$$

The complement $\widehat{M} \setminus M$ identifies with the total space of E via the map

$$E \ni e \mapsto \tfrac{1}{1+\|e\|^2}(1,e) \in \mathbb{R} \oplus E.$$

If $\iota \colon M \to N$ is an embedding of G-compact proper G-manifolds, and if the normal bundle to the embedding is identified with E, then the *Gysin map*

$$\iota_! \colon K^0_G(M) \to K^0_G(N)$$

is the composition

$$K^0_G(M) \xrightarrow{\text{Thom}} K^0_G(\widehat{M}, M) \xrightarrow{\cong} K^0_G(N, N') \longrightarrow K^0_G(N),$$

where $N' \subseteq N$ is the complement of a tubular neighborhood and the middle map is given by excision and the identification of the tubular neighborhood with E.

The Gysin map is functorial for compositions of embeddings. It is well-defined for embeddings of manifolds with boundary, as long as the embedding is transverse to the boundary of N, and carries the boundary of M into the boundary of N.

7. The Technical Theory

In this section we shall construct the homology groups $k_*^G(X)$ that were described in the introduction. They are obtained as direct limits of certain bordism groups.

7.1. DEFINITION. Let Z be a proper G-space and let E be a G-bundle on Z. A *stable* (Z, E)-*manifold* is a G-compact proper G-manifold M (possibly with boundary) together with an equivalence class of pairs (h, φ), where:

(a) $h\colon M \to Z$ is a continuous and G-equivariant map.

(b) φ is an isomorphism of topological real G-bundles

$$\varphi\colon \mathbb{R}^r \oplus TM \longrightarrow \mathbb{R}^s \oplus h^*E,$$

for some $r, s \geq 0$. Here $\mathbb{R}^r$ and $\mathbb{R}^s$ denote the trivial bundles of ranks r and s (with trivial action of G on the fibers).

The equivalence relation is stable homotopy: (h_0, φ_0) and (h_1, φ_1) are equivalent if there is a homotopy $h\colon M \times [0,1] \to Z$ between h_0 and h_1 and an isomorphism of real G-bundles over $M \times [0,1]$

$$\varphi\colon \mathbb{R}^r \oplus TM \longrightarrow \mathbb{R}^s \oplus h^*E,$$

with $r \geq r_0, r_1$, that restricts to $I_{r-r_0} \oplus \varphi_0$ and $I_{r-r_1} \oplus \varphi_1$ at the two endpoints of $[0,1]$.

If M is a stable (Z, E)-manifold with boundary, then its boundary may be equipped with a stable (Z, E)-structure by forming the composition

$$\mathbb{R}^r \oplus \mathbb{R} \oplus T\partial M \xrightarrow[\cong]{} \mathbb{R}^r \times TM\big|_{\partial M} \xrightarrow[\varphi]{} \mathbb{R}^s \oplus f^*E,$$

in which $\mathbb{R} \oplus T\partial M$ is identified with $TM|_{\partial M}$ by the "exterior normal first" convention.

7.2. DEFINITION. Let X be a G-finite proper G-CW-complex and let Y be a G-subcomplex of X. For $j = 0, 1, 2, \ldots$ we define $\Omega_j^{(Z,E)}(X, Y)$ to be the group of equivalence classes of triples (M, a, f), where

(a) M is a smooth, proper, G-compact G-manifold of dimension j with a stable (Z, E)-structure.

(b) a is a class in the group $K_G^0(M)$.

(c) $f\colon M \to X$ is a continuous, G-equivariant map such that $f[\partial M] \subseteq Y$.

The equivalence relation is the obvious notion of bordism.

7.3. REMARK. Of course, the relation of bordism is arranged to incorporate the classes $a \in K_G^0(M)$, so that if (M, a, f) is the boundary of (W, b, g), then not only do we have that $M = \partial W$ and $f = g|_M$, but also the restriction map $K_G^0(W) \to K_G^0(M)$ takes b to a. One approach to the concept of bordism between manifolds with boundary is reviewed in [**BHS07**, Definition 5.5].

The sets $\Omega^{(Z,E)}(X, Y)$ are abelian groups. The group operation given by disjoint union and the additive inverse of (M, a, f) is $(-M, a, f)$. Here the *opposite* $-M$ is obtained by composing the bundle isomorphism φ with an orientation-reversing automorphism of the trivial bundle $\mathbb{R}^s$. (See Lemma 8.3 for another description of the inverse.)

The groups $\Omega_j^{(Z,E)}(X,Y)$ constitute a homology theory on G-finite proper G-CW-complexes. The boundary maps

$$\partial\colon \Omega_j^{(Z,E)}(X,Y) \longrightarrow \Omega_{j-1}^{(Z,E)}(Y)$$

take (M,a,f) to $(\partial M, a|_{\partial M}, f|_{\partial M})$. They fit into sequences

$$\cdots \longrightarrow \Omega_j^{(Z,E)}(X) \longrightarrow \Omega_j^{(Z,E)}(X,Y) \longrightarrow \Omega_{j-1}^{(Z,E)}(Y) \longrightarrow \Omega_{j-1}^{(Z,E)}(X) \longrightarrow \cdots$$

whose exactness follows from direct manipulations with cycles, using two facts. First if M_1 and M_2 are stable (Z,E)-manifolds, and if $\partial M_1 = -\partial M_2$, then there is a stable (Z,E)-structure on the manifold M obtained by joining M_1 and M_2 together along their common boundary that restricts to the given structures on M_1 and M_2. Second, if a_1 and a_2 are equivariant K-theory classes on M_1 and M_2 that restrict to a common class on the boundary, then there is a K-theory class on M that restricts to a_1 and a_2. This follows from the Mayer-Vietoris sequence for equivariant K-theory, and therefore from the results of Lück and Oliver reviewed in Sections 5 and 6. A map of pairs of G-CW-complexes $(X_1,Y_1) \to (X_2,Y_2)$ that is a homeomorphism from $X_1 \setminus Y_1$ to $X_2 \setminus Y_2$ induces an isomorphism on relative groups.

As they stand, the bordism groups $\Omega_j^{(Z,E)}(X,Y)$ are rather far from equivariant K-homology groups, most obviously because they are not 2-periodic. We shall obtain the technical groups $k_j^G(X,Y)$ by simultaneously forcing periodicity and removing dependence of the bordism groups on the pair (Z,E).

Let M be a stable (Z,E)-manifold with structure maps h and φ, as in Definition 7.1 and let F be a complex Hermitian G-bundle on Z of rank k. The pullback of F to M has a unique smooth structure, and so we may form the sphere bundle $S(\mathbb{R}\oplus h^*F)$, which is a G-compact proper G-manifold. It is also a stable $(Z, E\oplus F)$-manifold. Indeed, if $B(\mathbb{R}\oplus h^*F)$ is the unit ball bundle, and if p is the projection to M, then

$$TB(\mathbb{R}\oplus h^*F) \cong \mathbb{R}\oplus p^*T^*M \oplus h^*F$$

(once a complement to the vertical tangent bundle is chosen). So we obtain an isomorphism

$$\begin{aligned} \mathbb{R}^r \oplus TB(\mathbb{R}\oplus h^*F) &\cong \mathbb{R}^r \oplus \mathbb{R} \oplus p^*T^*M \oplus p^*h^*F \\ &\cong \mathbb{R}^s \oplus \mathbb{R} \oplus p^*h^*E \oplus p^*h^*F \end{aligned}$$

using the given stable (Z,E)-structure on M. We can then equip the sphere bundle with the stable $(Z, E\oplus F)$-structure it inherits as the boundary of the ball bundle.

Suppose now that (M,a,f) is a cycle for the bordism group $\Omega_j^{(Z,E)}(X,Y)$. We can form from it the cycle $(\widehat{M}, \iota_!(a), f\circ\pi)$ for the group $\Omega_{j+2k}^{(Z,E\oplus F)}(X,Y)$, where:

(a) $\widehat{M}$ is the sphere bundle for $\mathbb{R}\oplus h^*F$ with the stable $(Z,E\oplus F)$-structure just described.
(b) $\iota\colon M \to \widehat{M}$ is the inclusion of M into the sphere bundle given by the formula $m \mapsto (1,0) \in \mathbb{R}\oplus F_{h(m)}$ and $\iota_!\colon K_G^0(M) \to K_G^0(\widehat{M})$ is the Gysin map.
(c) π is the projection from $\widehat{M}$ to M.

Since this construction may also be carried out on bordisms between manifolds, we obtain a well-defined map on bordism classes.

7.4. DEFINITION. Let $k = \dim_{\mathbb{C}}(F)$. Denote by

$$\beta^F\colon \Omega_j^{(Z,E)}(X,Y) \longrightarrow \Omega_{j+2k}^{(Z,E\oplus F)}(X,Y)$$

the map determined by the above construction.

7.5. LEMMA. *If* F_1 *and* F_2 *are* G*-bundles on* Z *of ranks* k_1 *and* k_2 *respectively, then*

$$\beta^{F_2} \circ \beta^{F_1} = \beta^F \colon \Omega_j^{(Z,E)}(X,Y) \longrightarrow \Omega_{j+2k}^{(Z,E\oplus F)}(X,Y),$$

where $F = F_1 \oplus F_2$ *and* $k = k_1 + k_2$.

PROOF. Let $c = (M, a, f)$ be a cycle for the bordism group $\Omega_j^{(Z,E)}(X,Y)$. The manifolds $\widehat{M}_1$ and $\widehat{M}_2$ obtained from M by the modification processes underlying $\beta^{F_2} \circ \beta^{F_1}$ and β^F are fiber bundles over M whose fibers are the product of spheres in $\mathbb{R} \oplus F_1$ and $\mathbb{R} \oplus F_2$ in the first case and the sphere in $\mathbb{R} \oplus F$ in the second. The product embeds as a hypersurface in the unit ball of $\mathbb{R} \oplus F$, for example via the map

$$\big((s_1, f_1), (s_2, f_2)\big) \mapsto \tfrac{1}{6}\big((2 + s_1)s_2, f_1, (2 + s_1)f_2\big),$$

and we obtain from this construction a bordism $\widehat{W}$ between $\widehat{M}_1$ and $\widehat{M}_2$. The map

$$j\colon t \mapsto \big(\tfrac{1}{2}(1 + t), 0, 0\big)$$

embeds $M\times[0, 1]$ into $\widehat{W}$, transversely to the boundary of $\widehat{W}$, and on the boundary components of $M\times[0, 1]$ the embedding restricts to the given embeddings of M into $\widehat{M}_1$ and $\widehat{M}_2$. The class $a \in K_G^0(M)$ determines a class $a \in K_G^0(M\times[0, 1])$ by homotopy invariance, and the triple $(\widehat{W}, j_!(a), f \circ \pi)$ gives a bordism between the cycles representing $\beta^{F_2}(\beta^{F_1}(c))$ and $\beta^F(c)$, as required. □

Now fix a universal space $\underline{E}G$ as in Section 2. Let Z be a G-finite G-subcomplex of $\underline{E}G$. In order to cope with the contingency that G might be finite we shall modify the notion of G-standard bundle as advertised in Remark 5.7, so that standard G-bundles are now taken to be suitable subbundles of $Z \times \mathbb{C}[G_\infty]$.

Let E be a standard G-bundle over Z, as considered in Section 5. Form a partial order on the set of pairs (Z, E) by inclusion:

$$(Z_1, E_1) \le (Z_2, E_2) \quad \Leftrightarrow \quad Z_1 \subseteq Z_2 \quad \text{and} \quad E_1 \subseteq E_2|_{Z_1}.$$

According to the results of Section 5 this is a directed set.

7.6. DEFINITION. For $j \in \mathbb{Z}$ define the groups $k_j^G(X,Y)$ to be the direct limits

$$k_j^G(X,Y) = \varinjlim_{(Z,E)} \Omega_{j+2\,\mathrm{rank}(E)}^{(Z,E)}(X,Y)$$

over the directed set of all pairs (Z, E), as above.

8. Proof of the Main Theorem

We aim to prove Theorem 4.6, that the geometric equivariant K-homology groups of Section 3 are isomorphic to the analytic groups of Section 4. We shall do so by comparing the technical groups of the previous section first to equivariant KK-theory and then to geometric K-homology.

The equivariant KK-groups determine a homology theory on G-finite proper G-CW pairs (or indeed on arbitrary second-countable G-compact proper G-CW pairs) if one defines the relative groups for a pair (X, Y) to be $KK_j^G(C_0(X \setminus Y), \mathbb{C})$. The boundary maps are provided by the boundary maps of the six-term exact sequence in KK-theory.

If (M, a, f) is a cycle for $\Omega^{(Z,E)}(X, Y)$, then an element of the Kasparov group $KK^G_j(C_0(X \setminus Y), \mathbb{C})$ may be defined as follows. Form the Dirac operator D on the interior of M using the Spin^c-structure associated to the given stable (Z, E)-structure on M. It determines a class

$$[D] \in KK^G_j(C_0(M \setminus \partial M), \mathbb{C}).$$

For example we may equip $M \setminus \partial M)$ with a complete G-invariant Riemannian metric and then form a KK-class using $F = D(I + D^2)^{-\frac{1}{2}}$ as in Section 4 (it does not depend on the choice of metric). Compare [**HR00**, Ch. 10], where the non-equivariant case is handled; the G-compact proper G-manifold situation is the same. We can then form the Kasparov product $a \otimes [D] \in KK^G_j(C_0(X \setminus Y), \mathbb{C})$ and hence the class

$$f_*(a \otimes [D]) \in KK^G_j(C_0(X \setminus Y), \mathbb{C})$$

more or less exactly as we did in Section 4.

8.1. THEOREM. *The correspondence that associates to a cycle* (M, a, f) *the element* $f_*(a \otimes [D])$ *above is a natural transformation*

$$k^G_*(X, Y) \longrightarrow KK^G_*(C_0(X \setminus Y), \mathbb{C})$$

between homology theories. □

This is a mild elaboration of Theorem 4.5 and the equivariant counterpart of [**BHS07**, Theorem 6.1]. The equivariant case may be handled exactly as in [**BHS07**].

Our first goal to show that this natural transformation is an isomorphism on G-finite proper G-CW-complexes. To do so it suffices to show that it is an isomorphism on the "points" $X = G/H$ corresponding to finite subgroups H of G. The following observation clarifies what needs to be done in this case.

8.2. LEMMA. *Let* H *be a finite subgroup of* G. *There is a commutative diagram*

$$\begin{tikzcd} k^G_*(G/H) \arrow[r, "\mu"] & KK^G_*(C_0(G/H), \mathbb{C}) \\ k^H_*(pt) \arrow[u, "\mathrm{Ind}"] \arrow[r, "\mu"'] & KK^H_*(\mathbb{C}, \mathbb{C}) \arrow[u, "\mathrm{Ind}"'] \end{tikzcd}$$

in which the vertical maps are isomorphisms.

PROOF. The right-hand induction map is defined as follows. If $\mathcal{H}$ is a Hilbert space, or Hilbert module, with unitary H-action, define $\mathrm{Ind}\,\mathcal{H}$ to be the space of square-integrable sections of $G \times_H \mathcal{H}$. It carries a natural representation of $C_0(G/H)$, and if F is an H-equivariant Fredholm operator on $\mathcal{H}$, then the operator $\mathrm{Ind}\,F$ on $\mathrm{Ind}\,\mathcal{H}$ given by the pointwise action of F determines a cycle for $KK^G_j(C_0(G/H), \mathbb{C})$.

The inverse of the induction map defined in this way is given by compression to the range of the projection operator determined by the indicator function of the identity coset in G/H (this function being viewed as an element of $C_0(G/H)$.

The left-hand induction map is defined in a similar fashion. We choose our model for $\underline{E}H$ to be a point, which we can include into $\underline{E}G$ as an H-fixed 0-cell, and we use the induced manifolds $\mathrm{Ind}\,M = G \times_H M$, which map canonically to $G/H \subseteq \underline{E}G$. Note that any G-manifold that maps to G/H has this form. The construction of an inverse of induction and the proof that induction is an isomorphism are immediate upon noting that any G-map $f\colon \mathrm{Ind}\,M \to \underline{E}G$ is G-equivariantly homotopic to one that factors through $G/H \subseteq \underline{E}G$. □

It suffices, therefore, to prove that the map

$$\mu\colon k_j^H(\mathrm{pt}) \longrightarrow KK_j^H(\mathbb{C},\mathbb{C})$$

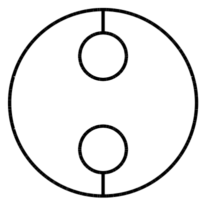
is an isomorphism. The right hand group is isomorphic to the representation ring $R(H)$ when j is even and is zero when j is odd.

8.3. LEMMA. *Let (M, a, f) be a cycle for $k_j^G(X)$. If $a = a_1 + a_2$ in $K_G^0(M)$, then*

$$[(M, a, f)] = [(M, a_1, f)] + [(M, a_2, f)]$$

in $k_j^G(X)$.

PROOF. Suppose that $(M, a, f) \in \Omega_{j+2k}^{(Z,E)}(X)$. Fix a bordism W between S^2 and $S^2 \sqcup S^2$ by situating two copies of the 2-sphere of radius $\frac{1}{4}$ inside the unit sphere. There are smooth paths I_1 and I_2 embedded into the bounding manifold W that connect the north and south poles of the large sphere to the south and north poles of the small spheres, that meet the spheres transversally, and that have trivial normal bundles in W.



The class $a_1 \in K_G^0(M)$ pulls back to a class $\widetilde{a}_1 \in K_G^0(M \times I_1)$. Similarly, the class a_2 pulls back to a class $\widetilde{a}_2 \in K_G^0(M \times I_2)$. We obtain

$$\widetilde{a} := \widetilde{a}_1 \sqcup \widetilde{a}_2 \in K_G^0(M \times I),$$

where $I = I_1 \sqcup I_2$. Now form the class

$$j_!(\widetilde{a}) \in K_G^0(M \times W),$$

where j is the inclusion of $M \times I$ into $M \times W$. If $\tilde{f}\colon M \times W \to X$ is the projection from $M \times W$ to M, followed by f, then the class $(M \times W, j_!(\tilde{a}), \tilde{f})$ is a bordism between the images of (M, a, f) and $(M, a_1, f) \sqcup (M, a_2, f)$ under the map

$$\beta\colon \Omega_j(X)^{(Z,E)} \to \Omega_{j+2(k+1)}^{(Z,E\oplus\mathbb{C})}(X).$$

In view of the definition of $k_j^G(X)$ the lemma is proved. $\square$

8.4. REMARK. The lemma shows that $-[(M, a, f)] = [(M, -a, f)]$ in $k_j^G(X)$.

8.5. PROPOSITION. *Let H be a finite group. The homomorphism*

$$\mu\colon k_*^H(\mathrm{pt}) \longrightarrow KK_*^H(\mathbb{C},\mathbb{C})$$

is an isomorphism.

PROOF. We shall prove that the homomorphism

$$R(H) \longrightarrow k_0^H(\mathrm{pt})$$
$$a \mapsto (\mathrm{pt}, a, \mathrm{Id})$$

is surjective, and that in addition, the group $k^H_1(\mathrm{pt})$ is zero. This will suffice because $KK^H_0(\mathbb{C},\mathbb{C}) \cong R(H)$, with the isomorphism being given by the above correspondence and the map μ, while $KK^H_1(\mathbb{C},\mathbb{C}) = 0$.

Fix an element of $k^H_0(\mathrm{pt})$ and represent it by a cycle (M, a) for $\Omega^{(\mathrm{pt},E)}_{2n}(\mathrm{pt})$ (we drop the map $f\colon M \to \mathrm{pt}$ from our notation here and below). Thus M is a $2n$-dimensional Spin^c-manifold with a given stable isomorphism from its tangent bundle to a trivial bundle $M\times E$, where E is a complex representation of H. The manifold M may be equivariantly embedded in a finite-dimensional complex representation V of H. By composing with the subspace embedding $V \to E\oplus V$ and by adding to V a multiple of the trivial representation, if necessary, we arrive at an embedding into $E\oplus V$ with trivial normal bundle $M\times V$.

Using the map

$$\beta^F\colon \Omega^{(\mathrm{pt},E)}_{2n}(\mathrm{pt}) \longrightarrow \Omega^{(\mathrm{pt},E\oplus V)}_{2n+2k}(\mathrm{pt})$$

we find that the given element of $k^H_0(\mathrm{pt})$ is represented by the cycle $(N, b) = (\widehat{M}, \iota_*(a))$. Here $N = \widehat{M}$ is a codimension one submanifold of $E\oplus V\oplus\mathbb{R}$ and is the boundary of some compact X (namely the ball bundle associated to $\widehat{M}$). By enclosing X in a large ball we construct a bordism Y between $\widehat{M}$ and a sphere $S \subseteq V\oplus F\oplus\mathbb{R}$. The union of X and Y is the ball bounded by the sphere S. By applying the Mayer-Vietoris sequence in H-equivariant K-theory to the decomposition $X \cup Y$ of the ball, we find that the class $b \in K^0_H(N)$ may be written as a sum $b_X + b_Y$, where b_X is the restriction of a class in $K^0_H(X)$ and b_Y is the restriction of a class in $K^0_H(Y)$. By the previous lemma,

$$[(N, b)] = [(N, b_X)] + [(N, b_Y)].$$

The first class is zero in $\Omega^{(\mathrm{pt},E\oplus F)}_{2n+2k}(\mathrm{pt})$, while the second is equal to some class $[(S, c)]$ thanks to the bordism Y. We have therefore shown that every class in $k^H_0(\mathrm{pt})$ is represented by a sphere in some $W\oplus\mathbb{R}$, where W is a complex representation of H.

To complete the proof we invoke Bott periodicity. The class c is a sum $c_0 + c_1$ where c_0 is represented by a trivial bundle (one that extends over the ball) and c_1 is in the image of the Gysin map associated to the inclusion

$$\mathrm{pt} \mapsto (0, 1) \in S \subseteq W \times \mathbb{R}.$$

This completes the computation of $k^H_0(\mathrm{pt})$.

The proof that $k^H_1(\mathrm{pt}) = 0$ is essentially the same. Every cycle is equivalent to one of the form $[(S, c)]$, where S is the sphere in a complex representaion of H. But the sphere is odd-dimensional and by Bott periodicity every class in $K^0_H(S)$ extends over the ball. So (S, c) is a boundary. □

8.6. COROLLARY. *If X is a G-finite proper G-CW-complex, then the map*

$$\mu\colon k^G_*(X) \longrightarrow KK^G_*(C_0(X), \mathbb{C})$$

is an isomorphism. □

Next we need to relate the technical groups $k^G_j(X)$ to the geometric groups $K^G_j(X)$.

8.7. LEMMA. *The class in $K^G_j(X)$ of an equivariant K-cycle (M, E, f) depends only on M, f and the class of the G-bundle E in the Grothendieck group $K^0_G(X)$.*

PROOF. Fixing M and f, the map that associates to a G-bundle E on M the class of (M, E, f) in $K^G_j(X)$ is additive, and so extends to a map from the Grothendieck group $K^0_G(X)$ into $K^G_j(X)$. □

Thanks to the lemma, we can attach a meaning to the class in the geometric group $K^G_j(X)$ of any triple (M, a, f), whenever M is a G-compact proper G-Spin^c-manifold, f is an equivariant map from M to X, and $a \in K_G(M)$. In particular, we can do so when M is a stable (Z, E)-manifold. We obtain in this way a natural transformation

$$\Omega^{(Z,E)}_j(X) \longrightarrow K^G_j(X)$$

8.8. LEMMA. *If* F *is any* G*-bundle over* Z *of rank* k*, then the diagram*

$$\begin{tikzcd} \Omega^{(Z,E)}_j(X) \arrow[r] \arrow[d, "\beta^F"'] & K^G_j(X) \arrow[d, equal] \\ \Omega^{(Z,E\oplus F)}_{j+2k}(X) \arrow[r] & K^G_{j+2k}(X) \end{tikzcd}$$

is commutative.

PROOF. This follows from the definitions of the Gysin homomorphism and vector bundle modification. □

We obtain therefore a natural transformation

$$k^G_j(X) \longrightarrow K^G_j(X)$$

that fits into a commutative diagram

$$\begin{tikzcd} k^G_j(X) \arrow[r] \arrow[d, equal] & K^G_j(X) \arrow[d, "\mu"] \\ k^G_j(X) \arrow[r, "\mu"'] & KK^G_j(C_0(X), \mathbb{C}). \end{tikzcd}$$

8.9. LEMMA. *For every* j *the map from* $k^G_j(X)$ *into the equivariant geometric* K*-homology group* $K^G_j(X)$ *is surjective.*

PROOF. Let (M, E, f) be an equivariant K-cycle for X. The manifold M maps equivariantly to a G-finite subcomplex Z of the universal space $\underline{E}G$ and so may be regarded as a closed subspace of $M \times Z$ by the diagonal embedding.

The tangent bundle for M (indeed its complexification) embeds as a summand of a G-bundle that is pulled back from a standard G-bundle on a G-finite subcomplex Z of $\underline{E}G$ along some map $h\colon X \to Z$ (see Corollary 5.2 and the comment following it). Thus there is an isomorphism of real bundles

$$TM \oplus F_0 \cong h^*E$$

where F is a real G-bundle on M and E is a complex G-bundle on Z. By adding a trivial bundle if necessary, we obtain an isomorphism

$$TM \oplus F_1 \cong h^*E \oplus \mathbb{R}^s,$$

where $F_1 = F \oplus \mathbb{R}^s$ has *even* fiber dimension. By the two out of three principle for Spin^c-structures from Section 3, the bundle F_1 carries a Spin^c-structure whose direct sum with the given Spin^c-structure on TM is the direct sum of the Spin^c-structure on h^*E associated with its complex structure and the trivial Spin^c-structure on $\mathbb{R}$. If we carry out a vector bundle modification using F_1, then we obtain an equivariant K-cycle $(M, E, f)^\wedge$ that is equivalent to (M, E, f) and for which $\widehat{M}$ carries a stable (Z, E)-structure compatible with its Spin^c-structure. □

With this, as we pointed out in the introduction, the proof of Theorem 4.6 is complete.

9. The Baum-Connes Assembly Map

Let G be a countable discrete group. The essence of the Baum-Connes conjecture for G is the assertion that every class in the K-theory of the reduced group C^*-algebra $C_r^*(G)$ arises as the index of an elliptic operator on a G-compact proper G-manifold, and that in addition the only relations among these indices arise from geometric relations (such as for example bordism) between the operators. The conjecture arose from a K-theoretic analysis of Lie groups and of crossed product algebras related to foliations. However here we shall discuss only the C^*-algebras of discrete groups.

To make their conjecture precise, Baum and Connes constructed in [**BC00**] geometric groups $K^j(G)$ from cycles related to the symbols of equivariant elliptic pseudodifferential operators, and an equivalence relation related to the Gysin map in K-theory. They then defined an index map

$$\mu\colon K^j(G) \longrightarrow K_j(C_r^*(G))$$

that they conjectured to be an isomorphism.

Although the Dirac operator on a Spin^c-manifold did not play a central role in [**BC00**], it is nonetheless a fairly routine matter to identify the geometric group defined there with the group generated from cycles (M, E), where:

(a) M is a G-compact proper G-Spin^c-manifold, all of whose components have either even or odd dimension, according as j is 0 or 1, and
(b) E is a complex G-bundle on M.

The equivalence relation between cycles is generated by bordism, direct sum/disjoint union and vector bundle modification, exactly as in Section 3, except that here there is no reference space X, nor any map from M to X. Compare [**Bau04**], where the conjecture for countable discrete groups is formulated in precisely this way.

It follows from the universal property of the space $\underline{E}G$ that there is an isomorphism

$$K^j(G) \cong \varinjlim_{X\subseteq \underline{E}G} K_j^G(X),$$

where on the right is the direct limit of the geometric K-homology groups of the G-finite subcomplexes of the G-CW-complex $\underline{E}G$.

In a later paper [**BCH94**], Baum, Connes and Higson defined an assembly map

$$\mu\colon \varinjlim_{X\subseteq \underline{E}G} KK_j^G(C(X),\mathbb{C}) \longrightarrow K_j(C_r^*(G))$$

Its relation to the original Baum-Connes map is summarized by the commutative diagram

$$\begin{tikzcd}
K^j(G) \arrow[r, "\cong"] \arrow[d, "\mu"'] & \varinjlim_{X\subseteq \underline{E}G} K_j^G(X) \arrow[r, "\mu"] & \varinjlim_{X\subseteq \underline{E}G} KK_j^G(C(X),\mathbb{C}) \arrow[d, "\mu"] \\
K_j(C_r^*(G)) \arrow[rr, "="'] & & K_j(C_r^*(G)),
\end{tikzcd}$$

where the horizontal map labelled (yet again) μ is the one analyzed in this paper, and shown to be an isomorphism. Because it is an isomorphism, the reformulation of the Baum-Connes conjecture in [**BCH94**] is equivalent to the original in [**BC00**] for discrete groups.

Despite the discovery some years ago of counterexamples to various extensions of the Baum-Connes conjecture (see [**HLS02**]), there is, as of today, no known counterexample to the conjecture as reviewed here.

References

[Bau04] P. Baum. On the index of equivariant elliptic operators. In *Operator algebras, quantization, and noncommutative geometry*, volume 365 of *Contemp. Math.*, pages 41–49. Amer. Math. Soc., Providence, RI, 2004.

[BC00] P. Baum and A. Connes. Geometric K-theory for Lie groups and foliations. *Enseign. Math. (2)*, 46(1-2):3–42, 2000.

[BCH94] P. Baum, A. Connes, and N. Higson. Classifying space for proper actions and K-theory of group C^*-algebras. In *C^*-algebras: 1943–1993 (San Antonio, TX, 1993)*, volume 167 of *Contemp. Math.*, pages 240–291. Amer. Math. Soc., Providence, RI, 1994.

[BD82] P. Baum and R. G. Douglas. K homology and index theory. In *Operator algebras and applications, Part I (Kingston, Ont., 1980)*, volume 38 of *Proc. Sympos. Pure Math.*, pages 117–173. Amer. Math. Soc., Providence, R.I., 1982.

[BHS07] P. Baum, N. Higson, and T. Schick. On the equivalence of geometric and analytic K-homology. *Pure Appl. Math. Q.*, 3(1, part 3):1–24, 2007.

[Bla98] B. Blackadar. *K-theory for operator algebras*, volume 5 of *Mathematical Sciences Research Institute Publications*. Cambridge University Press, Cambridge, second edition, 1998.

[HLS02] N. Higson, V. Lafforgue, and G. Skandalis. Counterexamples to the Baum-Connes conjecture. *Geom. Funct. Anal.*, 12(2):330–354, 2002.

[HR00] N. Higson and J. Roe. *Analytic K-homology*. Oxford Mathematical Monographs. Oxford University Press, Oxford, 2000. Oxford Science Publications.

[Kas88] G. G. Kasparov. Equivariant KK-theory and the Novikov conjecture. *Invent. Math.*, 91(1):147–201, 1988.

[LO01] W. Lück and R. Oliver. The completion theorem in K-theory for proper actions of a discrete group. *Topology*, 40(3):585–616, 2001.

[Lüc05] W. Lück. Survey on classifying spaces for families of subgroups. In *Infinite groups: geometric, combinatorial and dynamical aspects*, volume 248 of *Progr. Math.*, pages 269–322. Birkhäuser, Basel, 2005.

DEPARTMENT OF MATHEMATICS, PENN STATE UNIVERSITY, UNIVERSITY PARK, PA 16802
E-mail address: baum@math.psu.edu

DEPARTMENT OF MATHEMATICS, PENN STATE UNIVERSITY, UNIVERSITY PARK, PA 16802
E-mail address: higson@math.psu.edu

MATHEMATISCHES INSTITUT, UNIVERSITÄT GÖTTINGEN, BUNSENSTR. 3, D-37073 GÖTTINGEN, GERMANY
E-mail address: thomas.schick@uni-math.gwdg.de

Clay Mathematics Proceedings
Volume 11, 2010

The Planar Algebra of Diagonal Subfactors

Dietmar Bisch, Paramita Das, and Shamindra Kumar Ghosh

Dedicated to Alain Connes on the occasion of his 60th birthday.

Abstract. There is a natural construction which associates to a finitely generated, countable, discrete group G and a 3-cocycle ω of G an inclusion of II_1 factors, the so-called diagonal subfactors (with cocycle). In the case when the cocycle is trivial these subfactors are well studied and their standard invariant (or planar algebra) is known. We give a description of the planar algebra of these subfactors when a cocycle is present. The action of Jones' planar operad involves the 3-cocycle ω explicitly and some interesting identities for 3-cocycles appear when naturality of the action is verified.

1. Introduction

The theory of subfactors ([**10**]) has experienced several new developments in the last few years through the introduction of planar algebra technology ([**11**]). Every subfactor comes with a very rich mathematical object, the *standard invariant* or *planar algebra* of the subfactor, which in nice situations is a complete invariant of the subfactor ([**20**], [**21**]). It can be described in many interesting ways, as for instance a certain category of bimodules ([**18**], see also [**2**]), as lattices of multi-matrix algebras ([**19**]), or as a planar algebra ([**11**]). The planar algebra approach is particularly powerful since it allows one to use algebraic-combinatorial methods in conjunction with topological ones to investigate the structure of subfactors. A number of examples of explicit planar algebras associated to subfactors have been computed (see for instance [**3**], [**4**], [**8**], [**11**], [**14**], [**15**], [**16**]) but there is a need for more concrete examples. This is what we accomplish in this paper. We give a description of the planar algebra of the diagonal subfactors associated to a G-kernel.

Let P be a II_1 factor and let $\theta_1, \ldots, \theta_n$ be automorphisms of P (we may assume without loss of generality that $\theta_1 = \mathrm{id}$). Consider the subfactor $N = \{\sum_{i=1}^n \theta_i(x)e_{ii} \,|\, x \in P\} \subset M = P \otimes M_n(\mathbb{C})$, where $(e_{ij})_{1\leq i,j\leq n}$ denote matrix units in $M_n(\mathbb{C})$. $N \subset M$ is then called the *diagonal subfactor* associated to $\{\theta_i\}_{1\leq i\leq n}$. These subfactors were proposed by Jones in 1985 to provide examples of potentially non-classifiable subfactors, since this construction allows one to translate problems

2010 *Mathematics Subject Classification.* Primary: 46L37.
Key words and phrases. G-kernel, group cocycle, planar algebra, planar operad, subfactor.
The authors were supported by NSF under Grant No. DMS-0301173 and DMS-0653717.

on classification of group actions into problems on subfactors. Popa used them to prove vanishing of 2-cohomology results for cocycle actions of finitely generated, strongly amenable groups on an arbitrary II_1 factor ([**23**]). Ocneanu had proved such a result for cocycle actions of amenable groups on the hyperfinite II_1 factor using different techniques ([**17**]). The diagonal subfactors are of course reducible and have Jones index n^2. They provide a wealth of simple examples of infinite depth subfactors whose structure theory is well understood. In particular, the standard invariant or planar algebra of these subfactors has been determined in ([**20**], [**11**], [**1**]).

Let G be the group generated by the images g_i of θ_i, $1 \leq i \leq n$, in $\mathrm{Out}\, P = \mathrm{Aut}\, P/\mathrm{Int}\, P$. Popa showed that analytical properties of these subfactors are reflected in the corresponding properties of the group G. For instance, if P is hyperfinite, then the diagonal subfactor is amenable (in the sense of Popa) if and only if G is an amenable group ([**20**]). The subfactor has property (T) in the sense of Popa if and only if G has property (T) of Kazhdan ([**22**]). The principal graphs of these subfactors are then Cayley-like graphs of G with respect to the generators $g_1, \ldots, g_n$ and their inverses (see [**20**] or [**1**] for the precise statement). The higher relative commutants, Jones projections and conditional expectations have all been worked out in ([**20**], [**11**], [**1**]).

A well-known variation of the diagonal subfactors is obtained as follows (see e.g. [**20**]). Consider a G-kernel, that is, an injective homomorphism χ from a (countable, discrete, finitely generated) group G into $\mathrm{Out}\, P$. Denote by $\epsilon : \mathrm{Aut}\, P \to \mathrm{Out}\, P$ the canonical homomorphism, and let $\alpha : G \to \mathrm{Aut}\, P$ be a lift of χ such that $\epsilon \circ \alpha_s = \chi(s)$, for all $s \in G$. It follows that $\alpha_s \alpha_t = \mathrm{Ad}\, u(s,t) \alpha_{st}$, for all s, $t \in G$, and some unitaries $u(s,t) \in P$. Associativity of composition of automorphisms implies that $\mathrm{Ad}\,(u(r,s)u(rs,t)) = \mathrm{Ad}\,(\alpha_r(u(s,t))u(r,st))$. Hence there is a function $\omega : G \times G \times G \to \mathbb{T}$ with $u(r,s)u(rs,t) = \omega(r,s,t)\alpha_r(u(s,t))u(r,st)$. It is easy to check that ω is a 3-cocycle and that its class in $H^3(G, \mathbb{T})$ does not depend on the choices made. One usually denotes the class by $\mathrm{Ob}(G)$ or $\mathrm{Ob}(\chi)$, the *obstruction* of χ. It is an obstruction to lifting G to an *action* on the II_1 factor P. Clearly, if two G-kernels χ and η are conjugate (in $\mathrm{Out}\, P$), then $\mathrm{Ob}(\chi) = \mathrm{Ob}(\eta)$. It was shown in [**17**] that Ob is a complete conjugacy invariant for G-kernels if P is the hyperfinite II_1 factor and G is a countable, discrete, amenable group. Note that in general, even if $\mathrm{Ob}(G) = 0$, there may be no lifting of the G-kernel to an action on P. Connes and Jones found in [**7**] the first such example of a non-liftable G-kernel with vanishing obstruction, where G is a group with property (T). Vanishing of the obstruction is a necessary and sufficient condition for G to lift to a *cocycle action* on the II_1 factor P. See [**5**], [**12**], [**17**], [**23**], [**24**] for more on this.

Connes showed that if G is cyclic, one can construct G-kernels in $\mathrm{Out}\, R$ with arbitrary obstructions, where R denotes the *hyperfinite* II_1 factor ([**6**]). It was an open problem whether this result is true for more general groups, and Jones showed in [**9**] that this is indeed the case for G an arbitrary discrete group. Thus, given a discrete group G and a class $\pi \in H^3(G, \mathbb{T})$, there is G-kernel $\chi : G \to \mathrm{Out}\, R$ with $\mathrm{Ob}(\chi) = \pi$. Sutherland constructed G-kernels with arbitrary obstructions in non-hyperfinite II_1 factors ([**24**]).

Given a finitely generated, countable group G and a 3-cocycle ω, we can associate a (hyperfinite) subfactor to (G, ω) as follows. Let $\chi : G \to \mathrm{Out}\, P$ be a G-kernel with $\mathrm{Ob}(G) = [\omega] \in H^3(G, \mathbb{T})$ (choose for instance $P = R$, and use [**9**]). Fix generators $\{g_1, \ldots, g_n\}$ of G, let α be a lift of χ, and consider the diagonal subfactor associated to the automorphisms $\alpha_{g_1}, \ldots, \alpha_{g_n}$ (let us choose $g_1 = e$, the identity of G, and $\alpha_e = \mathrm{id}$). If η is another G-kernel with lift β, and automorphisms $\beta_{g_1}, \ldots,$ β_{g_n}, then the diagonal subfactors associated to $(\alpha_{g_i})_i$ and $(\beta_{g_j})_j$ are isomorphic if and only if there is an automorphism θ of the underlying II_1 factor P such that $\alpha_{g_{\pi(i)}} = \theta\beta_{g_i}\theta^{-1} \mod \mathrm{Int}\, P$, where π is a permutation of the indices (fixing 1) (see e.g. [**23**]). Thus, in particular, isomorphism of these diagonal subfactors implies that $\mathrm{Ob}(\chi) = \mathrm{Ob}(\eta)$ (up to a possible modification of χ by the permutation π). Isomorphism of standard invariants is weaker than isomorphism of subfactors, but we still have the following. If (G, χ) and (G, η) are two G-kernels as above with $\mathrm{Ob}(\eta) = \mathrm{Ob}(\chi)$, then the standard invariants of the associated diagonal subfactors are isomorphic (see ([**20**], page 228 ff.). The converse is not true due to the fact that group isomorphisms can change the class of a 3-cocycle. If one constructs diagonal subfactors where the automorphisms are repeated with distinct multiplicities, one can show a converse, e.g. if G is strongly amenable and P is hyperfinite, using Popa's classification results of amenable subfactors ([**20**], page 229 ff.). Clearly, the 3-cocycle ω giving rise to the obstruction of the G-kernel will appear in the standard invariant of these diagonal subfactors (with cocycle) and the purpose of this paper is to give a precise description of this occurrence.

Here is a more detailed outline of the sections of this paper. We review group cocycles in section 2. In section 3, we define an abstract planar algebra $P^{\langle g_i : i \in I \rangle, \omega}$ associated to a finitely generated group G with a fixed finite generating set $\{g_i\}_{i \in I}$, and a 3-cocycle $\omega \in Z^3(G, \mathbb{T})$. The vector spaces underlying this planar algebra are spanned by multi-indices in I^{2n} such that the corresponding alternating word on generators and their inverses is the identity in G. The action of Jones' planar operad is defined explicity, and of course ω appears prominently in this definition. It should be noted that the definition of the action of a tangle involves a labelling of the *strings* in the tangle whereas the planar algebra description of the group-type subfactors in [**3**] involved a labelling of *boundary segments* (called "openings" in [**3**]) of the internal and external discs of the tangle. We would like to point out that the 3-cocycle ω does not appear in our definition of the action of the multiplication, inclusion, Jones projection and right conditional expectation tangles. It does appear in the definition of the action of the left conditional expectation tangles (and hence the rotation tangles). We verify in this section that our definition of the action of tangles is indeed natural with respect to composition of tangles. This takes a little work, but some interesting identities for 3-cocycles appear along the way.

In section 4 we give a model for the higher relative commutants of the diagonal subfactor with cocycle and describe the associated concrete planar algebra. We choose an appropriate basis of the higher relative commutants which allows us to identify this concrete planar algebra with the abstractly defined one in section 3. This isomorphism is obtained in the usual way by constructing a filtered $*$-algebra isomorphism between the abstract planar algebra of section 3 and the concrete one of section 4, that preserves Jones projection and conditional expectation tangles. The main feature of this planar algebra is the fact that the distinguished basis of the higher relative commutants that we found here matches with the one coming from the description of the planar algebra as a path algebra associated to the principal graphs of the subfactor (see e.g. [**13**], [**11**]). Conversely, we prove that any finite index extremal subfactor whose standard invariant is given by the abstract planar algebra (in section 3), must necessarily be (isomorphic to) a diagonal subfactor.

2. A brief review of group cocycles

For the convenience of the reader, we recall in this brief section the definition of a cocycle of a group G. G will denote throughout this article a countable, discrete group, and we will denote the identity of G by e. Define $C^n = \mathrm{Fun}(G^n, \mathbb{T})$, the space of functions from G^n to $\mathbb{T}$, and $\partial^n : C^n \to C^{n+1}$ by

$$\begin{aligned}&\partial^n(\phi)(g_1, \cdots, g_{n+1})\\ &= \phi(g_2, \cdots, g_{n+1})\, \phi(g_1 g_2, g_3, \cdots, g_{n+1})^{-1}\, \phi(g_1, g_2 g_3, g_4, \cdots, g_{n+1}) \cdots\\ &\quad \cdots \phi(g_1, \cdots, g_{n-1}, g_n g_{n+1})^{(-1)^n}\, \phi(g_1, \cdots, g_n)^{(-1)^{n+1}}\end{aligned}$$

It follows that $\left(\partial^{n+1} \circ \partial^n\right)(\cdot) = 1_{C^{n+2}}$ where 1 denotes the constant function 1. Denote $\ker(\partial)$ by $Z^n(G, \mathbb{T})$ (whose elements are called n-cocycles) and $\mathrm{Im}(\partial^{n-1})$ by $B^n(G, \mathbb{T})$ (whose elements are called coboundaries). Note that $B^n(G, \mathbb{T}) \subset Z^n(G, \mathbb{T})$.

In this paper we will be dealing mostly with a 3-cocycle ω. Thus ω satisfies the identity

$$\omega(g_1, g_2, g_3)\, \omega(g_1, g_2 g_3, g_4)\, \omega(g_2, g_3, g_4) = \omega(g_1 g_2, g_3, g_4)\, \omega(g_1, g_2, g_3 g_4) \tag{2.1}$$

We call ω *normalized* if $\omega(g_1, g_2, g_3) = 1$ whenever either of g_1, g_2, g_3 is e.

Any cocycle ω is coboundary equivalent to a normalized cocycle. In particular, $(\omega \cdot \partial^2(\phi))$ is a normalized 3-cocycle, where $\phi \in C^2$ is defined as $\phi(g_1, g_2) = \omega(g_1, e, e)\, \overline{\omega}(e, e, g_2)$ for all $g_1, g_2 \in G$.

3. An abstract planar algebra

In this section we will construct an abstract planar algebra which, in section 4, will be shown to be isomorphic to the planar algebra of a diagonal subfactor with cocycle.

Let G be a countable, discrete group generated by a finite subset $\{g_i\}_{i\in I}$, and let $\omega \in Z^3(G, \mathbb{T})$ be a normalized 3-cocycle. We will construct a planar algebra $P^{\langle g_i : i \in I\rangle, \omega}$ $(= P)$ as follows. Let e denote the identity of G. We define first a map *alt* from multi-indices $\coprod_{n\geq 0} I^n$ to G by

$$\left(\coprod_{n\geq 0} I^n\right) \ni \underline{i} = (i_1, \cdots, i_n) \overset{alt}{\longmapsto} g_{i_1}^{-1} g_{i_2} \cdots g_{i_n}^{(-1)^n} = alt(\underline{i}) \in G$$

where *alt* of the empty multi-index is defined to be the identity element of the group. To define the planar algebra we need to define vector spaces P_n and an action of Jones' planar operad on these vector spaces. We refer to [**11**] for the planar algebra terminology used in this paper.

The vector spaces. Define $P_n = \begin{cases} \mathbb{C}\{\underline{i} \in I^{2n} : alt(\underline{i}) = e\} & \text{if } n > 0, \\ \mathbb{C} & \text{if } n = 0. \end{cases}$

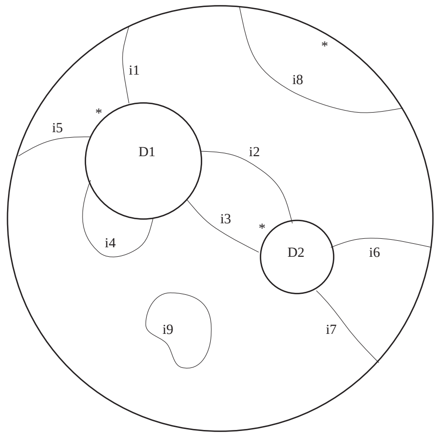


FIGURE 1. Example of a state on a tangle

Action of tangles. Let T be an n_0-tangle having internal discs $D_1, \cdots, D_b$ with colors $n_1, \cdots, n_b$ respectively (or no internal discs of course). A *state* σ on T is a map from $\{$strings in $T\}$ to I such that $alt(\sigma|_{\partial D_c}) = e$ for all $1 \le c \le b$, where $\sigma|_{\partial D_c}$ denotes the element of I^{2n_c} obtained by reading the elements of I induced at the marked points on the boundary of D_c by the strings via the map σ, starting from the first marked point and moving clockwise. This has been illustrated in Figure 1, where $alt(\sigma|_{\partial D_1})$ and $alt(\sigma|_{\partial D_2})$ are just the products $g_{i_1}^{-1} g_{i_2} g_{i_3}^{-1} g_{i_4} g_{i_4}^{-1} g_{i_5}$ and $g_{i_2}^{-1} g_{i_6} g_{i_7}^{-1} g_{i_3}$ respectively, and are thus required to be the identity element. It is a consequence that $alt(\sigma|_{\partial D_0})(= g_{i_8}^{-1} g_{i_6} g_{i_7}^{-1} g_{i_5} g_{i_1}^{-1} g_{i_8} = e$ in the figure) holds for the external disc. Let $\mathcal{S}(T)$ denote the states on T.

In order to define the action Z_T of T, it is enough to define the coefficient $\langle Z_T(\underline{k}^1, \cdots, \underline{k}^b) | \underline{k}^0 \rangle$ of $\underline{k}^0$ in the linear expansion of $Z_T(\underline{k}^1, \cdots, \underline{k}^b)$, where $\underline{k}^c \in I^{2n_c}$ such that $alt(\underline{k}^c) = e$ for $1 \le c \le b$. For this, we choose a picture T_1 in the isotopy class of T and then choose a simple path p_c in $D_0 \setminus [\coprod_{c=1}^b Int(D_c)]$ starting from the $*$ [1] of D_0 to that of D_c for $1 \le c \le b$ such that:

(i) p_c intersects the strings of T_1 transversally for $1 \le c \le b$,

(ii) p_{c_1} and p_{c_2} intersect exactly at the $*$ of D_0 for $1 \le c_1 \ne c_2 \le b$.

Note that any state σ on T gives an element $\sigma|_{p_c} \in I^{m_c}$ obtained by reading the elements of I induced by σ at the crossings of the path p_c and the strings along the direction of the path where m_c (necessarily even) is the number of strings cut by p_c.

Define

$$\langle Z_T(\underline{k}^1, \cdots, \underline{k}^b) \mid \underline{k}^0 \rangle = \sum_{\substack{\sigma \in \mathcal{S}(T) \text{ s.t.} \\ \sigma|_{\partial D_d} = \underline{k}^d \\ \text{for } 0 \le d \le b}} \prod_{c=1}^{b} \lambda_{\sigma|_{p_c}}(\underline{k}^c)$$

where $\lambda_{\underline{j}}(\underline{i}) = \prod_{s=1}^n \lambda_{\underline{j}}(\underline{i}, s)$ and

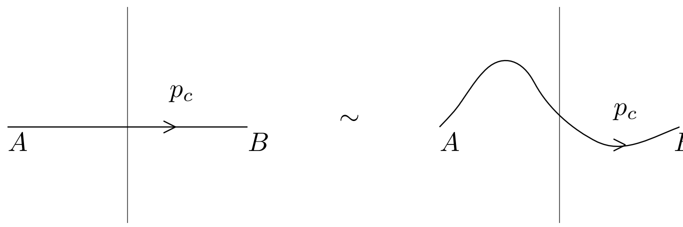
[1]Recall $*$ of a disc D is a point chosen on the boundary of D strictly between the last and the first marked points, moving clockwise.

$$\lambda_{\underline{j}}(\underline{i}, s) = \begin{cases} \overline{\omega}(alt(\underline{j}), alt(i_1, \cdots, i_s), g_{i_s}) \text{ if } s \text{ is odd} \\ \omega(alt(\underline{j}), alt(i_1, \cdots, i_{s-1}), g_{i_s}) \text{ if } s \text{ is even} \end{cases}$$

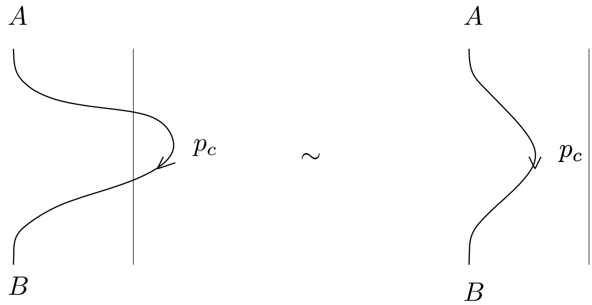
for $\underline{i} \in I^n$, $\underline{j} \in I^m$. If there is no compatible state on T, then we take the coefficient to be 0 and if there is no internal disc in T, then the scalar inside the sum is considered to be 1. (Note that $\lambda_{\underline{j}}(\underline{i})$ depends only on $alt(\underline{j})$ and $\underline{i}$.)

We need to show first that the multilinear map Z_T is well-defined. Two configurations of paths $\{p_c\}_{c=1}^b$ and $\{p'_c\}_{c=1}^b$ in T_1 can be obtained from each other using a finite sequence of the following moves:

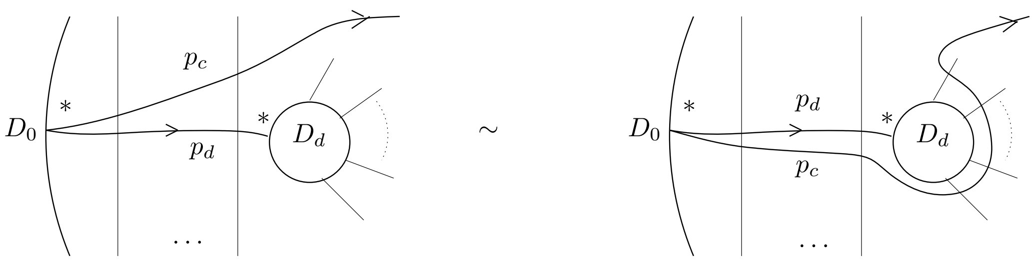
I. *isotopy*



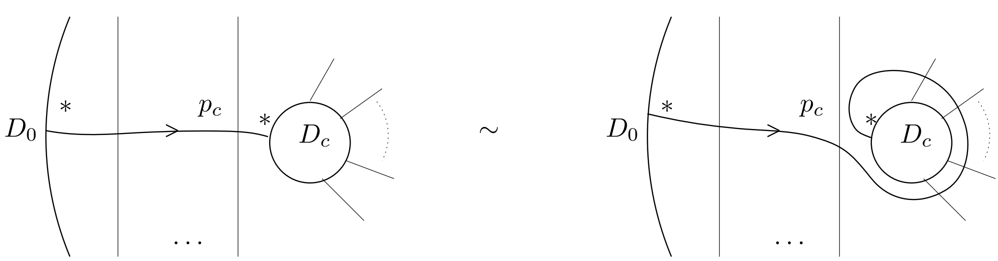
II. *cap-sliding moves*



III. *disc-sliding moves*



IV. *rotation moves*



It is enough to check that the definition of the action is independent under each of the above moves. Invariance under isotopy moves are the easiest to check since $\sigma|_{p_c} = \sigma|_{p'_c}$ for all $c \in \{1, \cdots, b\}$. To see invariance under the remaining three moves, we show that $alt(\sigma|_{p_c}) = alt(\sigma|_{p'_c})$ for $1 \leq c \leq b$. For a cap-sliding move, note that the cap induces the same index at the two consecutive crossing with the path but after applying the alt map, the corresponding group elements cancel each other since they inverses of each other. For the disc-sliding (resp., rotation moves), the invariance follows from the fact that $alt(\sigma|_{\partial D_d}) = e$ (resp., $alt(\sigma|_{\partial D_c}) = e$).

We will show next that the action is compatible with composition of tangles.

Action is natural with respect to composition of tangles. Let S be an m_0-tangle containing the internal discs $D'_1, \cdots, D'_a$ with colors $m_1, \cdots, m_a$ and T be an m_1-tangle containing internal discs $D_1, \cdots, D_b$ with colors $n_1, \cdots, n_b$. Let D'_0 and D_0 denote the external discs of S and T respectively. We need to show that for all $\underline{i}^{c'} \in I^{2m_{c'}}$ and $\underline{j}^c \in I^{2n_c}$, where $c' \in \{0, 2, \cdots, a\}$ and $c \in \{1, \cdots, b\}$,

$$\langle Z_S(Z_T(\underline{j}^1, \cdots, \underline{j}^b), \underline{i}^2, \cdots, \underline{i}^a)|\underline{i}^0\rangle = \langle Z_{S\circ_{D'_1} T}(\underline{j}^1, \cdots, \underline{j}^b, \underline{i}^2, \cdots, \underline{i}^a)|\underline{i}^0\rangle \tag{3.1}$$

The left-hand side of (3.1) can be expanded as

$$\sum_{\substack{\underline{i}^1 \in I^{2m_1} \text{ s.t.} \\ alt(\underline{i}^1) = e}} \langle Z_S(\underline{i}^1, \underline{i}^2, \cdots, \underline{i}^a)|\underline{i}^0\rangle \ \langle Z_T(\underline{j}^1, \cdots, \underline{j}^b)|\underline{i}^1\rangle$$

$$= \sum_{\substack{\sigma \in \mathcal{S}(S),\, \tau \in \mathcal{S}(T) \text{ s.t. } \sigma|_{\partial D'_{c'}} = \underline{i}^{c'},\, \tau|_{\partial D_c} = \underline{j}^c \text{ for} \\ c' \in \{0,2,\cdots,a\},\, c \in \{1,\cdots,b\} \text{ and } \sigma|_{\partial D'_1} = \tau|_{\partial D_0} = \underline{i}^1}} \left(\prod_{c'=1}^{a} \lambda_{\sigma|_{p'_{c'}}}(\underline{i}^{c'})\right) \left(\prod_{c=1}^{b} \lambda_{\tau|_{p_c}}(\underline{j}^c)\right)$$

where $p'_{c'}$'s and p_c's are paths in the tangles S and T respectively, required to define their actions. For the action of $S \circ_{D'_1} T$, we consider the paths $p'_{c'}$ for $2 \leq c' \leq a$ and $p'_1 \circ p_c$ for $1 \leq c \leq b$. (Strictly speaking, in order to define the action of $S \circ_{D'_1} T$, one has to make the p'_1-portion of the paths $(p'_1 \circ p_c)$ disjoint for different values of c.) So, the right-hand side of (3.1) becomes

$$\sum_{\substack{\gamma \in \mathcal{S}(S \circ_{D'_1} T) \text{ s.t. } \gamma|_{\partial D'_{c'}} = \underline{i}^{c'},\, \gamma|_{\partial D_c} = \underline{j}^c \\ \text{for } c' \in \{0,2,\cdots,a\},\, c \in \{1,\cdots,b\}}} \left(\prod_{c'=2}^{a} \lambda_{\gamma|_{p'_{c'}}}(\underline{i}^{c'})\right) \left(\prod_{c=1}^{b} \lambda_{\gamma|_{p'_1 \circ p_c}}(\underline{j}^c)\right)$$

Observe that $\left\{(\sigma, \tau) \in \mathcal{S}(S) \times \mathcal{S}(T) : \begin{array}{c} \sigma|_{\partial D'_{c'}} = \underline{i}^{c'} \text{ for } c' \in \{0, 2, \cdots, a\} \\ \tau|_{\partial D_c} = \underline{j}^c \text{ for } 1 \leq c \leq b,\ \sigma|_{\partial D'_1} = \tau|_{\partial D_0} \end{array}\right\}$ is clearly in bijection with $\left\{\gamma \in \mathcal{S}(S \circ_{D'_1} T) : \begin{array}{c} \gamma|_{\partial D'_{c'}} = \underline{i}^{c'} \text{ for } c' \in \{0, 2, \cdots, a\}, \\ \gamma|_{\partial D_c} = \underline{j}^c \text{ for } 1 \leq c \leq b \end{array}\right\}$.
A bijection is obtained by sending (σ, τ) to the state defined by σ (resp. τ) on the S-part (resp. T-part) of $S \circ_{D'_1} T$ and the well-definedness of such a state is a consequence of the condition $\sigma|_{\partial D'_1} = \tau|_{\partial D_0}$; we denote this state by $\sigma \circ \tau$. If these sets are empty, equation (3.1) holds trivially since both sides have value 0. Let us assume that the sets are nonempty. It is enough to prove for $\sigma \in \mathcal{S}(S)$ and

$\tau \in \mathcal{S}(T)$ such that $\sigma|_{\partial D'_{c'}} = \underline{i}^{c'}$ for $0 \leq c' \leq a$ and $\tau|_{\partial D_c} = \underline{j}^c$ for $1 \leq c \leq b$ and $\tau|_{\partial D_0} = \underline{i}^1$ we have

$$\lambda_{\sigma|_{p'_1}}(\underline{i}^1) \prod_{c=1}^{b} \lambda_{\tau|_{p_c}}(\underline{j}^c) = \prod_{c=1}^{b} \lambda_{(\sigma\circ\tau)|_{(p'_1 \circ p_c)}}(\underline{j}^c) \tag{3.2}$$

We prove this in two cases.

Case 1: T has no internal disc or closed loop, that is, T is a Temperley-Lieb diagram. Then the right-hand side of equation (3.2) is 1. It remains to show $\lambda_{\sigma|_{p'_1}}(\underline{i}^1) = 1$. This follows from the next lemma and the fact that $\underline{i}^1 = Z_T(1)$ is a sequence of non-crossing matched pairings of indices from I.

Lemma 3.1. *If $i \in I$, $\underline{j} \in I^m$, $\underline{i} = (i_1, \cdots, i_n) \in I^n$ and $0 \leq s \leq n$, then we have $\lambda_{\underline{j}}(\underline{i}) = \lambda_{\underline{j}}(i_1, \cdots, i_s, i, i, i_{s+1}, \cdots, i_n)$.*

Proof: Note that
(i) $\lambda_{\underline{j}}(\underline{i}, r) = \lambda_{\underline{j}}((i_1, \cdots, i_s, i, i, i_{s+1}, \cdots, i_n), r)$ for $1 \leq r \leq s$, and
(ii) $\lambda_{\underline{j}}(\underline{i}, r) = \lambda_{\underline{j}}((i_1, \cdots, i_s, i, i, i_{s+1}, \cdots, i_n), r+2)$ for $s+1 \leq r \leq n$.

We compute then,

$$\lambda_{\underline{j}}((i_1, \cdots, i_s, i, i, i_{s+1}, \cdots, i_n), s+1)\ \lambda_{\underline{j}}((i_1, \cdots, i_s, i, i, i_{s+1}, \cdots, i_n), s+2)$$
$$= \begin{cases} \overline{\omega}(alt(\underline{j}), alt(i_1, \cdots, i_s, i), g_i)\ \omega(alt(\underline{j}), alt(i_1, \cdots, i_s, i), g_i) & \text{if } s \text{ is even} \\ \omega(alt(\underline{j}), alt(i_1, \cdots, i_s), g_i)\ \overline{\omega}(alt(\underline{j}), alt(i_1, \cdots, i_s, i, i), g_i) & \text{if } s \text{ is odd} \end{cases}$$
$$=1.$$

□

Remark 3.2. In Lemma 3.1, if $\underline{i}$ is a sequence of indices with non-crossing matched pairings, then we can apply the lemma several times to reduce all the consecutive matched pairings to get $\lambda_{\underline{j}}(\underline{i}) = 1$. □

Case 2: T has at least one internal disc. Any unlabelled tangle T can be expressed as composition of elementary annular tangles of four types as described in [**3**], namely, capping, cap-inclusion, left-inclusion and disc-inclusion tangles. It is enough to prove equation (3.2) for any tangle S and compatible tangle T in $\mathcal{E}$ (= the set of all elementary tangles). If T is a capping or cap-inclusion type annular tangle, the proof directly follows from Lemma 3.1 and is left to the reader.

If T is a left-inclusion annular tangle, equation (3.2) is implied from the following lemma.

Lemma 3.3. *For all $\underline{i} \in I^{2m}$, $\underline{j} \in I^{2n}$ and $\underline{k} \in I^{2n_1}$ such that $alt(\underline{i}) = e$, we have*

$$\lambda_{(\underline{k},\underline{j})}(\underline{i}) = \lambda_{\underline{k}}(\underline{j}, \underline{i}, \widetilde{\underline{j}})\ \lambda_{\underline{j}}(\underline{i})$$

where $\widetilde{\underline{j}}$ is the sequence of indices from $\underline{j}$ in the reverse order.

Proof: We rearrange the terms of the right-hand side of the identity in the lemma in the following way:

$$\begin{aligned}&\lambda_{\underline{k}}(\underline{j},\underline{i},\widetilde{\underline{j}})\ \lambda_{\underline{j}}(\underline{i})\\ =&\left(\prod_{r=1}^{2n}\lambda_{\underline{k}}((\underline{j},\underline{i},\widetilde{\underline{j}}),r)\ \lambda_{\underline{k}}((\underline{j},\underline{i},\widetilde{\underline{j}}),2(m+2n)-r+1)\right)\left(\prod_{s=1}^{2m}\lambda_{\underline{k}}((\underline{j},\underline{i},\widetilde{\underline{j}}),2n+s)\ \lambda_{\underline{j}}(\underline{i},s)\right)\end{aligned}$$

Let $\underline{i}=(i_1,\cdots,i_{2m})$ and $\underline{j}=(j_1,\cdots,j_{2n})$. Note that for $1\le r\le 2n$,

$$\begin{aligned}&\lambda_{\underline{k}}((\underline{j},\underline{i},\widetilde{\underline{j}}),2(m+2n)-r+1)\\ =&\begin{cases}\omega(alt(\underline{k}),alt(j_1,\cdots,j_r),g_{j_r}) & \text{if } r \text{ is odd}\\ \overline{\omega}(alt(\underline{k}),alt(j_1,\cdots,j_{r-1}),g_{j_r}) & \text{if } r \text{ is even}\end{cases}\\ =&\overline{\lambda}_{\underline{k}}((\underline{j},\underline{i},\widetilde{\underline{j}}),r)\end{aligned}$$

since $alt(\underline{i})=e$ and $alt(\underline{j},\underline{i},\widetilde{\underline{j}})=alt(\underline{j})alt(\underline{i})(alt(\underline{j}))^{-1}=e$. Thus the first product of the terms in the rearrangement vanishes. For the second product, if $s\in\{1,\cdots,2m\}$ is odd, then

$$\begin{aligned}&\lambda_{\underline{k}}((\underline{j},\underline{i},\widetilde{\underline{j}}),2n+s)\ \lambda_{\underline{j}}(\underline{i},s)\\ =&\,\overline{\omega}(alt(\underline{k}),alt(\underline{j})alt(i_1,\cdots,i_s),g_{i_s})\ \overline{\omega}(alt(\underline{j}),alt(i_1,\cdots,i_s),g_{i_s})\\ =&\,\overline{\omega}(alt(\underline{k}),alt(\underline{j}),alt(i_1,\cdots,i_{s-1}))\ \overline{\omega}(alt(\underline{k},\underline{j}),alt(i_1,\cdots,i_s),g_{i_s})\cdot\\ &\omega(alt(\underline{k}),alt(\underline{j}),alt(i_1,\cdots,i_s))\\ =&\,\overline{\omega}(alt(\underline{k}),alt(\underline{j}),alt(i_1,\cdots,i_{s-1}))\ \lambda_{(\underline{k},\underline{j})}(\underline{i},s)\ \omega(alt(\underline{k}),alt(\underline{j}),alt(i_1,\cdots,i_s))\end{aligned}$$

where we used the defining equation (2.1) of the 3-cocycle ω for the second equality.

Similarly, if $s\in\{1,\cdots,2m\}$ is even, then

$$\begin{aligned}&\lambda_{\underline{k}}((\underline{j},\underline{i},\widetilde{\underline{j}}),2n+s)\ \lambda_{\underline{j}}(\underline{i},s)\\ =&\,\overline{\omega}(alt(\underline{k}),alt(\underline{j}),alt(i_1,\cdots,i_{s-1}))\ \lambda_{(\underline{k},\underline{j})}(\underline{i},s)\ \omega(alt(\underline{k}),alt(\underline{j}),alt(i_1,\cdots,i_s))\end{aligned}$$

Thus,

$$\begin{aligned}&\prod_{s=1}^{2m}\lambda_{\underline{k}}((\underline{j},\underline{i},\widetilde{\underline{j}}),2n+s)\ \lambda_{\underline{j}}(\underline{i},s)\\ =&\prod_{t=1}^{m}\Big(\lambda_{\underline{k}}((\underline{j},\underline{i},\widetilde{\underline{j}}),2n+2t-1)\ \lambda_{\underline{j}}(\underline{i},2t-1)\Big)\Big(\lambda_{\underline{k}}((\underline{j},\underline{i},\widetilde{\underline{j}}),2n+2t)\ \lambda_{\underline{j}}(\underline{i},2t)\Big)\\ =&\prod_{t=1}^{m}\left(\begin{array}{l}\overline{\omega}(alt(\underline{k}),alt(\underline{j}),alt(i_1,\cdots,i_{2t-2}))\ \lambda_{(\underline{k},\underline{j})}(\underline{i},2t-1)\\ \lambda_{(\underline{k},\underline{j})}(\underline{i},2t)\ \omega(alt(\underline{k}),alt(\underline{j}),alt(i_1,\cdots,i_{2t}))\end{array}\right)\\ =&\,\overline{\omega}(alt(\underline{k}),alt(\underline{j}),e)\left(\prod_{t=1}^{m}\lambda_{(\underline{k},\underline{j})}(\underline{i},2t-1)\ \lambda_{(\underline{k},\underline{j})}(\underline{i},2t)\right)\omega(alt(\underline{k}),alt(\underline{j}),alt(\underline{i}))\\ =&\,\lambda_{(\underline{k},\underline{j})}(\underline{i})\end{aligned}$$

since $alt(\underline{i})=e$ and ω is normalized. $\square$

REMARK 3.4. The proof of Lemma 3.3 also implies the following identity:

$$\lambda_{\underline{k}}(\underline{j}, \underline{i}, \underline{\widetilde{j}}) = \prod_{s=2n+1}^{2m+2n} \lambda_{\underline{k}}((\underline{j}, \underline{i}, \underline{\widetilde{j}}), s)$$

$\square$

Now, suppose T is a disc-inclusion tangle as shown in Figure 2. Note that $n_1 = m_1$. Without loss of generality, we can assume that T be given by the following picture in which we also indicate the paths p_1 and p_2.

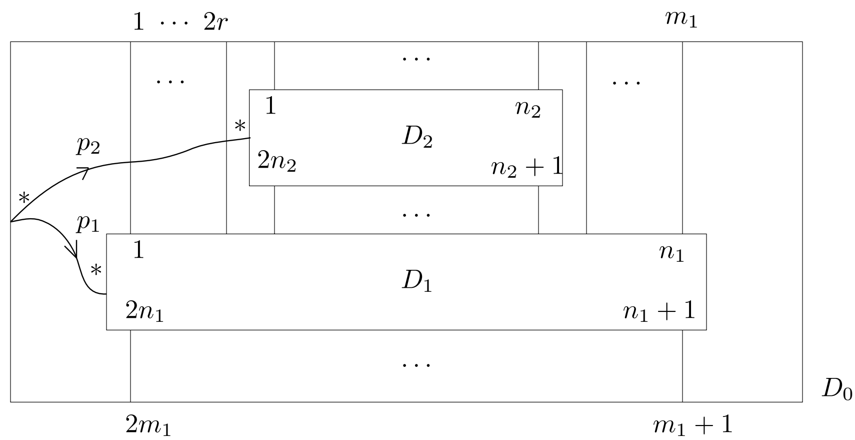


FIGURE 2. Disc inclusion tangle

Observe that since p_1 does not intersect any string, we have that $\tau|_{p_1}$ is the empty multi-index. So it is enough to prove

$$\lambda_{\sigma|_{p_1'}}(\underline{i}^1)\, \lambda_{\tau|_{p_2}}(\underline{j}^2) \;=\; \lambda_{\sigma|_{p_1'}}(\underline{j}^1)\, \lambda_{(\sigma|_{p_1'}, \tau|_{p_2})}(\underline{j}^2) \tag{3.3}$$

Let us denote $\tau|_{p_2}$ by $\underline{k} \in I^{2r}$. Since τ is a state, the following relations clearly follow from Figure 2:

(i) $i^1_s = j^1_s$ for $1 \leq s \leq 2r$ and $2r + n_2 + 1 \leq s \leq 2n_1$,
(ii) $i^1_{2r+s} = j^2_s$ for $1 \leq s \leq n_2$,
(iii) $i^1_s = k_s = j^1_s$ for $1 \leq s \leq 2r$,
(iv) $j^1_{2r+s} = j^2_{2n_2-s+1}$ for $1 \leq s \leq n_2$.

We now express $\lambda_{\sigma|_{p_1'}}(\underline{i}^1)$ as a product of three terms with which we work separately.

$$\lambda_{\sigma|_{p_1'}}(\underline{i}^1) = \left(\prod_{s=1}^{2r} \lambda_{\sigma|_{p_1'}}(\underline{i}^1, s)\right)\left(\prod_{s=2r+1}^{2r+n_2} \lambda_{\sigma|_{p_1'}}(\underline{i}^1, s)\right)\left(\prod_{s=2r+n_2+1}^{2n_1} \lambda_{\sigma|_{p_1'}}(\underline{i}^1, s)\right)$$

First term: For $1 \leq s \leq 2r$,

$$\begin{aligned}
\lambda_{\sigma|_{p'_1}}(\underline{i}^1, s) &= \begin{cases} \overline{\omega}(alt(\sigma|_{p'_1}), alt(i^1_1, \cdots, i^1_s), g_{i^1_s}) & \text{if } s \text{ is odd} \\ \omega(alt(\sigma|_{p'_1}), alt(i^1_1, \cdots, i^1_{s-1}), g_{i^1_s}) & \text{if } s \text{ is even} \end{cases} \\
&= \begin{cases} \overline{\omega}(alt(\sigma|_{p'_1}), alt(j^1_1, \cdots, j^1_s), g_{j^1_s}) & \text{if } s \text{ is odd} \\ \omega(alt(\sigma|_{p'_1}), alt(j^1_1, \cdots, j^1_{s-1}), g_{j^1_s}) & \text{if } s \text{ is even} \end{cases} \quad \text{(applying (i))} \\
&= \lambda_{\sigma|_{p'_1}}(\underline{j}^1, s).
\end{aligned}$$

Second term: For $2r+1 \leq s \leq 2r+n_2$,

$$\begin{aligned}
\lambda_{\sigma|_{p'_1}}(\underline{i}^1, s) &= \begin{cases} \overline{\omega}(alt(\sigma|_{p'_1}), alt(i^1_1, \cdots, i^1_s), g_{i^1_s}) & \text{if } s \text{ is odd} \\ \omega(alt(\sigma|_{p'_1}), alt(i^1_1, \cdots, i^1_{s-1}), g_{i^1_s}) & \text{if } s \text{ is even} \end{cases} \\
&= \begin{cases} \overline{\omega}(alt(\sigma|_{p'_1}), alt(\underline{k})alt(j^2_1, \cdots, j^2_{s-2r}), g_{j^2_{s-2r}}) & \text{if } s \text{ is odd} \\ \omega(alt(\sigma|_{p'_1}), alt(\underline{k})alt(j^2_1, \cdots, j^2_{s-2r-1}), g_{j^2_{s-2r}}) & \text{if } s \text{ is even} \end{cases} \\
&\text{(applying (ii) and (iii))} \\
&= \lambda_{\sigma|_{p'_1}}((\underline{k}, \underline{j}^2, \underline{\widetilde{k}}), s).
\end{aligned}$$

Third term: Note that

$$\text{(v)} \qquad \begin{aligned}
alt(i^1_{2r+1}, \cdots, i^1_{2r+n_2}) &= alt(j^2_1, \cdots, j^2_{n_2}) \text{ (using (ii))} \\
&= alt(j^2_{2n_2}, \cdots, j^2_{n_2+1}) \text{ (since } alt(\underline{j}^2) = e) \\
&= alt(j^1_{2r+1}, \cdots, j^1_{2r+n_2}) \text{ (using (iv))}.
\end{aligned}$$

Thus, for $2r+n_2+1 \leq s \leq 2n_1$,

$$\begin{aligned}
&\lambda_{\sigma|_{p'_1}}(\underline{i}^1, s) \\
&= \begin{cases} \overline{\omega}(alt(\sigma|_{p'_1}), alt(i^1_1, \cdots, i^1_{2r})alt(i^1_{2r+1}, \cdots, i^1_{2r+n_2}, \cdots i^1_s, g_{i^1_s}) & \text{if } s \text{ is odd} \\ \omega(alt(\sigma|_{p'_1}), alt(i^1_1, \cdots, i^1_{2r})alt(i^1_{2r+1}, \cdots, i^1_{2r+n_2}, \cdots i^1_{s-1}, g_{i^1_s}) & \text{if } s \text{ is even} \end{cases} \\
&= \begin{cases} \overline{\omega}(alt(\sigma|_{p'_1}), alt(j^1_1, \cdots, j^1_s), g_{j^1_s}) & \text{if } s \text{ is odd} \\ \omega(alt(\sigma|_{p'_1}), alt(j^1_1, \cdots, j^1_{s-1}), g_{j^1_s}) & \text{if } s \text{ is even} \end{cases} \quad \text{(using (i) and (v))} \\
&= \lambda_{\sigma|_{p'_1}}(\underline{j}^1, s).
\end{aligned}$$

Combining the three terms, we get

$$\begin{aligned}
\lambda_{\sigma|_{p'_1}}(\underline{i}^1) &= \left(\prod_{s=1}^{2r} \lambda_{\sigma|_{p'_1}}(\underline{j}^1, s)\right)\left(\prod_{s=2r+1}^{2r+n_2} \lambda_{\sigma|_{p'_1}}((\underline{k}, \underline{j}^2, \underline{\widetilde{k}}), s)\right)\left(\prod_{s=2r+n_2+1}^{2n_1} \lambda_{\sigma|_{p'_1}}(\underline{j}^1, s)\right) \\
&= \lambda_{\sigma|_{p'_1}}(\underline{j}^1)\left(\prod_{s=2r+1}^{2r+n_2} \overline{\lambda}_{\sigma|_{p'_1}}(\underline{j}^1, s)\right)\left(\prod_{s=2r+1}^{2r+n_2} \lambda_{\sigma|_{p'_1}}((\underline{k}, \underline{j}^2, \underline{\widetilde{k}}), s)\right).
\end{aligned}$$

Now, for $2r+1 \leq s \leq 2r+n_2$, we compute

$$
\begin{aligned}
&\overline{\lambda}_{\sigma|_{p_1'}}(\underline{j}^1, s)\\
=&\begin{cases} \omega(alt(\sigma|_{p_1'}), alt(j_1^1, \cdots, j_{2r}^1, \cdots, j_s^1), g_{j_s^1}) & \text{if } s \text{ is odd}\\ \overline{\omega}(alt(\sigma|_{p_1'}), alt(j_1^1, \cdots, j_{2r}^1, \cdots, j_{s-1}^1), g_{j_s^1}) & \text{if } s \text{ is even}\end{cases}\\
=&\begin{cases} \omega(alt(\sigma|_{p_1'}), alt(\underline{k}) alt(j_{2n_2}^2, \cdots, j_{2n_2+2r-s+1}^2), g_{j^2_{2n_2+2r-s+1}}) & \text{if } s \text{ is odd}\\ \overline{\omega}(alt(\sigma|_{p_1'}), alt(\underline{k}) alt(j_{2n_2}^2, \cdots, j_{2n_2+2r-s+2}^2), g_{j^2_{2n_2+2r-s+1}}) & \text{if } s \text{ is even}\end{cases}\\
&\text{(using (iii) and (iv))}\\
=&\begin{cases} \omega(alt(\sigma|_{p_1'}), alt(\underline{k}) alt(j_1^2, \cdots, j_{2n_2+2r-s}^2), g_{j^2_{2n_2+2r-s+1}}) & \text{if } s \text{ is odd}\\ \overline{\omega}(alt(\sigma|_{p_1'}), alt(\underline{k}) alt(j_1^2, \cdots, j_{2n_2+2r-s+1}^2), g_{j^2_{2n_2+2r-s+1}}) & \text{if } s \text{ is even}\end{cases}\\
&\text{(since } alt(\underline{j}^2) = e)\\
=&\lambda_{\sigma|_{p_1'}}((\underline{k}, \underline{j}^2, \widetilde{\underline{k}}), 2n_2 + 4r - s + 1).
\end{aligned}
$$

Hence, we obtain

$$
\begin{aligned}
\lambda_{\sigma|_{p_1'}}(\underline{i}^1) &= \lambda_{\sigma|_{p_1'}}(\underline{j}^1) \left(\prod_{s=2r+n_2+1}^{2r+2n_2} \lambda_{\sigma|_{p_1'}}((\underline{k}, \underline{j}^2, \widetilde{\underline{k}}), s) \right) \left(\prod_{s=2r+1}^{2r+n_2} \lambda_{\sigma|_{p_1'}}((\underline{k}, \underline{j}^2, \widetilde{\underline{k}}), s) \right)\\
&= \lambda_{\sigma|_{p_1'}}(\underline{j}^1)\, \lambda_{\sigma|_{p_1'}}(\underline{k}, \underline{j}^2, \widetilde{\underline{k}}) \ \ \text{(by Remark 3.4).}
\end{aligned}
$$

We can now proceed with the proof of equation (3.3):

$$
\begin{aligned}
\lambda_{\sigma|_{p_1'}}(\underline{i}^1)\, \lambda_{\tau|_{p_2}}(\underline{j}^2) &= \lambda_{\sigma|_{p_1'}}(\underline{j}^1)\, \lambda_{\sigma|_{p_1'}}(\underline{k}, \underline{j}^2, \widetilde{\underline{k}})\, \lambda_{\underline{k}}(\underline{j}^2)\\
&= \lambda_{\sigma|_{p_1'}}(\underline{j}^1)\ \ \lambda_{(\sigma|_{p_1'}, \underline{k})}(\underline{j}^2),
\end{aligned}
$$

where we applied Lemma 3.3 for the last identity. This completes the proof that the action of tangles defined above is compatible with composition of tangles (called *naturality of composition* in [**11**]). Hence, $P^{\langle g_i : i \in I\rangle, \omega}$ is a planar algebra.

We will define next a $*$-*structure* on $P^{\langle g_i : i \in I\rangle, \omega}$. Note that if $\underline{i} \in I^{2n}$, then $alt(\underline{i}) = e$ if and only if $alt(\widetilde{\underline{i}}) = e$. Extend the operation $\sim$ conjugate linearly to define a $*$-structure on $P_n^{\langle g_i : i \in I\rangle, \omega}$. Clearly, this is an involution. We need to check whether the action of tangles preserves $*$, that is, $Z_{T^*} \circ (* \times \cdots \times *) = * \circ Z_T$. It is enough to check this identity when T has no internal discs or closed loops, and when T is an elementary annular tangle.

If T has no internal discs or closed loops, then it is a Temperley-Lieb diagram and hence Z_T is the sum of all sequences of indices from I which have non-crossing matched pairings where the positions of the pairings are given by the numberings of the marked points on the boundary of T which are connected by a string. Now, in the tangle T^*, the m-th and the n-th marked points are connected by a string if and only if the m-th and the n-th marked points starting from the last point in T, reading anticlockwise, are connected. So, Z_{T^*} is indeed the sum of all sequences featuring in the linear expansion of Z_T in the reverse order (that is, applying $\sim$).

If T is an elementary annular tangle of capping (resp. cap-inclusion) type with m-th and $(m+1)$-st marked points of the internal (resp. external) disc being connected by a string, then T^* is also the same kind of elementary annular tangle but the 'capping' occurs at the m-th and $(m+1)$-st marked points of the internal (resp. external) disc starting from the last point reading anticlockwise. The identity easily follows from this observation.

If T is an elementary tangle of left-inclusion type, then $T = T^*$. The identity will then follow from the next lemma.

LEMMA 3.5. *If $\underline{j} \in I^m$ and $\underline{i} = (i_1, \cdots, i_{2n}) \in I^{2n}$ such that $alt(\underline{i}) = e$, then $\lambda_{\underline{j}}(\widetilde{\underline{i}}) = \overline{\lambda}_{\underline{j}}(\underline{i})$.*

Proof: The proof is an immediate consequence of $alt(i_1, \cdots, i_s) = alt(i_{2n}, \cdots, i_{s+1})$, which holds since $alt(\underline{i}) = e$. ☐

Lastly, if T is an elementary tangle of disc-inclusion type given by Figure 2, then we need to show $\langle Z_{T^*}(\widetilde{\underline{i}}, \widetilde{\underline{j}}), \widetilde{\underline{k}}\rangle = \overline{\langle Z_T(\underline{i}, \underline{j}), \underline{k}\rangle}$ for $\underline{i}, \underline{k} \in I^{2n_1}, \underline{j} \in I^{2n_2}$ with $alt(\underline{i}) = alt(\underline{j}) = alt(\underline{k}) = e$. It is easy to verify that $\underline{i}$, $\underline{j}$ and $\underline{k}$ define a state on T if and only if $\widetilde{\underline{i}}$, $\widetilde{\underline{j}}$ and $\widetilde{\underline{k}}$ define the same on T^*. If they fail to define a state, then both sides are zero. If they define a state, then the scalars appearing on both sides can be made equal by applying Lemma 3.5.

REMARK 3.6. If ω is trivial, that is, a coboundary, then $P^{\langle g_i : i \in I\rangle, \omega}$ is isomorphic to example 2.7 in [**11**]. Jones constructed this example of a planar algebra by considering a certain subspace of the tensor planar algebra (TPA) over the vector space with the indexing set I as a basis. He then showed that this subspace is closed under the TPA-action of tangles, thus showing that the subspace is indeed a planar algebra. One can view our planar algebra $P^{\langle g_i : i \in I\rangle, \omega}$ as a subspace of the TPA in an obvious way but the action of planar tangles induced by TPA will not be the same as our action which involves the extra data of a 3-cocycle. It does not seem clear if our planar algebra $P^{\langle g_i : i \in I\rangle, \omega}$ can be viewed as a planar subalgebra of the tensor planar algebra over the vector space with basis I if ω is nontrivial.

4. The planar algebra of the diagonal subfactor with cocycle

In this section we will compute the relative commutants of the diagonal subfactor associated to a G-kernel. We will determine the filtered $*$-algebra structure, Jones projections and the conditional expectations. Note that some of this can already be found in [**1**], [**20**], [**11**]. The main point here is that we are able to choose an appropriate basis of the higher relative commutants such that the action of planar tangles on these allows us to identify this concrete planar algebra with the abstract one defined in the previous section.

Let N be a II_1 factor, I be a finite set and for $i \in I$, choose $\theta_i \in \mathrm{Aut}\, N$. Set $M = M_I \otimes N$ where M_I denotes the matrix algebra whose rows and columns are indexed by I. As in the introduction, consider the *diagonal subfactor* $N \hookrightarrow M$ given by

$$N \ni x \longmapsto \sum_{i \in I} E_{i,i} \otimes \theta_i(x) \in M$$

that is, an element x of N sits in M as a diagonal matrix whose ith diagonal is $\theta_i(x)$. If we have another collection of automorphisms $\theta'_i \in \mathrm{Aut}\, N$ for $i \in I$ such that $\theta_i = \theta'_i \mod \mathrm{Int}\, N$, for all $i \in I$ (up to a permutation of I), then the associated diagonal subfactors are isomorphic. It is therefore natural to associate a diagonal subfactor to a collection of elements $g_i \in \mathrm{Out}\, N$, $i \in I$. We consider the subgroup

$G = \langle g_i : i \in I \rangle$ of $\operatorname{Out} N$, which can be viewed as a G-kernel in the obvious way, and choose a lift

$$G \ni g \stackrel{\alpha}{\longmapsto} \alpha_g \in \operatorname{Aut} N$$

such that $\alpha_e = \mathrm{id}_N$. Set $\alpha_i = \alpha_{g_i}$ for all $i \in I$. Consider the diagonal subfactor $N \subset M = M_I \otimes N$ where the ith diagonal entry of an element of N viewed in M is twisted by the action of α_i . The index of this subfactor is $|I|^2$. Set $M_n = M_{I^{n+1}} \otimes N$ for $n \geq 0$. We will often identify $E_{\underline{i},\underline{j}} \otimes E_{\underline{k},\underline{l}} \in M_{I^m} \otimes M_{I^n}$ with $E_{(\underline{i},\underline{k}),(\underline{j},\underline{l})} \in M_{I^{m+n}}$ for $\underline{i}, \underline{j} \in I^m$ and $\underline{k}, \underline{l} \in I^n$. Now, M_{n-1} is included in M_n in the following way:

$$M_{n-1} = M_{I^n} \otimes N \ni E_{\underline{i},\underline{j}} \otimes x \longmapsto E_{\underline{i},\underline{j}} \otimes \psi_n(x) \in M_{I^n} \otimes M_I \otimes N = M_{I^{n+1}} \otimes N = M_n$$

for all $\underline{i}, \underline{j} \in I^n$ and $x \in N$ where $\psi_n : N \to M_I \otimes N$ is defined as:

$$\psi_n(x) = \begin{cases} \sum_{k \in I} E_{k,k} \otimes \alpha_k(x) & \text{if } n \text{ is even} \\ \sum_{k \in I} E_{k,k} \otimes \alpha_k^{-1}(x) & \text{if } n \text{ is odd} \end{cases}$$

It is easy to check (see [**20**], [**1**]) that $N \subset M \subset M_1 \subset M_2 \subset \cdots$ is isomorphic to the Jones tower of II_1 factors associated to $N \subset M$ where the Jones projections and conditional expectations are given by:

$$e_n = |I|^{-1} \sum_{\underline{k} \in I^{n-1}, i,j \in I} E_{(\underline{k},i,i),(\underline{k},j,j)} \otimes 1 \in M_n$$

$$\mathbb{E}^{M_n}_{M_{n-1}}(E_{(\underline{i},k),(\underline{j},l)} \otimes x) = \delta_{k,l} \, |I|^{-1} \, E_{\underline{i},\underline{j}} \otimes \alpha_k^{(-1)^{n-1}}(x)$$

for all $\underline{i}, \underline{j} \in I^n$, $k, l \in I$, $x \in N$. Moreover, $\left\{\sqrt{|I|}(E_{i,j} \otimes 1) : i, j \in I\right\}$ forms a Pimsner-Popa basis of M over N. This basis will be used to write the conditional expectation of commutant of N onto the commutant of M (see [**2**]).

To find the relative commutant $N' \cap M_{n-1}$, first note that N is included in M_{n-1} by the following map:

$$N \ni x \longmapsto \sum_{\underline{i} \in I^n} E_{\underline{i},\underline{i}} \otimes alt_\alpha^{-1}(\underline{i})(x) \in M_{n-1}$$

where $alt_\alpha(\underline{i}) = \alpha_{i_1}^{-1} \alpha_{i_2} \cdots \alpha_{i_n}^{(-1)^n} \in \operatorname{Aut} N$ for $\underline{i} = (i_1, \cdots, i_n) \in I^n$. Now, if $\sum_{\underline{i},\underline{j} \in I^n} x_{\underline{i},\underline{j}} (E_{\underline{i},\underline{j}} \otimes 1) \in N' \cap M_{n-1}$, then

$$\sum_{\underline{i},\underline{j} \in I^n} x_{\underline{i},\underline{j}} (E_{\underline{i},\underline{j}} \otimes 1) \, y = y \sum_{\underline{i},\underline{j} \in I^n} x_{\underline{i},\underline{j}} (E_{\underline{i},\underline{j}} \otimes 1) \text{ for all } y \in N$$

$$\Leftrightarrow x_{\underline{i},\underline{j}} \left(alt_\alpha(\underline{i}) alt_\alpha^{-1}(\underline{j})\right)(y) = y \, x_{\underline{i},\underline{j}} \text{ for all } y \in N, \, \underline{i}, \underline{j} \in I^n$$

$$\Leftrightarrow x_{\underline{i},\underline{j}} alt_\alpha(\underline{i}, \widetilde{\underline{j}})(y) = y \, x_{\underline{i},\underline{j}} \text{ for all } y \in N, \, \underline{i}, \underline{j} \in I^n$$

Thus, for $\underline{i}, \underline{j} \in I^n$, if $x_{\underline{i},\underline{j}} \neq 0$, then $x_0 = \frac{x_{\underline{i},\underline{j}}}{\|x_{\underline{i},\underline{j}}\|} \in \mathcal{U}(N)$ and $\operatorname{Ad} x_0 \circ alt_\alpha(\underline{i}, \widetilde{\underline{j}}) = \mathrm{id}_N$ which implies $alt(\underline{i}, \widetilde{\underline{j}}) = e$. Similarly, if there exist $\underline{i}, \underline{j} \in I^n$ such that $alt(\underline{i}, \widetilde{\underline{j}}) = e$, then $u(E_{\underline{i},\underline{j}} \otimes 1) \in N' \cap M_{n-1}$ where $u \in \mathcal{U}(N)$ satisfies $\operatorname{Ad} u \circ alt_\alpha(\underline{i}, \widetilde{\underline{j}}) = \mathrm{id}_N$. Thus,

$$N' \cap M_{n-1} = \operatorname{span} \left\{ u(E_{\underline{i},\underline{j}} \otimes 1) \in M_{n-1} \;\middle|\; \begin{array}{c} \underline{i}, \underline{j} \in I^n \text{ and } u \in \mathcal{U}(N) \\ \text{s.t. } \operatorname{Ad} u \circ alt_\alpha(\underline{i}, \widetilde{\underline{j}}) = \mathrm{id}_N \end{array} \right\}.$$

The elements in this set do not yet form a basis since the unitary u can be modified by a scalar of absolute value 1. To get a good basis of $N' \cap M_{n-1}$ we need to choose u in such a way that the planar algebra associated to $N \subset M$ can easily be identified with the abstract one defined in Section 3.

We now digress a little bit to set up some notations. Let $u : G \times G \to \mathcal{U}(N)$ be a unitary defined by

$$\alpha_{g_1}\alpha_{g_2} = \operatorname{Ad} u(g_1, g_2) \circ \alpha_{g_1 g_2} \quad \text{for all } g_1, g_2 \in G$$

such that $u(g_1, g_2) = 1$ whenever either of g_1 or g_2 is e. For $\underline{i} = (i_1, \cdots, i_n) \in I^n$, define

$$v_m(\underline{i}) = \begin{cases} u^*(alt(i_1, \cdots, i_m), g_{i_m}) & \text{if } m \text{ is odd} \\ u(alt(i_1, \cdots, i_{m-1}), g_{i_m}) & \text{if } m \text{ is even} \end{cases}$$

and set $v(\underline{i}) = v_1(\underline{i}) \cdots v_n(\underline{i})$. Next, we prove a several useful lemmas involving v. The first lemma motivates our choice of the basis.

Lemma 4.1. *$alt_\alpha(\underline{i}) = Ad\, v(\underline{i}) \circ \alpha_{alt(\underline{i})}$ for all $\underline{i} \in I^n$.*

Proof: Using the definition of u, note that for all $m \geq 1$,

$$\operatorname{Ad} v_m(\underline{i}) = \begin{cases} \alpha_{alt(i_1,\cdots,i_{m-1})}\alpha_{i_m}^{-1}\alpha_{alt(i_1,\cdots,i_m)}^{-1} & \text{if } m \text{ is odd} \\ \alpha_{alt(i_1,\cdots,i_{m-1})}\alpha_{i_m}\alpha_{alt(i_1,\cdots,i_m)}^{-1} & \text{if } m \text{ is even} \end{cases}$$

Hence

$$\operatorname{Ad} v(\underline{i}) = \operatorname{Ad} v_1(\underline{i}) \cdots \operatorname{Ad} v_n(\underline{i}) = \alpha_e \alpha_{i_1}^{-1}\alpha_{i_2} \cdots \alpha_{i_n}^{(-1)^n} \alpha_{alt(i_1,\cdots,i_n)}^{-1} = alt_\alpha(\underline{i})\alpha_{alt(\underline{i})}^{-1}.$$

$\square$

Lemma 4.2. *For any $\underline{k} = (k_1, \cdots, k_{2n}) \in I^{2n}$ such that $alt(\underline{k}) = e$, we have $v(\widetilde{\underline{k}})v(\underline{k}) = 1$.*

Proof: First we expand $v(\widetilde{\underline{k}})$ and $v(\underline{k})$ into products of unitaries arising from the definition of v, and then we consider the product of the p-th unitary of $v(\widetilde{\underline{k}})$ from the right and p-th unitary of $v(\underline{k})$ from the left, that is,

$$\begin{aligned} v_{2n-p+1}(\widetilde{\underline{k}})v_p(\underline{k}) &= \begin{cases} u(alt(k_{2n}, \cdots, k_{p+1}), k_p)\, u^*(alt(k_1, \cdots, k_p), k_p) & \text{if } p \text{ is odd} \\ u^*(alt(k_{2n}, \cdots, k_p), k_p)\, u(alt(k_1, \cdots, k_{p-1}), k_p) & \text{if } p \text{ is even} \end{cases} \\ &= \begin{cases} u(alt(k_1, \cdots, k_p), k_p)\, u^*(alt(k_1, \cdots, k_p), k_p) & \text{if } p \text{ is odd} \\ u^*(alt(k_1, \cdots, k_{p-1}), k_p)\, u(alt(k_1, \cdots, k_{p-1}), k_p) & \text{if } p \text{ is even} \end{cases} \\ &= 1, \end{aligned}$$

for $1 \leq p \leq 2n$. $\square$

Lemma 4.3. *For $\underline{i} = (i_1, \cdots, i_n)$, $\underline{j} = (j_1, \cdots, j_n)$, $\underline{k} = (k_1, \cdots, k_n) \in I^n$ such that $alt(\underline{i}, \widetilde{\underline{j}}) = e = alt(\underline{j}, \widetilde{\underline{k}})$, we have $v(\underline{i}, \widetilde{\underline{j}})\, v(\underline{j}, \widetilde{\underline{k}}) = v(\underline{i}, \widetilde{\underline{k}})$.*

Proof: Using an argument similar to the proof of Lemma 4.2, one can prove that the product of the p-th unitary of $v(\underline{i}, \widetilde{\underline{j}})$ from the right and the p-th unitary of

$v(\underline{j},\widetilde{\underline{k}})$ from the left, is 1 for $1 \le p \le n$. Again, for $n+1 \le p \le 2n$,

$$\begin{aligned} v_p(\underline{j},\widetilde{\underline{k}}) &= \begin{cases} u^*(alt(\underline{j})k_{2n}^{(-1)^{n+1}} \cdots k_{2n-p+1}^{-1}, k_{2n-p+1}) & \text{if } p \text{ is odd} \\ u(alt(\underline{j})k_{2n}^{(-1)^{n+1}} \cdots k_{2n-p+2}^{-1}, k_{2n-p+1}) & \text{if } p \text{ is even} \end{cases} \\ &= \begin{cases} u^*(alt(\underline{i})k_{2n}^{(-1)^{n+1}} \cdots k_{2n-p+1}^{-1}, k_{2n-p+1}) & \text{if } p \text{ is odd} \\ u(alt(\underline{i})k_{2n}^{(-1)^{n+1}} \cdots k_{2n-p+2}^{-1}, k_{2n-p+1}) & \text{if } p \text{ is even} \end{cases} \quad (alt(\underline{i},\widetilde{\underline{j}}) = e) \\ &= v_p(\underline{i},\widetilde{\underline{k}}). \end{aligned}$$

Thus,

$$\begin{aligned} v(\underline{i},\widetilde{\underline{j}})\ v(\underline{j},\widetilde{\underline{k}}) &= v_1(\underline{i},\widetilde{\underline{j}}) \cdots v_n(\underline{i},\widetilde{\underline{j}}) v_{n+1}(\underline{j},\widetilde{\underline{k}}) \cdots v_{2n}(\underline{j},\widetilde{\underline{k}}) \\ &= v_1(\underline{i},\widetilde{\underline{k}}) \cdots v_n(\underline{i},\widetilde{\underline{k}}) v_{n+1}(\underline{i},\widetilde{\underline{k}}) \cdots v_{2n}(\underline{i},\widetilde{\underline{k}}) = v(\underline{i},\widetilde{\underline{k}}) \end{aligned}$$

□

By applying Lemma 4.1, we see that the set $\left\{ v^*(\underline{i},\widetilde{\underline{j}})(E_{\underline{i},\underline{j}} \otimes 1) \;\middle|\; \begin{array}{l} \underline{i},\underline{j} \in I^n \text{ s.t.} \\ alt(\underline{i},\widetilde{\underline{j}}) = e \end{array} \right\}$ is a basis for $N' \cap M_{n-1}$. We will use this basis to establish an isomorphism between the planar algebra associated to $N \subset M$ and the abstract planar algebra defined in Section 3. Let $\omega : G \times G \times G \to \mathbb{T}$ be the 3-cocycle associated to G, that is, for all $g_1, g_2, g_3 \in G$ we have

$$u(g_1,g_2)u(g_1g_2,g_3) = \omega(g_1,g_2,g_3)\alpha_{g_1}(u(g_2,g_3))u(g_1,g_2g_3).$$

As before, this is a consequence of associativity of the multiplication in G.

We prove next a useful lemma relating v and ω.

Lemma 4.4. *If $i \in I$ and $\underline{k} = (k_1, \cdots, k_{2n}) \in I^{2n}$ such that $alt(\underline{k}) = e$ and $k_1 = k_{2n}$, then $alt_\alpha(i,k_1)(v(\underline{k})) = \overline{\lambda}_{(i,k_1)}(\underline{k})\ v(i,k_2,\cdots,k_{2n-1},i)$.*

Proof: We expand $alt_\alpha(i,k_1)(v(\underline{k}))$ as a product of unitaries and work with them separately. For $1 \le p \le n$, we have

(i) $alt_\alpha(i,k_1)(v_{2p-1}(\underline{k}))$

$= \left(\alpha_{g_i}^{-1}\alpha_{k_1}\right)(u^*(alt(k_1,\cdots,k_{2p-1}),g_{k_{2p-1}}))$

$= \left(\mathrm{Ad}\, u^*(g_i^{-1},g_i)u(g_i^{-1},g_{k_1}) \circ \alpha_{g_i^{-1}g_{k_1}}\right)(u^*(alt(k_1,\cdots,k_{2p-1}),g_{k_{2p-1}}))$

(using the definition of u)

$= \omega(alt(i,k_1),alt(k_1,\cdots,k_{2p-1}),g_{k_{2p-1}})$

$\mathrm{Ad}\, u^*(g_i^{-1},g_i)u(g_i^{-1},g_{k_1}) \begin{pmatrix} u(alt(i,k_1),alt(k_1,\cdots,k_{2p-2},k_{2p-1},k_{2p-1})) \\ \cdot\, u^*(alt(i,k_1,k_1,k_2,\cdots,k_{2p-1}),g_{k_{2p-1}}) \\ \cdot\, u^*(alt(i,k_1),alt(k_1,\cdots,k_{2p-1})) \end{pmatrix}$

(using the definition of ω)

(ii) $alt_\alpha(i,k_1)(v_{2p}(\underline{k}))$

$= \left(\alpha_{g_i}^{-1}\alpha_{k_{i_1}}\right)(u(alt(k_1,\cdots,k_{2p-1}),g_{k_{2p}}))$

$= \left(\mathrm{Ad}\, u^*(g_i^{-1},g_i)u(g_i^{-1},g_{k_1}) \circ \alpha_{g_i^{-1}g_{k_1}}\right)(u(alt(k_1,\cdots,k_{2p-1}),g_{k_{2p}}))$

(using the definition of u)

$= \overline{\omega}(alt(i,k_1),alt(k_1,\cdots,k_{2p-1}),g_{k_{2p}})$

$\mathrm{Ad}\, u^*(g_i^{-1},g_i)u(g_i^{-1},g_{k_1}) \begin{pmatrix} u(alt(i,k_1),alt(k_1,\cdots,k_{2p-1})) \\ \cdot\, u(alt(i,k_1,k_1,k_2,\cdots,k_{2p-1}),g_{k_{2p}}) \\ \cdot\, u^*(alt(i,k_1),alt(k_1,\cdots,k_{2p})) \end{pmatrix}$

(using the definition of ω)

Multiplying (i) and (ii), we get,

$$
\begin{aligned}
&alt_\alpha(i,k_1)((v_{2p-1}(\underline{k}))\ (v_{2p}(\underline{k})))\\
=&\overline{\lambda}_{(i,k_1)}(\underline{k},2p-1)\ \overline{\lambda}_{(i,k_1)}(\underline{k},2p)\\
&\operatorname{Ad} u^*(g_i^{-1},g_i)u(g_i^{-1},g_{k_1})\begin{pmatrix} u(alt(i,k_1),alt(k_1,\cdots,k_{2p-2},k_{2p-1},k_{2p-1}))\\ \cdot\, u^*(alt(i,k_1,k_1,k_2,\cdots,k_{2p-1}),g_{k_{2p-1}})\\ \cdot\, u(alt(i,k_1,k_1,k_2,\cdots,k_{2p-1}),g_{k_{2p}})\\ \cdot\, u^*(alt(i,k_1),alt(k_1,\cdots,k_{2p}))\end{pmatrix}\\
=&\overline{\lambda}_{(i,k_1)}(\underline{k},2p-1)\ \overline{\lambda}_{(i,k_1)}(\underline{k},2p)\\
&\begin{cases}\operatorname{Ad} u^*(g_i^{-1},g_i)u(g_i^{-1},g_{k_1})\begin{pmatrix} u(alt(i,k_1),e)\ u^*(g_i^{-1},g_{k_1})\\ \cdot\, v_2(i,k_2,\cdots,k_{2n-1},i)\\ \cdot\, u^*(alt(i,k_1),alt(k_1,k_2))\end{pmatrix} & \text{if } p=1\\[2ex]
\operatorname{Ad} u^*(g_i^{-1},g_i)u(g_i^{-1},g_{k_1})\begin{pmatrix} u(alt(i,k_1),alt(k_1,\cdots,k_{2p-2}))\\ \cdot\, v_{2p-1}(i,k_2,\cdots,k_{2n-1},i)\\ \cdot\, v_{2p}(i,k_2,\cdots,k_{2n-1},i)\\ \cdot\, u^*(alt(i,k_1),alt(k_1,\cdots,k_{2p}))\end{pmatrix} & \text{if } 2\le p\le n-1\\[2ex]
\operatorname{Ad} u^*(g_i^{-1},g_i)u(g_i^{-1},g_{k_1})\begin{pmatrix} u(alt(i,k_1),alt(k_1,\cdots,k_{2n-2}))\\ \cdot\, v_{2n-1}(i,k_2,\cdots,k_{2n-1},i)\\ \cdot\, u(g_i^{-1},g_{k_1})\ u^*(alt(i,k_1),e)\end{pmatrix} & \text{if } p=n\end{cases}\\
\overset{\text{def}}{=}&W_p.
\end{aligned}
$$

Thus,

$$
\begin{aligned}
&alt_\alpha(i,k_1)(v(\underline{k}))=W_1\cdots W_n=\overline{\lambda}_{(i,k_1)}(\underline{k})\cdot\\
&\operatorname{Ad} u^*(g_i^{-1},g_i)u(g_i^{-1},g_{k_1})\begin{pmatrix} u^*(g_i^{-1},g_{k_1})\\ \cdot\, v_2(i,k_2,\cdots,k_{2n-1},i)\cdots v_{2n-1}(i,k_2,\cdots,k_{2n-1},i)\\ \cdot\, u(g_i^{-1},g_{k_1})\end{pmatrix}\\
=&\ \overline{\lambda}_{(i,k_1)}(\underline{k})u^*(g_i^{-1},g_i)v_2(i,k_2,\cdots,k_{2n-1},i)\cdots v_{2n-1}(i,k_2,\cdots,k_{2n-1},i)u(g_i^{-1},g_i)\\
=&\ \overline{\lambda}_{(i,k_1)}(\underline{k})\ v(i,k_2,\cdots,k_{2n-1},i).
\end{aligned}
$$

□

Let us recall the following well-known fact about isomorphisms of two planar algebras which will be used in the next theorem ([**11**]). Let P^1 and P^2 be two planar algebras. Then $P^1 \cong P^2$ (as planar algebras) if and only if there exist a vector space isomorphism $\psi : P^1_n \to P^2_n$ such that:

(i) ψ preserves the filtered algebra structure,

(ii) ψ preserves the actions of all Jones projection tangles and the (two types of) conditional expectation tangles.

If P^1 and P^2 are $*$-planar algebras, then we require ψ to be $*$-preserving.

Theorem 4.5. *The planar algebra P^{sf} associated to the diagonal subfactor obtained from a II_1 factor N and a finite collection of automorphisms $\alpha_i \in Aut\,N$ for $i \in I$, is isomorphic to $P^{\langle g_i : i\in I\rangle,\omega}$, where $g_i = [\alpha_i] \in Out\,N$ for all $i \in I$, and ω is the normalized 3-cocycle associated to $G = \langle g_i : i \in I\rangle \subseteq Out\,N$ as above.*

Proof: Let $G = \langle g_i : i \in I\rangle$ and without loss of generality, let us assume that α is a lift of G such that $\alpha_i = \alpha_{g_i}$. By [**11**] we have $P^{sf}_n = N' \cap M_{n-1}$ for all $n \ge 0$.

Define the map $\phi : P^{sf} \to P \stackrel{\text{def}}{=} P^{\langle g_i : i \in I \rangle, \omega}$ by first defining it on basis elements as $\phi(v^*(\underline{i}, \widetilde{\underline{j}})(E_{\underline{i},\underline{j}} \otimes 1)) = (\underline{i}, \widetilde{\underline{j}})$ for all $\underline{i}, \underline{j} \in I^n$ such that $alt(\underline{i}, \widetilde{\underline{j}}) = e$, and then extending it linearly. Clearly, ϕ is a vector space isomorphism. We will show that ϕ is $*$-planar algebra isomorphism. We make first the following observation:

For $i \in I$, $\underline{i} = (i_1, \cdots, i_n) \in I^n$ and $0 \leq s \leq n$, we have the identity $v(\underline{i}) = v(i_1, \cdots, i_s, i, i, i_{s+1}, \cdots, i_n)$. The proof is similar to that of Lemma 3.1. Thus, if $\underline{i}$ is a sequence of indices with non-crossing matched pairings, then using this identity several times to reduce all consecutive matched pairings, we get $v(\underline{i}) = 1$.

We show now that ϕ is indeed a planar algebra isomorphism following the remark just before the theorem.

(a) ϕ *is unital*: Since $1_{P_n^{sf}} = \sum_{\underline{i} \in I^n} (E_{\underline{i},\underline{i}} \otimes 1)$ and $v(\underline{i}, \widetilde{\underline{i}}) = 1$ by the above observation, we get $\phi(1_{P_n^{sf}}) = \sum_{\underline{i} \in I^n} (\underline{i}, \widetilde{\underline{i}}) = 1_{P_n}$.

(b) ϕ *preserves Jones projection tangles*: By Theorem 4.2.1 in [**11**], the n-th Jones projection tangle E_n acts as $Z^{P^{sf}}_{E_n} = |I| e_n = \sum_{\underline{k} \in I^{n-1}, i,j \in I} (E_{(\underline{k},i,i),(\underline{k},j,j)} \otimes 1)$. Since $(\underline{k}, i, i, j, j, \widetilde{\underline{k}})$ is a sequence of indices with non-crossing matched pairings, by the above note $v(\underline{k}, i, i, j, j, \widetilde{\underline{k}}) = 1$. So, $\phi(Z^{P^{sf}}_{E_n}) = \sum_{\underline{k} \in I^{n-1}, i,j \in I} (\underline{k}, i, i, j, j, \widetilde{\underline{k}}) = Z^P_{E_n}$.

(c) ϕ *preserves the action of the conditional expectation tangle*: By Theorem 4.2.1 in [**11**], the action of the conditional expectation tangle $\mathcal{E}_n^{n+1}$ is given by $Z^{P^{sf}}_{\mathcal{E}_n^{n+1}} = |I| \, \mathbb{E}^{M_n}_{M_{n-1}}|_{N' \cap M_n}$. We compute

$$\begin{aligned} Z^{P^{sf}}_{\mathcal{E}_n^{n+1}}(v^*(\underline{i}, k, l, \widetilde{\underline{j}})(E_{(\underline{i},k),(\underline{j},l)} \otimes 1)) &= v^*(\underline{i}, k, l, \widetilde{\underline{j}}) \mathbb{E}^{M_n}_{M_{n-1}}(E_{(\underline{i},k),(\underline{j},l)} \otimes 1) \\ &= \delta_{k,l} \; v^*(\underline{i}, k, k, \widetilde{\underline{j}})(E_{\underline{i},\underline{j}} \otimes 1) \\ &= \delta_{k,l} \; v^*(\underline{i}, \widetilde{\underline{j}})(E_{\underline{i},\underline{j}} \otimes 1) \stackrel{\phi}{\longmapsto} \delta_{k,l} \; (\underline{i}, \widetilde{\underline{j}}) \in P_n \end{aligned}$$

for all $\underline{i}, \underline{j} \in I^n$, $k, l \in I$ such that $alt(\underline{i}, k, l, \widetilde{\underline{j}}) = e$. From the action of tangles defined in Section 3, it is easy to check $Z^P_{\mathcal{E}_n^{n+1}}(\underline{i}, k, l, \widetilde{\underline{j}}) = \delta_{k,l} \; (\underline{i}, \widetilde{\underline{j}})$.

(d) ϕ *preserves* $*$: Applying $*$ on a basis element of P_n^{sf}, we get

$$\begin{aligned} \Big(v^*(\underline{i}, \widetilde{\underline{j}})(E_{\underline{i},\underline{j}} \otimes 1)\Big)^* &= (E_{\underline{j},\underline{i}} \otimes 1) v(\underline{i}, \widetilde{\underline{j}}) = \big(alt_\alpha(\underline{j}) alt_\alpha^{-1}(\underline{i})\big)\,(v(\underline{i}, \widetilde{\underline{j}}))(E_{\underline{j},\underline{i}} \otimes 1) \\ &= alt_\alpha(\underline{j}, \widetilde{\underline{i}})(v(\underline{i}, \widetilde{\underline{j}}))(E_{\underline{j},\underline{i}} \otimes 1) \\ &= v(\underline{j}, \widetilde{\underline{i}}) \; v(\underline{i}, \widetilde{\underline{j}}) \; v^*(\underline{j}, \widetilde{\underline{i}})(E_{\underline{j},\underline{i}} \otimes 1) \\ &= v^*(\underline{j}, \widetilde{\underline{i}})(E_{\underline{j},\underline{i}} \otimes 1) \stackrel{\phi}{\longmapsto} (\underline{j}, \widetilde{\underline{i}}) = (\underline{i}, \widetilde{\underline{j}})^* \end{aligned}$$

for $\underline{i}, \underline{j} \in I^n$ such that $alt(\underline{i}, \widetilde{\underline{j}}) = e$. Note that we used Lemma 4.1 for the third equality and Lemma 4.2 for the fourth one.

(e) ϕ *preserves multiplication*: Suppose $\underline{i}, \underline{j}, \underline{k}, \underline{l} \in I^n$ such that $alt(\underline{i}, \widetilde{\underline{j}}) = e = alt(\underline{k}, \widetilde{\underline{l}})$. Then,

$$\begin{aligned}
&\left(v^*(\underline{i}, \widetilde{\underline{j}})(E_{\underline{i},\underline{j}} \otimes 1)\right) \cdot \left(v^*(\underline{k}, \widetilde{\underline{l}})(E_{\underline{k},\underline{l}} \otimes 1)\right) \\
&= v^*(\underline{i}, \widetilde{\underline{j}}) \left(alt_\alpha(\underline{i}, \widetilde{\underline{j}})\right) (v^*(\underline{k}, \widetilde{\underline{l}}))(E_{\underline{i},\underline{j}} \otimes 1)(E_{\underline{k},\underline{l}} \otimes 1) \\
&= v^*(\underline{i}, \widetilde{\underline{j}}) v(\underline{i}, \widetilde{\underline{j}}) v^*(\underline{k}, \widetilde{\underline{l}})) v^*(\underline{i}, \widetilde{\underline{j}}) \delta_{\underline{j},\underline{k}} \; (E_{\underline{i},\underline{l}} \otimes 1) \;\; \text{(using Lemma 4.1)} \\
&= \delta_{\underline{j},\underline{k}} \left(v(\underline{i}, \widetilde{\underline{j}}) v(\underline{j}, \widetilde{\underline{l}}))\right)^* (E_{\underline{i},\underline{l}} \otimes 1) = \delta_{\underline{j},\underline{k}} \; v^*(\underline{i}, \widetilde{\underline{l}})(E_{\underline{i},\underline{l}} \otimes 1) \;\; \text{(using Lemma 4.3)}
\end{aligned}$$

On the other hand, one can easily deduce from the action of the multiplication tangle in P that $(\underline{i}, \widetilde{\underline{j}}) \cdot (\underline{k}, \widetilde{\underline{l}}) = \delta_{\underline{j},\underline{k}} \; (\underline{i}, \widetilde{\underline{l}})$.

(f) ϕ *preserves the action of the left conditional expectation tangle*: By Theorem 4.2.1 in [**11**], the action of the left conditional expectation tangle $\mathcal{E}'_n$ is given by $Z^{P^{sf}}_{\mathcal{E}'_n} = |I| \; \mathbb{E}^{N' \cap M_{n-1}}_{M' \cap M_{n-1}}$. Using the basis of M over N mentioned before, the conditional expectation onto $M' \cap N_{n-1}$ can be expressed as (see [**2**])

$$\begin{aligned}
\mathbb{E}^{N' \cap M_{n-1}}_{M' \cap M_{n-1}}(x) &= |I|^{-2} \sum_{i,j \in I} \left(\sqrt{|I|}(E_{i,j} \otimes 1)\right) x \left(\sqrt{|I|}(E_{j,i} \otimes 1)\right) \\
&= |I|^{-1} \sum_{i,j \in I} (E_{i,j} \otimes 1) x (E_{j,i} \otimes 1)
\end{aligned}$$

for $x \in N' \cap M_{n-1}$. Hence, for $\underline{i} = (i_1, \cdots, i_{n-1})$, $\underline{j} = (j_1, \cdots, j_{n-1}) \in I^{n-1}$ and $k, l \in I$ such that $alt(k, \underline{i}, \widetilde{\underline{j}}, l) = e$, we have

$$\begin{aligned}
&Z^{P^{sf}}_{\mathcal{E}'_n} \left(v^*(k, \underline{i}, \widetilde{\underline{j}}, l)(E_{(k,\underline{i}),(l,\underline{j})} \otimes 1)\right) \\
&= \sum_{i,j \in I} (E_{i,j} \otimes 1) v^*(k, \underline{i}, \widetilde{\underline{j}}, l)(E_{(k,\underline{i}),(l,\underline{j})} \otimes 1)(E_{j,i} \otimes 1) \\
&= \sum_{i,j \in I} alt_\alpha(i,j)(v^*(k, \underline{i}, \widetilde{\underline{j}}, l))(E_{i,j} \otimes 1)(E_{(k,\underline{i}),(l,\underline{j})} \otimes 1)(E_{j,i} \otimes 1) \\
&= \delta_{k,l} \sum_{i \in I} alt_\alpha(i,k)(v^*(k, \underline{i}, \widetilde{\underline{j}}, k))(E_{(i,\underline{i}),(i,\underline{j})} \otimes 1) \\
&= \delta_{k,l} \sum_{i \in I} \lambda_{i,k}(k, \underline{i}, \underline{j}, k) v^*(i, \underline{i}, \widetilde{\underline{j}}, i)(E_{(i,\underline{i}),(i,\underline{j})} \otimes 1) \;\; \text{(using Lemma 4.4)}
\end{aligned}$$

On the other hand, one can easily check that the action of $\mathcal{E}'_n$ is given by

$$Z^P_{\mathcal{E}'_n}(k, \underline{i}, \widetilde{\underline{j}}, l) = \delta_{k,l} \sum_{i \in I} \lambda_{i,k}(k, \underline{i}, \underline{j}, k) \; (i, \underline{i}, \widetilde{\underline{j}}, i).$$

□

Corollary 4.6. *Given a group G generated by a finite collection g_i for $i \in I$ and given a normalized 3-cocycle $\omega \in Z^3(G, \mathbb{T})$, there exists a hyperfinite subfactor whose associated planar algebra is isomorphic to $P^{\langle g_i : i \in I \rangle, \omega}$.*

Proof: The proof follows from [**9**] and Theorem 4.5. □

Remark 4.7. Note that the isomorphism ϕ in the proof of Theorem 4.5 uses the 3-cocycle ω only in the step involving the conditional expectation onto the commutant of M. In particular, the filtered $*$-algebra structure does not involve ω.

Analyzing the filtered $*$-algebra structure of our planar algebra, one can easily find that the principal graph Γ of $N \subset M$ is a Cayley-like graph. More precisely, if $G_n = \{alt(\underline{i}) : \underline{i} \in I^n\}$ for $n \geq 1$, and $G_0 = \{e\}$, then $V_n(\Gamma) = G_n \setminus G_{n-2}$ denotes the set of vertices of Γ which are at a distance n from the distinguished vertex for $n \geq 1$, and $V_0(\Gamma) = \{e\}$. The number of edges between $g \in V_n(\Gamma)$ and $h \in V_{n+1}(\Gamma)$ is $\sum_{i \in I} \delta_{g,hg_i}$ (resp. $\sum_{i \in I} \delta_{gg_i,h}$) if n is odd (resp. even). Note that this is well known.

The most elegant feature of the planar algebra $P^{\langle g_i : i \in I \rangle, \omega}$ is that the distinguished basis forms the 'loop-basis' of the filtered $*$-algebra arising from paths on the principal graph. Note that the 3-cocycle ω does not enter in the definition of the actions of multiplication, inclusion and unit tangles (defined in Section 3) or in the $*$-operation. Of course, we found the abstract planar algebra by first computing the action of tangles on the relative commutants. We then deduced from it an abstract prescription of the planar algebra associated to a G-kernel and a 3-cocycle, which is the one presented in Section 3.

The path algebra associated to the principal graph can always be used to obtain a description of the filtered $*$-algebra structure of a subfactor planar algebra (see for instance [**13**]). The extra information encoded in the planar algebra, which the principal graph cannot provide, is the action of the left conditional expectation tangle (or equivalently, the rotation tangle or the left-inclusion tangle with an even number of strings). The main issue in this paper was the choice of the unitaries $v(\underline{i})$ satisfying the conclusion of Lemma 4.1 in such a way that the basis elements of $N' \cap M_n$ correspond to loops on the principal graph with the correspondence extending to a filtered $*$-algebra isomorphism. Note that this choice of $v(\underline{i})$ is unique up to a scalar in $\mathbb{T}$. It is a delicate choice in the sense that another choice would very likely make the 3-cocycle ω appear in the description of the filtered $*$-algebra structure, whereas ω does not feature in the path algebra on the principal graph Γ.

We will prove next a converse of Theorem 4.5. We will refer to the abstract planar algebra defined in Section 3 as *diagonal planar algebra.*

THEOREM 4.8. *Any finite index extremal subfactor $N \subset M$ whose standard invariant is given by a diagonal planar algebra is a diagonal subfactor. Moreover, if the associated group and its generators of the diagonal planar algebra is given by G and $\{g_i : i \in I\}$ repectively, then for every $i_0 \in I$, there exists $\alpha_i \in Aut N$, $i \in I$, such that:*

(i) $(N \subset M) \cong (N \hookrightarrow M_I \otimes N)$ where $N \hookrightarrow M_I \otimes N$ is the diagonal subfactor with respect to the automorphisms α_i for $i \in I$.

(ii) There exists a group isomorphism $\psi : \langle \alpha_i : i \in I \rangle_{Out\, N} \to \langle g_{i_0}^{-1} g_i : i \in I \rangle \leq G$ sending α_i to $g_{i_0}^{-1} g_i$ for all $i \in I$.

Proof: Let P be the planar algebra associated to $N \subset M$, P^{Δ} be a diagonal planar algebra associated to G and $\{g_i : i \in I\}$, and $\phi : P^{\Delta} \to P$ a $*$-planar algebra isomorphism.

Setting up matrix units:
For all $i, j \in I$ such that $g_i = g_j$, set $q^i_j := \phi((i,j)) \in P_1 = N' \cap M$. Note that

$\sum\limits_{i\in I} q^i_i = 1$. So, to create other off-diagonal matrix units q^i_j, we partition $I = \coprod\limits_{n=0}^{m} I_n$ such that:

(i) $i_0 \in I_0$,

(ii) $g_i = g_j \Leftrightarrow i, j \in I_n$ for some $0 \le n \le m$.

For each $n \in \{1, 2, \cdots, m\}$, choose $i_n \in I^n$ and a partial isometry $q^{i_0}_{i_n} \in M$ such that $q^{i_0}_{i_n}\left(q^{i_0}_{i_n}\right)^* = q^{i_0}_{i_0}$ and $\left(q^{i_0}_{i_n}\right)^* q^{i_0}_{i_n} = q^{i_n}_{i_n}$. (Note that $q^{i_0}_{i_0} = \phi((i_0, i_0))$ and $q^{i_n}_{i_n} = \phi((i_n, i_n))$ have the same trace $|I|^{-1}$). Extend q by defining $q : I \times I \to M$ by

$$I \times I \supset I^s \times I^t \ni (i, j) \overset{q}{\mapsto} q^i_j := q^i_{i_s}\left(q^{i_0}_{i_s}\right)^* q^{i_0}_{i_t} q^{i_t}_j = \phi((i, i_s))\left(q^{i_0}_{i_s}\right)^* q^{i_0}_{i_t}\phi((i_t, j)) \in M$$

It is completely routine to check (using properties of partial isometry and action of multiplication tangle in P^{Δ}) that (i) q is well-defined, (ii) $q^i_j q^k_l = \delta_{j,k} q^i_l$, and (iii) $\left(q^i_j\right)^* = q^j_i$.

Finding automorphisms:
Using extremality of $N \subset M$, for each $i \in I$, we get

$$\begin{aligned}
[q^i_i M q^i_i : N q^i_i] = 1 &\Rightarrow N q^i_i = q^i_i M q^i_i \\
&\Rightarrow N q^i_j \subset q^i_i M q^j_j \subset q^i_j N \text{ and } N q^i_j \supset q^i_i M q^j_j \supset q^i_j N \\
&\Rightarrow N q^i_j = q^i_i M q^j_j = q^i_j N \Rightarrow M \cong \bigoplus_{i,j\in I} q^i_j N
\end{aligned}$$

where we use $[q^i_i, N] = 0$ in the second implication.

For each $i \in I$, define $\alpha_i : N \to N$ by $q^{i_0}_i x = \alpha_i(x) q^{i_0}_i$ for all $x \in N$. Since N is a factor and $[q^j_j, N] = 0$ for $j \in I$, α_i is well-defined and injective; surjectivity follows from $N q^i_j = q^i_j N$. Linearity and homomorphism property of α_i follow immediately, and we also have the identity $x q^{i_0}_i = q^{i_0}_i \alpha_i^{-1}(x)$. To show $q^i_j x = \left(\alpha_i^{-1}\alpha_j\right)(x) q^i_j$, note that

$$\begin{aligned}
&q^i_{i_0} x = q^i_{i_0} x q^{i_0}_i q^i_{i_0} = q^i_i \alpha_i^{-1}(x) q^i_{i_0} = \alpha_i^{-1}(x) q^i_{i_0} \\
\Rightarrow\ &q^i_j x = q^i_{i_0}\alpha_j(x) q^{i_0}_j = \left(\alpha_i^{-1}\alpha_j\right)(x) q^i_j, \text{ for all } x \in N, \text{ for all } i, j \in I.
\end{aligned}$$

Using this relation, it is easy to show that α_i is $*$-preserving. Two other easy consequences are $\alpha_i = \alpha_j$ if and only if $g_i = g_j$ and $\alpha_{i_0} = \mathrm{id}_N$.

Structure of diagonal subfactor:
Define:

(i) $\kappa : \tilde{M} := M_I \otimes N \to M$ by $\kappa(E_{i,j} \otimes x) = q^i_j \alpha_j^{-1}(x) = \alpha_i^{-1}(x) q^i_j$,

(ii) $\lambda : M \to \tilde{M}$ by $\lambda(x) = \sum\limits_{i,j\in I} E_{i,j} \otimes \lambda_{i,j}(x)$,

where $\lambda_{i,j} : M \to N$ is the map given by the relation $q^i_i x q^j_j = q^i_j \alpha_j^{-1}(\lambda_{i,j}(x))$ for all $x \in M$.

Clearly, $\kappa \circ \lambda = \mathrm{id}_M$, $\lambda \circ \kappa = \mathrm{id}_{\tilde{M}}$ and κ is a $*$-isomorphism. Set $\tilde{N} := \lambda(N) = \left\{ \sum\limits_{i\in I} E_{i,i} \otimes \alpha_i(x) \,\middle|\, x \in N \right\} \subset \tilde{M}$.

This proves that $N \subset M$ is a diagonal subfactor as claimed. The rest of the proof pertains to 4.8 (ii).

Matrix units for the tower of the basic construction:
Let M_n denote the II_1 factor obtained from $N \subset M$ by iterating the basic construction n times. We will first define $q^{\underline{i}}_{\underline{j}} \in M_{n-1}$ for $\underline{i}, \underline{j} \in I^n$ and $n \geq 2$ satisfying:

(i) $\left(q^{\underline{i}}_{\underline{j}}\right)^* = q^{\underline{j}}_{\underline{i}}$ for all $\underline{i}, \underline{j} \in I^n$,

(ii) $q^{\underline{i}}_{\underline{j}}\, x \;=\; alt_\alpha(\underline{i}, \tilde{\underline{j}})(x)\; q^{\underline{i}}_{\underline{j}}$ for all $\underline{i}, \underline{j} \in I^n$ and $x \in N$,

(iii) $q^{\underline{i}}_{\underline{i}} = \phi((\underline{i}, \tilde{\underline{i}}))$ for all $\underline{i} \in I^n$,

(iv) $q^{\underline{i}}_{\underline{j}}\, q^{\underline{k}}_{\underline{l}} \;=\; \delta_{\underline{j},\underline{k}}\, q^{\underline{i}}_{\underline{l}}$ for all $\underline{i}, \underline{j}, \underline{k}, \underline{l} \in I^n$,

(v) $q^{(\underline{s},s)}_{(\underline{t},t)}\, q^{\underline{u}}_{\underline{v}} \;=\; \delta_{\underline{t},\underline{u}}\, q^{(\underline{s},s)}_{(\underline{v},t)}$ for all $\underline{s}, \underline{t}, \underline{u}, \underline{v} \in I^{n-1}$ and $s, t \in I$

by induction on n, where alt_α is defined by $alt_\alpha(i_1, \cdots, i_m) = \alpha^{-1}_{i_1}\alpha_{i_2}\alpha^{-1}_{i_3}\cdots\alpha^{(-1)^m}_{i_m}$.

Suppose we have defined such $q^{\underline{i}}_{\underline{j}}$'s for all $\underline{i}, \underline{j} \in I^m$ and $m \leq n$. Now, for $i, j \in I$ and $\underline{i}, \underline{j} \in I^n$, set $q^{(\underline{i},i)}_{(\underline{j},j)} := q^{\underline{i}}_{(\underline{s},i)}\; E_n\; q^{(\underline{s},j)}_{\underline{j}} \in M_n$ for some $\underline{s} \in I^{n-1}$, where $E_n = |I|e_n$ is the element in M_n given by the n-th Jones projection tangle. To show that the definition of $q^{(\underline{i},i)}_{(\underline{j},j)}$ is independent of $\underline{s} \in I^{n-1}$, observe that

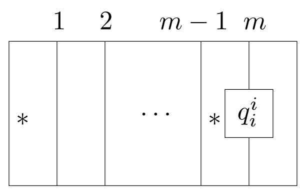
$$q^{(\underline{i},i)}_{(\underline{j},j)} = q^{\underline{i}}_{(\underline{s},i)}\; \phi(\underline{s}, \tilde{\underline{s}})\; E_n\; q^{(\underline{s},j)}_{\underline{j}} = q^{\underline{i}}_{(\underline{s},i)}\; q^{\underline{s}}_{\underline{t}}\; E_n\; q^{\underline{t}}_{\underline{s}}\; q^{(\underline{s},j)}_{\underline{j}} = q^{\underline{i}}_{(\underline{t},i)}\; E_n\; q^{(\underline{t},j)}_{\underline{j}}$$

for all $\underline{t} \in I^{n-1}$. Properties (i) and (v) hold trivially. For (ii), note that

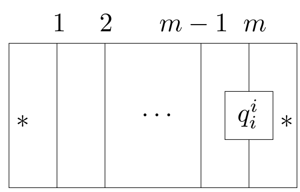
$$q^{(\underline{i},i)}_{(\underline{j},j)}\; x = alt_\alpha(\underline{i}, i, \tilde{\underline{s}}, \underline{s}, j, \tilde{\underline{j}})(x)\; q^{(\underline{i},i)}_{(\underline{j},j)} = alt_\alpha((\underline{i}, i), \widetilde{(\underline{j}, j)})(x)\; q^{(\underline{i},i)}_{(\underline{j},j)} \text{ for all } x \in N.$$

Next, we prove property (iv). For $i \in I$ and $m \geq 1$, let $\eta_m(i)$ denote the element

$\eta_m(i) =$



if m is odd



if m is even

in $M'_{m-2} \cap M_{m-1}$ (where $M_{-1} := N$, $M_0 := M$). Two important relations which we will often use, are $\eta_m(i)E_m = \eta_{m+1}(i)E_m$ and $E_m\eta_m(i) = E_m\eta_{m+1}(i)$. Getting back to property (iv), we have

$$\begin{aligned}
q^{(\underline{i},i)}_{(\underline{j},j)}\; q^{(\underline{k},k)}_{(\underline{l},l)} &= q^{\underline{i}}_{(\underline{s},i)}\; E_n\; q^{(\underline{s},j)}_{\underline{j}}\; q^{\underline{k}}_{(\underline{s},k)}\; E_n\; q^{(\underline{s},l)}_{\underline{l}} = \delta_{\underline{j},\underline{k}}\; q^{\underline{i}}_{(\underline{s},i)}\; E_n\; q^{(\underline{s},j)}_{(\underline{s},k)}\; E_n\; q^{(\underline{s},l)}_{\underline{l}} \\
&= \delta_{\underline{j},\underline{k}}\; q^{\underline{i}}_{(\underline{s},i)}\; E_n\; \eta_n(j)\; q^{(\underline{s},j)}_{(\underline{s},k)}\; E_n\; q^{(\underline{s},l)}_{\underline{l}} = \delta_{\underline{j},\underline{k}}\; q^{\underline{i}}_{(\underline{s},i)}\; E_n\; \eta_{n+1}(j)\; q^{(\underline{s},j)}_{(\underline{s},k)}\; E_n\; q^{(\underline{s},l)}_{\underline{l}} \\
&= \delta_{\underline{j},\underline{k}}\; q^{\underline{i}}_{(\underline{s},i)}\; E_n\; q^{(\underline{s},j)}_{(\underline{s},k)}\; \eta_{n+1}(j)\; E_n\; q^{(\underline{s},l)}_{\underline{l}} = \delta_{\underline{j},\underline{k}}\; q^{\underline{i}}_{(\underline{s},i)}\; E_n\; q^{(\underline{s},j)}_{(\underline{s},k)}\; \eta_n(j)\; E_n\; q^{(\underline{s},l)}_{\underline{l}} \\
&= \delta_{\underline{j},\underline{k}}\; \delta_{j,k}\; q^{\underline{i}}_{(\underline{s},i)}\; E_n\; q^{(\underline{s},j)}_{(\underline{s},j)}\; E_n\; q^{(\underline{s},l)}_{\underline{l}} = \delta_{\underline{j},\underline{k}}\; \delta_{j,k}\; q^{\underline{i}}_{(\underline{s},i)}\; E_n\; \phi((\underline{s}, j, j, \tilde{\underline{s}}))\; E_n\; q^{(\underline{s},l)}_{\underline{l}} \\
&= \delta_{\underline{j},\underline{k}}\; \delta_{j,k}\; q^{\underline{i}}_{(\underline{s},i)}\; \phi((\underline{s}, \tilde{\underline{s}}))\; E_n\; q^{(\underline{s},l)}_{\underline{l}} = \delta_{\underline{j},\underline{k}}\; \delta_{j,k}\; q^{\underline{i}}_{(\underline{s},i)}\; E_n\; q^{(\underline{s},l)}_{\underline{l}} = \delta_{\underline{j},\underline{k}}\; \delta_{j,k}\; q^{(\underline{i},i)}_{(\underline{l},l)}.
\end{aligned}$$

It remains to establish property (iii). Now, for $\underline{i} \in I^n$, $i \in I$ and $\underline{s} \in I^{n-1}$,

$$\begin{aligned}
&\phi((\underline{i}, i, i, \tilde{\underline{i}}))\ q^{(\underline{i},i)}_{(\underline{i},i)}\ \phi((\underline{i}, i, i, \tilde{\underline{i}})) = \phi((\underline{i}, \tilde{\underline{i}}))\ \eta_{n+1}(i)\ q^{i}_{\underline{s},i}\ E_n\ q^{\underline{s},i}_{\underline{i}}\ \eta_{n+1}(i)\ \phi((\underline{i}, \tilde{\underline{i}})) \\
&= \phi((\underline{i}, \tilde{\underline{i}}))\ q^{i}_{\underline{s},i}\ \eta_{n+1}(i)\ E_n\ \eta_{n+1}(i)\ q^{\underline{s},i}_{\underline{i}}\ \phi((\underline{i}, \tilde{\underline{i}})) = q^{i}_{\underline{s},i}\ \eta_n(i)\ E_n\ \eta_n(i)\ q^{\underline{s},i}_{\underline{i}} \\
&= q^{i}_{\underline{s},i}\ E_n\ q^{\underline{s},i}_{\underline{i}} = q^{(\underline{i},i)}_{(\underline{i},i)}.
\end{aligned}$$

Since $q^{(\underline{i},i)}_{(\underline{i},i)} \in \mathcal{P}(N' \cap M_n)$ (by property (ii)) and $\phi((\underline{i}, i, i, \tilde{\underline{i}}))$ is a minimal projection of $P_{n+1} = N' \cap M_n$, therefore $q^{(\underline{i},i)}_{(\underline{i},i)} = \phi((\underline{i}, i, i, \tilde{\underline{i}}))$.

The proof for the initial case of $n = 2$ is similar and is left to reader.

Correspondence between relations satisfied by α_i's and g_i's:
In this part, we will prove that for $\underline{i}, \underline{j} \in I^n$, $alt(\underline{i}, \tilde{\underline{j}}) = e$ if and only if $alt_\alpha(\underline{i}, \tilde{\underline{j}}) \in \operatorname{Int} N$.

Following the construction of the isomorphism λ between M and $\tilde{M}$, one can define an isomorphism $\lambda^{(n)}$ as follows:

$$M_{n-1} \ni x \overset{\lambda^{(n)}}{\longmapsto} \lambda^{(n)}(x) = \sum_{\underline{k},\underline{l} \in I^n} E_{\underline{k},\underline{l}} \otimes \lambda^{(n)}_{\underline{k},\underline{l}}(x) \in \tilde{M}_{n-1} := M_{I^n} \otimes N$$

where $\lambda^{(n)}_{\underline{k},\underline{l}} : M_{n-1} \to N$ is the map given by the relation

$$q^{k}_{\underline{k}}\ x\ q^{l}_{\underline{l}} = q^{k}_{\underline{l}}\ alt_\alpha(\underline{l}) \left(\lambda^{(n)}_{\underline{k},\underline{l}}(x)\right).$$

Thus, $\lambda^{(n)}(N) = \left\{ \sum_{\underline{k} \in I^n} E_{\underline{k},\underline{k}} \otimes alt^{-1}_\alpha(\underline{k})(x)\ :\ x \in N \right\}$, for $n \geq 1$. Note that $\lambda^{(1)} = \lambda$.

Let $alt(\underline{i}, \tilde{\underline{j}}) = e$ for $\underline{i}, \underline{j} \in I^n$. Note that $\lambda^{(n)}(q^{i}_{\underline{j}}) = E_{\underline{i},\underline{j}} \otimes 1$ and $\lambda^{(n)}\left(\phi((\underline{i}, \tilde{\underline{j}}))\right)$ are partial isometries between $\lambda^{(n)}(q^{i}_{\underline{i}}) = \lambda^{(n)}\left(\phi((\underline{i}, \tilde{\underline{i}}))\right) = E_{\underline{i},\underline{i}} \otimes 1$ and $\lambda^{(n)}(q^{j}_{\underline{j}}) = \lambda^{(n)}\left(\phi((\underline{j}, \tilde{\underline{j}}))\right) = E_{\underline{j},\underline{j}} \otimes 1$. So, there exists $u \in \mathcal{U}(N)$ such that $\lambda^{(n)}\left(\phi((\underline{i}, \tilde{\underline{j}}))\right) = E_{\underline{i},\underline{j}} \otimes u = \lambda^{(n)}\left(q^{i}_{\underline{j}}\ v\right)$ where $v = alt_\alpha(\underline{j})(u) \in \mathcal{U}(N)$. Hence,

$$\begin{aligned}
&q^{i}_{\underline{j}}\ v = \phi((\underline{i}, \tilde{\underline{j}})) \in N' \cap M_{n-1} \\
\Rightarrow\ &y\ q^{i}_{\underline{j}}\ v = q^{i}_{\underline{j}}\ v\ y = \left(alt_\alpha((\underline{i}, \tilde{\underline{j}})) \circ \operatorname{Ad} v\right)(y)\ q^{i}_{\underline{j}}\ v \ \text{ for all } y \in N \\
\Rightarrow\ &\left(alt_\alpha((\underline{i}, \tilde{\underline{j}})) \circ \operatorname{Ad} v\right)(y) = y \ \text{ for all } y \in N \\
\Rightarrow\ &alt_\alpha((\underline{i}, \tilde{\underline{j}})) \in \operatorname{Int} N.
\end{aligned}$$

Conversely, if $alt_\alpha((\underline{i}, \tilde{\underline{j}})) \in \operatorname{Int} N$, that is, $alt_\alpha((\underline{i}, \tilde{\underline{j}})) \circ \operatorname{Ad} v = \mathrm{id}_N$ for some $v \in \mathcal{U}(N)$, then $\left(E_{\underline{i},\underline{j}} \otimes alt^{-1}_\alpha(\underline{j})(v)\right) \in \left(\left(\lambda^{(n)}(N)\right)' \cap \tilde{M}_{n-1}\right)$. Now, $alt((\underline{i}, \tilde{\underline{j}})) \neq e$ implies $\dim\left(\left(\lambda^{(n)}(N)\right)' \cap \tilde{M}_{n-1}\right) > \dim\left(N' \cap M_{n-1}\right)$ which is a contradiction. Hence, $alt((\underline{i}, \tilde{\underline{j}})) = e$.

The group generated by α_i's:
Let $H := \langle \theta_i = [\alpha_i]_{\text{Out N}} : i \in I \rangle \leq \operatorname{Out} N$, $\tilde{H} := \langle g^{-1}_{i_0} g_i : i \in I \rangle \leq G$, $J :=$

$I \times \{1, -1\}$. Define maps $w_H : \coprod_{n \geq 1} J^n \to H$ and $w_{\tilde{H}} : \coprod_{n \geq 1} J^n \to \tilde{H}$ by

$$J^n \ni ((i_1, \epsilon_1), (i_2, \epsilon_2), \cdots (i_n, \epsilon_n)) \overset{w_H}{\longmapsto} \theta_{i_1}^{\epsilon_1} \theta_{i_2}^{\epsilon_2} \cdots \theta_{i_n}^{\epsilon_n} \in H$$

$$J^n \ni ((i_1, \epsilon_1), (i_2, \epsilon_2), \cdots (i_n, \epsilon_n)) \overset{w_{\tilde{H}}}{\longmapsto} (g_{i_0}^{-1} g_{i_1})^{\epsilon_1} (g_{i_0}^{-1} g_{i_2})^{\epsilon_2} \cdots (g_{i_0}^{-1} g_{i_n})^{\epsilon_n} \in \tilde{H}$$

Define $\gamma : H \to \tilde{H}$ by $\gamma \left(w_H(\underline{j}) \right) = w_{\tilde{H}}(\underline{j})$. For γ to be an isomorphism, it is enough to show that γ is well-defined and injective. Suppose the map $\rho : J \to I^2$ sends $(i, 1)$ (resp. $(i, -1)$) to (i_0, i) (resp. (i, i_0)). Extend ρ to $\rho : J^n \to I^{2n}$ entrywise. Note that $w_H(\underline{j}) = alt_H \left(\rho(\underline{j}) \right)$ and $w_{\tilde{H}}(\underline{j}) = alt \left(\rho(\underline{j}) \right)$. This implies

$$w_H(\underline{j}) = 1_H \Leftrightarrow alt_\alpha \left(\rho(\underline{j}) \right) \in \operatorname{Int} N \Leftrightarrow alt \left(\rho(\underline{j}) \right) = e \Leftrightarrow w_{\tilde{H}}(\underline{j}) = e.$$

Hence, H and $\tilde{H}$ are isomorphic. $\square$

References

[1] D. Bisch, *Entropy of groups and subfactors*, Jour. Funct. Anal. 103 (1992), 190-208.

[2] D. Bisch, *Bimodules, higher relative commutants and the fusion algebra associated to a subfactor*, Operator algebras and their applications, 13-63, Fields Inst. Commun., 13, Amer. Math. Soc., Providence, RI, 1997.

[3] D. Bisch, P. Das, S. Ghosh, *The planar algebra of the group-type subfactors*, J. Funct. Anal. 257 (1), (2009), 20-46.

[4] D. Bisch, V. Jones, *Algebras associated to intermediate subfactors* Invent. Math. 128 (1997), no. 1, 89-157.

[5] , D. Bisch, R. Nicoara, S. Popa, *Continuous families of hyperfinite subfactors with the same standard invariant*, Internat. Journal Math. 18 (2007), no. 3, 255-267.

[6] A. Connes, *Periodic automorphisms of the hyperfinite factor of type II*, Acta Sci. Math. 39 (1977), 39-66.

[7] A. Connes, V. Jones, *Property T for von Neumann algebras*, Bull. London Math. Soc., 17 (1985), 57-62.

[8] S. Ghosh, *Higher exchange relations*, Pacific J. Math. 210 (2003), no. 2, 299–316.

[9] V. Jones, *An invariant for group actions*, Algèbres d'opérateurs (Sém., Les Plans-sur-Bex, 1978), pp. 237–253, Lecture Notes in Math., 725, Springer, Berlin, 1979.

[10] V. Jones, *Index for subfactors*, Invent. Math., 72, (1983), 1-25.

[11] V. Jones, *Planar Algebras*, preprint, arxiv:math/9909027v1.

[12] V. Jones, *Actions of finite groups on the hyperfinite type II_1 factor*, Mem. Amer. Math. Soc., 237 (1980).

[13] V. Jones, V.S. Sunder, *Introduction to subfactors*, LMS Lecture Note Series 234, 1997.

[14] V. Kodiyalam, Z. Landau, V.S. Sunder, *The planar algebra associated to a Kac algebra*, Proc. Indian Acad. Sci. Math. Sci. 113 (2003), no. 1, 15-51.

[15] Z. Landau, *Exchange relation planar algebras*, Geom. Dedicata 95 (2002), 183-214.

[16] S. Morrison, E. Peters, N. Snyder, *Skein theory for the D_{2n} planar algebras*, J. Pure Appl. Alg. 214 (2010), 117-139.

[17] A. Ocneanu, *Actions of discrete amenable groups on von Neumann algebras*, Lecture Notes in Math. 1138, Springer-Verlag, Berlin 1985.

[18] A. Ocneanu, *Quantized groups, string algebras and Galois theory for algebras*, Operator algebras and applications, Vol. 2, 119-172, London Math. Soc. Lecture Note Ser., 136, Cambridge Univ. Press, Cambridge, 1988.

[19] S. Popa, *An axiomatization of the lattice of higher relative commutants of a subfactor*, Invent. Math., 120, (1995), 427-445.

[20] S. Popa, *Classification of amenable subfactors of type II*, Acta Math. 172, (1994), 163-255.

[21] S. Popa, *Classification of subfactors: the reduction to commuting squares*, Invent. Math., 101, (1990), 19-43.

[22] S. Popa, *Some properties of the symmetric enveloping algebra of a subfactor, with applications to amenability and property T*, Doc. Math. 4 (1999), 665-744.

[23] S. Popa, *Sousfacteurs, actions des groupes et cohomologie*, C.R. Acad. Sci. Paris 309 (I) (1989) 771-776.

[24] C. Sutherland, *Cohomology and extensions of von Neumann algebras II*, Publ. RIMS, 16 (1980), 135-176.

Vanderbilt University, Department of Mathematics, SC 1326, Nashville, TN 37240, USA

E-mail address: dietmar.bisch@vanderbilt.edu

Vanderbilt University, Department of Mathematics, SC 1326, Nashville, TN 37240, USA

E-mail address: paramita.das@vanderbilt.edu

Vanderbilt University, Department of Mathematics, SC 1326, Nashville, TN 37240, USA

E-mail address: shamindra.k.ghosh@vanderbilt.edu

Clay Mathematics Proceedings
Volume 11, 2010

K_1-Injectivity for Properly Infinite C*-Algebras

Étienne Blanchard

Dedicated to Alain Connes on the occasion of his 60th birthday

1. Introduction

One of the main tools to classify C*-algebras is the study of its projections and its unitaries. It was proved by Cuntz in [**Cun81**] that if A is a *purely infinite* simple C*-algebra, then the kernel of the natural map for the unitary group $\mathcal{U}(A)$ to the K-theory group $K_1(A)$ is reduced to the connected component $\mathcal{U}^0(A)$, *i.e.* A is K_1*-injective* (see §3). We study in this note a finitely generated C*-algebra, the K_1-injectivity of which would imply the K_1-injectivity of all unital *properly infinite* C*-algebras.

Note that such a question was already considered in [**Blac07**], [**BRR08**].

The author would like to thank H. Lin, R. Nest, M. Rørdam and W. Winter for helpful comments.

2. Preliminaries

Let us first review briefly the theory introduced by Cuntz ([**Cun78**]) of comparison of positive elements in a C*-algebra.

Definition 2.1. ([**Cun78**], [**Rør92**]) Given two positive elements a, b in a C*-algebra A, one says that:
– a is *dominated* by b (written $a \precsim b$) if and only if there is a sequence $\{d_k; k \in \mathbb{N}\}$ in A such that $\|d_k^* b d_k - a\| \to 0$ when $k \to \infty$,
– a is *properly infinite* if $a \neq 0$ and $a \oplus a \precsim a \oplus 0$ in the C*-algebra $M_2(A) := M_2(\mathbb{C}) \otimes A$.

This leads to the following notions of infiniteness for C*-algebras.

Definition 2.2. ([**Cun78**], [**Cun81**], [**KR00**]) A unital C*-algebra A is said to be:
– *properly infinite* if its unit 1_A is properly infinite in A,
– *purely infinite* if all the non-zero positive elements in A are properly infinite in A.

2010 *Mathematics Subject Classification.* Primary: 46L80; Secondary: 46L06, 46L35.
Key words and phrases. K_1-injectivity, Proper Infiniteness, C*-algebras.

REMARK 2.3. Kirchberg and Rørdam have proved in [**KR00**, Theorem 4.16] that a C*-algebra A is purely infinite (in the above sense) if and only if there is no character on the C*-algebra A and any positive element a in A which lies in the closed two-sided ideal generated by another positive element b in A satisfies $a \precsim b$.

The first examples of such C*-algebras were given by Cuntz in [**Cun81**]: For any integer $n \geq 2$, the C*-algebra $\mathcal{T}_n$ is the universal unital C*-algebra generated by n isometries $s_1, \ldots, s_n$ satisfying the relation

$$s_1 s_1^* + \cdots + s_n s_n^* \leq 1 \tag{2.1}$$

Then, the closed two sided ideal in $\mathcal{T}_n$ generated by the *minimal* projection $p_{n+1} := 1 - s_1 s_1^* - \ldots - s_n s_n^*$ is isomorphic to the C*-algebra $\mathcal{K}$ of compact operators on an infinite dimension separable Hilbert space and one has an exact sequence

$$0 \to \mathcal{K} \to \mathcal{T}_n \xrightarrow{\pi} \mathcal{O}_n \to 0\,, \tag{2.2}$$

where the quotient $\mathcal{O}_n$ is a purely infinite *simple* unital nuclear C*-algebra ([**Cun81**]).

REMARK 2.4. A unital C*-algebra A is properly infinite if and only if there exists a unital $*$-homomorphism $\mathcal{T}_2 \to A$.

3. K_1-injectivity of $\mathcal{T}_n$

Given a unital C*-algebra A with unitary group $\mathcal{U}(A)$, denote by $\mathcal{U}^0(A)$ the connected component of 1_A in $\mathcal{U}(A)$. For each strictly positive integer $k \geq 1$, the upper diagonal embedding $u \in \mathcal{U}(M_k(A)) \mapsto (u \oplus 1_A) \in \mathcal{U}(M_{k+1}(A))$ sends the connected component $\mathcal{U}^0(M_k(A))$ into $\mathcal{U}^0(M_{k+1}(A))$, whence a canonical homomorphism Θ_A from $\mathcal{U}(A)/\mathcal{U}^0(A)$ to $K_1(A) := \lim\limits_{k\to\infty} \mathcal{U}(M_k(A))/\mathcal{U}^0(M_k(A))$. As noticed by Blackadar in [**Blac07**], this map is (1) neither injective, (2) nor surjective in general:

(1) If $\mathfrak{U}_2$ denotes the compact unitary group of the matrix C*-algebra $M_2(\mathbb{C})$, $A := C(\mathfrak{U}_2 \times \mathfrak{U}_2, M_2(\mathbb{C}))$ and $u, v \in \mathcal{U}(A)$ are the two unitaries $u(x,y) = x$ and $v(x,y) = y$, then $z := uvu^*v^*$ is not unitarily homotopic to 1_A in $\mathcal{U}(A)$ but the unitary $z \oplus 1_A$ belongs to $\mathcal{U}^0(M_2(A))$ ([**AJT60**]).
(2) If $A = C(\mathbb{T}^3)$, then $\mathcal{U}(A)/\mathcal{U}^0(A) \cong \mathbb{Z}^3$ but $K_1(A) \cong \mathbb{Z}^4$.

DEFINITION 3.1. The unital C*-algebra A is said to be K_1*-injective* if the map Θ_A is injective.

Cuntz proved in [**Cun81**] that Θ_A is surjective as soon as the C*-algebra A is properly infinite and that it is also injective if the C*-algebra A is simple and purely infinite. Now, the K-theoretical six-term cyclic exact sequence associated to the exact sequence (2.2) implies that $K_1(\mathcal{T}_n) = 0$ since $K_1(\mathcal{K}) = K_1(\mathcal{O}_n) = 0$. Thus, the map $\Theta_{\mathcal{T}_n}$ is zero.

PROPOSITION 3.2. *For all $n \geq 2$, the C*-algebra $\mathcal{T}_n$ is K_1-injective, i.e. any unitary $u \in \mathcal{U}(\mathcal{T}_n)$ is unitarily homotopic to $1_{\mathcal{T}_n}$ in $\mathcal{U}(\mathcal{T}_n)$ (written $u \sim_h 1_{\mathcal{T}_n}$).*

Proof. The C*-algebras $\mathcal{T}_n$ have real rank zero by Proposition 2.3 of [**Zha90**]. And Lin proved that any unital C*-algebra of real rank zero is K_1-injective ([**Lin01**, Corollary 4.2.10]). $\square$

COROLLARY 3.3. *If $\alpha : \mathcal{T}_3 \to \mathcal{T}_3$ is a unital $*$-endomorphism, then its restriction to the unital copy of $\mathcal{T}_2$ generated by the two isometries s_1, s_2 is unitarily homotopic to $id_{\mathcal{T}_2}$ among all unital $*$-homomorphisms $\mathcal{T}_2 \to \mathcal{T}_3$.*

Proof. The isometry $\sum_{k=1,2} \alpha(s_k)s_k^*$ extends to a unitary $u \in \mathcal{U}(\mathcal{T}_3)$ such that $\alpha(s_k) = us_k$ for $k = 1, 2$ ([**BRR08**, Lemma 2.4]). But Proposition 3.2 yields that $\mathcal{U}(\mathcal{T}_3) = \mathcal{U}^0(\mathcal{T}_3)$, whence a homotopy $u \sim_h 1$ in $\mathcal{U}(\mathcal{T}_3)$, and so the desired result holds. □

REMARK 3.4. The unital map $\iota : \mathbb{C} \to \mathcal{T}_2$ induces an isomorphism in K-theory. Indeed, $[1_{\mathcal{T}_2}] = [s_1s_1^*] + [s_2s_2^*] + [p_3] = 2\,[1_{\mathcal{T}_2}] + [p_3]$ and so $[1_{\mathcal{T}_2}] = -[p_3]$ is invertible in $K_0(\mathcal{T}_2)$.

4. K_1-injectivity of properly infinite C*-algebras

Denote by $\mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2$ the universal unital free product with amalgamation over $\mathbb{C}$ (in the sequel called full unital free product) of two copies of $\mathcal{T}_2$ amalgamated over $\mathbb{C}$ and let $\jmath_0$, $\jmath_1$ be the two canonical unital inclusions of $\mathcal{T}_2$ in $\mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2$. We show in this section that the K_1-injectivity of $\mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2$ is equivalent to the K_1-injectivity of all the unital properly infinite C*-algebras. The proof is similar to that of the universality of the full unital free product $\mathcal{O}_\infty *_{\mathbb{C}} \mathcal{O}_\infty$ (see Theorem 5.5 of [**BRR08**]).

DEFINITION 4.1. ([**Blan09**], [**BRR08**, §2]) If X is a compact Hausdorff space, a unital $C(X)$-algebra is a unital C*-algebra A endowed with a unital $*$-homomorphism from the C*-algebra $C(X)$ of continuous functions on X to the centre of A.

For any non-empty closed subset Y of X, we denote by π_Y^A (or simply by π_Y if no confusion is possible) the quotient map from A to the quotient A_Y of A by the (closed) ideal $C_0(X \setminus Y) \cdot A$. For any point $x \in X$, we also denote by A_x the quotient $A_{\{x\}}$ and by π_x the quotient map $\pi_{\{x\}}$.

PROPOSITION 4.2. *The following assertions are equivalent.*

(i) *$\mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2$ is K_1-injective.*
(ii) *$\mathcal{D} := \{ f \in C([0,1], \mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2)\,;\, f(0) \in \jmath_0(\mathcal{T}_2)$ and $f(1) \in \jmath_1(\mathcal{T}_2)\,\}$ is properly infinite.*
(iii) *There exists a unital $*$-homomorphism $\theta : \mathcal{T}_2 \to \mathcal{D}$.*
(iv) *There exists a projection $q \in \mathcal{D}$ with $\pi_0(q) = \jmath_0(s_1s_1^*)$ and $\pi_1(q) = \jmath_1(s_1s_1^*)$.*
(v) *Any unital properly infinite C*-algebra A is K_1-injective.*

Proof. (i)⇒(ii) We have a pull-back diagram

$$\begin{tikzcd} & \mathcal{D} \arrow[dl] \arrow[dr] & \\ \mathcal{D}_{[0,\frac{1}{2}]} \arrow[dr, "\pi_{\frac{1}{2}}"'] & & \mathcal{D}_{[\frac{1}{2},1]} \arrow[dl, "\pi_{\frac{1}{2}}"] \\ & \mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2 & \end{tikzcd}$$

and the two C*-algebras $\mathcal{D}_{[0,\frac{1}{2}]}$ and $\mathcal{D}_{[\frac{1}{2},1]}$ are properly infinite (Remark 2.4). Hence, the implication follows from [**BRR08**, Proposition 2.7].

(ii)⇒(iii) is Remark 2.4 applied to the C*-algebra $\mathcal{D}$.

(iii)⇒(iv) The two full, properly infinite projections $\jmath_0(s_1s_1^*)$ and $\pi_0\circ\theta(s_1s_1^*)$ are unitarily equivalent in $\jmath_0(\mathcal{T}_2)$ by [**LLR00**, Lemma 2.2.2] and [**BRR08**, Proposition 2.3]. Thus, they are homotopic among the projections in the C*-algebra $\jmath_0(\mathcal{T}_2)$ (written $\jmath_0(s_1s_1^*)\sim_h\pi_0\circ\theta(s_1s_1^*)$) by Proposition 3.2. Similarly, $\pi_1\circ\theta(s_1s_1^*)\sim_h \jmath_1(s_1s_1^*)$ in $\jmath_1(\mathcal{T}_2)$. Further, $\pi_0\circ\theta(s_1s_1^*)\sim_h\pi_1\circ\theta(s_1s_1^*)$ in $\mathcal{T}_2*_{\mathbb{C}}\mathcal{T}_2$ by hypothesis, whence the result by composition.

(iv)⇒(v) By [**BRR08**, Proposition 5.1], it is equivalent to prove that if p and p' are two properly infinite full projections in A, then there exist full properly infinite projections p_0, and p_0' in A such that $p_0\le p$, $p_0'\le p'$ and $p_0\sim_h p_0'$.

Fix two such projections p and p' in A. Then, there exist unital $*$-homomorphisms $\sigma:\mathcal{T}_2\to pAp$, $\sigma':\mathcal{T}_2\to p'Ap'$ and isometries $t,t'\in A$ such that $1_A=t^*pt=(t')^*p't'$. Now, one thoroughly defines unital $*$-homomorphisms $\alpha_0:\mathcal{T}_2\to A$ and $\alpha_1:\mathcal{T}_2\to A$ by

$$\alpha_0(s_k):=\sigma(s_k)\cdot t \quad\text{and}\quad \alpha_1(s_k):=\sigma'(s_k)\cdot t' \qquad \text{for } k=1,2\,,$$

whence a unital $*$-homomorphism $\alpha:=\alpha_0*\alpha_1:\mathcal{T}_2*_{\mathbb{C}}\mathcal{T}_2\to A$ such that $\alpha\circ\jmath_0=\alpha_0$ and $\alpha\circ\jmath_1=\alpha_1$.

The two full properly infinite projections $p_0=\alpha_0(s_1s_1^*)$ and $p_0'=\alpha_1(s_1s_1^*)$ satisfy $p_0\le p$ and $p_0'\le p'$. Further, the projection $(id\otimes\alpha)(q)$ gives a continuous path of projections in A from p_0 to p_0'. □

Remark 4.3. The C*-algebra $M_2(\mathcal{D})$ is properly infinite by [**BRR08**, Proposition 2.7].

Lemma 4.4. *$K_0(\mathcal{T}_2*_{\mathbb{C}}\mathcal{T}_2)=\mathbb{Z}$ and $K_1(\mathcal{T}_2*_{\mathbb{C}}\mathcal{T}_2)=0$*

Proof. The commutative diagram
$$\begin{tikzcd} \mathbb{C} \arrow[r, "\imath_1"] \arrow[d, "\imath_0"'] & \mathcal{T}_2 \arrow[d, "\jmath_1"] \\ \mathcal{T}_2 \arrow[r, "\jmath_0"] & \mathcal{T}_2*_{\mathbb{C}}\mathcal{T}_2 \end{tikzcd}$$
yields by [**Ger97**, Theorem 2.2] a six-term cyclic exact sequence

$$\begin{array}{ccccc} K_0(\mathbb{C})=\mathbb{Z} & \overset{(\imath_0\oplus\imath_1)_*}{\longrightarrow} & K_0(\mathcal{T}_2\oplus\mathcal{T}_2)=\ \mathbb{Z}\oplus\mathbb{Z} & \overset{(\jmath_0)_*-(\jmath_1)_*}{\longrightarrow} & K_0(\mathcal{T}_2*_{\mathbb{C}}\mathcal{T}_2) \\ \uparrow & & & & \downarrow \\ K_1(\mathcal{T}_2*_{\mathbb{C}}\mathcal{T}_2) & \longleftarrow & K_1(\mathcal{T}_2\oplus\mathcal{T}_2)=0\oplus 0 & \longleftarrow & K_1(\mathbb{C})=0 \end{array}$$

Now, Remark 3.4 implies that the map $(\imath_0\oplus\imath_1)_*$ is injective, whence the equalities. □

Remark 4.5. Skandalis noticed that the C*-algebra $\mathcal{T}_2$ is KK-equivalent to $\mathbb{C}$ and so $\mathcal{T}_2*_{\mathbb{C}}\mathcal{T}_2$ is KK-equivalent to $\mathbb{C}*_{\mathbb{C}}\mathbb{C}=\mathbb{C}$.

This Lemma entails that the K_1-injectivity question for unital properly infinite C*-algebras boils down to knowing whether $\mathcal{U}(\mathcal{T}_2*_{\mathbb{C}}\mathcal{T}_2)=\mathcal{U}^0(\mathcal{T}_2*_{\mathbb{C}}\mathcal{T}_2)$. Note that Proposition 3.2 already yields that $\mathcal{U}(\mathcal{T}_2)*_{\mathbb{T}}\mathcal{U}(\mathcal{T}_2)\subset\mathcal{U}^0(\mathcal{T}_2*_{\mathbb{C}}\mathcal{T}_2)$.

But the following holds.

Proposition 4.6. *Set $p_3=1-s_1s_1^*-s_2s_2^*$ in the Cuntz algebra $\mathcal{T}_2$ and let u be the canonical unitary generating $C^*(\mathbb{Z})$.*
(i) The relations $\jmath_0(s_k)\mapsto s_k$ and $\jmath_1(s_k)\mapsto u\,s_k$ ($k=1,2$) uniquely define a unital

$*$-homomorphism $\mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2 \to \mathcal{T}_2 *_{\mathbb{C}} C^*(\mathbb{Z})$ which is injective but not K_1-surjective.
(ii) The two projections $\jmath_0(p_3)$ and $\jmath_1(p_3)$ satisfy $\jmath_1(p_3) \not\sim \jmath_0(p_3)$ in $\mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2$.
(iii) There is no $v \in \mathcal{U}(\mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2)$ such that $\jmath_1(s_1s_1^* + s_2s_2^*) = v\,\jmath_0(s_1s_1^* + s_2s_2^*)\,v^*$.
(iv) There is a unitary $v \in \mathcal{U}(\mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2)$ such that $\jmath_1(s_1s_1^*) = v\,\jmath_0(s_1s_1^*)\,v^*$.

Proof. (i) The unital C*-subalgebra of $\mathcal{O}_3$ generated by the two isometries s_1 and s_2 is isomorphic to $\mathcal{T}_2$, whence a unital C*-embedding $\mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2 \subset \mathcal{O}_3 *_{\mathbb{C}} \mathcal{O}_3$ ([**ADEL04**]). Let Φ be the $*$-homomorphism from $\mathcal{O}_3 *_{\mathbb{C}} \mathcal{O}_3$ to the free product $\mathcal{O}_3 *_{\mathbb{C}} C^*(\mathbb{Z}) = C^*\big(s_1, s_2, s_3, u\big)$ fixed by the relations

$$\Phi(\jmath_0(s_k)) = s_k \quad \text{and} \quad \Phi(\jmath_1(s_k)) = u\,s_k \quad \text{for } k = 1, 2, 3$$

and let $\Psi : \mathcal{O}_3 *_{\mathbb{C}} C^*(\mathbb{Z}) \to \mathcal{O}_3 *_{\mathbb{C}} \mathcal{O}_3$ be the only $*$-homomorphism such that

$$\Psi(u) = \sum_{l=1}^{3} \jmath_1(s_l)\jmath_0(s_l)^* \quad \text{and} \quad \Psi(s_k) = \jmath_0(s_k) \quad \text{for } k = 1, 2, 3.$$

For all $k = 1, 2, 3$, we have the equalities:
– $\Psi \circ \Phi(\jmath_0(s_k)) = \Psi(s_k) = \jmath_0(s_k)$,
– $\Psi \circ \Phi(\jmath_1(s_k)) = \Psi(us_k) = \jmath_1(s_k)$,
– $\Phi \circ \Psi(s_k) = \Phi(\jmath_0(s_k)) = s_k$.

Also, $\Psi(u)^*\Psi(u) = \sum_{l,l'} \jmath_0(s_{l'})\jmath_1(s_{l'})^*\jmath_1(s_l)\jmath_0(s_l)^* = 1_{\mathcal{O}_3 *_{\mathbb{C}} \mathcal{O}_3} = \Psi(u)\Psi(u)^*$, *i.e.* $\Psi(u)$ is a unitary in $\mathcal{O}_3 *_{\mathbb{C}} \mathcal{O}_3$ which satisfies:
– $\Phi \circ \Psi(u) = \sum_{l=1,2,3} \Phi(\jmath_1(s_l))\Phi(\jmath_0(s_l)^*) = u$.

Thus, Φ is an invertible unital $*$-homomorphism with inverse Ψ ([**Blac07**]), and the restriction of Φ to the C*-subalgebra $\mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2$ takes values in $\mathcal{T}_2 *_{\mathbb{C}} C^*(\mathbb{Z}) \subset \mathcal{O}_3 *_{\mathbb{C}} C^*(\mathbb{Z})$.

Now, there is (see [**Ger97**]) a six-term cyclic exact sequence

$$\begin{array}{ccccc} K_0(\mathbb{C}) = \mathbb{Z} & \hookrightarrow & K_0\big(\mathcal{T}_2 \oplus C^*(\mathbb{Z})\big) = \mathbb{Z} \oplus \mathbb{Z} & \to & K_0(\mathcal{T}_2 *_{\mathbb{C}} C^*(\mathbb{Z})\,) \\ \uparrow & & & & \downarrow \\ K_1(\mathcal{T}_2 *_{\mathbb{C}} C^*(\mathbb{Z})\,) & \leftarrow & K_1\big(\mathcal{T}_2 \oplus C^*(\mathbb{Z})\big) = 0 \oplus \mathbb{Z} & \leftarrow & K_1(\mathbb{C}) = 0 \end{array}$$

and so, $K_1(\mathcal{T}_2 *_{\mathbb{C}} C^*(\mathbb{Z})\,) = \mathbb{Z}$, whereas $K_1(\mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2) = 0$ by Lemma 4.4.

(ii) Let $\pi_0 : \mathcal{T}_2 \to L(H)$ be a unital $*$-representation on a separable Hilbert space H such that $\pi_0(p_3)$ is a rank one projection, let $\pi_1 : \mathcal{T}_2 \to L(H)$ be a unital $*$-representation such that $\pi_1(p_3)$ is a rank two projection and consider the induced unital $*$-representation $\pi = \pi_0 * \pi_1$ of the unital free product $\mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2$.

Then the two projections $\pi[\jmath_0(p_3)] = \pi_0(p_3)$ and $\pi[\jmath_1(p_3)] = \pi_1(p_3)$ have distinct ranks and so cannot be equivalent in $L(H)$. Hence, $\jmath_0(p_3) \not\sim \jmath_1(p_3)$ in $\mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2$.

(iii) This is just a rewriting of the previous assertion since $s_1s_1^* + s_2s_2^* = 1 - p_3$. Indeed, the partial isometry $b = \jmath_1(s_1)\jmath_0(s_1)^* + \jmath_1(s_2)\jmath_0(s_2)^*$ defines a Murray-von Neumann equivalence in $\mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2$ between the projections $\jmath_0(s_1s_1^* + s_2s_2^*) = 1 - \jmath_0(p_3)$ and $\jmath_1(s_1s_1^* + s_2s_2^*) = 1 - \jmath_1(p_3)$. Thus, they are unitarily equivalent in $\mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2$ if and only if $\jmath_0(p_3) \sim \jmath_1(p_3)$ in $\mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2$ ([**LLR00**, Proposition 2.2.2]).

(iv) There exists a unitary $v \in \mathcal{U}(\mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2)$ (which is necessarily K_1-trivial by Lemma 4.4) such that $\jmath_1(s_1s_1^*) = v\,\jmath_0(s_1s_1^*)\,v^*$. Indeed, we have the inequalities

$$1 > s_2s_2^* + p_3 > s_2s_2^* > s_2s_1(s_2s_2^* + p_3)s_1^*s_2^* + s_2s_2(s_2s_2^* + p_3)s_2^*s_2^* \qquad \text{in } \mathcal{T}_2\,.$$

Thus, if we set $w := \jmath_1(s_1)\jmath_0(s_1)^*$, then $1 - w^*w = \jmath_0(s_2s_2^* + p_3)$ and $1 - ww^* = \jmath_1(s_2s_2^* + p_3)$ are two properly infinite and full K_0-equivalent projections in $\mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2$. Thus, there is a partial isometry $a \in \mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2$ with $a^*a = 1 - w^*w$ and $aa^* =$

$1 - ww^*$ ([**Cun81**]). The sum $v = a + w$ has the required properties ([**BRR08**, Lemma 2.4]). □

REMARKS 4.7. (i) The equivalence (iv)⇔(v) in Proposition 4.2 implies that all unital properly infinite C*-algebras are K_1-injective if and only if the unitary $v \in \mathcal{U}(\mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2)$ constructed in Proposition 4.6.(iv) belongs to the connected component $\mathcal{U}^0(\mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2)$.

Note that $v \oplus 1 \sim_h 1 \oplus 1$ in $\mathcal{U}(M_2(\mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2))$ by [**LLR00**, Exercice 8.11].

(ii) Let $\sigma \in \mathcal{U}(\mathcal{T}_2)$ be the symmetry $\sigma = s_1 s_2^* + s_2 s_1^* + p_3$, let $v \in \mathcal{U}(\mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2)$ be a unitary such that $\jmath_1(s_1 s_1^*) = v \jmath_0(s_1 s_1^*) v^*$ (Proposition 4.6.(iv)) and set $z := v^* \jmath_1(\sigma) v \jmath_0(\sigma)$.

Then, $q_1 = \jmath_0(s_1 s_1^*)$, $q_2 = \jmath_0(s_2 s_2^*)$ and $q_3 = z \jmath_0(s_2 s_2^*) z^*$ are three properly infinite full projections in $\mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2$ which satisfy:

– $q_1 q_3 = \jmath_0(s_1 s_1^*)\, v^*\, \jmath_1(s_2 s_2^*)\, v = v^*\, \jmath_1(s_1 s_1^*) \jmath_1(s_2 s_2^*)\, v = 0 = q_1 q_2$,

– $q_2 \sim_h q_1 \sim_h q_3$ in $\mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2$ since $\sigma \in \mathcal{U}^0(\mathcal{T}_2)$ and so $z \sim_h v^* v = 1$ in $\mathcal{U}(\mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2)$,

– $q_1 + q_3 = v^* \jmath_1(s_1 s_1^* + s_2 s_2^*) v \not\sim_u \jmath_0(s_1 s_1^* + s_2 s_2^*) = q_1 + q_2$ in $\mathcal{T}_2 *_{\mathbb{C}} \mathcal{T}_2$ by Proposition 4.6.(iii).

References

[AJT60] S. Araki, I. M. James, E. Thomas, *Homotopy-abelian Lie groups*, Bull. Amer. Math. Soc. **66** (1960), 324-326.

[ADEL04] S. Armstrong, K. Dykema, R. Exel, H. Li, *On embeddings of full amalgamated free product C*-algebras*, Proc. Amer. Math. Soc. **132** (2004), 2019-2030.

[Blac07] B. Blackadar, *Non stable K-theory for unital free products of C*-algebras*, talk in Barcelona (2007).

[Blan09] E. Blanchard, *Amalgamated free products of C*-bundles*, Proc. Edinburgh Math. Soc., **52** (2009), 23–36.

[BRR08] E. Blanchard, R. Rohde, M. Rørdam, *Properly infinite $C(X)$-algebras and K_1-injectivity*, J. Noncommut. Geom. **2** (2008), 263–282.

[Cun78] J. Cuntz, *Dimension functions on simple C*-algebras*, Math. Ann. **233** (1978), 145-153.

[Cun81] J. Cuntz, *K-theory for certain C*-algebras*, Ann. of Math. **113** (1981), 181-197.

[Ger97] E. Germain, *KK-theory of the full free product of unital C*-algebras*, J. Reine Angew. Math. **485** (1997), 1–10.

[KR00] E. Kirchberg, M. Rørdam, *Non-simple purely infinite C*-algebras*, American J. Math. **122** (2000), 637-666.

[LLR00] F. Larsen, N. J. Laustsen, M. Rørdam, *An Introduction to K-theory for C*-algebras*, London Mathematical Society Student Texts **49** (2000) CUP, Cambridge.

[Lin01] *An introduction to the classification of amenable C*-algebras*, World Scientific 2001.

[Rør92] M. Rørdam, *On the structure of simple C*-algebras tensored with a UHF-algebra, II*, J. Funct. Anal. **107** (1992), 255-269.

[Zha90] S. Zhang, *Certain C*-algebras with real rank zero and the internal structure of their corona and multiplier algebras, III*, Canad. J. Math. **42** (1990), 159-190.

CNRS IMJ, 175, rue du Chevaleret, F–75013 Paris, and UPMC Univ Paris 06, UMR 7586, F–75005 Paris

E-mail address: Etienne.Blanchard@math.jussieu.fr

Clay Mathematics Proceedings
Volume 11, 2010

Families of Monads and Instantons from a Noncommutative ADHM Construction

Simon Brain and Giovanni Landi

Dedicated to Alain Connes

Abstract. We give a θ-deformed version of the ADHM construction of instantons with arbitrary topological charge on the sphere S^4. Classically, the instanton gauge fields are constructed from suitable monad data; we show that in the deformed case the set of monads is itself a noncommutative space. We use these monads to construct noncommutative 'families' of instantons (i.e. noncommutative families of anti-self-dual connections) on the deformed sphere S^4_θ. We also compute the topological charge of each of the families. Finally we discuss what it means for such families to be gauge equivalent.

Contents

1. Introduction

The purpose of the present article is to generalise the ADHM method for constructing instantons on the four-sphere S^4 to the framework of noncommutative geometry, by giving a construction of instantons on the noncommutative four-sphere S^4_θ of [**9**].

2010 *Mathematics Subject Classification.* Primary 58B34; Secondary 14D21, 81T13, 58B32, 81R60.

Partially supported by:
the 'Italian project Cofin06 - Noncommutative geometry, quantum groups and applications'.

Instantons arise in physics as anti-self-dual solutions of the Yang-Mills equations. Mathematically they are connections with anti-self-dual curvature on smooth G-bundles over a four-dimensional compact manifold. Since the very beginning they have been of central importance for both disciplines, an importance that has only grown over the years.

Of particular interest are instantons on SU(2)-bundles over the Euclidean four-sphere S^4. Thanks to the ADHM method of [**2**], the full solution to the problem of constructing such instantons on S^4 has long been known and, as a consequence, the moduli space $\mathcal{M}_k$ of instantons with topological charge equal to k is known to be a manifold of dimension $8k-3$. Starting with a trivial vector bundle over S^4, the ADHM strategy is to construct an orthogonal projection to some (non-trivial) sub-bundle E in such a way that the projection of the trivial connection to E has anti-self-dual curvature.

The geometric ingredient which implements the classical ADHM construction is the Penrose twistor fibration $\mathbb{CP}^3 \to S^4$. The total space $\mathbb{CP}^3$ of the fibration is called the twistor space of S^4 and may be thought of as the bundle of projective spinors over S^4 (although it has its origins elsewhere [**21**]). The pull-back of an instanton bundle along this fibration is a holomorphic vector bundle over $\mathbb{CP}^3$ equipped with a set of reality conditions which identify it as such a pull-back [**24**]. In this way, the construction of instantons is equivalent to the construction of holomorphic bundles over twistor space.

Using powerful results from algebraic geometry, one gives an explicit description of all relevant holomorphic vector bundles over a complex projective space ([**12, 4**], *cf.* also [**20**]). Each of them arises as the cohomology of a *monad*: a suitable complex of vector bundles

$$0 \to \mathcal{A} \xrightarrow{\sigma} \mathcal{B} \xrightarrow{\tau} \mathcal{C} \to 0$$

such that σ is injective and τ is surjective. The ADHM construction tells us how to convert a given monad into an orthogonal projection of vector bundles as described above and guarantees that the resulting connection has anti-self-dual curvature.

Following the general strategy of the classical case, our goal is to give a deformed version of the ADHM method and hence a construction of instantons on the noncommutative four-sphere S^4_θ. The techniques involved lend themselves rather neatly to the framework of noncommutative geometry; the construction of vector bundles and connections by orthogonal projection is particularly natural in light of the Serre-Swan theorem [**11**], which trades vector bundles for finitely generated projective modules.

The paper is organised as follows. Sect. 2 reviews the noncommutative spaces in question, namely the θ-deformed versions of the four-sphere S^4_θ and its twistor space $\mathbb{CP}^3_\theta$. We recall also the construction of the basic instanton and the principal bundle on which it is defined, as well as the details of the noncommutative twistor fibration. Sect. 3 recalls the construction of the quantum group $\mathrm{SL}_\theta(2,\mathbb{H})$ of conformal transformations of S^4_θ and the quantum subgroup $\mathrm{Sp}_\theta(2)$ of isometries. The main purpose of these two sections is to gather together into one place the relevant contributions from [**9, 14, 15, 16, 5**] and to establish notation; in doing so we also make some novel improvements to previous versions. Sect. 4 presents the deformed ADHM construction itself. We show that in the deformed case the set of all monads is parameterised by a collection of noncommutative spaces $\widetilde{\mathcal{M}}_{\theta;k}$ indexed by k a positive integer. We use each of these spaces to construct a noncommutative 'family' of instantons whose topological charge we show to be equal to k. Finally in Sect. 5 we discuss what it means for families of instantons to be gauge equivalent. In particular, we show that the quantum symmetries of the sphere S^4_θ generate gauge degrees of freedom, a feature which is a consequence of the noncommutativity and is not present in the classical construction. For further discussion in this direction we refer to [**6**].

2. The Twistor Fibration

The use of the twistor fibration in the ADHM construction is crucial: this fibration captures in its geometry the very nature of the anti-self-duality equations, with the result that an instanton bundle is reinterpreted *via* pull-back in terms of holomorphic data on twistor space [**24**] (*cf.* also [**1**]). In particular, this means that twistor space plays the role of an 'auxiliary space' on which the ADHM construction takes place, before passing back down to the base space S^4 (we refer to [**19**] for more on the ADHM construction from a twistor perspective).

We start by recalling the details of the algebra inclusion $\mathcal{A}(S^4_\theta) \hookrightarrow \mathcal{A}(S^7_\theta)$ as a noncommutative principal bundle with undeformed structure group SU(2); associated to this principal bundle there is in particular a basic instanton bundle [**14**]. Next we give a description of the noncommutative twistor space in terms of its coordinate algebra $\mathcal{A}(\mathbb{CP}^3_\theta)$, as well as a dualised description of the twistor fibration, now appearing [**5**] as an algebra inclusion $\mathcal{A}(S^4_\theta) \hookrightarrow \mathcal{A}(\mathbb{CP}^3_\theta)$.

2.1. The noncommutative Hopf fibration. With $\lambda = \exp(2\pi \mathrm{i}\theta)$ the deformation parameter, the coordinate algebra $\mathcal{A}(S^4_\theta)$ of the noncommutative four-sphere S^4_θ is the $*$-algebra generated by a central real element x and elements α, β, α^*, β^*, modulo the relations

$$\alpha\beta = \lambda\beta\alpha, \quad \alpha^*\beta^* = \lambda\beta^*\alpha^*, \quad \beta^*\alpha = \lambda\alpha\beta^*, \quad \beta\alpha^* = \lambda\alpha^*\beta, \tag{1}$$

together with the sphere relation

$$\alpha^*\alpha + \beta^*\beta + x^2 = 1. \tag{2}$$

Similarly, the coordinate algebra of the noncommutative seven-sphere $\mathcal{A}(S^7_\theta)$ is generated as a $*$-algebra by the elements $\{z_j, z_j^* \mid j = 1, \dots, 4\}$ and is subject to the commutation relations

$$z_j z_l = \eta_{jl} z_l z_j, \quad z_j z_l^* = \eta_{lj} z_l^* z_j, \quad z_j^* z_l^* = \eta_{jl} z_l^* z_j^*, \tag{3}$$

as well as the sphere relation

$$z_1^* z_1 + z_2^* z_2 + z_3^* z_3 + z_4^* z_4 = 1. \tag{4}$$

Compatibility with the SU(2) principal bundle structure requires the deformation matrix (η_{jk}) be given by

$$(\eta_{jk}) = \begin{pmatrix} 1 & 1 & \bar\mu & \mu \\ 1 & 1 & \mu & \bar\mu \\ \mu & \bar\mu & 1 & 1 \\ \bar\mu & \mu & 1 & 1 \end{pmatrix}, \qquad \mu = \exp(\mathrm{i}\pi\theta). \tag{5}$$

The values of the deformation parameters λ, μ are precisely those which allow an embedding of the classical group SU(2) into the group $\operatorname{Aut}\mathcal{A}(S^7_\theta)$. We denote by $\mathcal{A}(\mathbb{C}^4_\theta)$ the algebra generated by the $\{z_j,\ z_j^*\}$ subject to the relations (3); the quotient by the additional sphere relation yields the algebra $\mathcal{A}(S^7_\theta)$. The algebra inclusion $\mathcal{A}(S^4_\theta) \hookrightarrow \mathcal{A}(S^7_\theta)$ is given explicitly by

$$(6)\ \alpha = 2(z_1 z_3^* + z_2^* z_4), \quad \beta = 2(z_2 z_3^* - z_1^* z_4), \quad x = z_1 z_1^* + z_2 z_2^* - z_3 z_3^* - z_4 z_4^*.$$

One easily verifies that for the right SU(2)-action on $\mathcal{A}(S^7_\theta)$ given on generators by

$$(7) \qquad (z_1, z_2^*, z_3, z_4^*) \mapsto (z_1, z_2^*, z_3, z_4^*) \begin{pmatrix} w & 0 \\ 0 & w \end{pmatrix}, \qquad w = \begin{pmatrix} w^1 & -\bar{w}^2 \\ w^2 & \bar{w}^1 \end{pmatrix} \in \mathrm{SU}(2),$$

the invariant subalgebra is generated as expected by α, β, x and their conjugates, so one indeed has

$$\mathrm{Inv}_{\mathrm{SU}(2)} \mathcal{A}(S^7_\theta) = \mathcal{A}(S^4_\theta).$$

When $\theta = 0$ we recover the usual algebras of functions on the classical spheres S^4 and S^7. The inclusion $\mathcal{A}(S^4) \hookrightarrow \mathcal{A}(S^7)$ is just a dualised description of the standard SU(2) Hopf fibration $S^7 \to S^4$.

These noncommutative spheres have canonical differential calculi arising as deformations of the classical ones. Explicitly, one has a first order differential calculus $\Omega^1(S^7_\theta)$ on $\mathcal{A}(S^7_\theta)$ spanned as an $\mathcal{A}(S^7_\theta)$-bimodule by $\{\mathrm{d}z_j,\, \mathrm{d}z_j^*,\, j = 1, \ldots, 4\}$, subject to the relations

$$z_i \mathrm{d}z_j = \eta_{ij} \mathrm{d}z_j z_i, \qquad z_i \mathrm{d}z_j^* = \eta_{ji} \mathrm{d}z_j^* z_i,$$

with η_{ij} as before. One also has relations

$$\mathrm{d}z_i \mathrm{d}z_j + \eta_{ij} \mathrm{d}z_j \mathrm{d}z_i = 0, \qquad \mathrm{d}z_i \mathrm{d}z_j^* + \eta_{ji} \mathrm{d}z_j^* \mathrm{d}z_i = 0,$$

allowing one to extend the first order calculus to a differential graded algebra $\Omega(S^7_\theta)$. There is a unique differential d on $\Omega(S^7_\theta)$ such that $\mathrm{d} : z_j \mapsto \mathrm{d}z_j$. Furthermore, $\Omega(S^7_\theta)$ has an involution given by the graded extension of the map $z_j \mapsto z_j^*$. The story is similar for the four-sphere, in that the differential graded algebra $\Omega(S^4_\theta)$ is generated in degree one by $\mathrm{d}\alpha$, $\mathrm{d}\alpha^*$, $\mathrm{d}\beta$, $\mathrm{d}\beta^*$, $\mathrm{d}x$, subject to the relations

$$\alpha \mathrm{d}\beta = \lambda(\mathrm{d}\beta)\alpha, \qquad \beta^* \mathrm{d}\alpha = \lambda(\mathrm{d}\alpha)\beta^*,$$

$$\mathrm{d}\alpha \mathrm{d}\beta + \lambda \mathrm{d}\beta \mathrm{d}\alpha = 0, \qquad \mathrm{d}\beta^* \mathrm{d}\alpha + \lambda \mathrm{d}\alpha \mathrm{d}\beta^* = 0.$$

The above are the same as the relations (1) or (3) but with d inserted. As vector spaces, the graded components $\Omega^k(S^7_\theta)$ and $\Omega^k(S^4_\theta)$ of k-forms on the noncommutative spheres are identical to their classical counterparts, although the algebra relations between forms are twisted. In particular this means that the Hodge $*$-operator on S^4_θ,

$$*_\theta : \Omega^k(S^4_\theta) \to \Omega^{4-k}(S^4_\theta),$$

is defined by the same formula as it is classically. One still has that $*_\theta^2 = 1$, whence there is a direct sum decomposition of two-forms

$$\Omega^2(S^4_\theta) = \Omega^2_+(S^4_\theta) \oplus \Omega^2_-(S^4_\theta),$$

with $\Omega^2_\pm(S^4_\theta) := \{\omega \in \Omega^2(S^4_\theta) \mid *_\theta \omega = \pm\omega\}$ the spaces of self-dual and anti-self-dual two-forms.

2.2. The basic instanton. Amongst the nice properties of the classical Hopf fibration is that its canonical connection is an anti-instanton: its curvature is a self-dual two-form with values in the Lie algebra $\mathfrak{su}(2)$ of the structure group. This property holds also in the noncommutative case, giving a simple example of a noncommutative instanton. It has an elegant description [**14**] in terms of the function algebras $\mathcal{A}(S^7_\theta)$, $\mathcal{A}(S^4_\theta)$ as follows. One takes the pair of elements of the right $\mathcal{A}(S^7_\theta)$-module $\mathcal{A}(S^7_\theta)^4 := \mathbb{C}^4 \otimes \mathcal{A}(S^7_\theta)$ given by

$$|\psi_1\rangle = \begin{pmatrix} z_1 & z_2 & z_3 & z_4 \end{pmatrix}^{\mathrm{t}}, \qquad |\psi_2\rangle = \begin{pmatrix} -z_2^* & z_1^* & -z_4^* & z_3^* \end{pmatrix}^{\mathrm{t}}.$$

With the natural Hermitian structure on $\mathcal{A}(S^7_\theta)^4$ given by $\langle \xi | \eta \rangle = \sum_i \xi_i^* \eta_i$, one sees that $\langle \psi_j | \psi_l \rangle = \delta_{jl}$. It is convenient to introduce the matrix-valued function Ψ on S^7_θ given by

$$\Psi = \begin{pmatrix} |\psi_1\rangle & |\psi_2\rangle \end{pmatrix} = \begin{pmatrix} z_1 & z_2 & z_3 & z_4 \\ -z_2^* & z_1^* & -z_4^* & z_3^* \end{pmatrix}^{\mathrm{t}}. \tag{8}$$

From orthonormality of the columns one has that $\Psi^*\Psi = 1$ and hence the matrix

$$\mathsf{q} := \Psi\Psi^* = \frac{1}{2}\begin{pmatrix} 1+x & 0 & \alpha & -\bar{\mu}\,\beta^* \\ 0 & 1+x & \beta & \mu\,\alpha^* \\ \alpha^* & \beta^* & 1-x & 0 \\ -\mu\,\beta & \bar{\mu}\,\alpha & 0 & 1-x \end{pmatrix} \tag{9}$$

is a self-adjoint idempotent of rank two, *i.e.* $\mathsf{q}^* = \mathsf{q} = \mathsf{q}^2$ and $\operatorname{Tr}\mathsf{q} = 2$. The action (7) of SU(2) on $\mathcal{A}(S^7_\theta)$ now takes the form

$$\Psi \mapsto \Psi w, \qquad w \in \mathrm{SU}(2),$$

from which the SU(2)-invariance of the entries of q is immediately deduced. We may also write the commutation relations of $\mathcal{A}(S^7_\theta)$ in the useful form

$$\Psi_{ia}\Psi_{jb} = \eta_{ij}\Psi_{jb}\Psi_{ia}, \qquad a,b = 1,2 \quad i,j = 1,2,3,4. \tag{10}$$

If ρ is the defining representation of SU(2) on $\mathbb{C}^2$, the finitely generated projective right $\mathcal{A}(S^4_\theta)$-module $\mathcal{E} := \mathsf{q}\mathcal{A}(S^4_\theta)^4$ is isomorphic to the module of equivariant maps from $\mathcal{A}(S^7_\theta)$ to $\mathbb{C}^2$,

$$\mathcal{E} \cong \{\phi \in \mathcal{A}(S^7_\theta) \otimes \mathbb{C}^2 \mid (w \otimes \mathrm{id})\phi = (\mathrm{id} \otimes \rho(w^{-1}))\phi \text{ for all } w \in \mathrm{SU}(2)\}.$$

The module $\mathcal{E}$ has the role of the module of sections of the 'associated bundle' $E = S^7_\theta \times_{\mathrm{SU}(2)} \mathbb{C}^2$. With the projection $\mathsf{q} = \Psi\Psi^*$ there comes the canonical Grassmann connection defined on the module $\mathcal{E}$ by

$$\nabla := \mathsf{q} \circ \mathrm{d} : \mathcal{E} \to \mathcal{E} \otimes_{\mathcal{A}(S^4_\theta)} \Omega^1(S^4_\theta).$$

The curvature of ∇ is $\nabla^2 = \mathsf{q}(\mathrm{d}\mathsf{q})^2$, which may be shown to be self-dual with respect to the Hodge operator,

$$*_\theta(\mathsf{q}(\mathrm{d}\mathsf{q})^2) = \mathsf{q}(\mathrm{d}\mathsf{q})^2.$$

The complementary projector $\mathsf{p} = 1 - \mathsf{q}$ yields a connection whose curvature is anti-self-dual, $*_\theta(\mathsf{p}(\mathrm{d}\mathsf{p})^2) = -\mathsf{p}(\mathrm{d}\mathsf{p})^2$, and hence an instanton on the noncommutative four-sphere, which we call the *basic instanton*. Noncommutative index theory computes its 'topological charge' to be equal to -1.

Using the standard basis (e_1, e_2) of $\mathbb{C}^2$, equivariant maps are written as $\phi = \sum_a \phi_a \otimes e_a$. On them, one has explicitly that

$$\nabla(\phi_a) = \mathrm{d}\phi_a + \sum\nolimits_b \omega_{ab}\phi_b,$$

where the connection one-form $\omega = \omega_{ab}$ is found to be

$$\omega_{ab} = \tfrac{1}{2} \sum\nolimits_j \left((\Psi^*)_{aj} \mathrm{d}\Psi_{jb} - \mathrm{d}(\Psi^*)_{aj} \Psi_{jb} \right). \tag{11}$$

From this it is easy to see that $\omega_{ab} = -(\omega^*)_{ba}$ and $\sum_a \omega_{aa} = 0$, so that ω is an element of $\Omega^1(S^7_\theta) \otimes \mathfrak{su}(2)$.

2.3. Noncommutative twistor space. It is well-known that, as a real six-dimensional manifold, the space $\mathbb{CP}^3$ may be identified with the set of all 4×4 Hermitian projector matrices of rank one: this is because each such matrix uniquely determines and is uniquely determined by a one-dimensional subspace of $\mathbb{C}^4$. Thus the coordinate algebra $\mathcal{A}(\mathbb{CP}^3)$ of $\mathbb{CP}^3$ has a defining matrix of generators

$$Q = \begin{pmatrix} t_1 & x_1 & x_2 & x_3 \\ x_1^* & t_2 & y_3 & y_2 \\ x_2^* & y_3^* & t_3 & y_1 \\ x_3^* & y_2^* & y_1^* & t_4 \end{pmatrix}, \tag{12}$$

with $t_j^* = t_j$, $j = 1, \ldots, 4$ and $\mathrm{Tr}\, Q = \sum_j t_j = 1$, as well as the relations coming from the condition $Q^2 = Q$, that is to say $\sum_j Q_{kj}Q_{jl} = Q_{kl}$. The noncommutative twistor algebra $\mathcal{A}(\mathbb{CP}^3_\theta)$ is obtained by deforming these relations: with deformation parameter $\lambda = \exp(2\pi \mathrm{i}\theta)$, one has that $t_1, \ldots, t_4$ are central, that

$$x_1x_3 = \bar{\lambda}x_3x_1, \quad x_2x_1 = \bar{\lambda}x_1x_2, \quad x_2x_3 = \bar{\lambda}x_3x_2$$

as well as the auxiliary relations

$$y_1y_2 = \bar{\lambda}y_2y_1, \quad y_1y_3 = \bar{\lambda}y_3y_1, \quad y_2y_3 = \bar{\lambda}y_3y_2, \quad x_1(y_1, y_2, y_3) = (\bar{\lambda}^2 y_1, \bar{\lambda}y_2, \lambda y_3)x_1,$$
$$x_2(y_1, y_2, y_3) = (\bar{\lambda}y_1, y_2, \lambda y_3)x_2, \quad x_3(y_1, y_2, y_3) = (\bar{\lambda}y_1, \bar{\lambda}y_2, y_3)x_3,$$

and similar relations obtained by taking the adjoint under $*$ of those above (we refer to [**5**] for further details). To proceed further it is useful to note that classically $\mathbb{CP}^3$ is the quotient of the sphere S^7 by the action of the diagonal U(1) subgroup of SU(2). This remains true in the noncommutative case and one identifies the generators of $\mathcal{A}(\mathbb{CP}^3_\theta)$ as

$$Q_{jl} = z_j z_l^*, \tag{13}$$

via the generators $\{z_j, z_j^*\}$ of $\mathcal{A}(S^7_\theta)$. Indeed, from equation (13) one could infer the relations on the generators of $\mathcal{A}(\mathbb{CP}^3_\theta)$ from those on the generators of $\mathcal{A}(S^7_\theta)$. By its very definition $\mathcal{A}(\mathbb{CP}^3_\theta)$ is the invariant subalgebra of $\mathcal{A}(S^7_\theta)$ under this U(1)-action and equation (13) defines an inclusion of algebras

$$\mathcal{A}(\mathbb{CP}^3_\theta) \hookrightarrow \mathcal{A}(S^7_\theta),$$

giving a noncommutative principal bundle with structure group U(1). We thus have algebra inclusions

$$\mathcal{A}(S^4_\theta) \hookrightarrow \mathcal{A}(\mathbb{CP}^3_\theta) \hookrightarrow \mathcal{A}(S^7_\theta), \tag{14}$$

with the left-hand arrow still to be determined. As in the classical case, this inclusion is not a principal fibration (the 'typical fibre' is a copy of the undeformed $\mathbb{CP}^1$) but we may nevertheless express the generators of $\mathcal{A}(\mathbb{CP}^3_\theta)$ in terms of the generators of $\mathcal{A}(S^4_\theta)$. For this we need the non-degenerate map on $\mathcal{A}(\mathbb{C}^4_\theta)$ given on generators by

$$J(z_1, z_2, z_3, z_4) := (-z_2^*, z_1^*, -z_4^*, z_3^*) \tag{15}$$

and extended as an anti-algebra map. Classically, in doing so we would be identifying the set of quaternions $\mathbb{H}$ with the set of 2×2 matrices over $\mathbb{C}$ of the form

$$c_1 + c_2 j \in \mathbb{H} \;\mapsto\; \begin{pmatrix} c_1 & -\bar{c}_2 \\ c_2 & \bar{c}_1 \end{pmatrix} \in \mathrm{M}_2(\mathbb{C}),$$

and the map J corresponds to right multiplication by the quaternion j. In the deformed case, this very same identification defines the algebra $\mathcal{A}(\mathbb{H}^2_\theta)$ to be equal to the algebra $\mathcal{A}(\mathbb{C}^4_\theta)$ equipped with the map J [**16**].

Using the identification of generators (13) the map J extends to an automorphism of $\mathcal{A}(\mathbb{CP}^3_\theta)$, given in terms of the matrix generators in equation (12) by

$$\begin{array}{llll} J(t_1) = t_2, & J(t_2) = t_1, & J(t_3) = t_4, & J(t_4) = t_3, \\ J(x_1) = -x_1, & J(y_1) = -y_1, & J(x_1^*) = -x_1^*, & J(y_1^*) = -y_1^*, \\ J(x_2) = \mu\, y_2^*, & J(x_3) = -y_3^*, & J(x_2^*) = \bar{\mu}\, y_2, & J(x_3^*) = -y_3, \\ J(y_2) = \bar{\mu}\, x_2^*, & J(y_3) = -x_3^*, & J(y_2^*) = \mu\, x_2, & J(y_3^*) = -x_3, \end{array}$$

as required for J to respect the algebra relations of $\mathcal{A}(\mathbb{CP}^3_\theta)$. The subalgebra fixed by the map J is precisely $\mathcal{A}(S^4_\theta)$; in fact one has an algebra inclusion $\mathcal{A}(S^4_\theta) \hookrightarrow \mathcal{A}(\mathbb{CP}^3_\theta)$ given on generators by

$$x \mapsto 2(t_1 + t_2 - 1), \quad \alpha \mapsto 2(x_2 + \mu\, y_2^*), \quad \beta \mapsto 2(-x_3^* + y_3), \tag{16}$$

with $\mu = \sqrt{\lambda} = \exp(\pi \mathrm{i} \theta)$. In the notation of equation (8) we have $Q = |\psi_1\rangle\langle\psi_1|$, and we note also that $|\psi_2\rangle = |J\psi_1\rangle$, so that equation (16) is just the statement that

$$\mathsf{q} = |\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2| = |\psi_1\rangle\langle\psi_1| + |J\psi_1\rangle\langle J\psi_1| = Q + J(Q).$$

This gives us the promised algebraic description of the twistor fibration (14): the generators of $\mathcal{A}(S^4_\theta)$ are identified with the degree one elements of $\mathcal{A}(\mathbb{CP}^3_\theta)$ of the form $Z + J(Z)$.

3. The Quantum Conformal Group

Next, we briefly review the construction of the quantum groups which describe the symmetries of the spheres S^4_θ and S^7_θ (and the symmetries of the Hopf fibration defined in Sect. 2.1).

3.1. The quantum groups $\mathbf{SL}_\theta(2, \mathbb{H})$ and $\mathbf{Sp}_\theta(2)$. To begin, we need a noncommutative analogue of the set of all linear transformations of the quaternionic vector space $\mathbb{H}^2_\theta$ defined above. To this end, we define a transformation bialgebra for the algebra $\mathcal{A}(\mathbb{H}^2_\theta)$ to be a bialgebra $\mathcal{B}$ such that there is a $*$-algebra map $\Delta_L : \mathcal{A}(\mathbb{C}^4_\theta) \to \mathcal{B} \otimes \mathcal{A}(\mathbb{C}^4_\theta)$ commuting with the map J of equation (15). The set of all transformation bialgebras for $\mathcal{A}(\mathbb{H}^2_\theta)$ forms a category in the natural way; we define the bialgebra $\mathcal{A}(\mathrm{M}_\theta(2, \mathbb{H}))$ as the universal initial object in the category, meaning that whenever $\mathcal{B}$ is a transformation bialgebra for $\mathcal{A}(\mathbb{H}^2_\theta)$ there is a morphism of transformation bialgebras $\mathcal{A}(\mathrm{M}_\theta(2, \mathbb{H})) \to \mathcal{B}$ [**16**]. Using the universality property, one finds that $\mathcal{A}(\mathrm{M}_\theta(2,\mathbb{H}))$ is the associative algebra generated by the entries of the following 4×4 matrix:

$$A=\begin{pmatrix} a_{ij} & b_{ij} \\ c_{ij} & d_{ij}\end{pmatrix}=\begin{pmatrix} a_1 & -a_2^* & b_1 & -b_2^* \\ a_2 & a_1^* & b_2 & b_1^* \\ c_1 & -c_2^* & d_1 & -d_2^* \\ c_2 & c_1^* & d_2 & d_1^* \end{pmatrix}. \tag{17}$$

With our earlier notation, we think of this matrix as generated by four quaternion-valued functions, writing

$$a=(a_{ij})=\begin{pmatrix} a_1 & -a_2^* \\ a_2 & a_1^*\end{pmatrix}$$

and similarly for the other entries b,c,d. The coalgebra structure on $\mathcal{A}(\mathrm{M}_\theta(2,\mathbb{H}))$ is given by

$$\Delta(A_{ij})=\sum\nolimits_l A_{il}\otimes A_{lj}, \qquad \epsilon(A_{ij})=\delta_{ij}$$

for $i,j=1,\ldots,4$, and its $*$-structure is evident from the matrix (17). The coaction Δ_L is determined to be

$$\Delta_L:\mathcal{A}(\mathbb{C}^4_\theta)\to\mathcal{A}(\mathrm{M}_\theta(2,\mathbb{H}))\otimes\mathcal{A}(\mathbb{C}^4_\theta), \qquad \Delta_L(\Psi_{ia})=\sum\nolimits_j A_{ij}\otimes\Psi_{ja}, \tag{18}$$

where Ψ is the matrix in equation (8) (although here we do not assume the sphere relation and instead think of the entries of Ψ as generators of the algebra $\mathcal{A}(\mathbb{C}^4_\theta)$). The relations between the generators of $\mathcal{A}(\mathrm{M}_\theta(2,\mathbb{H}))$ are found from the requirement that Δ_L make $\mathcal{A}(\mathbb{C}^4_\theta)$ into an $\mathcal{A}(\mathrm{M}_\theta(2,\mathbb{H}))$-comodule algebra. One computes

$$\Delta_L(\Psi_{ia}\Psi_{jb})=\sum\nolimits_{km}(A_{im}A_{jl}-\eta_{ij}\eta_{lm}A_{jl}A_{im})\otimes\Psi_{ma}\Psi_{lb} \tag{19}$$

and, since the products $\Psi_{ma}\Psi_{lb}$ may be taken to be all independent as k,l,a,b vary, we must have that

$$A_{im}A_{jl}=\eta_{ij}\eta_{lm}A_{jl}A_{im} \tag{20}$$

for $i,j,l,m=1,\ldots,4$. It is not difficult to see that the algebra generated by the a_{ij} is commutative, as are the algebras generated by the b_{ij}, c_{ij}, d_{ij}, although overall the algebra is noncommutative due to some non-trivial relations among components in different blocks.

Of course, $\mathcal{A}(\mathrm{M}_\theta(2,\mathbb{H}))$ is not quite a Hopf algebra since it does not have an antipode. We obtain a Hopf algebra by passing to the quotient of $\mathcal{A}(\mathrm{M}_\theta(2,\mathbb{H}))$ by the Hopf $*$-ideal generated by the element $D-1$, where $D=\det A$ is the formal determinant of the matrix A in (17). We denote the quotient by $\mathcal{A}(\mathrm{SL}_\theta(2,\mathbb{H}))$, the coordinate algebra on the quantum group $\mathrm{SL}_\theta(2,\mathbb{H})$ of matrices in $\mathrm{M}_\theta(2,\mathbb{H})$ with determinant one, and continue to write the generators of the quotient as A_{ij}. The algebra $\mathcal{A}(\mathrm{SL}_\theta(2,\mathbb{H}))$ inherits a $*$-bialgebra structure from that of $\mathcal{A}(\mathrm{M}_\theta(2,\mathbb{H}))$ and we use the determinant to define an antipode $S:\mathcal{A}(\mathrm{SL}_\theta(2,\mathbb{H}))\to\mathcal{A}(\mathrm{SL}_\theta(2,\mathbb{H}))$ as in [**16**]. The datum $(\mathcal{A}(\mathrm{SL}(2,\mathbb{H})),\Delta,\epsilon,S)$ constitutes a Hopf $*$-algebra.

The Hopf algebra $\mathcal{A}(\mathrm{Sp}_\theta(2))$ is the quotient of $\mathcal{A}(\mathrm{SL}_\theta(2,\mathbb{H}))$ by the two-sided $*$-Hopf ideal generated by

$$\sum\nolimits_l (A^*)_{li}A_{lj}-\delta_{ij}, \qquad i,j=1,\ldots,4.$$

In this algebra we have the relations $A^*A=AA^*=1$, or equivalently $S(A)=A^*$. This Hopf algebra is the coordinate algebra on the quantum group $\mathrm{Sp}_\theta(2)$, the subgroup of $\mathrm{SL}_\theta(2,\mathbb{H})$ of unitary matrices.

Finally there is an inclusion of algebras $\mathcal{A}(S^7_\theta) \hookrightarrow \mathcal{A}(\mathrm{Sp}_\theta(2))$ given on generators by the $*$-algebra map

$$z_1 \mapsto a_1, \quad z_2 \mapsto a_2, \quad z_3 \mapsto c_1, \quad z_4 \mapsto c_2. \tag{21}$$

This means that we may identify the first two columns of the matrix A with the matrix Ψ of equation (8). Similarly there is an algebra inclusion $\mathcal{A}(S^4_\theta) \hookrightarrow \mathcal{A}(\mathrm{Sp}_\theta(2))$ given by

$$x \mapsto a_1 a_1^* - a_2 a_2^* + c_1 c_1^* - c_2 c_2^*, \quad \alpha \mapsto a_1 c_1^* - a_2^* c_2, \quad \beta \mapsto -a_1^* c_2 + a_2 c_1^*. \tag{22}$$

These inclusions yield algebra isomorphisms of $\mathcal{A}(S^7_\theta)$ and $\mathcal{A}(S^4_\theta)$ with certain subalgebras of $\mathcal{A}(\mathrm{Sp}_\theta(2))$ of coinvariants under coactions by appropriate sub-Hopf algebras, thus realising the noncommutative spheres as quantum homogeneous spaces for $\mathrm{Sp}_\theta(2)$. We refer to [**16**] for details of these constructions.

3.2. Quantum conformal transformations. We now review how the quantum groups obtained in the previous section (co)act on the spheres S^7_θ and S^4_θ as 'quantum symmetries'. The coaction

$$\Delta_L : \mathcal{A}(\mathbb{C}^4_\theta) \to \mathcal{A}(\mathrm{SL}_\theta(2,\mathbb{H})) \otimes \mathcal{A}(\mathbb{C}^4_\theta), \qquad \Delta_L(\Psi_{ia}) = \sum\nolimits_j A_{ij} \otimes \Psi_{ja}, \tag{23}$$

is by construction a $*$-algebra map and so, if we assume that the quantity

$$r^2 := \sum\nolimits_j z_j^* z_j$$

is invertible with inverse r^{-2}, then we may also define an inverse for the quantity

$$\rho^2 := \Delta_L \Big(\sum\nolimits_j z_j^* z_j \Big)$$

by $\rho^{-2} := \Delta_L(r^{-2})$. Inverting r^2 corresponds to deleting the origin in $\mathbb{C}^4_\theta$ and we define the coordinate algebra of the corresponding subset of $\mathbb{C}^4_\theta$ by

$$\mathcal{A}_0(\mathbb{C}^4_\theta) := \mathcal{A}(\mathbb{C}^4_\theta)[r^{-2}],$$

the algebra $\mathcal{A}(\mathbb{C}^4_\theta)$ with r^{-2} adjoined. Extending Δ_L as a $*$-algebra map gives a well-defined coaction

$$\Delta_L : \mathcal{A}_0(\mathbb{C}^4_\theta) \to \mathcal{A}(\mathrm{SL}_\theta(2,\mathbb{H})) \otimes \mathcal{A}_0(\mathbb{C}^4_\theta)$$

for which $\mathcal{A}_0(\mathbb{C}^4_\theta)$ is an $\mathcal{A}(\mathrm{SL}_\theta(2,\mathbb{H}))$-comodule algebra.

Writing $\mathcal{A}_0(\widetilde{\mathbb{C}}^4_\theta) := \Delta_L(\mathcal{A}_0(\mathbb{C}^4_\theta))$ for the image of $\mathcal{A}_0(\mathbb{C}^4_\theta)$ under Δ_L, both ρ^2 and ρ^{-2} are central in the algebra $\mathcal{A}_0(\widetilde{\mathbb{C}}^4_\theta)$, since r^2 and r^{-2} are central in $\mathcal{A}_0(\mathbb{C}^4_\theta)$.

Now the coaction Δ_L descends to a coaction of the Hopf algebra $\mathcal{A}(\mathrm{Sp}_\theta(2))$,

$$\Delta_L : \mathcal{A}_0(\mathbb{C}^4_\theta) \to \mathcal{A}(\mathrm{Sp}_\theta(2)) \otimes \mathcal{A}_0(\mathbb{C}^4_\theta), \tag{24}$$

by the same formula (18) now viewed for the quotient $\mathcal{A}(\mathrm{Sp}_\theta(2))$. In particular, for this coaction one has

$$(\Psi^*\Psi)_{ab} \mapsto \sum\nolimits_{ijl} (A^*)_{li} A_{ij} \otimes (\Psi^*)_{al} \Psi_{jb} = \sum\nolimits_{jl} \delta_{lj} \otimes (\Psi^*)_{al} \Psi_{jb} = 1 \otimes (\Psi^*\Psi)_{ab},$$

since the generators A_{ij} satisfy the relations $\sum_i (A^*)_{li} A_{ij} = \delta_{lj}$ in the algebra $\mathcal{A}(\mathrm{Sp}_\theta(2))$. Then both $\mathcal{A}(S^4_\theta)$ and $\mathcal{A}(S^7_\theta)$ are $\mathcal{A}(\mathrm{Sp}_\theta(2))$-comodule algebras, since this coaction preserves the sphere relations (2) and (4).

In contrast, the spheres S^7_θ and S^4_θ are not preserved under the coaction of the larger quantum group $\mathrm{SL}_\theta(2,\mathbb{H})$. Although defined on the algebra $\mathcal{A}_0(\mathbb{C}^4_\theta)$, the coaction Δ_L of $\mathcal{A}(\mathrm{SL}_\theta(2,\mathbb{H}))$ is not well-defined on the seven-sphere $\mathcal{A}(S^7_\theta)$ since it does not preserve the sphere relation $r^2 = 1$ of equation (4). By definition, we have instead that $\Delta_L(r^2) = \rho^2$, meaning that the coaction of $\mathcal{A}(\mathrm{SL}_\theta(2,\mathbb{H}))$ 'inflates' the sphere $\mathcal{A}(S^7_\theta)$ [**16**]. Since r^2 is a central element of $\mathcal{A}_0(\mathbb{C}^4_\theta)$, we may evaluate it as a positive real number. The result is the coordinate algebra of a noncommutative sphere $S^7_{\theta,r}$ of radius r; as this radius varies in $\mathcal{A}_0(\mathbb{C}^4_\theta)$, it sweeps out a family of seven-spheres. Similarly, evaluation of the central element ρ^2 in $\mathcal{A}_0(\widetilde{\mathbb{C}}^4_\theta)$ yields the coordinate algebra of a noncommutative sphere $\widetilde{S}^7_{\theta,\rho}$ of radius ρ and, as the value of ρ varies in $\mathcal{A}_0(\widetilde{\mathbb{C}}^4_\theta)$, it sweeps out another family of seven-spheres. The coaction Δ_L of $\mathcal{A}(\mathrm{SL}_\theta(2,\mathbb{H}))$ on $\mathcal{A}_0(\mathbb{C}^4_\theta)$ serves to map the family parameterised by r^2 onto the family parameterised by ρ^2.

A similar fact is found for the generators α, β, x of the four-sphere algebra $\mathcal{A}(S^4_\theta)$. The coaction of $\mathcal{A}(\mathrm{SL}_\theta(2,\mathbb{H}))$ does not preserve the sphere relation but gives instead that

$$\Delta_L(\alpha^*\alpha + \beta^*\beta + x^2) = \rho^4,$$

and the four-sphere S^4_θ is also inflated. Let us write $\mathcal{A}(\mathcal{Q}_\theta)$ for the subalgebra of $\mathcal{A}_0(\mathbb{C}^4_\theta)$ generated by α, β, x and their conjugates. Then as r^4 varies in $\mathcal{A}(\mathcal{Q}_\theta)$, we get a family of noncommutative four-spheres. Similarly, we define $\tilde{\alpha} := \Delta_L(\alpha)$, $\tilde{\beta} := \Delta_L(\beta)$, $\tilde{x} := \Delta_L(x)$ and so forth, and write $\mathcal{A}(\widetilde{\mathcal{Q}}_\theta)$ for the subalgebra of $\mathcal{A}_0(\widetilde{\mathbb{C}}^4_\theta)$ that they generate. It is precisely the SU(2)-invariant subalgebra of $\mathcal{A}_0(\widetilde{\mathbb{C}}^4_\theta)$, and as ρ^4 varies in $\mathcal{A}(\widetilde{\mathcal{Q}}_\theta)$ we get another family of noncommutative four-spheres. The coaction of the quantum group $\mathcal{A}(\mathrm{SL}_\theta(2,\mathbb{H}))$ maps the family parameterised by r^4 onto the family parameterised by ρ^4.

Thus there is a family of SU(2)-principal fibrations given by the algebra inclusion $\mathcal{A}(\mathcal{Q}_\theta) \hookrightarrow \mathcal{A}_0(\mathbb{C}^4_\theta)$, the family being parameterised by the function r^2. For a fixed value of r^2 we get an SU(2) principal bundle $S^7_{\theta,r} \to S^4_{\theta,r^2}$. Similarly, the algebra inclusion $\mathcal{A}(\widetilde{\mathcal{Q}}_\theta) \hookrightarrow \mathcal{A}_0(\widetilde{\mathbb{C}}^4_\theta)$ defines a family of SU(2)-principal fibrations parameterised by the function ρ^2. The above construction shows that the coaction of the quantum group $\mathcal{A}(\mathrm{SL}_\theta(2,\mathbb{H}))$ carries the former family of principal fibrations onto the latter.

All of this means that, as things stand, we cannot use the presentations of $\mathcal{A}(S^4_\theta)$ and $\mathcal{A}(S^7_\theta)$ of Sect. 2.1 to give a well-defined coaction of $\mathcal{A}(\mathrm{SL}_\theta(2,\mathbb{H}))$, since the sphere relations we use to define them are not preserved by the coaction. Rather we should work with the families of spheres all at once (this is the price we have to pay for working with the coaction of a Hopf algebra rather than the action of a group). To do this, we note that the algebra $\mathcal{A}(S^4_\theta)$ may be identified with the subalgebra of $\mathcal{A}_0(\mathbb{C}^4_\theta)$ generated by $r^{-2}\alpha$, $r^{-2}\beta$, $r^{-2}x$, together with their conjugates, since the sphere relation

$$(r^{-2}\alpha)(r^{-2}\alpha)^* + (r^{-2}\beta)(r^{-2}\beta)^* + (r^{-2}x)^2 = 1 \tag{25}$$

is automatically satisfied in $\mathcal{A}_0(\mathbb{C}^4_\theta)$. The result of doing so is that we have a well-defined coaction,

$$\Delta_L : \mathcal{A}(S^4_\theta) \to \mathcal{A}(\mathrm{SL}_\theta(2,\mathbb{H})) \otimes \mathcal{A}(S^4_\theta),$$

defined on the generators $r^{-2}\alpha$, $r^{-2}\beta$, $r^{-2}x$ and their conjugates, with the sphere relation (25) now preserved by Δ_L. In this way, we think of $\mathrm{SL}_\theta(2,\mathbb{H}))$ as the quantum group of conformal transformations of S^4_θ.

In these new terms, the construction of the defining projector for $\mathcal{A}(S^4_\theta)$ needs to be modified only slightly. We now take the normalised matrix

$$\Psi = r^{-1} \begin{pmatrix} z_1 & z_2 & z_3 & z_4 \\ -z_2^* & z_1^* & -z_4^* & z_3^* \end{pmatrix}^{\mathrm{t}}, \tag{26}$$

at the price of including the generator r^{-1} as well (not a problem in the smooth closure [**16**]). Thanks to the relation (25), we still have $\Psi^*\Psi = 1$ and the required projector is

$$\mathsf{q} := \Psi\Psi^* = \tfrac{1}{2} r^{-2} \begin{pmatrix} r^2 + x & 0 & \alpha & -\bar{\mu}\,\beta^* \\ 0 & r^2 + x & \beta & \mu\,\alpha^* \\ \alpha^* & \beta^* & r^2 - x & 0 \\ -\mu\,\beta & \bar{\mu}\,\alpha & 0 & r^2 - x \end{pmatrix}. \tag{27}$$

By the above discussion, the coaction Δ_L of $\mathcal{A}(\mathrm{SL}_\theta(2,\mathbb{H}))$ is now well-defined on the algebra generated by the entries of this matrix. Writing $\widetilde{\Psi}_{ia} := \Delta_L(\Psi_{ia})$, the image of q under Δ_L is computed to be

$$\tilde{\mathsf{q}} := \widetilde{\Psi}\widetilde{\Psi}^* = \tfrac{1}{2} \rho^{-2} \begin{pmatrix} \rho^2 + \tilde{x} & 0 & \tilde{\alpha} & -\bar{\mu}\,\tilde{\beta}^* \\ 0 & \rho^2 + x & \tilde{\beta} & \mu\,\tilde{\alpha}^* \\ \tilde{\alpha}^* & \tilde{\beta}^* & \rho^2 - x & 0 \\ -\mu\,\tilde{\beta} & \bar{\mu}\,\tilde{\alpha} & 0 & \rho^2 - \tilde{x} \end{pmatrix}. \tag{28}$$

The entries of these projectors generate respectively subalgebras of $\mathcal{A}_0(\mathbb{C}^4_\theta)$ and $\mathcal{A}_0(\widetilde{\mathbb{C}}^4_\theta)$, each parameterising the families of noncommutative four-spheres discussed above.

Finally, we observe that similar statements may be made about the U(1)-principal fibration $S^7_\theta \to \mathbb{CP}^3_\theta$. We do not need a sphere relation in order to define the coordinate algebra $\mathcal{A}(\mathbb{CP}^3_\theta)$: in Sect. 2.3 it was merely convenient to do so. Instead, we may identify $\mathcal{A}(\mathbb{CP}^3_\theta)$ as the U(1)-invariant subalgebra of $\mathcal{A}_0(\mathbb{C}^4_\theta)$ generated by elements $t_1 = r^{-2}z_1z_2^*$, $x_1 = r^{-2}z_1z_2^*$, $x_2 = r^{-2}z_1z_3^*$, $x_3 = r^{-2}z_1z_4^*$ and so forth.

4. A Noncommutative ADHM construction

There is a well-known solution to the problem of constructing instantons on the classical four-sphere S^4 which goes under the name of ADHM construction. Techniques of linear algebra are used to construct vector bundles over twistor space $\mathbb{CP}^3$, which are in turn put together to construct a vector bundle over S^4 equipped with an instanton connection. It is known that all such connections are obtained in this way [**2, 3**].

Our goal here is to generalise the ADHM method to a deformed version which constructs instantons on the noncommutative sphere S^4_θ. The classical construction may be obtained from our deformed version by setting $\theta = 0$. As usual our approach stems from writing the classical construction in a dualised language which does not depend on the commutativity of the available function algebras, although here the situation is not as straightforward as one might first expect. The deformed construction is rather more subtle than it is in the commutative case and produces noncommutative 'families' of instantons.

4.1. A noncommutative space of monads. The algebra $\mathcal{A}(\mathbb{C}^4_\theta)$ has a natural $\mathbb{Z}$-grading given by assigning to its generators the degrees

$$\deg(z_j) = 1, \quad \deg(z_j^*) = -1, \qquad j = 1, \ldots, 4,$$

which results in a decomposition $\mathcal{A}(\mathbb{C}^4_\theta) = \oplus_{n\in\mathbb{Z}}\mathcal{A}_n$. Then for each $r \in \mathbb{Z}$ there is a 'degree shift' map from $\mathcal{A}(\mathbb{C}^4_\theta)$ to itself whose image we denote $\mathcal{A}(\mathbb{C}^4_\theta)(r)$; by definition the degree n component of $\mathcal{A}(\mathbb{C}^4_\theta)(r)$ is $\mathcal{A}_{r+n}$.

Similarly, if a given $\mathcal{A}(\mathbb{C}^4_\theta)$-module $\mathcal{E}$ is $\mathbb{Z}$-graded, we denote the degree-shifted modules by $\mathcal{E}(r)$, $r \in \mathbb{Z}$. In particular, for each finite dimensional vector space H the corresponding free right module $H \otimes \mathcal{A}(\mathbb{C}^4_\theta)$ is $\mathbb{Z}$-graded by the grading on $\mathcal{A}(\mathbb{C}^4_\theta)$, and the shift maps on $\mathcal{A}(\mathbb{C}^4_\theta)$ induce the shift maps on $H \otimes \mathcal{A}(\mathbb{C}^4_\theta)$.

The input data for the classical ADHM construction of SU(2) instantons with topological charge k is a *monad*, by which we mean a sequence of free right modules over the algebra $\mathcal{A}(\mathbb{C}^4)$,

$$H \otimes \mathcal{A}(\mathbb{C}^4)(-1) \xrightarrow{\sigma_z} K \otimes \mathcal{A}(\mathbb{C}^4) \xrightarrow{\tau_z} L \otimes \mathcal{A}(\mathbb{C}^4)(1), \tag{29}$$

where H, K and L are complex vector spaces of dimensions k, $2k+2$ and k respectively. The arrows σ_z and τ_z are $\mathcal{A}(\mathbb{C}^4)$-module homomorphisms assumed to be such that σ_z is injective, τ_z is surjective and that the composition $\tau_z\sigma_z = 0$. This is the usual approach in algebraic geometry [**20**], although here we work with $\mathcal{A}(\mathbb{C}^4)$-modules, *i.e.* global sections of vector bundles, rather than with locally-free sheaves.

The degree shifts signify that we think of σ_z and τ_z respectively as elements of $H^* \otimes K \otimes \mathcal{A}_1$ and $K^* \otimes L \otimes \mathcal{A}_1$, where $\mathcal{A}_1$ is the degree one component of $\mathcal{A}(\mathbb{C}^4)$ (the vector space spanned by the generators $z_1, \ldots, z_4$). This means that alternatively we may think of them as linear maps

$$\sigma_z : H \times \mathbb{C}^4 \to K, \qquad \tau_z : K \times \mathbb{C}^4 \to L, \tag{30}$$

thus recovering the more explicit geometric approach of [**2**].

Our goal in this section is to give a description of a monad of the form (29) in an algebraic framework which allows the possibility of the algebra $\mathcal{A}(\mathbb{C}^4_\theta)$ being noncommutative. In this setting, we require the maps σ_z and τ_z to be parameterised by the noncommutative space $\mathbb{C}^4_\theta$ rather than by the classical space $\mathbb{C}^4$, as was the case in equation (30). Our first task then is to find an analogue of the space of linear module maps $H \otimes \mathcal{A}(\mathbb{C}^4_\theta)(-1) \to K \otimes \mathcal{A}(\mathbb{C}^4_\theta)$.

Following a general strategy [**23**], we define $\mathcal{A}(\widetilde{\mathcal{M}}_\theta(H,K))$ to be the universal algebra for which there is a morphism of right $\mathcal{A}(\mathbb{C}^4_\theta)$-modules,

$$\sigma_z : H \otimes \mathcal{A}(\mathbb{C}^4_\theta)(-1) \to \mathcal{A}(\widetilde{\mathcal{M}}_\theta(H,K)) \otimes K \otimes \mathcal{A}(\mathbb{C}^4_\theta),$$

which is linear in the generators $z_1, \ldots, z_4$ of $\mathcal{A}(\mathbb{C}^4_\theta)$. By this we mean that whenever $\mathcal{B}$ is an algebra satisfying these properties there exists a morphism of algebras $\phi : \mathcal{A}(\widetilde{\mathcal{M}}_\theta(H,K)) \to \mathcal{B}$ and a commutative diagram

$$\begin{array}{ccc} H \otimes \mathcal{A}(\mathbb{C}^4_\theta)(-1) & \xrightarrow{\sigma_z} & \mathcal{A}(\widetilde{\mathcal{M}}_\theta(H,K)) \otimes K \otimes \mathcal{A}(\mathbb{C}^4_\theta) \\ \downarrow{\scriptstyle \mathrm{id}} & & \downarrow{\scriptstyle \phi\otimes\mathrm{id}} \\ H \otimes \mathcal{A}(\mathbb{C}^4_\theta)(-1) & \xrightarrow{\sigma'_z} & \mathcal{B} \otimes K \otimes \mathcal{A}(\mathbb{C}^4_\theta) \end{array}$$

of right $\mathcal{A}(\mathbb{C}^4_\theta)$-modules, with σ'_z denoting the corresponding map for the algebra $\mathcal{B}$.

Choosing a basis $(u_1, \ldots, u_k)$ for the vector space H and a basis $(v_1, \ldots, v_{2k+2})$ for the vector space K, the algebra $\mathcal{A}(\widetilde{\mathcal{M}}_\theta(H,K))$ is generated by the matrix elements

$$\{M^\alpha_{ab} \mid a = 1, \ldots, 2k+2,\ b = 1, \ldots, k,\ \alpha = 1, \ldots, 4\},$$

which define a map σ_z, expressed on simple tensors by

$$\sigma_z : u_b \otimes Z \mapsto \sum\nolimits_{a,\alpha} M^\alpha_{ab} \otimes v_a \otimes z_\alpha Z, \qquad Z \in \mathcal{A}(\mathbb{C}^4_\theta). \tag{31}$$

In more compact notation, for each α we arrange these elements into a $(2k+2)\times k$ matrix $M^\alpha = (M^\alpha_{ab})$, so that with respect to the above bases, σ_z may be written

$$\sigma_z = \sum\nolimits_\alpha M^\alpha \otimes z_\alpha. \tag{32}$$

To find the relations in the algebra $\mathcal{A}(\widetilde{\mathcal{M}}_\theta(H,K))$, let us write $(\hat{u}_1, \ldots, \hat{u}_k)$ for the basis of H^* which is dual to $(u_1, \ldots, u_k)$ and write $(\hat{v}_1, \ldots, \hat{v}_{2k+2})$ for the basis of K^* dual to $(v_1, \ldots, v_{2k+2})$. Then the map (31) has an equivalent dual description (also denoted σ_z) in terms of the dual vector spaces H^*, K^* as

$$\sigma_z : \hat{v}_a \otimes Z \mapsto \sum\nolimits_{b,\alpha} M^\alpha_{ab} \otimes \hat{u}_b \otimes z_\alpha Z, \tag{33}$$

and extended as an $\mathcal{A}(\mathbb{C}^4_\theta)$-module map. The functionals $\hat{u}_b$, $\hat{v}_a$ together with their conjugates $\hat{u}^*_b$, $\hat{v}^*_a$ generate the coordinate algebras of H and K respectively. It is only natural to require that (33) be an algebra map.

PROPOSITION 4.1. *With $(\eta_{\alpha\beta})$ the matrix (5) of deformation parameters, the matrix elements M^α_{ab} enjoy the relations*

$$M^\alpha_{ab} M^\beta_{cd} = \eta_{\beta\alpha} M^\beta_{cd} M^\alpha_{ab} \tag{34}$$

for each $a, c = 1, \ldots, 2k+2$, each $b, d = 1, \ldots, k$ and each $\alpha, \beta = 1, \ldots, 4$.

PROOF. The requirement that (33) is an algebra map means that in degree one we need $\sigma_z(\hat{v}_a\hat{v}_c) = \sigma_z(\hat{v}_c\hat{v}_a)$ for all $a, c = 1, \ldots, 2k+2$, which translates into the statement that

$$\sum\nolimits_{b,d,\alpha,\beta} M^\alpha_{ab} M^\beta_{cd} \otimes \hat{u}_b\hat{u}_d \otimes z_\alpha z_\beta = \sum\nolimits_{b,d,\alpha,\beta} M^\beta_{cd} M^\alpha_{ab} \otimes \hat{u}_d\hat{u}_b \otimes z_\beta z_\alpha$$

for all $a, c = 1, \ldots, 2k+2$. Using in turn the relations (3) and the fact that the generators $\hat{u}_b$, $\hat{u}_d$ commute for all values of b, d, this equation may be rearranged to give

$$\sum\nolimits_{b,d,\alpha,\beta} \left(M^\alpha_{ab} M^\beta_{cd} - \eta_{\beta\alpha} M^\beta_{cd} M^\alpha_{ab} \right) \otimes \hat{u}_b\hat{u}_d \otimes z_\alpha z_\beta = 0.$$

Since for $b \le d$ and $\alpha \le \beta$ the quantities $\hat{u}_b\hat{u}_d \otimes z_\alpha z_\beta$ may all be taken to be independent, we must have that their coefficients are all zero, leading to the stated relations. $\square$

The above proposition simply says that the entries of a given matrix M^α all commute, whereas the relations between the entries of the matrices M^α and M^β are determined by the deformation parameter $\eta_{\beta\alpha}$. Hence the algebra $\mathcal{A}(\widetilde{\mathcal{M}}_\theta(H,K))$ is generated by the M^α_{ab} subject to the relations (34). The algebra $\mathcal{A}(\widetilde{\mathcal{M}}_{\theta=0}(H,K))$ is commutative and parameterises the space of all possible maps σ_z, since for each point $x \in \widetilde{\mathcal{M}}_{\theta=0}(H,K)$ there is an evaluation map,

$$\mathrm{ev}_x : \mathcal{A}(\widetilde{\mathcal{M}}_{\theta=0}(H,K)) \to \mathbb{C},$$

which yields an $\mathcal{A}(\mathbb{C}^4)$-module homomorphism

$$(\mathrm{ev}_x \otimes \mathrm{id})\sigma_z : H \otimes \mathcal{A}(\mathbb{C}^4)(-1) \to K \otimes \mathcal{A}(\mathbb{C}^4),$$
$$(\mathrm{ev}_x \otimes \mathrm{id})\sigma_z := \sum \mathrm{ev}_x(M^\alpha_{ab}) \otimes z_\alpha.$$

When θ is different from zero, there need not be enough evaluation maps available. Nevertheless, we think of $\mathcal{A}(\widetilde{\mathcal{M}}_\theta(H,K))$ as a noncommutative family of maps parameterised by the noncommutative space $\widetilde{\mathcal{M}}_\theta(H,K)$.

REMARK 4.2. Since we constructed $\mathcal{A}(\widetilde{\mathcal{M}}_\theta(H,K))$ through the minimal requirement that σ_z is an algebra map, it is indeed the universal algebra with the required properties. This means that our interpretation of $\mathcal{A}(\widetilde{\mathcal{M}}_\theta(H,K))$ as a noncommutative family of maps is in agreement with the approaches of [**23, 25, 22**] for quantum families of maps parameterised by noncommutative spaces. Moreover, it also agrees with the definition of algebras of rectangular quantum matrices discussed in [**17**]. It may also be viewed as a kind of 'comeasuring' as introduced in [**18**], but now for modules instead of algebras.

Thus we have a noncommutative analogue of the space of all maps σ_z. A similar construction works for the maps τ_z: there is a universal algebra $\mathcal{A}(\widetilde{\mathcal{M}}_\theta(K,L))$ generated by matrix elements N^α_{ba} for labels $b = 1,\dots,k$, $a = 1,\dots,2k+2$ and $\alpha = 1,\dots,4$, here coming from a map

$$\tau_z : v_a \otimes Z \mapsto \sum\nolimits_{b,\alpha} N^\alpha_{ba} \otimes w_b \otimes z_\alpha Z, \tag{35}$$

having chosen a basis $(w_1,\dots,w_k)$ for the vector space L. Dually, the requirement that τ_z be an algebra map from the coordinate algebra of L to the coordinate algebra of K results in relations for the generators of the algebra $\mathcal{A}(\widetilde{\mathcal{M}}_\theta(K,L))$,

$$N^\alpha_{ba} N^\beta_{dc} = \eta_{\beta\alpha} N^\beta_{dc} N^\alpha_{ba}, \tag{36}$$

which are the parallel of conditions (34) for the algebra $\mathcal{A}(\widetilde{\mathcal{M}}_\theta(H,K))$.

To complete the monad picture we finally require that the composition of the maps σ_z and τ_z be zero. In the dualised format the composition is easily dealt with as the composition as a map from the coordinate algebra of L to that of H, with the product appearing as part of a general procedure for 'gluing' quantum matrices [**17**]. By this we mean that the composition $\vartheta_z := \tau_z \circ \sigma_z$ is given in terms of an algebra-valued $k\times k$ matrix, the product of a $k\times(2k+2)$ matrix with a $(2k+2)\times k$ matrix. Explicitly, the map is

$$\vartheta_z : H \otimes \mathcal{A}(\mathbb{C}^4_\theta)(-1) \to \mathcal{A}(\widetilde{\mathcal{M}}_\theta(H,L)) \otimes L \otimes \mathcal{A}(\mathbb{C}^4_\theta)(1),$$
$$\vartheta_z : \hat{w}_a \otimes Z \mapsto \sum\nolimits_{b,\alpha,\beta} T^{\alpha,\beta}_{ab} \otimes \hat{w}_b \otimes z_\alpha z_\beta Z,$$

where $\mathcal{A}(\widetilde{\mathcal{M}}_\theta(H,L))$ is the coordinate algebra generated by the matrix elements $T^{\alpha,\beta}_{ab}$ for $\alpha,\beta = 1,\dots,4$ and $a,b = 1,\dots,k$. The matrix multiplication $(\tau_z,\sigma_z) \mapsto \vartheta_z$ now appears as a 'coproduct'

$$\mathcal{A}(\widetilde{\mathcal{M}}_\theta(H,L)) \to \mathcal{A}(\widetilde{\mathcal{M}}_\theta(K,L)) \otimes \mathcal{A}(\widetilde{\mathcal{M}}_\theta(H,K)),$$
$$T^{\alpha,\beta}_{cd} := \sum\nolimits_b N^\alpha_{cb} \otimes M^\beta_{bd}, \qquad \alpha,\beta = 1,\dots,4,\ c,d = 1,\dots,k.$$

The condition $\tau_z\sigma_z = 0$ is therefore that the image of this map in $\mathcal{A}(\widetilde{\mathcal{M}}_\theta(K,L)) \otimes \mathcal{A}(\widetilde{\mathcal{M}}_\theta(H,K))$ is zero; this is established by the following proposition.

PROPOSITION 4.3. *The condition $\tau_z\sigma_z = 0$ is equivalent to the requirement that*

$$\sum\nolimits_r (N^\alpha_{br} M^\beta_{rd} + \eta_{\beta\alpha} N^\beta_{br} M^\alpha_{rd}) = 0 \tag{37}$$

for all $b, d = 1, \ldots, k$ and all $\alpha, \beta = 1, \ldots, 4$.

PROOF. In terms of algebra-valued matrices the map $\tau_z\sigma_z$ is computed as the composition of the duals of the maps (31) and (35), following the discussion above, to be equal to

$$(\tau_z\sigma_z)_{bd} = \sum\nolimits_{r,\alpha,\beta} N^\alpha_{br} M^\beta_{rd} \otimes z_\alpha z_\beta.$$

Equating to zero the coefficients of the linearly independent generators $z_\alpha z_\beta$ for $\alpha \le \beta$ gives the relations as stated. □

The conditions in equation (37) may be expressed more compactly in terms of products of matrices as

$$N^\alpha M^\beta + \eta_{\beta\alpha} N^\beta M^\alpha = 0,$$

for $\alpha, \beta = 1, \ldots, 4$ (and as in (36) there is no sum over α and β in this expression).

DEFINITION 4.4. Define $\mathcal{A}(\widetilde{\mathcal{M}}_{\theta;k})$ to be the algebra generated by the matrix elements M^α_{ab} and N^β_{ba} subject to the relations

$$M^\alpha_{ab} M^\beta_{cd} = \eta_{\beta\alpha} M^\beta_{cd} M^\alpha_{ab}, \quad N^\alpha_{ba} N^\beta_{dc} = \eta_{\beta\alpha} N^\beta_{dc} N^\alpha_{ba},$$

as well as the relations

$$\sum\nolimits_r (N^\alpha_{dr} M^\beta_{rb} + \eta_{\beta\alpha} N^\beta_{br} M^\alpha_{rd}) = 0$$

for all $\alpha, \beta = 1, \ldots, 4$, all $b, d = 1, \ldots, k$ and all $a, c = 1, \ldots, 2k+2$.

The noncommutative algebra $\mathcal{A}(\widetilde{\mathcal{M}}_{\theta;k})$ is by construction universal amongst all algebras having the property that the resulting maps σ_z and τ_z are algebra maps which compose to zero. Our interpretation is that for fixed k the collection of monads over $\mathbb{C}^4_\theta$ is parameterised by the noncommutative space which is 'dual' to this algebra.

4.2. The subspace of self-dual monads. In the classical case, the input datum of a monad is by itself insufficient to construct bundles over the four-sphere S^4. To achieve this, one must incorporate the quaternionic structure afforded by the map J as in (15) (in the classical limit) and ensure that the monad is compatible with this extra structure. The same is true in the noncommutative case, as we shall see presently.

Given the pair of maps constructed in the previous section,

$$\sigma_z : H \otimes \mathcal{A}(\mathbb{C}^4_\theta)(-1) \to \mathcal{A}(\widetilde{\mathcal{M}}_\theta(H,K)) \otimes K \otimes \mathcal{A}(\mathbb{C}^4_\theta),$$
$$\tau_z : K \otimes \mathcal{A}(\mathbb{C}^4_\theta) \to \mathcal{A}(\widetilde{\mathcal{M}}_\theta(K,L)) \otimes L \otimes \mathcal{A}(\mathbb{C}^4_\theta)(1),$$

we firstly note that the anti-algebra map J in (15) induces a new pair of maps,

$$\sigma_{J(z)} : H \otimes J\left(\mathcal{A}(\mathbb{C}^4_\theta)(-1)\right) \to \mathcal{A}\left(\widetilde{\mathcal{M}}_\theta(H,K)\right) \otimes K \otimes J\left(\mathcal{A}(\mathbb{C}^4_\theta)\right),$$

$$\sigma_{J(z)} := \sum_\alpha M^\alpha \otimes J(z_\alpha), \tag{38}$$

and

$$\tau_{J(z)} : K \otimes J\left(\mathcal{A}(\mathbb{C}^4_\theta)\right) \to \mathcal{A}\left(\widetilde{\mathcal{M}}_\theta(K,L)\right) \otimes L \otimes J\left(\mathcal{A}(\mathbb{C}^4_\theta)(1)\right),$$

$$\tau_{J(z)} := \sum_\alpha N^\alpha \otimes J(z_\alpha). \tag{39}$$

Here, $J\left(\mathcal{A}(\mathbb{C}^4_\theta)\right)$ is the left $\mathcal{A}(\mathbb{C}^4_\theta)$-module induced by the anti-algebra map J and $\sigma_{J(z)}$, $\tau_{J(z)}$ are homomorphisms of left $\mathcal{A}(\mathbb{C}^4_\theta)$-modules. We may also take the adjoints of the above maps. To make sense of this, we need to add to our picture the matrix elements ${M^\alpha_{ab}}^*$, so that the adjoint of σ_z is

$$\sigma^\star_z : v_a \otimes Z \mapsto \sum\nolimits_{b,\alpha} {M^\alpha_{ab}}^* \otimes u_b \otimes z^*_\alpha Z, \qquad Z \in \mathcal{A}(\mathbb{C}^4_\theta), \tag{40}$$

where $a = 1, \dots, 2k+2$, $b = 1, \dots, k$ and $\alpha = 1, \dots, 4$. Let us denote by $M^{\alpha\dagger}$ the $k \times (2k+2)$ matrix with entries $(M^{\alpha\dagger})_{ba} = {M^\alpha_{ab}}^*$. Then with respect to the above choice of bases, the adjoint map $\sigma^\star_z$ may be written more compactly as

$$\sigma^\star_z = \sum\nolimits_\alpha M^{\alpha\dagger} \otimes z^*_\alpha.$$

Similarly, we add the matrix elements ${N^\alpha_{dc}}^*$ and write $(N^{\alpha\dagger})_{cd} = {N^\alpha_{dc}}^*$, so that the adjoint of τ_z is

$$\tau^\star_z : w_b \otimes Z \mapsto \sum\nolimits_{a,\alpha} {N^\alpha_{ba}}^* \otimes v_a \otimes z^*_\alpha Z,$$

or $\tau^\star_z = \sum_\alpha N^{\alpha\dagger} \otimes z^*_\alpha$ in compact notation. The elements ${M^\alpha_{ab}}^*$ are the generators of the algebra $\mathcal{A}(\widetilde{\mathcal{M}}_\theta(K^*, H^*))$, whereas the elements ${N^\alpha_{dc}}^*$ are the generators the algebra $\mathcal{A}(\widetilde{\mathcal{M}}_\theta(L^*, K^*))$. Applied to equations (38) and (39), all of this yields a pair of homomorphisms of right $\mathcal{A}(\mathbb{C}^4_\theta)$-modules

$$\sigma^\star_{J(z)} : K^* \otimes J\left(\mathcal{A}(\mathbb{C}^4_\theta)\right)^* \to \mathcal{A}\left(\widetilde{\mathcal{M}}_\theta(K^*, H^*)\right) \otimes H^* \otimes J\left(\mathcal{A}(\mathbb{C}^4_\theta)\right)^*(1),$$

$$\tau^\star_{J(z)} : L^* \otimes J\left(\mathcal{A}(\mathbb{C}^4_\theta)\right)^*(-1) \to \mathcal{A}\left(\widetilde{\mathcal{M}}_\theta(L^*, K^*)\right) \otimes K^* \otimes J\left(\mathcal{A}(\mathbb{C}^4_\theta)\right)^*,$$

defined respectively by

$$\sigma^\star_{J(z)} = \sum_\alpha M^{\alpha\dagger} \otimes J(z_\alpha)^*, \qquad \tau^\star_{J(z)} = \sum_\alpha N^{\alpha\dagger} \otimes J(z_\alpha)^*.$$

Of course, we may identify the vector spaces H and L^* through the basis isomorphism $u_b \mapsto \hat{w}_b$ for each $b = 1, \dots, k$. Similarly the isomorphism $v_a \mapsto \hat{v}_a$ for $a = 1, \dots, 2k+2$ gives an identification of the vector space K with its dual K^*. Also, the right module $J(\mathcal{A}(\mathbb{C}^4_\theta))^*$ may be identified with $\mathcal{A}(\mathbb{C}^4_\theta)$ by the composition of the map J with the involution $*$ (noting that this identification is not the identity map). Through these identifications, we may think of $\sigma^\star_{J(z)}$ and $\tau^\star_{J(z)}$ as module homomorphisms

$$\tau^\star_{J(z)} : H \otimes \mathcal{A}(\mathbb{C}^4_\theta)(-1) \to \mathcal{A}\left(\widetilde{\mathcal{M}}_\theta(H, K)\right) \otimes K \otimes \mathcal{A}(\mathbb{C}^4_\theta),$$

$$\sigma^\star_{J(z)} : K \otimes \mathcal{A}(\mathbb{C}^4_\theta) \to \mathcal{A}\left(\widetilde{\mathcal{M}}_\theta(K, L)\right) \otimes L \otimes \mathcal{A}(\mathbb{C}^4_\theta)(1).$$

It is straightforward to check that we now have $\sigma^\star_{J(z)}\tau^\star_{J(z)} = 0$ and so all of this means that the maps $\sigma^\star_{J(z)}$ and $\tau^\star_{J(z)}$ also give a parameterisation of the noncommutative space of monads, albeit a different parameterisation from the one we started with. In the classical case the above procedure applied to a given monad again yields a monad, although it is not necessarily the one we started with. If fact, in the classical case, one is interested only in the subset of monads which are invariant under the above construction, namely the monad obtained by applying J and dualising is required to be isomorphic to the one we start with (this is the sense in which we require monads to be compatible with J). We call such monads *self-dual*. In our algebraic framework, where we work not with specific monads but rather with the (possibly noncommutative) space $\widetilde{\mathcal{M}}_{\theta;k}$ of all monads, this extra requirement is encoded as follows.

PROPOSITION 4.5. *The space of self-dual monads is parameterised by the algebra $\mathcal{A}(\widetilde{\mathcal{M}}^{SD}_{\theta;k})$, the quotient of the algebra $\mathcal{A}(\widetilde{\mathcal{M}}_{\theta;k})$ by the further relations*

$$N^1 = -M^{2\dagger}, \quad N^2 = M^{1\dagger}, \quad N^3 = -M^{4\dagger}, \quad N^4 = M^{3\dagger}. \tag{41}$$

PROOF. The condition that the maps σ_z and τ_z should parameterise self-dual monads is that $\sigma_z = \tau^\star_{J(z)}$, equivalently that $\tau_z = -\sigma^\star_{J(z)}$. In terms of the matrices M^α, N^α, the former condition reads

$$\sum\nolimits_\alpha M^\alpha \otimes z_\alpha = \sum\nolimits_\alpha N^{\alpha\dagger} \otimes J(z_\alpha)^*. \tag{42}$$

Equating coefficients of generators of $\mathcal{A}(\mathbb{C}^4_\theta)$ in each of these equations yields the extra relations as stated. □

REMARK 4.6. The identification of the vector space K with its dual K^* means that the module $K \otimes \mathcal{A}(\mathbb{C}^4_\theta)$ acquires a bilinear form given by

$$(\xi, \eta) := \langle J\xi | \eta \rangle = \sum\nolimits_a (J\xi)^*_a \eta_a \tag{43}$$

for $\xi = (\xi_a)$ and $\eta = (\eta_a) \in K \otimes \mathcal{A}(\mathbb{C}^4_\theta)$, with $\langle \cdot | \cdot \rangle$ the canonical Hermitian structure on $K \otimes \mathcal{A}(\mathbb{C}^4_\theta)$. The monad condition, which now reads

$$0 = \tau_z \sigma_z = -\sigma^\star_{J(z)} \sigma_z,$$

translates into the more practical condition that the columns of the matrix σ_z (equivalently the rows of τ_z) are orthogonal with respect to the form $(\,\cdot\,,\,\cdot\,)$.

Moreover, we see that

$$\begin{aligned} 0 = \tau_{z+J(z)}\sigma_{z+J(z)} &= \tau_z\sigma_z + \tau_z\sigma_{J(z)} + \tau_{J(z)}\sigma_z + \tau_{J(z)}\sigma_{J(z)} = \tau_z\sigma_{J(z)} + \tau_{J(z)}\sigma_z \\ &= -\sigma^\star_{J(z)}\sigma_{J(z)} + \sigma^\star_z\sigma_z \end{aligned}$$

so that in the matrix algebra $\mathrm{M}_k(\mathbb{C}) \otimes \mathcal{A}(\widetilde{\mathcal{M}}^{SD}_{\theta;k}) \otimes \mathcal{A}(\mathbb{C}^4_\theta)$ we have also

$$\sigma^\star_{J(z)}\sigma_{J(z)} = \sigma^\star_z\sigma_z.$$

REMARK 4.7. The above identifications of vector spaces $H \cong L^*$ and $K \cong K^*$ yield an identification of $\mathcal{A}(\widetilde{\mathcal{M}}_\theta(H, K))$ with $\mathcal{A}(\widetilde{\mathcal{M}}_\theta(L^*, K^*))$ and hence a reality structure on the generators M^α_{ab}. It follows that the space of self-dual monads is parameterised by a total of $4k(2k+2)$ generators M^α_{ab}. As already remarked, the condition $\sigma^\star_{J(z)}\sigma_z = 0$ is equivalent to demanding that the columns of σ_z are pairwise orthogonal with respect to the bilinear form $(\,\cdot\,,\,\cdot\,)$ and, since σ_z has k columns, this yields $\frac{1}{2}k(k-1)$ such orthogonality conditions. Now as in Prop. 4.3 we may equate to zero the coefficients of the products $z_\alpha z_\beta$ for $\alpha \leq \beta$, and we note that there are 10 such coefficients in each orthogonality condition. This yields a total of $5k(k-1)$ constraints on the generators M^α_{ab}.

4.3. ADHM construction of noncommutative instantons. We are ready for the construction of charge k noncommutative bundles with instanton connections. As in previous sections, we have the $(2k+2) \times k$ algebra-valued matrices

$$\sigma_z = M^1 \otimes z_1 + M^2 \otimes z_2 + M^3 \otimes z_3 + M^4 \otimes z_4,$$
$$\sigma_{J(z)} = -M^1 \otimes z_2^* + M^2 \otimes z_1^* - M^3 \otimes z_4^* + M^4 \otimes z_3^*$$

which, as already observed, have the properties $\sigma^\star_{J(z)}\sigma_z = 0$ and $\sigma^\star_{J(z)}\sigma_{J(z)} = \sigma^\star_z\sigma_z$.

LEMMA 4.8. *The entries of the matrix $\rho^2 := \sigma^\star_z\sigma_z = \sigma^\star_{J(z)}\sigma_{J(z)}$ commute with those of the matrix σ_z.*

PROOF. One finds that the (μ,ν) entry of ρ^2 is

$$(\rho^2)_{\mu\nu} = \sum\nolimits_{r,\alpha,\beta} (M^{\alpha\dagger})_{\mu r} M^\beta_{r\nu} \otimes z^*_\alpha z_\beta$$

and that the a,b entry of σ_z is

$$(\sigma_z)_{ab} = \sum\nolimits_\gamma M^\gamma_{ab} \otimes z_\gamma .$$

It is straightforward to check that these elements always commute using the relations (3) for $\mathcal{A}(\mathbb{C}^4_\theta)$ and the relations of Prop. 4.1 for $\mathcal{A}(\widetilde{\mathcal{M}}_\theta(H,K))$. The essential feature is that every factor of $\eta_{\beta\alpha}$ coming from the relations between the M^α's is cancelled by a factor of $\eta_{\alpha\beta}$ coming from the relations between the z_α's. $\square$

We need to enlarge slightly the matrix algebra $\mathrm{M}_k(\mathbb{C}) \otimes \mathcal{A}(\widetilde{\mathcal{M}}^{SD}_{\theta;k}) \otimes \mathcal{A}(\mathbb{C}^4_\theta)$ by adjoining an inverse element ρ^{-2} for ρ^2, together with a square root ρ^{-1}. That these matrices may be inverted is an assumption, even in the commutative case where doing so corresponds to the deletion of the non-generic points of the moduli space; these correspond to so-called 'instantons of zero size'.

From the previous lemma the matrix ρ^2, which is self-adjoint by construction, has entries in the centre of the algebra $\mathcal{A}(\widetilde{\mathcal{M}}^{SD}_{\theta;k}) \otimes \mathcal{A}(\mathbb{C}^4_\theta)$, so these new matrices ρ^{-1} and ρ^{-2} must also be self-adjoint with central entries. We collect the matrices σ_z, $\sigma_{J(z)}$ together into the $(2k+2) \times 2k$ matrix

$$\mathrm{V} := \begin{pmatrix} \sigma_z & \sigma_{J(z)} \end{pmatrix} \tag{44}$$

and we have by construction that

$$\mathrm{V}^*\mathrm{V} = \rho^2 \begin{pmatrix} \mathbb{I}_k & 0 \\ 0 & \mathbb{I}_k \end{pmatrix},$$

where $\mathbb{I}_k$ denotes the $k \times k$ identity matrix. This of course means that the quantity

$$\mathsf{Q} := \mathrm{V}\rho^{-2}\mathrm{V}^* = \sigma_z\rho^{-2}\sigma^\star_z + \sigma_{J(z)}\rho^{-2}\sigma^\star_{J(z)} \tag{45}$$

is automatically a projection: $\mathsf{Q}^2 = \mathsf{Q} = \mathsf{Q}^*$. For convenience we denote

$$\mathsf{Q}_z := \sigma_z\rho^{-2}\sigma^\star_z, \qquad \mathsf{Q}_{J(z)} := \sigma_{J(z)}\rho^{-2}\sigma^\star_{J(z)},$$

which are themselves projections, in fact orthogonal ones, $\mathsf{Q}_{J(z)}\mathsf{Q}_z = 0$, due to the fact that $\sigma^\star_{J(z)}\sigma_z = 0$.

LEMMA 4.9. *The trace of the projection Q_z is equal to k; likewise for $\mathsf{Q}_{J(z)}$.*

PROOF. We compute the trace as follows:

$$\begin{aligned}\operatorname{Tr} \mathsf{Q}_z &= \sum\nolimits_{\mu} (\sigma_z \rho^{-2} \sigma_z^{\star})_{\mu\mu} = \sum\nolimits_{\mu,r,s} (\sigma_z)_{\mu r} (\rho^{-2})_{rs} (\sigma_z^{\star})_{s\mu} \\ &= \sum\nolimits_{\mu,r,s} (\rho^{-2})_{rs} (\sigma_z)_{\mu r} (\sigma_z)^{\star}_{s\mu} = \sum\nolimits_{\mu,r,s} (\rho^{-2})_{rs} (\sigma_z)^{\star}_{s\mu} (\sigma_z)_{\mu r} \\ &= \sum\nolimits_{r,s} (\rho^{-2})_{rs} (\sigma_z^{\star} \sigma_z)_{sr} = \operatorname{Tr} \mathbb{I}_k = k.\end{aligned}$$

In the third equality we have used the fact that, as said, the entries of ρ^{-2} commute with those of σ_z, whereas in the fourth equality we have used the fact that every element of $\mathcal{A}(\widetilde{\mathcal{M}}^{SD}_{\theta;k}) \otimes \mathcal{A}(\mathbb{C}^4_\theta)$ commutes with its own adjoint. An analogous chain of equality establishes the same result for the projection $\mathsf{Q}_{J(z)}$. $\square$

As a consequence the projection Q has trace $2k$.

PROPOSITION 4.10. *The operator*

$$\mathsf{P} := \mathbb{I}_{2k+2} - \mathsf{Q}$$

is a projection in the algebra $\mathrm{M}_{2k+2}\left(\mathcal{A}(\widetilde{\mathcal{M}}^{SD}_{\theta;k}) \otimes \mathcal{A}(S^4_\theta)\right)$ *with trace equal to* 2.

PROOF. The entries of the projection Q_z are in the subalgebra of $\mathcal{A}(\widetilde{\mathcal{M}}^{SD}_{\theta;k}) \otimes \mathcal{A}(\mathbb{C}^4_\theta)$ made up of U(1)-invariants which, by the discussion of Sect. 3.2, is precisely $\mathcal{A}(\widetilde{\mathcal{M}}^{SD}_{\theta;k}) \otimes \mathcal{A}(\mathbb{CP}^3_\theta)$. Now recall from Sect. 2.3 that the degree one elements of $\mathcal{A}(\mathbb{CP}^3_\theta)$ of the form $Z + J(Z)$ generate the J-invariant subalgebra, which may be identified with $\mathcal{A}(S^4_\theta)$. The entries of Q_z being linear in the generators of $\mathcal{A}(\mathbb{CP}^3_\theta)$, it follows that the projection Q has entries in $\mathcal{A}(\widetilde{\mathcal{M}}^{SD}_{\theta;k}) \otimes \mathcal{A}(S^4_\theta)$. The same is true of the complementary projection P as well. Finally, since the projection Q has trace $2k$, the trace of the projector P is just 2. $\square$

We think of the projective right $\mathcal{A}(S^4_\theta)$-module $\mathcal{E} := \mathsf{P}\mathcal{A}(S^4_\theta)^{2k+2}$ as defining a family of rank two vector bundles over S^4_θ parameterised by the noncommutative space $\widetilde{\mathcal{M}}^{SD}_{\theta;k}$. We equip this family of vector bundles with the associated family of Grassmann connections $\nabla := \mathsf{P} \circ (\mathrm{id} \otimes \mathrm{d})$, after extending the exterior derivative from $\mathcal{A}(S^4_\theta)$ to $\mathcal{A}(\widetilde{\mathcal{M}}^{SD}_{\theta;k}) \otimes \mathcal{A}(S^4_\theta)$ by $\mathrm{id} \otimes \mathrm{d}$. Moreover, we need also to extend the Hodge $*$-operator as $\mathrm{id} \otimes *_\theta$.

PROPOSITION 4.11. *The curvature* $F = \mathsf{P}((\mathrm{id} \otimes \mathrm{d})\mathsf{P})^2$ *of the Grassmann connection* $\nabla = \mathsf{P} \circ (\mathrm{id} \otimes \mathrm{d})$ *is anti-self-dual, that is to say* $(\mathrm{id} \otimes *_\theta)F = -F$.

PROOF. When $\theta = 0$ this construction is the usual ADHM construction and it is known [**2**] (*cf.* also [**19**]) that it produces connections whose curvature is an anti-self-dual two-form:

$$(\mathrm{id} \otimes *_\theta)\, \mathsf{P}((\mathrm{id} \otimes \mathrm{d})\mathsf{P})^2 = -\mathsf{P}((\mathrm{id} \otimes \mathrm{d})\mathsf{P})^2.$$

As observed in Sect. 2.1, the Hodge $*$-operator is defined by the same formula as it is classically and, as vector spaces, the self-dual and anti-self-dual two-forms $\Omega^2_\pm(S^4_\theta)$ are the same as their undeformed counterparts $\Omega^2_\pm(S^4)$. This identification survives the tensoring by $\mathcal{A}(\widetilde{\mathcal{M}}^{SD}_{\theta;k})$ which yields $\mathcal{A}(\widetilde{\mathcal{M}}^{SD}_{\theta;k}) \otimes \Omega^2_\pm(S^4_\theta)$ to be isomorphic, as vector spaces, to $\mathcal{A}(\widetilde{\mathcal{M}}^{SD}_{k}) \otimes \Omega^2_\pm(S^4)$. Thus the anti-self-duality holds also when $\theta \neq 0$. $\square$

REMARK 4.12. One may alternatively verify the anti-self-duality via a complex structure. Indeed, there is an (almost) complex structure $\gamma : \Omega^1(\mathbb{CP}^3_\theta) \to \Omega^1(\mathbb{CP}^3_\theta)$ given by $\gamma(\mathrm{d}z_l) := (\mathrm{d} \circ J)(z_l)$, $l = 1, \ldots, 4$, the operator J being the one defined in (15), for which we declare the forms $\mathrm{d}z_l$ to be holomorphic and the forms $\mathrm{d}z_l^*$ to be anti-holomorphic. For instance, on generators of $\mathcal{A}(\mathbb{CP}^3_\theta)$ we have

$$\mathrm{d}(z_j z_l^*) = \eta_{lj} z_l^* \mathrm{d}z_j + z_j \mathrm{d}z_l^*$$

from the Leibniz rule and the relations (3), and we write $\mathrm{d} = \partial + \bar{\partial}$ with respect to this decomposition into holomorphic and anti-holomorphic forms. Since as vector spaces the various graded components of the differential algebra $\Omega(\mathbb{CP}^3_\theta)$ are undeformed, these operators ∂, $\bar{\partial}$ extend to a full Dolbeault complex with $\partial^2 = \bar{\partial}^2 = \bar{\partial}\partial + \partial\bar{\partial} = 0$. The algebra inclusion $\mathcal{A}(S^4_\theta) \hookrightarrow \mathcal{A}(\mathbb{CP}^3_\theta)$ extends to an inclusion of differential graded algebras $\Omega(S^4_\theta) \hookrightarrow \Omega(\mathbb{CP}^3_\theta)$ and the Hodge operator $*_\theta$ is, as in the classical case, defined in such a way that a two-form $\omega \in \Omega^2(S^4_\theta)$ is anti-self-dual if and only if its image in $\Omega^2(\mathbb{CP}^3_\theta)$ is of type $(1, 1)$. Thus, to check that the curvature $\mathsf{P}((\mathrm{id} \otimes \mathrm{d})\mathsf{P})^2$ is anti-self-dual, we use this inclusion of forms (i.e. we express everything in terms of $\mathrm{d}z_j$, $\mathrm{d}z_j^*$) and check that each component $F_{ad} = \mathsf{P}_{ab}((\mathrm{id} \otimes \mathrm{d})\mathsf{P}_{bc}) \wedge (\mathrm{id} \otimes \mathrm{d})\mathsf{P}_{cd})$ of the curvature is a sum of terms of type $(1, 1)$. This approach to noncommutative twistor theory, including a more explicit description of the noncommutative Penrose-Ward Transform, will be discussed in more detail elsewhere.

We next turn to the computation of the topological charge of the family of bundles $\mathcal{E} := \mathsf{P}\mathcal{A}(S^4_\theta)^{2k+2}$ given above. To this end we observe that the matrix σ_z has k linearly independent columns (since if not, it would not be injective) and that the columns of $\sigma_{J(z)}$ are obtained from those of σ_z by applying the map J. Clearly we are free to rearrange the columns of the matrix V (since this will not alter the class of the projection P), whence we may as well arrange them as

$$\mathrm{V} = \begin{pmatrix} {\sigma_z}^{(1)} & J({\sigma_z}^{(1)}) & {\sigma_z}^{(2)} & J({\sigma_z}^{(2)}) & \cdots & {\sigma_z}^{(k)} & J({\sigma_z}^{(k)}) \end{pmatrix},$$

where ${\sigma_z}^{(l)}$ denotes the l-th column of σ_z and $J({\sigma_z}^{(l)})$ denotes the l-th column of $\sigma_{J(z)}$. For fixed l, we denote the entries of the column ${\sigma_z}^{(l)}$ (together with their conjugates) by

$${w_\mu}^{(l)} := \sum\nolimits_\alpha M^\alpha_{\mu l} \otimes z_\alpha, \quad ({w_\mu}^{(l)})^* := \sum\nolimits_\alpha M^{\alpha\,*}_{\mu l} \otimes z^*_\alpha, \qquad \mu = 1, \ldots, 2k+2.$$

The entries of the column $J({\sigma_z}^{(l)})$ are obtained from those of ${\sigma_z}^{(l)}$ by applying the map J, and one clearly has $J(({w_\mu}^{(l)})^*) = (J({w_\mu}^{(l)}))^*$. In the classical limit $\theta = 0$, we could evaluate the parameters M^α_{ab} as fixed numerical values: this would identify the columns $\sigma_z^{(l)}$ and $J(\sigma_z^{(l)})$ as spanning a quaternionic line in $\mathbb{H}^{k+1}$, where the latter is defined by the $2k + 2$ complex coordinates $w_\mu^{(l)}$ and their conjugates, equipped with an anti-involution J. In the noncommutative case, although we lack the evaluation of the parameters M^α_{ab}, we continue to interpret the columns $\sigma_z^{(l)}$ and $J(\sigma_z^{(l)})$ as spanning a 'one-dimensional' quaternionic line.

As already observed in Rem. 4.6, the columns of σ_z are orthogonal, as are the columns of $\sigma_{J(z)}$; whence the rank $2k$ projection Q in (45) decomposes as a sum of projections

$$\mathsf{Q} = \mathsf{Q}_1 + \cdots + \mathsf{Q}_k,$$

where $\mathsf{Q}_l := \widetilde{\Psi}_i \widetilde{\Psi}_l{}^*$ is the rank two projection defined by the $(2k+2)\times 2$ matrix $\widetilde{\Psi}_l$ comprised of the columns $\sigma_z{}^{(l)}$ and $J(\sigma_z{}^{(l)})$, appropriately normalised by ρ^{-1}. Explicitly, this matrix is

$$\widetilde{\Psi}_l := \big(\textstyle\sum_{r,\alpha}(M^\alpha_{\mu r}\otimes z_\alpha)(\rho^{-1})_{rl} \quad \sum_{s,\beta}(M^\beta_{\mu s}\otimes J(z_\beta))(\rho^{-1})_{sl}\big)_{\mu=1,\dots,2k+2},$$

and a direct check yields $\widetilde{\Psi}_l{}^*\widetilde{\Psi}_l = \mathbb{I}_2$ so that Q_l is indeed a projection for each $l = 1,\dots,k$. Hence the matrix V in (44) has $2k$ columns which we interpret as spanning k quaternionic lines, with the same being true of the normalised matrix $\mathrm{V}\rho^{-1}$. The computation of the topological charge of the projection Q therefore boils down to the computation of the charge of each of the projections Q_l, for $l = 1,\dots,k$.

LEMMA 4.13. *For each $l = 1,\dots,k$ the projection Q_l is Murray-von Neumann equivalent to the projection $1\otimes\mathsf{q}$ in the algebra $\mathrm{M}_{2k+2}(\mathcal{A}(\widetilde{\mathcal{M}}^{SD}_{\theta;k})\otimes\mathcal{A}(S^4_\theta))$, where q is the basic projection defined in equation (27).*

PROOF. From equations (26) and (27) we know that $\mathsf{q} = \Psi\Psi^*$. Then, for each $l = 1,\dots,k$ define a partial isometry V_l in $\mathrm{M}_{2k+2,4}\left(\mathcal{A}(\widetilde{\mathcal{M}}^{SD}_{\theta;k})\otimes\mathcal{A}(S^4_\theta)\right)$ by

$$V_l := \widetilde{\Psi}_l\,(1\otimes\Psi^*), \qquad V_l^* := (1\otimes\Psi)\,\widetilde{\Psi}_l^*.$$

Straightforward computations show that $V_lV_l{}^* = \mathsf{Q}_l$ and $V_l{}^*V_l = 1\otimes\mathsf{q}$. □

We invoke the strategy of [**16**] to compute the topological charge of the family of bundles defined by each Q_l. Indeed, the charge of the projection q was shown in [**14**] be equal to 1, given as a pairing between the second Chern class $\mathrm{ch}_2(\mathsf{q})$, which lives in the cyclic homology group $\mathrm{HC}_4(\mathcal{A}(S^4_\theta))$, with the fundamental class of S^4_θ, which lives in the cyclic cohomology $\mathrm{HC}^4(\mathcal{A}(S^4_\theta))$. Although the class $\mathrm{ch}_2(\mathsf{Q}_l)$, being an element in $\mathrm{HC}_4(\mathcal{A}(\widetilde{\mathcal{M}}^{SD}_{\theta;k})\otimes\mathcal{A}(S^4_\theta))$, may not *a priori* be paired with the fundamental class of S^4_θ, Kasparov's KK-theory is used to show that in fact there is a well-defined pairing between the K-theory $\mathrm{K}_0\left(\mathcal{A}(\widetilde{\mathcal{M}}^{SD}_{\theta;k})\otimes\mathcal{A}(S^4_\theta)\right)$ and the K-homology $\mathrm{K}^0(\mathcal{A}(S^4_\theta))$. Since by the previous lemma the projections $1\otimes\mathsf{q}$ and Q_l define the same class in the K-theory of $\mathcal{A}(\widetilde{\mathcal{M}}^{SD}_{\theta;k})\otimes\mathcal{A}(S^4_\theta)$, it follows as in [**16**] that the topological charge of each projection Q_l is equal to 1.

PROPOSITION 4.14. *The family of bundles $\mathcal{E} = \mathsf{P}\mathcal{A}(S^4_\theta)^{2k+2}$ has topological charge equal to $-k$.*

PROOF. By the argument given above, the projections Q_l have topological charge equal to 1 for each $l = 1,\dots,k$. The projection $\mathsf{Q} = \mathsf{Q}_1+\cdots+\mathsf{Q}_k$ therefore has charge k, whence P must have charge $-k$. □

We finish this section by remarking that the construction given above in the section has an interpretation in terms of 'universal connections', as described in [**3**]. As already said, the classical quaternion vector space $\mathbb{H}^{k+1}$ may be identified with the complex vector space $\mathbb{C}^{2k+2}$ equipped with the quaternionic structure J. Points of the Grassmannian manifold $\mathrm{Gr}_k(\mathbb{H}^{k+1})$ of quaternionic k-dimensional subspaces of $\mathbb{H}^{k+1}$ may thus be identified with $2k$-dimensional subspaces of $\mathbb{C}^{2k+2}$ which are invariant under the involution J. Following the general strategy of [**5**] for the coordinatisation of Grassmannians, the algebra of functions on $\mathrm{Gr}_k(\mathbb{H}^{k+1})$

is given by functions taking values in the set of rank $2k$ projectors $P = (P^\mu{}_\nu)$ on $\mathbb{C}^{2k+2}$ which are J-invariant, viz.

$$\mathcal{A}(\mathrm{Gr}_k(\mathbb{H}^{k+1})) :=$$
$$\mathbb{C}\left[P^\mu{}_\nu \mid \sum\nolimits_\lambda P^\mu{}_\lambda P^\lambda{}_\nu = P^\mu{}_\nu,\ (P^\mu{}_\nu)^* = P^\nu{}_\mu,\ \sum\nolimits_\mu P^\mu{}_\mu = 2k,\ J(P^\mu{}_\nu) = P^\mu{}_\nu\right],$$

where $\mu, \nu = 1, \ldots, 2k+2$. In the classical case, when $\theta = 0$, the projection Q in (45) realises $\mathcal{A}(S^4_{\theta=0})$ as a subalgebra of $\mathcal{A}(\mathrm{Gr}_k(\mathbb{H}^{k+1}))$, whence this construction should be viewed as the dual of an embedding $S^4 \hookrightarrow \mathrm{Gr}_k(\mathbb{H}^{k+1})$, as given in [**3**]. We expect that, in the deformed case, the projection Q views $\mathcal{A}(S^4_\theta)$ as a subalgebra of a suitably-deformed version of $\mathcal{A}(\mathrm{Gr}_k(\mathbb{H}^{k+1}))$. For fixed k, the set of monads is bound to parameterise the set of such 'algebra embeddings'.

4.4. ADHM construction of charge one instantons. As a way of illustration we briefly verify that the above ADHM construction of noncommutative families of instantons gives back the family constructed in [**16**] when performed for the charge one case.

The starting point is the basic instanton on S^4_θ described in Sect. 2.2 and which arises via a monad construction as follows. The monad we consider is the sequence

$$\mathcal{A}(\mathbb{C}^4_\theta)(-1) \xrightarrow{\sigma_z} \mathbb{C}^4 \otimes \mathcal{A}(\mathbb{C}^4_\theta) \xrightarrow{\tau_z} \mathcal{A}(\mathbb{C}^4_\theta)(1), \tag{46}$$

where the arrows are the maps

$$\sigma_z = \begin{pmatrix} z_1 & z_2 & z_3 & z_4 \end{pmatrix}^{\mathrm{t}}, \qquad \tau_z = \sigma^\star_{J(z)} = \begin{pmatrix} -z_2 & z_1 & -z_4 & z_3 \end{pmatrix}.$$

Since $\tau_z \sigma_z = \sigma^\star_{J(z)}\sigma_z = 0$, it is clear that this is a monad with $k = 1$; by construction it is self-dual. In the present case $\rho^2 = \sigma^\star_z \sigma_z = \sum_j z^*_j z_j = r^2$, which we already assumed was invertible (corresponding to the deletion of the origin in $\mathbb{C}^4_\theta$). One computes that

$$\mathrm{VV}^* = \tfrac{1}{2} r^{-2} \begin{pmatrix} r^2 + x & 0 & \alpha & \beta \\ 0 & r^2 + x & -\mu\beta^* & \bar{\mu}\alpha^* \\ \alpha^* & -\bar{\mu}\beta & r^2 - x & 0 \\ \beta^* & \mu\alpha & 0 & r^2 - x \end{pmatrix}$$

which is just the projector q of equation (27). This is the 'tautological' monad construction given in [**5**]. The anti-self-dual version is the projector $\mathsf{P} = 1 - \mathrm{VV}^*$, in agreement with the ADHM construction above.

The monad (46) may be rewritten in the form

$$\sigma_z = (1,0,0,0)^{\mathrm{t}} \otimes z_1 \ +\ (0,1,0,0)^{\mathrm{t}} \otimes z_2 \ +\ (0,0,1,0)^{\mathrm{t}} \otimes z_3 \ +\ (0,0,0,1)^{\mathrm{t}} \otimes z_4, \tag{47}$$

with τ_z defined as its dual. With the strategy of [**16**] one generates new instantons by coacting on the generators $z_1, \ldots, z_4$ with the quantum conformal group $\mathcal{A}(\mathrm{SL}_\theta(2, \mathbb{H}))$. Using the formula (18) for the coaction, the monad map (47) transforms into

$$\sigma_{\Delta_L(z)} = \begin{pmatrix} a_1 \\ a_2 \\ c_1 \\ c_2 \end{pmatrix} \otimes z_1 \ +\ \begin{pmatrix} -a_2^* \\ a_1^* \\ -c_2^* \\ c_1^* \end{pmatrix} \otimes z_2 \ +\ \begin{pmatrix} b_1 \\ b_2 \\ d_1 \\ d_2 \end{pmatrix} \otimes z_3 \ +\ \begin{pmatrix} -b_2^* \\ b_1^* \\ -d_2^* \\ d_1^* \end{pmatrix} \otimes z_4, \tag{48}$$

and these four column vectors are the columns of the matrix (17) which defines the algebra $\mathcal{A}(\mathrm{SL}_\theta(2,\mathbb{H}))$. If we write

$$\widehat{M}^1 = \begin{pmatrix} a_1 & a_2 & c_1 & c_2 \end{pmatrix}^{\mathrm{t}}, \qquad \widehat{M}^2 = \begin{pmatrix} -a_2^* & a_1^* & -c_2^* & c_1^* \end{pmatrix}^{\mathrm{t}},$$
$$\widehat{M}^3 = \begin{pmatrix} b_1 & b_2 & d_1 & d_2 \end{pmatrix}^{\mathrm{t}}, \qquad \widehat{M}^4 = \begin{pmatrix} -b_2^* & b_1^* & -d_2^* & d_1^* \end{pmatrix}^{\mathrm{t}},$$

then we have the algebra relations $\widehat{M}_j^\alpha \widehat{M}_l^\beta = \eta_{jl}\eta_{\beta\alpha}\widehat{M}_l^\beta \widehat{M}_j^\alpha$ coming from the relations (20) for $\mathcal{A}(\mathrm{SL}_\theta(2,\mathbb{H}))$. We thus think of the algebra generated by the $\widehat{M}_j^\alpha$ as parameterising the set of charge one instantons, since the map $\sigma_{\Delta_L(z)}$ may be used to construct the family (28) of projections with topological charge equal to 1 and hence a family of Grassmann connections with anti-self-dual curvature, just as in [**16**].

In contrast, the ADHM construction of Sect. 4.3 for the case $k = 1$ says that the charge one monads are parameterised by the algebra $\mathcal{A}(\widetilde{\mathcal{M}}_{\theta;k})$ generated by the matrix elements M_j^α, with $j, \alpha = 1, \ldots, 4$, subject in particular to the relations $M_j^\alpha M_l^\beta = \eta_{\beta\alpha} M_l^\beta M_j^\alpha$.

We see that these two approaches seem to give different parameterisations of the set of monads for the case $k = 1$, and hence of the set of charge one instantons. The discrepancy has its root in the fact that the ADHM construction requires generators lying in the same row of the matrix (A_{ij}) to commute, whereas the 'coaction approach' given above says that such generators do not commute.

However, the discrepancy fades away when we pass to the 'true' parameter space for the families. On the one hand, as observed in [**16**], the coaction (24) of the quantum subgroup $\mathcal{A}(\mathrm{Sp}_\theta(2))$ of $\mathcal{A}(\mathrm{SL}_\theta(2,\mathbb{H}))$ leaves the basic one-form (11) invariant. We think of the latter coaction as generating gauge-equivalent instantons, so that the 'true' parameter space for this family is rather the subalgebra of $\mathcal{A}(\mathrm{SL}_\theta(2,\mathbb{H}))$ of coinvariants under the coaction of $\mathcal{A}(\mathrm{Sp}_\theta(2))$. The generators of this algebra are computed to be

$$\widehat{m}_{\alpha\beta} := \sum\nolimits_l \widehat{M}_l^{\alpha *} \widehat{M}_l^\beta, \qquad \alpha, \beta = 1\ldots, 4,$$

whose relations are easily found to be

$$\widehat{m}_{\alpha\beta}\widehat{m}_{\mu\nu} = \eta_{\beta\mu}\eta_{\nu\beta}\eta_{\mu\alpha}\eta_{\alpha\nu}\widehat{m}_{\mu\nu}\widehat{m}_{\alpha\beta}, \tag{49}$$

and which certainly do not depend on the rows of the matrix (A_{ij}). On the other hand, gauge equivalence for the ADHM family parameterised by the M_i^α is generated by the action of the classical group $\mathrm{Sp}(2)$ (we borrow this result from Prop. 5.2 in the next section), and here the invariant subalgebra is generated by elements of the form

$$m_{\alpha\beta} := \sum\nolimits_l M_l^{\alpha *} M_l^\beta, \qquad \alpha, \beta = 1\ldots, 4.$$

The relations in this algebra are just as in equation (49), so that these two families of charge one instantons are just the same.

5. Gauge Equivalence of Noncommutative Instantons

Classically, a way to think of a gauge transformation of a vector bundle E over S^4 is as a unitary change of basis in each fibre E_x in a way which depends smoothly on $x \in S^4$. Two connections on E are said to be gauge equivalent if they are related by a gauge transformation in this way. Now, rather than being interested in the set of all instantons on S^4, one is interested in the collection of gauge equivalence classes, that is to say classes of instantons modulo gauge transformations.

It is therefore necessary to have an analogue of the notion of gauge equivalence also for the noncommutative families of instantons constructed previously. In fact, noncommutative geometry is a very natural setting for the study of gauge transformations, as we shall see in this section; we refer in particular to [**7, 8**] (*cf.* also [**13**]).

5.1. Gauge equivalence for families of instantons. Recall that a first order differential calculus on a unital $*$-algebra A is a pair $(\Omega^1 A, \mathrm{d}_A)$, where $\Omega^1 A$ is an A-A-bimodule giving the space of one-forms and $\mathrm{d}_A : A \to \Omega^1 A$ is a linear map satisfying the Leibniz rule,

$$\mathrm{d}_A(xy) = x(\mathrm{d}_A y) + (\mathrm{d}_A x)y \qquad \text{for all} \quad x, y \in A.$$

One also assumes that the map $x \otimes y \to x(\mathrm{d}_A y)$ is surjective. One names $\Omega^1 A$ a $*$-calculus if for $x_j, y_j \in A$ one has that $\sum_j x_j \mathrm{d} y_j = 0$ implies $\sum_j \mathrm{d}(y_j^*) x_j^* = 0$: it follows from this condition that there is [**26**] a unique $*$-structure on $\Omega^1 A$ such that $(\mathrm{d}_A a)^* = \mathrm{d}_A(a^*)$ for all $a \in A$. The differential calculi on $\mathcal{A}(S^4_\theta)$ and $\mathcal{A}(S^7_\theta)$ in Sect. 2.1 are examples of first order differential $*$-calculi on noncommutative spaces.

Let us fix a choice of $*$-calculus on A. Then let $\mathcal{E}$ be a finitely generated projective right A-module endowed with an A-valued Hermitian structure denoted by $\langle\cdot|\cdot\rangle$. A connection on $\mathcal{E}$ is a linear map $\nabla : \mathcal{E} \to \mathcal{E} \otimes_A \Omega^1 A$ satisfying the Leibniz rule

$$\nabla(\xi x) = (\nabla \xi)x + \xi \otimes \mathrm{d}_A x \qquad \text{for all } \xi \in \mathcal{E},\ x \in A.$$

The connection ∇ is said to be compatible with the Hermitian structure on $\mathcal{E}$ if it obeys

$$\langle \nabla \xi | \eta \rangle + \langle \xi | \nabla \eta \rangle = \mathrm{d}_A \langle \xi | \eta \rangle \qquad \text{for all } \xi \in \mathcal{E},\ x \in A.$$

On $\mathcal{E}$ there is at least one compatible connection, the so-called Grassmann connection ∇_0. If $P \in \mathrm{End}_A(\mathcal{E})$, $P^2 = P = P^*$, is the projection which defines $\mathcal{E}$ as a direct summand of a free module, that is, $\mathcal{E} = P(\mathbb{C}^N \otimes A)$, then $\nabla_0 = P \circ \mathrm{d}$. Any other connection on $\mathcal{E}$ is of the form $\nabla = \nabla_0 + \omega$, where ω is an element of $\mathrm{Hom}_A(\mathcal{E}, \mathcal{E} \otimes_A \Omega^1 A)$.

The gauge group of $\mathcal{E}$ is defined to be

$$\mathcal{U}(\mathcal{E}) := \{ U \in \mathrm{End}_A(\mathcal{E}) \mid \langle U\xi | U\eta \rangle = \langle \xi | \eta \rangle \text{ for all } \xi, \eta \in \mathcal{E} \}.$$

If ∇ is a compatible connection on $\mathcal{E}$, each element U of the gauge group $\mathcal{U}(\mathcal{E})$ induces a 'new' connection by the action

$$\nabla^U := U \nabla U^*.$$

Of course, ∇^U is not really a different connection, it simply expresses ∇ in terms of the transformed bundle $U\mathcal{E}$, hence one says that a pair of connections ∇_1, ∇_2 on $\mathcal{E}$ are *gauge equivalent* if they are related by such a gauge transformation U. In terms of the decomposition $\nabla = \nabla_0 + \omega$, one finds that $\nabla^U = \nabla_0 + \omega^U$, where

$$\omega^U := U(\nabla_0 U^*) + U \omega U^*.$$

A choice of gauge would be a choice of partial isometry $\Psi : \mathcal{E} \to \mathcal{A}^N$ such that $\Psi^* \Psi = \mathrm{Id}_{\mathcal{E}}$ and $\Psi \Psi^* = P$. Any other gauge is then given by an element U of the gauge group of $\mathcal{E}$: the partial isometry Ψ gets replaced by $U\Psi$, for which we indeed have

$$(U\Psi)^*(U\Psi) = \Psi^* \Psi = \mathrm{Id}_{\mathcal{E}}.$$

and the projection P gets transformed to

$$(U\Psi)(U\Psi)^* = U(\Psi\Psi^*)U^* = UPU^*,$$

an operation that does not change the equivalence class of P. In the fixed gauge the Grassmann connection $\nabla_0 = P\circ \mathrm{d}$ naturally acts on 'equivariant maps' $\varphi = \Psi F$ where $F \in \mathcal{A}^N$. The result is an 'equivariant one-form',

$$\nabla_0(\Psi F) = (\Psi\Psi^*)\mathrm{d}(\Psi F) = \Psi\Big(\mathrm{d}F + \Psi^*\mathrm{d}(\Psi)F\Big),$$

and identifies the *gauge potential* to be given by

$$A = \tfrac{1}{2}\,(\Psi^*(\mathrm{d}\Psi) - (\,\mathrm{d}\Psi^*)\Psi).$$

Under the transformation $\Psi \mapsto U\Psi$, the gauge potential transforms as expected:

$$\Psi^*\mathrm{d}\Psi \mapsto \Psi^*(\mathrm{d}\Psi) + \Psi^*U^*(\mathrm{d}U)\Psi.$$

We now turn back to the construction of instantons. Gauge equivalence being defined as above by unitary module endomorphisms means that we are free to act on the right $\mathcal{A}(\mathbb{C}^4_\theta)$-module $\mathcal{K} = K\otimes\mathcal{A}(\mathbb{C}^4_\theta)$ by a unitary element of the matrix algebra $\mathrm{M}_{2k+2}(\mathbb{C})\otimes\mathcal{A}(\mathbb{C}^4_\theta)$. In order to preserve the instanton construction, we must do so in a way preserving the bilinear form $(\,\cdot\,,\,\cdot\,)$ of equation (43) which comes from the identification of K with its dual K^*. Hence the map σ_z in (31) (or in (33)) is defined up to a transformation $A \in \mathrm{End}_{\mathcal{A}(\mathbb{C}^4_\theta)}(\mathcal{K})$, which is unitary and is required to commute with the quaternion structure J. Similarly, we are free to change basis in the modules $\mathcal{H} = H\otimes\mathcal{A}(\mathbb{C}^4_\theta)$ and $\mathcal{L} = L\otimes\mathcal{A}(\mathbb{C}^4_\theta)$, provided we preserve the fact that we identify $J(\mathcal{H})^\star$ and $\mathcal{L}$. This means that the map τ_z of (35) is defined up to an invertible transformation $B \in \mathrm{End}_{\mathcal{A}(\mathbb{C}^4_\theta)}(\mathcal{H})$.

All this is saying is that the monad maps σ_z and τ_z were expressed as matrices with respect to a choice of basis for each of the vector spaces H, K and L; it is natural to question the extent to which the resulting Grassmann connection ∇ depends on the choice of these bases. We denote by $\mathrm{GL}(\mathcal{H})$ the set of automorphisms of $\mathcal{H}$ and by $\mathrm{Sp}(\mathcal{K})$ the set of all unitary endomorphisms of $\mathcal{K}$ respecting the quaternion structure:

$$\mathrm{Sp}(\mathcal{K}) := \{A \in \mathrm{End}_{\mathcal{A}(\mathbb{C}^4_\theta)}(\mathcal{K}) \mid \langle A\xi|A\xi\rangle = \langle\xi|\xi\rangle,\ J(A\xi) = AJ(\xi) \text{ for all } \xi\in\mathcal{K}\}.$$

Given $A \in \mathrm{Sp}(\mathcal{K})$ and $B \in \mathrm{GL}(\mathcal{H})$, the gauge freedom is to map $\sigma_z \mapsto A\sigma_z B$.

PROPOSITION 5.1. *For all $B \in \mathrm{GL}(\mathcal{H})$, under the transformation $\sigma_z \mapsto \sigma_z B$ the projection P of Prop. 4.10 is left invariant.*

PROOF. One first checks that $\rho^2 \mapsto (\sigma_z B)^\star(\sigma_z B) = B^\star\rho^2 B$ under this transformation, so that

$$\mathsf{Q}_z \mapsto \sigma_z B(B^\star\rho^2 B)^{-1}B^\star\sigma_z^\star = \sigma_z B(B^{-1}\rho^{-2}(B^\star)^{-1})B^\star\sigma_z^\star = \mathsf{Q}_z,$$

whence the projection P is unchanged. □

PROPOSITION 5.2. *For all $A \in \mathrm{Sp}(\mathcal{K})$, under the transformation $\sigma_z \mapsto A\sigma_z$ the projection P of Prop. 4.10 transforms as $\mathsf{P} \mapsto A\mathsf{P}A^\star$.*

PROOF. Replacing σ_z by $A\sigma_z$ leaves ρ^2 invariant (since A is unitary) and so has the effect that

$$\mathsf{Q}_z \mapsto A\sigma_z\rho^{-2}\sigma_z^\star A^\star = A\mathsf{Q}_z A^\star,$$

whence it follows that P is mapped to $A\mathsf{P}A^\star$. □

These results give the general gauge freedom on monads, although from the point of view of computing the number of constraints on the algebra generators M^{α}_{ab} we need only consider the effect of these transformations on the vector spaces H, K and L, i.e. it is enough to consider the groups of 'constant' automorphisms. This means the group $\mathrm{Sp}(K) = \mathrm{Sp}(k+1) \subset \mathrm{Sp}(\mathcal{K})$ and the group $\mathrm{GL}(k, \mathbb{R}) \subset \mathrm{GL}(\mathcal{H})$ (the latter because we must preserve the identification of $J(\mathcal{H})^{\star}$ with $\mathcal{L}$, and complex linear transformations of H would interfere with the tensor product in $J(\mathcal{H})^{\star}$). In fact, it is known in the classical case that these constant transformations are sufficient to generate all gauge symmetries of the instanton bundles produced by the ADHM construction.

We conclude that in the noncommutative case as well the gauge equivalence imposes an additional $(k+1)(2(k+1)+1)$ constraints due to $\mathrm{Sp}(k+1)$ and a further k^2 constraints due to $\mathrm{GL}(k, \mathbb{R})$. From Rem. 4.7, the total number of generators minus the total number of constraints is thus computed to be

$$(8k^2 + 8k) - 5k(k-1) - (3k^2 + 5k + 3) = 8k - 3,$$

just as for the classical case, a result which is somehow reassuring.

5.2. Morita equivalent geometries and gauge theory. It is a known idea that Morita equivalent algebras describe the same topological space. The simplest case is that of a one-point space $X = \{*\}$: the matrix algebras $\mathrm{M}_n(\mathbb{C})$ for any positive integer n all have the same one-point spectrum. More generally, if X is a compact Hausdorff space, the algebras $C(X) \otimes \mathrm{M}_n(\mathbb{C})$ are all Morita equivalent and all have the same spectrum X.

With this in mind, gauge theory arises naturally out of the consideration of how to transfer differential structures between Morita equivalent algebras. If one takes such structures to be defined by a Dirac operator and associated spectral triple, then the method for doing this is discussed in [**7, 8**]. Here we discuss a more general framework, where algebras may be equipped with differential calculi not necessarily coming from a spectral triple.

Let A be a unital $*$-algebra and suppose that the $*$-algebra B is Morita equivalent to A *via* the B-A-bimodule $\mathcal{E}$, that is to say $B \simeq \mathrm{End}_A(\mathcal{E})$. In addition, on $\mathcal{E}$ there are compatible A-valued and B-valued Hermitian structures [1]. Then a choice of a connection ∇ on $\mathcal{E}$, viewed as a right A-module, yields a differential calculus on B. First of all, the operator on B given by

$$\mathrm{d}_B^{\nabla}(x) := [\nabla, x], \qquad x \in B,$$

is easily seen to be a derivation: $\mathrm{d}_B^{\nabla}(xy) = x(\mathrm{d}_B^{\nabla} y) + (\mathrm{d}_B^{\nabla} x)y$, for $x, y \in B$. The B-B-bimodule $\Omega^1 B$ of one-forms is then defined by

$$\Omega^1 B := B \left(\mathrm{d}_B^{\nabla}(B) \right) B.$$

For this to define a $*$-calculus we need that the connection ∇ be compatible with the A-valued Hermitian structure on $\mathcal{E}$ in the sense that

$$\langle \nabla \xi | \eta \rangle + \langle \xi | \nabla \eta \rangle = \mathrm{d}_A \langle \xi | \eta \rangle$$

[1] We shall also require the Hermitian structures to be self-dual, *i.e.* every right A-module homomorphism $\varphi : \mathcal{E} \to A$ is represented by an element of $\eta \in \mathcal{E}$ by the assignment $\varphi(\cdot) = \langle \eta | \cdot \rangle$. A similar property holds for the second Hermitian structure as well.

for all $a \in A$ and $\xi, \eta \in \mathcal{E}$. If this compatibility condition is satisfied, the assumption $\sum_j x_j \mathrm{d}_B^\nabla y_j = 0$ translates into $\sum_j (x_j \nabla y_j)\xi = \sum_j x_j y_j (\nabla \xi)$ for all $x_j, y_j \in B$ and all $\xi \in \mathcal{E}$. This implies, for all $\xi, \eta \in \mathcal{E}$ and all $x_j, y_j \in B$, that

$$\begin{aligned}\sum\nolimits_j \langle \mathrm{d}_B^\nabla (y_j^*) x_j^* \xi | \eta \rangle &= \sum\nolimits_j \langle \nabla (x_j^* y_j^* \xi) - x_j^* \nabla (y_j^* \xi) | \eta \rangle \\ &= \sum\nolimits_j -\langle x_j^* y_j^* \xi | \nabla \eta \rangle + \mathrm{d}_A \langle x_j^* y_j^* \xi | \eta \rangle \\ &\qquad\qquad + \langle y_j^* \xi | \nabla (x_j \eta) \rangle - \mathrm{d}_A \langle y_j^* \xi | x_j \eta \rangle \\ &= \sum\nolimits_j -\langle x_j^* y_j^* \xi | \nabla \eta \rangle + \langle \xi | y_j x_j \nabla \eta \rangle,\end{aligned}$$

whence it follows that $\sum_j \mathrm{d}_B^\nabla (y_j^*) x_j^* = 0$ as it should for a $*$-calculus. We interpret the passage $\mathrm{d}_A \to \mathrm{d}_B^\nabla$ as an *inner fluctuation* of the geometry which results in a 'Morita equivalent' first order calculus $(\Omega^1 B, \mathrm{d}_B^\nabla)$, now for the algebra B.

A natural application is to think of the algebra A as being Morita equivalent to itself, so that $\mathcal{E} = A$ as a right A-module and $B = A$. In this case, any Hermitian connection on $\mathcal{E}$ is necessarily of the form

$$\nabla \xi = \mathrm{d}_A \xi + \omega \xi, \qquad \text{for} \quad \xi \in \mathcal{E}, \tag{50}$$

with $\omega = -\omega^* \in \Omega^1 A$ a skew-adjoint one-form. The corresponding differential on $B = A$ is computed to be

$$(\mathrm{d}_A^\nabla b)\xi = [\nabla, b]\xi = \nabla (b\xi) - b\nabla \xi = \mathrm{d}_A (b\xi) + \omega b \xi - b \mathrm{d}_A \xi - b\omega \xi = (\mathrm{d}_A b)\xi + [\omega, b]\xi,$$

using the Leibniz rule for d_A. The passage

$$\mathrm{d}_A \to \mathrm{d}_A^\nabla = \mathrm{d}_A + [\omega, \ \cdot\ ]$$

is once again interpreted as an inner fluctuation of the geometry, although when A is commutative there are no non-trivial inner fluctuations and thus no new degrees of freedom generated by the above self-Morita mechanism. However, in the noncommutative situation there is an interesting special case where ω is taken to be of the form $\omega = u^* \mathrm{d}_A u$, for u a unitary element of the algebra A. Such a fluctuation is unitarily equivalent to acting on A by the inner automorphism

$$\alpha_u : A \to A, \qquad \alpha_u(a) = uau^*,$$

since for all $a \in A$ we have that $\mathrm{d}_A^\nabla (a) = u^* \mathrm{d}_A (\alpha_u(a)) u$. It therefore follows that inner fluctuations defined by inner automorphisms generate gauge theory on A.

5.3. Gauge theory from quantum symmetries. We now consider a slightly different type of gauge equivalence for our instanton construction which is not present in the classical case and is a purely quantum (*i.e.* noncommutative) phenomenon.

We consider the case where A is a comodule $*$-algebra under a left coaction of a Hopf algebra H, so that A is isomorphic to its image $B = \Delta_L(A)$. To transfer a calculus on A to one on B, a possible strategy is as follows. We take the B-A-bimodule to be $\mathcal{E} := B = \Delta_L(A)$ with left B-action and right A-action defined by

$$b \triangleright \xi := b\xi, \qquad \xi \triangleleft a = \xi \Delta_L(a)$$

for $\xi \in \mathcal{E}$, $a \in A$, $b \in B$. We also assume that the calculus $\Omega^1 A$ is left H-covariant, so that Δ_L extends to a coaction on $\Omega^1 A$ as a bimodule map such that d_A is an intertwiner, whence the above bimodule structure on $\mathcal{E}$ extends to one-forms in the natural way. This also canonically equips B with a $*$-calculus $\Omega^1 B$, where the differential is $\mathrm{d}_B = \mathrm{id} \otimes \mathrm{d}_A$.

We choose an arbitrary Hermitian connection on the right A-module $\mathcal{E}$ for the calculus $(\Omega^1 A, \mathrm{d}_A)$, which is necessarily of the form

$$\nabla\xi = (\mathrm{id} \otimes \mathrm{d}_A)\xi + \tilde{\omega}\xi, \qquad \xi \in \mathcal{E}$$

with $\tilde{\omega} = \Delta_L(\omega)$ for some $\omega = -\omega^* \in \Omega^1 A$ a skew-adjoint one-form. The corresponding differential on B is again defined by

$$(\mathrm{d}_B^\nabla b)\xi = [\nabla, b\triangleright]\xi = \nabla(b \triangleright \xi) - b \triangleright \nabla\xi = \mathrm{d}_A(b \triangleright \xi) + \omega(b \triangleright \xi) - b \triangleright (\mathrm{d}_A\xi + \omega\xi)\,,$$

and works out to be

$$\mathrm{d}_B b = (\mathrm{id} \otimes \mathrm{d}_A)b + [\tilde{\omega}, b].$$

Note also that for all $b \in B$ we have $b = \Delta_L(a)$ for some $a \in A$ and so it follows that

$$\mathrm{d}_B b = \Delta_L(\mathrm{d}_A a) + \Delta_L([\omega, a]),$$

so that the coaction commutes with inner fluctuations. Moreover, in the case where A is noncommutative, there are non-trivial inner automorphisms of A and hence non-trivial gauge degrees of freedom which carry over from A to $\Delta_L(A)$.

In particular, we apply this to the case $A = \mathcal{A}(S^4_\theta)$, with $H = \mathcal{A}(\mathrm{SL}_\theta(2, \mathbb{H}))$ the quantum conformal group of S^4_θ. The above discussion means that the coaction of $\mathcal{A}(\mathrm{SL}_\theta(2, \mathbb{H}))$ on $\mathcal{A}(S^4_\theta)$ by conformal transformations in itself generates gauge freedom. The natural way to extend the exterior derivative d_A on $\mathcal{A}(S^4_\theta)$ to $\Delta_L(\mathcal{A}(S^4_\theta))$ is as $\mathrm{id} \otimes \mathrm{d}_A$: this corresponds to taking $\tilde{\omega} = 0$ and is the choice made in [**16**]. However, in general we have the freedom to make the transition

$$\mathrm{d}_A \to (\mathrm{id} \otimes \mathrm{d}_A) + [\tilde{\omega},\ \cdot\ ]$$

for some $\tilde{\omega} = \Delta_L(u^*\mathrm{d}_A u)$, where u is some unitary element of $\mathcal{A}(S^4_\theta)$. Since the group of inner automorphisms of A is trivial when A is commutative, this is a feature of gauge theory which is certainly not present in the classical case and is unique to the noncommutative paradigm. More on this will be reported elsewhere.

References

[1] Atiyah M.F., Ward R.S.: Instantons and Algebraic Geometry. Commun. Math. Phys. **55**, 117–124 (1977)

[2] Atiyah M.F., Drinfel'd V.G., Hitchin N.J., Manin Yu.I.: Construction of Instantons. Phys. Lett. **65A**, 185–187 (1978)

[3] Atiyah M.F.: *Geometry of Yang-Mills Fields*. Fermi Lectures, Scuola Normale Pisa, 1979

[4] Barth W.: Moduli of Vector Bundles on the Projective Plane. Invent. Math. **42**, 63–91 (1977)

[5] Brain S.J., Majid S.: Quantisation of Twistor Theory by Cocycle Twist. Commun. Math. Phys. **284**, 713–774 (2008)

[6] Brain S., Landi G.: Moduli Spaces of Noncommutative Instantons: Gauging Away Noncommutative Parameters. [`arXiv:0909.4402`]

[7] Connes A.: Gravity Coupled with Matter and the Foundation of Noncommutative Geometry. Commun. Math. Phys. **182**, 155–176 (1996)

[8] Connes A.: On the Spectral Characterization of Manifolds. [`arXiv:0810.2088`]

[9] Connes A., Landi G.: Noncommutative Manifolds, the Instanton Algebra and Isospectral Deformations. Commun. Math. Phys. **221**, 141–159 (2001)

[10] Cuntz J., Quillen D.: Algebra Extensions and Nonsingularity. J. Amer. Math. Soc. **8**, 251–289 (1995)

[11] Dubois-Violette M., Georgelin Y.: Gauge Theory in Terms of Projector-Valued Fields. Phys. Lett. **82B**, 251–254 (1979)

[12] Horrocks G.: Vector Bundles on the Punctured Spectrum of a Ring. Proc. London Math. Soc. **14**, 689–713 (1964)
[13] Landi G.: *An Introduction to Noncommutative Spaces and their Geometry.* Lecture Notes in Physics **51**. Springer–Verlag, Berlin, 1997
[14] Landi G., van Suijlekom W.D.: Principal Fibrations from Noncommutative Spheres. Commun. Math. Phys. **260**, 203–225 (2005)
[15] Landi G., van Suijlekom W.D.: Noncommutative Instantons from Twisted Conformal Symmetries. Commun. Math. Phys. **271**, 591–634 (2007)
[16] Landi G., Pagani C., Reina C., van Suijlekom W.D.: Noncommutative Families of Instantons. Int. Math. Res. IMRN **12**, Art. ID rnn038 (2008)
[17] Majid S., Markl M.: Glueing Operations for R-Matrices, Quantum Groups and Link-Invariants of Hecke Type. Math. Proc. Camb. Phil. Soc. **119**, 139–166 (1996)
[18] Majid S.: Quantum and Braided Diffeomorphism Groups. J. Geom. Phys. **28**, 94–128 (1998)
[19] Mason L.J., Woodhouse N.M.J.: *Integrability, Self-Duality and Twistor Theory.* Oxford University Press, 1996
[20] Okonek C., Schneider M., Spindler H.: *Vector Bundles on Complex Projective Spaces.* Birkhauser, Boston, 1980
[21] Penrose R., Rindler W.: *Spinors in Space-Time* vol. 2. Cambridge University Press, 1986
[22] Soltan P.M.: Quantum Families of Maps and Quantum Semigroups on Finite Quantum Spaces. J. Geom. Phys. **59**, 354–368 (2009)
[23] Wang S.: Quantum Symmetry Groups of Finite Spaces. Commun. Math. Phys. **195**, 195–211 (1998)
[24] Ward R.S.: On Self-Dual Gauge Fields. Phys. Lett. **61A**, 81–82 (1977)
[25] Woronowicz S.L.: Pseudogroups, Pseudospaces and Pontryagin Duality. In *Proceedings of the International Conference on Mathematical Physics, Lausanne 1979*, Ed. K. Osterwalder. Lecture Notes in Physics **116**, Springer, 407–412, 1980
[26] Woronowicz S.L.: Differential Calculus on Compact Matrix Pseudogroups (Quantum Groups). Commun. Math. Phys. **122**, 125–170 (1989)

SISSA, Via Beirut 2-4, 34151 Trieste, Italy
E-mail address: brain@sissa.it

Dipartimento di Matematica e Informatica, Università di Trieste, Via A. Valerio 12/1, 34127 Trieste, Italy, and INFN, Sezione di Trieste, Trieste, Italy
E-mail address: landi@univ.trieste.it

Clay Mathematics Proceedings
Volume 11, 2010

Cyclic Theory for Commutative Differential Graded Algebras and S–Cohomology

Dan Burghelea

Dedicated to A. Connes for his 60th birthday

Abstract. In this paper we consider three homotopy functors on the category of manifolds , hH^*, cH^*, sH^*, and parallel them with three other homotopy functors on the category of connected commutative differential graded algebras, HH^*, CH^*, SH^*. If P is a smooth 1-connected manifold and the algebra is the de-Rham algebra of P the two pairs of functors agree but in general they do not. The functors HH^* and CH^* can also be derived as Hochschild resp. cyclic homology of a commutative differential graded algebra, but this is not the way they are introduced here. The third one SH^*, although inspired from negative cyclic homology, cannot be identified with any sort of cyclic homology of any algebra. The functor sH^* might play some role in topology. Important tools in the construction of the functors $HH^*, CH^*$and SH^*, in addition to the linear algebra suggested by cyclic theory, are Sullivan's minimal model theorem and the "free loop" construction described in this paper.

Contents

1. Introduction

This paper deals with commutative differential graded algebras and the results are of significance in "commutative" geometry/topology. However they were

2010 *Mathematics Subject Classification.* Primary 16E45; Secondary 16E40, 19D55, 55P50, 55P62.

inspired largely by the linear algebra underlying Connes' cyclic theory. The topological results formulated here, Theorem 2 and Theorem 3, were first established as a consequence of the identification of the cohomology, resp. S^1-equivariant cohomology, of the free loop spaces of a 1-connected smooth manifold with the Hochschild, resp. cyclic, homology of its de Rham algebra, cf. [**16**], [**8**], [**4**]. In this paper this identification is circumvented. Still, the results illustrate the powerful influence of Connes' mathematics in areas outside non–commutative geometry.

In this paper, inspired by the relationship between Hochschild, cyclic and negative cyclic homology of a unital algebra, we consider two systems of graded vector space valued homotopy functors $hH^*, cH^*, sH^*$and HH^*, CH^*, SH^* and investigate their relationship. The first three functors are defined on the category of smooth manifolds and smooth maps via the free loop space P^{S^1} of a smooth manifold P, which is a smooth S^1-manifold of infinite dimension. The next three functors are defined on the category of connected commutative differential graded algebras via an algebraic analogue of the "free loop" construction and via Sullivan's minimal model theorem, Theorem 1. The relationship between them is suggested by the general diagram Fig 2 in section 2.

When applied to the de Rham algebra of a 1-connected smooth manifold the last three functors take the same values as the first three. This is not the case when the smooth manifold is not 1-connected; the exact relationship will be addressed in a future work.

The first three functors are based on a formalism (manipulation with differential forms) which can be considered for any smooth (finite or infinite dimensional) manifold M and any smooth vector field X on M. However, it seems to be of relevance when the vector field X comes from a smooth S^1-action on M. This is of mild interest if the manifold is of finite dimension but more interesting when the manifold is of infinite dimension. In particular, it is quite interesting when $M = P^{S^1}$, the free loop space of P, and the action is the canonical S^1-action on P^{S^1}. Manipulation with differential forms on P^{S^1} leads to the graded vector spaces $hH^*(P)$, $cH^*(P)$, $sH^*(P)$ with the first two being the cohomology, resp. the S^1-equivariant cohomology, of P^{S^1} but sH^* being a new homotopy functor, referred to here as s–cohomology.

This functor was first introduced in [**3**], [**4**] but so far not been seriously investigated. The functor sH^* relates, at least in the case of a 1-connected manifold P, the Waldhausen algebraic $K-$theory of P and the Atiyah–Hirzebruch complex $K-$theory (based on complex vector bundles) of P. It has a rather simple description in terms of an infinite sequence of smooth invariant differential forms on P^{S^1}.

The additional structures on P^{S^1}, the power maps ψ_k $k = 1, 2, \cdots$, and the involution $\tau = \psi_{-1}$, provide endomorphisms of $hH^*(P)$, $cH^*(P), sH^*(P)$ whose eigenvalues and eigenspaces are interesting issues. They are clarified only when P is 1-connected. This is done in view of the relationship with the functors HH^*, CH^*, SH^*.

It might be only a coincidence but still an appealing observation that the symmetric, resp. antisymmetric, part of $sH^*(P)$ with respect to the canonical involution τ calculates, for a 1-connected manifold P and in the stability range, the vector space $\mathrm{Hom}(\pi_*(\mathrm{H}/\mathrm{Diff}(P),\ \kappa)$, $\kappa = \mathbb{R}, \mathbb{C}$; the symmetric part when $\dim P$ is even the antisymmetric part when $\dim P$ is odd, cf. [**2**], [**3**]. Here $\mathrm{H}/\mathrm{Diff}(P)$ denotes the (homotopy) quotient of the space of homotopy equivalences of P by the group of diffeomorphisms with the $C^\infty-$topology.

The functors HH^*, CH^*, SH^* are the algebraic versions of hH^*, cH^*, sH^* and are defined on the category of (homologically connected) commutative differential graded algebras. Their definition uses the "free loop" construction, an algebraic analogue of the free loop space, described in this paper only for free connected commutative differential graded algebras $(\Lambda[V], d_V)$. A priori these functors are defined only for free connected commutative differential graded algebras. Since they are homotopy functors they extend to all connected commutative differential graded algebras via Sullivan's minimal model theorem, Theorem 1.

Using the definition presented here one can take full advantage of the simple form that the algebraic analogue of the power maps take on the free loop construction. As a consequence one obtains a simple description of the eigenvalues and eigenspaces of the endomorphisms induced from the algebraic power maps on HH^* and CH^* and one can implicitly understand their additional structure.

The extension of the results of Sullivan–Vigué, cf. [**20**], to incorporate S^1-actions and the power maps in the minimal model of P^{S^1} summarized in Section 7 leads finally to results about hH^*, cH^*, sH^* when P is 1-connected, Theorem 3.

In addition to the algebraic definition of HH^*, CH^*, SH^* this paper contains the proof of the homotopy invariance of sH^*.

2. Mixed complexes, a formalism inspired from Connes' cyclic theory

A mixed complex (C^*, δ^*, β_*) consists of a graded vector space C^* ($*$ a non negative integer) and linear maps, $\delta^* : C^* \to C^{*+1}$, $\beta_{*+1} : C^{*+1} \to C^*$ which satisfy

$$\begin{aligned} \delta^{*+1}\cdot\delta^* &= 0\\ \beta_*\cdot\beta_{*+1} &= 0\\ \beta_{*+1}\cdot\delta^* + \delta^{*-1}\cdot\beta_* &= 0. \end{aligned}$$

When there is no risk of confusion the index $*$ will be deleted and we write (C, δ, β) instead of (C^*, δ^*, β_*). Using the terminology of [**4**], [**9**] a mixed complex can be viewed either as a cochain complex (C^*, d^*) equipped with an S^1-action β_*, or as a chain complex (C^*, β_*) equipped with an algebraic S^1-action δ^*.

To a mixed complex (C^*, δ^*, β_*) one associates a number of cochain, chain and $2-$periodic cochain complexes, and then their cohomologies, homologies and $2-$periodic cohomologies[1], as follows.

First denote by

$$\begin{aligned} {}^+C^r &= \prod_{k\geq 0} C^{r-2k} & {}^-C^r &:= \prod_{k\geq 0} C^{r+2k}\\ \mathbb{P}C^{2r+1} &= \prod_{k\geq 0} C^{2k+1} & \mathbb{P}C^{2r} &= \prod_{k\geq 0} C^{2k} \text{ for any } r\\ PC^{2r+1} &= \bigoplus_{k\geq 0} C^{2k+1} & PC^{2r} &= \bigoplus_{k\geq 0} C^{2k} \text{ for any } r. \end{aligned} \tag{1}$$

[1] We will use the word "homology" for a functor derived from a chain complex and "cohomology" for one derived from a cochain complex. The $2-$periodic chain and cochain complexes can be identified.

Since our vector spaces are $\mathbb{Z}_{\geq 0}$-graded the direct product ${}^{+}C^r$ involves only finitely many factors.

Next introduce

$$
\begin{aligned}
{}^{+}D^r_\beta(w_r, w_{r-2}, \cdots) &:= (\delta\omega_r, (\delta\omega_{r-2} + \beta\omega_r), \cdots)\\
{}^{+}D^\delta_r(w_r, w_{r-2}, \cdots) &:= ((\beta\omega_r + \delta\omega_{r-2}), (\beta\omega_{r-2} + \delta\omega_{r-4}), \cdots)\\
{}^{-}D^r_\beta(\cdots, w_{r+2}, w_r) &:= (\cdots, (\beta\omega_{r+4} + \delta\omega_{r+2}), (\beta\omega_{r+2} + \delta\omega_r))\\
{}^{-}D^\delta_r(\cdots, w_{r+2}, w_r) &:= (\cdots, (\delta\omega_r + \beta\omega_{r+2}), \beta\omega_r)\\
D^{2r}(\cdots, \omega_{2r+2}, \omega_{2r}, \cdots \omega_0) &= (\cdots, (\delta\omega_{2r} + \beta\omega_{2r+2}), \cdots)\\
D^{2r+1}(\cdots, \omega_{2r+3}, \omega_{2r+1} \cdots \omega_1) &= (\cdots, (\delta\omega_{2k+1} + \beta\omega_{2k+3}), \cdots).
\end{aligned}
\tag{2}
$$

Finally consider the cochain complexes

$$\mathcal{C} := (C^*, \delta^*), \quad {}^{+}\mathcal{C}_\beta := ({}^{+}C^*, {}^{+}D^*_\beta), \quad {}^{-}\mathcal{C}_\beta := ({}^{-}C^*, {}^{-}D^*_\beta),$$

the chain complexes

$$\mathcal{H} := (C^*, \beta_*), \quad {}^{+}\mathcal{H}^\delta := ({}^{+}C^*, {}^{+}D^\delta_*), \quad {}^{-}\mathcal{H}^\delta := ({}^{-}C*, {}^{-}D^\delta_*)$$

and the $2-$periodic cochain complexes [2]

$$P\mathcal{C} := (PC^*, D^*), \quad \mathbb{P}\mathcal{C} := (\mathbb{P}C^*, D^*)$$

whose cohomology, homology and $2-$periodic cohomology are denoted by

$$
\begin{aligned}
&H^* := H^*(C, \delta), \quad {}^{+}H^*_\beta := {}^{+}H^*_\beta(C, \delta, \beta), \quad {}^{-}H^*_\beta := {}^{-}H^*_\beta(C, \delta, \beta),\\
&H_* := H_*(C, \beta), \quad {}^{+}H^\delta_* := {}^{+}H^\delta_*(C, \delta, \beta), \quad {}^{-}H^\delta_* := {}^{-}H^\delta_*(C, \delta, \beta),\\
&PH^* := PH^*(C, \delta, \beta), \qquad \mathbb{P}H^* := \mathbb{P}H^*(C, \delta, \beta).
\end{aligned}
$$

In this paper the chain complexes $\mathcal{H}, {}^{\pm}\mathcal{H}^\delta$ will only be used to derive conclusions about the cochain complexes $\mathcal{C}, {}^{\pm}\mathcal{C}_\beta,\ \mathbb{P}\mathcal{C}$.

The obvious inclusions and projections lead to the following commutative diagrams of short exact sequences:

$$
\begin{tikzcd}
0 \arrow[r] & {}^{-}\mathcal{H}^\delta_* \arrow[r] \arrow[d] & \mathbb{P}\mathcal{C}^* \arrow[r] \arrow[d] & {}^{+}\mathcal{H}^\delta_{*-2} \arrow[r] \arrow[d, "id"] & 0\\
0 \arrow[r] & \mathcal{H}_* \arrow[r] & {}^{+}\mathcal{H}^\delta_* \arrow[r] & {}^{+}\mathcal{H}^\delta_{*-2} \arrow[r] & 0
\end{tikzcd}
$$

$$
\begin{tikzcd}
0 \arrow[r] & {}^{+}\mathcal{C}^{*-2}_\beta \arrow[r, "i^{*-2}"] \arrow[d, "id"] & {}^{+}\mathcal{C}^*_\beta \arrow[r] \arrow[d, "I^*"] & \mathcal{C}^* \arrow[r] \arrow[d] & 0\\
0 \arrow[r] & {}^{+}\mathcal{C}^{*-2}_\beta \arrow[r, "I^{*-2}"] & \mathbb{P}\mathcal{C}^* \arrow[r] & {}^{-}\mathcal{C}^*_\beta \arrow[r] & 0.
\end{tikzcd}
$$

They give rise to the following commutative diagram of long exact sequences.

---

[2]here $(\mathbb{P}C^*, D^*)$ is regarded as a cochain complex with D^* obtained from the degree $+1$ derivation δ perturbed by the degree -1 derivation β ; the same complex can be regarded as a chain complex with D^* obtained from the degree -1 derivation β perturbed by the degree $+1$ derivation δ; the cohomology for the first is the same as the homology for the second

$$\begin{tikzcd}
\cdots \arrow[r] & {}^{-}H^{\delta}_{r} \arrow[r] \arrow[d, "h_r"] & \mathbb{P}H(r) \arrow[r] \arrow[d] & {}^{+}H^{\delta}_{r-2} \arrow[r] \arrow[d, "id"] & {}^{-}H^{\delta}_{r-1} \arrow[r] \arrow[d, "h_{r-1}"] & \cdots \\
\cdots \arrow[r] & H_r \arrow[r, "J_r"] & {}^{+}H^{\delta}_{r} \arrow[r, "S_r"] & {}^{+}H^{\delta}_{r-2} \arrow[r, "B_{r-2}"] & H_{r-1} \arrow[r] & \cdots
\end{tikzcd}$$

Fig 1.

$$\begin{tikzcd}
\cdots \arrow[r] & {}^{+}H^{r-2}_{\beta} \arrow[r, "S^{r-2}"] \arrow[d, "id"] & {}^{+}H^{r}_{\beta} \arrow[r] \arrow[d, "\mathbb{I}^*"] & H^r \arrow[r, "B^r"] \arrow[d, "i^r"] & {}^{+}H^{r-1}_{\beta} \arrow[r] \arrow[d, "id"] & \cdots \\
\cdots \arrow[r] & {}^{+}H^{r-2}_{\beta} \arrow[r, "\mathbb{I}^{r-2}"] & \mathbb{P}H^r \arrow[r, "\mathbb{J}^r"] & {}^{-}H^{r}_{\beta} \arrow[r, "\mathbb{B}^r"] & {}^{+}H^{r-1}_{\beta} \arrow[r] & \cdots
\end{tikzcd}$$

Fig 2.

and

$$\begin{tikzcd}
{}^{+}H^{r-2}_{\beta} \arrow[r, "S^{r-2}"] \arrow[d, "\mathbb{I}^{r-2}"] & {}^{+}H^{r}_{\beta} \arrow[dl, "\mathbb{I}^r"] \arrow[d] \\
\mathbb{P}H^r = \mathbb{P}H^{r+2} & \varinjlim_{S} {}^{+}H^{r+2k}_{\beta} \arrow[l]
\end{tikzcd}$$

The diagram (Fig 1) is the one familiar in the homological algebra of Hochschild versus cyclic homologies, cf [**19**]. The diagram Fig 2 is the one we will use in this paper.

Note that the Hochschild, cyclic, periodic cyclic, negative cyclic homology of an associative unital algebra A as defined in [**19**], is $H_*, {}^{+}H^{\delta}_{*}$, $\mathbb{P}H^*, {}^{-}H^{\delta}_{*}$ of the Hochschild mixed complex with $C^r := A^{\otimes(r+1)}$, β the Hochschild boundary, and $\delta^r = (1-\tau_{r+1})\cdot s_r\cdot(1+\tau_r+\cdots\tau_r^r)$ where $\tau_r(a_0\otimes a_1\otimes\cdots\otimes a_r) = (a_r\otimes a_0\otimes\cdots\otimes a_{r-1})$ and $s_r(a_0\otimes a_1\otimes\cdots\otimes a_r) = (1\otimes a_0\otimes a_1\otimes\cdots\otimes a_r)$.

A morphism $f : (C^*_1, \delta^*_1, \beta^1_*) \to (C^*_2, \delta^*_2, \beta^2_*)$ is a degree preserving linear map which intertwines δ's and β's. It induces degree preserving linear maps between any of the homologies /cohomologies defined above. The following elementary observations will be used below.

Proposition 1. *Let (C, δ, β) be a mixed cochain complex.*

1.$PH^r = \varinjlim_{S} {}^{+}H^{r+2k}_{\beta}$, where $S^{k+2r} : {}^{+}H^{k+2r}_{\beta} \to {}^{+}H^{k+2r+2}_{\beta}$ is induced by the inclusion ${}^{+}\mathcal{C}^{*}_{\beta} \to {}^{+}\mathcal{C}^{*+2}_{\beta}$.

2. The following is an exact sequence

$$0 \longrightarrow \varprojlim_{S}{}' \; {}^{+}H^{\delta}_{r-1+2k} \longrightarrow \mathbb{P}H^r \longrightarrow \varprojlim_{S} \; {}^{+}H^{\delta}_{r+2k} \longrightarrow 0,$$

with $S_{k+2r} : {}^{+}H^{\delta}_{k+2r} \to {}^{+}H^{\delta}_{k+2r-2}$ induced by the projection ${}^{+}\mathcal{H}^{\delta}_{*} \to {}^{+}\mathcal{H}^{\delta}_{*-2}$.

Let $f^* : (C^*_1, \delta^*_1, \beta^1_*) \to (C^*_2, \delta^*_2, \beta^2_*)$ be a morphism of mixed complexes.

3. If $H^*(f)$ is an isomorphism then so is ${}^{+}H^{*}_{\beta}(f)$ and $PH^*(f)$.

4. If $H_*(f)$ is an isomorphism then so is ${}^{+}H^{\delta}_{*}(f)$ and $\mathbb{P}H^*(f)$.

5. If $H^*(f)$ and $H_*(f)$ are both isomorphisms then, in addition to the conclusions in (3) and (4), ${}^{-}H^{*}_{\beta}(f)$ is an isomorphism.

Proof. (1): Recall that a direct sequence of cochain complexes

$$\mathcal{C}_0^* \xrightarrow{i_0} \mathcal{C}_1^* \xrightarrow{i_1} \mathcal{C}_2^* \xrightarrow{i_2} \cdots$$

induces, by passing to cohomology, the direct sequence

$$H^*(\mathcal{C}_0^*) \xrightarrow{H(i_0)} H^*(\mathcal{C}_1^*) \xrightarrow{H(i_1)} H^*(\mathcal{C}_2^*) \xrightarrow{i_2} \cdots$$

and that $H^j(\lim\limits_{\rightarrow} \mathcal{C}_i^*) = \lim\limits_{\rightarrow} H^j(\mathcal{C}_i^*)$ for any j.

(2): Recall that an inverse sequence of chain complexes

$$\mathcal{H}_*^0 \xleftarrow{p_0} \mathcal{H}_*^1 \xleftarrow{p_1} \mathcal{H}_*^2 \xleftarrow{p_2} \cdots$$

induces, by passing to homology, the sequence

$$H_*(\mathcal{H}_0^*) \xleftarrow{H(p_0)} H_*(\mathcal{H}_1^*) \xleftarrow{H(p_1)} H_*(\mathcal{H}_2^*) \xleftarrow{p_2} \cdots$$

and the following short exact sequence cf.[**19**] 5.1.9.

$$0 \longrightarrow \lim\limits_{\leftarrow}{}' H_{j-1}(\mathcal{H}_*^i) \longrightarrow H_j(\lim\limits_{\leftarrow} \mathcal{H}_*^i) \longrightarrow \lim\limits_{\leftarrow} H_j(\mathcal{H}_*^i) \longrightarrow 0$$

for any j.

Item (3) follows by induction on degree from the naturality of the first exact sequence in the diagram Fig 2 and (1).

Item (4) follows by induction from the naturality of the second exact sequence of the diagram Fig 1 and from (2).

Item (5) follows from the naturality of the second exact sequence in diagram Fig 2 and from (3) and (4). $\square$

The mixed complex (C^*, δ^*, β_*) is called $\beta-$acyclic if β_1 is surjective and $\ker(\beta_r) = \operatorname{im}(\beta_{r+1})$. If so consider the diagram whose rows are short exact sequences of cochain complexes

$$\begin{tikzcd}
0 \arrow[r] & (\mathrm{Im}(\beta)^*, \delta^*) \arrow[r, "i"] \arrow[dr, "j"] & (C^*, \delta^*) \arrow[r, "\beta"] \arrow[dr, "\mathrm{id}"] & ((\mathrm{Im}(\beta))^{*-1}, \delta^{*-1}) \arrow[r] & 0 \\
0 \arrow[r] & ({}^+C_\beta^{*-2}, {}^+D_\beta^{*-2}) \arrow[r, "i^{*-2}"] & ({}^+C_\beta^*, {}^+D_\beta^*) \arrow[r] & (C^*, \delta^*) \arrow[r] & 0
\end{tikzcd}$$

Each row induces the long exact sequence in the diagram below and a simple inspection of boundary maps in these long exact sequences permits one to construct linear maps θ^r and to verify that the diagram below is commutative.

$$\begin{tikzcd}
H^{r-2}(\mathrm{Im}(\beta), \delta) \arrow[r] \arrow[d, "\theta^{r-2}"] & H^r(\mathrm{Im}(\beta), \delta) \arrow[r] \arrow[d, "\theta^r"] & H^r(C, \delta) \arrow[r] \arrow[d, "\mathrm{id}"] & H^{r-1}(\mathrm{Im}(\beta), \delta) \arrow[d, "\theta^{r-1}"] \\
{}^+H_\beta^{r-2}(C, \delta, \beta) \arrow[r] & {}^+H_\beta^r(C, \delta, \beta) \arrow[r] & H^r(C, \delta) \arrow[r] & {}^+H_\beta^{r-1}(C, \delta, \beta)
\end{tikzcd}$$

As a consequence one verifies by induction on degree that the inclusion $j : (\mathrm{Im}(\beta), \delta) \to ({}^+C, {}^+D_\beta)$ induces an isomorphism $H^*(\mathrm{Im}(\beta), \delta) \to {}^+H_\beta^*(C, \delta, \beta)$.

Mixed complex with power maps and involution

A collection of degree zero (degree preserving) linear maps $\Psi_k, k = 1, 2. \cdots, \tau := \Psi_{-1}$ which satisfy

(i) $\Psi_k \circ \delta = \delta \circ \Psi_k$,
(ii) $\Psi_k \circ \beta = k\beta \circ \Psi_k$,
(iii) $\Psi_k \circ \Psi_r = \Psi_r \circ \Psi_k = \Psi_{kr}, \quad \Psi_1 = \mathrm{id}$

will be referred to as "power maps and involution", or simpler "power maps", $\Psi_k, k = -1, 1, 2, \cdots$.[3]. They provide the morphisms of cochain complexes

$$\Psi_k : \mathcal{C} \to \mathcal{C},$$
$${}^{\pm}\Psi_k : {}^{\pm}\mathcal{C}_\beta \to {}^{\pm}\mathcal{C}_\beta$$
$$\mathbb{P}\Psi_k : \mathbb{P}\mathcal{C} \to \mathbb{P}\mathcal{C}$$

defined as follows

$$\begin{aligned} {}^{+}\Psi_k^r(w_r, w_{r-2}, \cdots) =& (\Psi^r(\omega_r), \frac{1}{k}\Psi^{r-2}(\omega_{r-2}), \cdots) \\ {}^{-}\Psi_k^r(\cdots, w_{r+2}, w_r) :=& (\cdots, k\Psi_k^{2r+2}(\omega_{r+2}), \Psi_k^r(\omega_r)) \end{aligned}$$

and

$$\begin{aligned} \mathbb{P}\Psi_k^{2r}(\cdots, \omega_{2r+2}, \omega_{2r}, \omega_{r-2}, \cdots, \omega_0) = \\ = (\cdots, k\Psi_k^{2r+2}(\omega_{r+2}), \Psi_k^{2r}(\omega_{2r}), \frac{1}{k}\Psi_k^{2r-2}(\omega_{2r-2}), \cdots, \frac{1}{k^r}\Psi_k^0(\omega_0)) \\ \mathbb{P}\Psi_k^{2r+1}(\cdots, \omega_{2r+3}, \omega_{2r+1}, \omega_{2r-1}, \cdots, \omega_1) = \\ = (\cdots, k\Psi_k^{2r+3}(\omega_{2r+3}), \Psi_k^{2r+1}(\omega_{2r+1}), \frac{1}{k}\Psi_k^{2r-1}(\omega_{2r-1}), \cdots, \frac{1}{k^r}\Psi_k^1(\omega_1)) \end{aligned}$$

Consequently they induce the endomorphisms

$$\underline{\Psi}_k^* : H^* \to H^*$$
$$\underline{{}^{\pm}\Psi}_k^* : {}^{\pm}H_\beta^* \to {}^{\pm}H_\beta^*$$
$$\underline{\mathbb{P}\Psi^*}_k : \mathbb{P}H^* \to \mathbb{P}H^*$$

Note that in diagram (Fig 2):
$\mathbb{J}^*$, J^* and the vertical arrows intertwine the endomorphisms induced by Ψ_k,
$\mathbb{B}^*$, resp. B^*, intertwines $k({}^{-}\underline{\Psi}_k)$, resp. $k\underline{\Psi}_k$, with ${}^{+}\underline{\Psi}_k$,
$\mathbb{I}^{*-2}$, resp. S^{*-2}, intertwines ${}^{+}\underline{\Psi}_k$ with $k\underline{\mathbb{P}\Psi}_k$ resp. $k({}^{+}\underline{\Psi}_k)$,
The above elementary linear algebra will be applied to CDGA's in the next sections.

3. Mixed commutative differential graded algebras

Let κ be a field of characteristic zero (for example $\mathbb{Q}, \mathbb{R}, \mathbb{C}$).

Definition 1. (i) A commutative graded algebra, abbreviated CGA, is an associative unital augmentable graded algebra $\mathcal{A}^*$, (the augmentation is not part of the data) which is commutative in the graded sense, i.e.

$$a_1 \cdot a_2 = (-1)^{r_1 r_2} a_2 \cdot a_1, \quad a_i \in \mathcal{A}^{r_i}, i = 1, 2.$$

(ii) An exterior differential $d_{\mathcal{A}}^* : \mathcal{A}^* \to \mathcal{A}^{*+1}$, is a degree $+1$ linear map which satisfies

$$d(a_1 \cdot a_2) = d(a_1) \cdot a_2 + (-1)^{r_1} a_1 \cdot d(a_2), a_1 \in \mathcal{A}^{r_1}, \quad d_{\mathcal{A}}^{*+1} d_{\mathcal{A}}^* = 0.$$

[3]We use the notation τ for Ψ_{-1} to emphasize that is an involution and to suggest consistency with other familiar involutions in homological algebra and topology.

(iii) An interior differential $\beta_*^{\mathcal{A}} : \mathcal{A}^* \to \mathcal{A}^{*-1}$ is a degree -1 linear map which satisfies

$$\beta(a_1 \cdot a_2) = \beta(a_1) \cdot a_2 + (-1)^{r_1} a_1 \cdot \beta(a_2), a_1 \in \mathcal{A}^{r_1}, \quad \beta_{*-1}^{\mathcal{A}} \beta_*^{\mathcal{A}} = 0.$$

(iv) The exterior and interior differentials d^* and β_* are compatible if

$$d^{*-1} \cdot \beta_* + \beta_{*+1} \cdot d^* = 0.$$

(v) A pair $(\mathcal{A}^*, d^*)$, $\mathcal{A}^*$ a CGA and d^* an exterior differential, is called CDGA and a triple $(\mathcal{A}^*, d^*, \beta_*)$, $\mathcal{A}^*$ a CGA, d^* an exterior differential and β_* an interior differential, with d^* and β_* compatible, is called a mixed CDGA.

A mixed CDGA is a mixed cochain complex.

A degree preserving linear map $f^* : \mathcal{A}^* \to \mathcal{B}^*$ is a morphism of CGA's, resp. CDGA's, resp. mixed CDGA's if is a unit preserving graded algebra homomorphism and intertwines d's and β's when appropriate.

We will consider the categories of CGA's, CDGA's and mixed CDGA's. In all these three categories there is a canonical tensor product and in the category of CDGA's a well defined concept of homotopy between two morphisms [4](cf. [**19**], [**14**]). The category of mixed CDGA's is a subcategory of mixed cochain complexes and all definitions and considerations in section 2 can be applied.

For a (commutative) differential graded algebra $(\mathcal{A}^*, d_{\mathcal{A}}^*)$, the graded vector space $H^*(\mathcal{A}^*, d^*) = \mathrm{Ker}(d^*)/\mathrm{Im}(d^{*-1})$ is a commutative graded algebra whose multiplication is induced by the multiplication in $\mathcal{A}^*$. A morphism $f = f^* : (\mathcal{A}^*, d_{\mathcal{A}}^*) \to (\mathcal{B}^*, d_{\mathcal{B}}^*)$ induces a degree preserving linear map, $H^*(f) : H^*(\mathcal{A}^*, d_{\mathcal{A}}^*) \to H^*(\mathcal{B}^*, d_{\mathcal{B}}^*)$, which is an algebra homomorphism.

Definition 2. A morphism of CDGA's f, with $H^k(f)$ an isomorphism for any k, is called a quasi–isomorphism.

The CDGA $(\mathcal{A}, d_{\mathcal{A}})$ is called homologically connected if $H^0(\mathcal{A}, d_{\mathcal{A}}) = \kappa$ and homologically 1-connected if it is homologically connected and $H^1(\mathcal{A}, d_{\mathcal{A}}) = 0$.

The full subcategory of homologically connected CDGA's will be denoted by c–CDGA. For all practical purposes (related to geometry and topology) it suffices to consider only c-CDGA' s.

Definition 3. 1. The CDGA $(\mathcal{A}, d)$ is called free if $\mathcal{A} = \Lambda[V]$, where $V = \sum_{i \geq 0} V^i$ is a graded vector space and $\Lambda[V]$ denotes the free commutative graded algebra generated by V. If in addition $V^0 = 0$ then it is called a free connected commutative differential graded algebra, abbreviated fc-CDGA.

2. The CDGA $(\mathcal{A}, d)$ is called minimal if it is a fc-CDGA and in addition

i. $d(V) \subset \Lambda^+[V] \cdot \Lambda^+[V]$, with $\Lambda^+[V]$ the ideal generated by V,

ii. $V^1 = \oplus_{\alpha \in I} V_\alpha$ with I a well ordered set and $d(V_\beta) \subset \Lambda[\oplus_{\alpha < \beta} V_\alpha]$ (the set I and its order are not part of the data)

---

[4]let $k(t, dt)$ be the free commutative graded algebra generated by the symbol t of degree zero and dt of degree one, equipped with the differential $d(t) = dt$. A morphism $F : (\mathcal{A}, d_{\mathcal{A}}) \to (\mathcal{B}, d_{\mathcal{B}}) \otimes_k (k(t, dt), d)$, is called an elementary homotopy from f to g , $f, g : (\mathcal{A}, d_{\mathcal{A}}) \to (\mathcal{B}, d_{\mathcal{B}})$, if $\rho_0 \cdot F = f$, and $\rho_1 \cdot F = g$ where

$$\begin{aligned} \rho_0(a \otimes p(t)) =& p(0)a, \ \rho_0(a \otimes p(t)dt) = 0, \\ \rho_1(a \otimes p(t)) =& p(1)a, \ \rho_1(a \otimes p(t)dt) = 0, \end{aligned}$$

The homotopy is the equivalence relation generated by elementary homotopy.

OBSERVATION 1. *If* $(\Lambda[V], d_V)$ *is minimal and 1-connected, then* $V^1 = 0$ *and, for* $v \in V^i$, $d_V(v)$*is a linear combination of products of elements* $v_j \in V^j$ *with* $j < i$. *In particular for* $v \in V^2$ *one has* $dv = 0$.

The interest of minimal algebras comes from the following result [**18**], [**14**].

THEOREM 1. *1 (D. Sullivan)*

1. A quasi–isomorphism between two minimal CDGA's is an isomorphism.

2. For any homologically connected CDGA, $(\mathcal{A}, d_{\mathcal{A}})$, *there exist quasi–isomorphisms* $\theta : (\Lambda[V], d_V) \to (\mathcal{A}, d_{\mathcal{A}})$ *with* $(\Lambda[V], d_V)$ *minimal. Such a* θ *will be called a minimal model of* $(\mathcal{A}, d_{\mathcal{A}})$.

3. Given a morphism $f : (\mathcal{A}, d_{\mathcal{A}}) \to (\mathcal{B}, d_{\mathcal{B}})$ *and the minimal models* $\theta_A : (\Lambda[V_A], d_{V_A}) \to (A, d_A)$ *and* $\theta_B : (\Lambda[V_B], d_{V_B}) \to (B, d_B)$, *there exist morphisms* $f' : (\Lambda[V_A], d_{V_A}) \to (\Lambda[V_B], d_{V_B})$ *such that* $f \cdot \theta_A$ *and* $\theta_B \cdot f'$ *are homotopic; moreover any two such* (f')*'s are homotopic.*

We can therefore consider the homotopy category of c–CDGA's, whose morphisms are homotopy classes of morphisms of CDGA's. By the above theorem the full subcategory of fc-CDGA's is a skeleton, and therefore any homotopy functor a priori defined on fc–CDGA's admits extensions to homotopy functors defined on the full homotopy category of c–CDGA's and all these extensions are isomorphic as functors. In particular any statement about a homotopy functor on the category c-CDGA need only be verified for fc–CDGA.

Precisely for a c–GDGA, $(\mathcal{A}, d_{\mathcal{A}})$, choose a minimal model $\theta_{\mathcal{A}} : (\Lambda[V_{\mathcal{A}}], d_{V_{\mathcal{A}}}) \to (\mathcal{A}, d_{\mathcal{A}})$ and for any $f : (\mathcal{A}, d_{\mathcal{A}}) \to (\mathcal{B}, d_{\mathcal{B}})$ choose a morphism $f' : (\Lambda[V_{\mathcal{A}}], d_{V_{\mathcal{A}}}) \to (\Lambda[V_{\mathcal{B}}], d_{V_{\mathcal{B}}})$ so that $\theta_B \cdot f'$ and $f \cdot \theta_A$ are homotopic. Define the value of the functor on $(\mathcal{A}, d_{\mathcal{A}})$ to be the value on $(\Lambda[V_{\mathcal{A}}], d_{V_{\mathcal{A}}})$ and the value on a morphism $f : (\mathcal{A}, d_{\mathcal{A}}) \to (\mathcal{B}, d_{\mathcal{B}})$ to be the value on the morphism $f' : (\Lambda[V_{\mathcal{A}}], d_{V_{\mathcal{A}}}) \to (\Lambda[V_{\mathcal{B}}], d_{V_{\mathcal{B}}})$.

There are two natural examples of mixed CDGA's; one is provided by a smooth manifold equipped with a smooth vector field, the other by a construction referred to as "the free loop", considered first by Sullivan-Vigué. The free loop construction applies directly only to an fc-CDGA but in view of Theorem 1 can be indirectly used for any c–CDGA.

The first will lead to (the de Rham version of) a new homotopy functor defined on the category of possibly infinite dimensional manifolds (hence on the homotopy category of all countable CW complexes) , the **s-cohomology**, and its relationship with other familiar homotopy functors [5], cf. section 4 below. The second leads to simple definitions of three homotopy functors defined on the full category of c–CDGA's (via the minimal model theorem) with values in the graded vector spaces endowed with weight decomposition, cf section 5 below. Their properties lead to interesting results about the cohomology of the free loop space of 1-connected spaces.

---

[5]this functor was called in [**4**] and [**6**] string cohomology for its unifying role explained below, cf. Observation 2. The name "string homology" was afterwards used by Sullivan and his school to designate the homology and equivariant homology of the free loop space of a closed manifold when endowed with additional structures induced by intersection theory and the Pontrjagin product in the chains of based pointed loops, cf. [**11**].

4. de Rham Theory in the presence of a smooth vector field

Let M be a smooth manifold, possibly of infinite dimension. In the latter case the manifold is modeled on a good Fréchet space [6], for which the differential calculus can be performed as expected.

Consider the CDGA of differential forms $\Omega^*(M)$ with exterior differential $d^* : \Omega^*(M) \to \Omega^{*+1}(M)$ and interior differential $i^X_* : \Omega^*(M) \to \Omega^{*-1}(M)$, the contraction along the vector field X. They are not compatible. However we can consider the Lie derivative $L_X := d \cdot i^X + i^X \cdot d$ and define $\Omega_X(M) := \{\omega \in \Omega(M) | L_X \omega = 0\}$; $\Omega_X(M)$ consists of the smooth forms invariant by the flow induced by X. The graded vector space $\Omega^*_X(M)$, a subalgebra of $\Omega^*(M)$, is a commutative graded algebra and the restriction of d^* and of i^X_* leave invariant $\Omega^*_X(M)$ and are compatible. Consequently $(\Omega^*_X(M), d^*, i^X_*)$ is a mixed CDGA.

Denote

(i) $H^*_X(M) := H^*(\Omega^*_X, d^*)$,
(ii) ${}^{\pm}H^*_X(M) :={}^{\pm} H^*(\Omega^*_X, d^*, i^X_*)$,
(iii) $PH^*_X(M) := PH^*(\Omega^*_X, d^*, i^X_*)$,
(iv) $\mathbb{P}H^*_X(M) := \mathbb{P}H^*(\Omega^*_X, d^*, i^X_*)$.

The diagram Fig 2 becomes

$$\begin{tikzcd}
\arrow[r] & {}^+H^r_X(M) \arrow[r, "S^r"] \arrow[d, "id"] & {}^+H^{r+2}_X(M) \arrow[r, "J^{r+2}"] \arrow[d, "\mathbb{I}^{r+2}"] & H^{r+2}(M) \arrow[r, "B^{r+2}"] \arrow[d] & {}^+H^{r+1}_X(M) \arrow[r] \arrow[d] & \cdots \\
\arrow[r] & {}^+H^r_X(M) \arrow[r, "\mathbb{I}^r"] & \mathbb{P}H^{r+2}(M) \arrow[r, "\mathbb{J}^{r+2}"] & {}^-H^{r+2}_X(M) \arrow[r, "\mathbb{B}^{r+2}"] & {}^+H^{r+1}_X(M) \arrow[r] & \cdots
\end{tikzcd}$$

Fig 3

The above diagram becomes more interesting if the vector field X is induced from an S^1 action $\mu : S^1 \times M \to M$ (i.e. if $x \in M$ then $X(x)$ is the tangent to the orbits through x). In this section we will explore particular cases of this diagram.

Observe that since μ is a smooth action, the subset F of fixed points is a smooth submanifold. For any $x \in F$ denote by $\rho_x : S^1 \times T_x(M) \to T_x(M)$ the linearization of the action at x, which is a linear representation. The inclusion $F \subset M$ induces the morphism $r^* : (\Omega^*_X(M), d^*, i^X_*) \to (\Omega^*(F), d^*, 0)$.

For a linear representation $\rho : S^1 \times V \to V$ on a good Fréchet space denote by V^f the fixed point set and by X the vector field associated to ρ when regarded as a smooth action.

Definition 4. A linear representation $\rho : S^1 \times V \to V$ on the good Fréchet space is good if the following conditions hold:

a. V^f, the fixed point set, is a good Fréchet space,
b. The map $r^* : \Omega^*(V) \to \Omega^*(V^f)$ induced by the inclusion is surjective,
c. $(\Omega^*_X(V, V^f), i^X_*)$ with $(\Omega^*_X(V, V^f) = \ker r^*$ is acyclic.

We have:

Proposition 2. *1. Any representation on a finite dimensional vector space is good.*

[6]this is Fréchet space with countable base which admits a smooth partition of unity. Note that if a Fréchet space V is good then the space of smooth maps $C^\infty(S^1, V)$ equipped with the $C^\infty-$ topology is also good.

2. If V is a good Fréchet space then the regular representation $\rho : S^1 \times C^\infty(S^1, V) \to C^\infty(S^1, V)$, with $C^\infty(S^1, V)$, the Fréchet space of smooth functions, is good.

For a proof consult Appendix [**3**]. The proof is based on an explicit formula for i^X in the case of an irreducible S^1- representation and on the writing of the elements of $C^\infty(S^1, V)$ as Fourier series.

Definition 5. A smooth action $\mu : S^1 \times M \to M$ is good if its linearization at any fixed point is a good representation.

Thus a smooth action on any finite dimensional manifold is good and so is the canonical smooth action of S^1 on P^{S^1}, the smooth manifold of smooth maps from S^1 to P where P is any smooth Fréchet manifold (in particular a finite dimensional manifold). In view of the definitions above observe the following.

Proposition 3. *If $\tilde{M} = (M, \mu)$ is a smooth S^1-manifold and X is the associated vector field, then:*

1. $H^*_X(M) = H^*(M)$,

2. ${}^+H^*_X(M) = H^*_{S^1}(\tilde{M})$ and $S : H^*(M) \to H^{*+2}(M)$ is identified with the multiplication by $u \in H^2_{S^1}(\mathrm{pt})$, the generator of the equivariant cohomology of the one–point space,

3. $PH^*_X(M) = \varinjlim_{S} H^{*+2k}_{S^1}(\tilde{M})$.

If the action is good then:

4. $\mathbb{P}H^*_X(M) = K^*(F)$ where

$$K^r(F) = \prod_k H^{2k}(F) \text{ if } r \text{ is even}$$
$$K^r(F) = \prod_k H^{2k+1}(F) \text{ if } r \text{ is odd.}$$

If M is a closed $n-$dimension manifold then :

5. ${}^-H^*_X(M) = H^{S^1}_{n-1-*}(\tilde{M}, \mathcal{O}_M)$, with $H^{S^1}_*(\tilde{M}, \mathcal{O}_M)$ the equivariant homology with coefficients in the orientation bundle [7] of M.

Proof. 1. The verification is standard since S^1 is compact and connected; one constructs $\mathrm{av}^* : (\Omega^*(M), d^*) \to (\Omega^*_X, d^*)$ by S^1- averaging using the compactness of S^1. The homomorphism induced in cohomology by av^* is obviously surjective. To check it is injective one has to show that any closed k differential form ω which becomes exact after applying av^* is already exact, precisely $\int_c \omega = 0$ for any smooth $k-$ cycle c. Indeed, since the connectivity of S^1 implies $\int_c \omega = \int_{\mu(-\theta, c)} \omega,\ \theta \in S^1$, one has:

[7]Recall that $H^{S^1}_*(M, \mathcal{O}_M) = H_*(M//S^1, \mathbb{O}_M)$ where $M//S^1$ is the homotopy quotient of this action. This equivariant homology can be derived from invariant currents in the same way as equivariant cohomology from invariant forms, cf. [**1**]. The complex of invariant currents (with coefficients in the orientation bundle) contains the complex $(\Omega^{n-*}_X(M, \mathcal{O}_M), \partial_{n-*})$ as a quasi–isomorphic subcomplex.

$$\int_c \omega = (1/2\pi)\int_{S^1}(\int_c \omega)d\theta =$$
$$(1/2\pi)\int_{S^1}(\int_{\mu(-\theta,c)}\omega)d\theta = (1/2\pi)\int_{S^1}(\int_c \mu^*_\theta(\omega))d\theta =$$
$$\int_c (1/2\pi)\int_{S^1}\mu^*_\theta(\omega)d\theta) = \int_c \mathrm{av}^*(\omega) = 0.$$

Here μ_θ denotes the diffeomorphism $\mu(\theta,\cdot): M \to M$.

2. Looking at the definition in section 2 one recognizes one of the most familiar definition of equivariant cohomology using invariant differential forms cf. [**1**].

3. The proof is a straightforward consequence of Proposition 1 in section 2 and (2) above.

4. Let F be the smooth submanifold of fixed points of μ. Clearly r^* : $(\Omega^*_X(M), d^*, i^X_*) \to (\Omega^*(F), d^*_F, 0)$ is a morphism of mixed CDGA, hence of mixed complexes. If the smooth action is good then the above morphism induces an isomorphism in homology $H_*(\Omega^*_X, i^X_*) \to H_*(\Omega^*(F), 0)$. To check this we have to show that $(\ker r^*, i^X_*)$, with $\ker r^* := \{\omega \in \Omega^*_X(M) | \omega|_F = 0\}$, is acyclic. This follows (by S^1-averageing) from the acyclicity of the chain complex $(\Omega^*(M,F), i^X_*)$ which in turn can be derived using the linearity with respect to functions of of i^X. Indeed, using a "partition of unity" argument it suffices to verify this acyclicity locally. For points outside F the acyclicity follows from the acyclicity of the complex

$\cdots \Lambda^{*-1}(V) \xrightarrow{i^e} \Lambda^*(V) \xrightarrow{i^e} \Lambda^{*-1}(V) \to \cdots$, where V is a Fréchet space , $\Lambda^k(F)$ the space of skew-symmetric $k-$linear maps from V to $\kappa = \mathbb{R}, \mathbb{C}$ and $e \in V \setminus 0$. For points $x \in F$ this follows from the fact that the linearization of the action at x is a good representation, as stated in Proposition 2.

5. If $\tilde{M}$ is a finite dimensional smooth S^1-manifold we can equip M with an invariant Riemannian metric g and consider $\star : \Omega^*(M) \to \Omega^{n-*}(M; \mathcal{O}_M)$ the Hodge star operator. Denote by $\omega \in \Omega^1(M)$ the $1-$form corresponding to X with respect to the metric g, by $e_\omega : \Omega^*(M;\mathcal{O}_M) \to \Omega^{*+1}(M;\mathcal{O}_M)$ the exterior multiplication with ω and by $\partial_* : \Omega^*(M;\mathcal{O}_M) \to \Omega^{*-1}(M;\mathcal{O}_M)$ the formal adjoint of d^{*-1} with respect to g, i.e. $\partial_* = \pm \star \cdot d^{n-*} \cdot \star^{-1}$. Note that $e_\omega = \pm \star \cdot i^X \cdot \star^{-1}$. All these operators leave Ω_X invariant since g is invariant. Clearly $(\Omega^*_X(M;\mathcal{O}_M), e^*_\omega, \partial_*)$ is a mixed cochain complex and we have

$$^-H^*_{i^X}(\Omega_X(M), d, i^X) = ^+ H^\partial_{n-*}(\Omega_X(M;\mathcal{O}_M), e^\omega, \partial).$$

The equivariant homology of $\tilde{M}$ with coefficients in the orientation bundle can be calculated from the complex of invariant currents which, if M is closed, contains the complex $(\Omega^{n-*}_X(M, \mathcal{O}_M), \partial_{n-*})$ as a quasi–isomorphic sub complex. As a consequence we have

$$H_{n-*}(\Omega_X(M;\mathcal{O}_M), \partial) = H_*(M;\mathcal{O}_M)$$
$$^+H^{e_\omega}_{n-*}(\Omega_X(M;\mathcal{O}_M), e_\omega, \partial) = H^{S^1}_*(M;\mathcal{O}_M)$$

(cf. section 2 for notations). □

As a consequence of Proposition 3 (1)-(4) for any smooth S^1 manifold with good S^1- action the second long exact sequence in the diagram Fig 3 becomes

$$
\begin{tikzcd}
\cdots \arrow[r] & H^{r-2}_{S^1}(\tilde M) \arrow[r, "\mathbb{I}^{r-2}"] \arrow[dr] & K^r(F) \arrow[r, "\mathbb{J}^r"] & {}^-H^r_{S^1}(\tilde M) \arrow[r, "\mathbb{B}^r"] & H^{r-1}_{S^1}(\tilde M) \arrow[r] & \cdots \\
& & \varinjlim_S H^{r+2k}_{S^1}(\tilde M) \arrow[u] & & &
\end{tikzcd}
$$

Fig 4

The sequence above is obviously natural in the sense that $f : \tilde M \to \tilde N$, an S^1- equivariant smooth map, induces a commutative diagram whose rows are the above exact sequence Fig 4 for $\tilde M$ and $\tilde N$. Then if f and its restriction to the fixed point set induce isomorphisms in cohomology it induces isomorphisms in $H^*_{S^1}$ and K^* and then all other types of equivariant cohomologies ${}^-H^*_{S^1}$, $PH^*_{S^1}$, $\mathbb{P}H^*_{S^1}$.

If $\tilde M$ is a compact smooth S^1- manifold, in view of Proposition 3 (5), one identifies ${}^-H^r_{S^1}(M)$ with $H^{S^1}_{n-r}(M;\mathcal{O}_M)$ and in view of this identification writes Pd_{n-r}. instead of $\mathbb{B}^r$. The long exact sequence becomes

$$
\cdots \longrightarrow K^r(F) \longrightarrow H^{S^1}_{n-r}(\tilde M;\mathcal{O}_M) \xrightarrow{\mathrm{Pd}_{n-r}} H^{r-1}_{S^1}(\tilde M) \longrightarrow K^{r+1}(F) \longrightarrow \cdots .
$$

Fig 5

In case the fixed point set is empty we conclude that

$$
\mathrm{Pd}_{n-r} : H^{S^1}_{n-r}(\tilde M, \mathcal{O}_M) \to H^{r-1}_{S^1}(\tilde M)
$$

is an isomorphism. In this case the orbit space M/S^1 is a $\mathbb{Q}-$homological manifold of dimension $(n-1)$, hence

$$
\begin{aligned}
H^{S^1}_{n-r}(\tilde M;\mathcal{O}_M) &= H_{n-r}(M/S^1;\mathcal{O}_{M/S^1})\\
H^{r-1}_{S^1}(\tilde M;\mathcal{O}_M) &= H^{r-1}(M/S^1;\mathcal{O}_{M/S^1})
\end{aligned}
$$

and Pd_* is nothing but the Poincaré duality isomorphism for $\mathbb{Q}-$ homology manifolds. In general the long exact sequence Fig 5 measures the failure of the Poincaré duality map, Pd_*, to be an isomorphism.

5. The free loop space and S-cohomology

A more interesting example is provided by the S^1-manifold $\tilde{P^{S^1}} := (P^{S^1}, \mu)$. Here P^{S^1} denotes the smooth manifold of smooth maps from S^1 to P modeled by the Fréchet space $C^\infty(S^1, V)$ where V is the model for P (finite or infinite dimensional Fréchet space) cf [**3**]. This smooth manifold is equipped with the canonical smooth S^1-action $\mu : S^1 \times P^{S^1} \to P^{S^1}$ defined by

$$
\mu(\theta,\alpha)(\theta') = \alpha(\theta+\theta'), \quad \alpha : S^1 \to P, \quad \theta,\ \theta' \in S^1 = \mathbb{R}/2\pi.
$$

The fixed points set of the action μ consist of the constant maps and is hence identified with P. This action is the restriction of the canonical action of $O(2)$, the group of isometries of S^1, to the subgroup of orientation preserving isometries identified with S^1 itself. For any $x \in P$ viewed as a fixed point of μ the linearization representation is the regular representation of S^1 on $V = T_x(P)$. In view of Proposition 2 the action μ is good. The space P^{S^1} is also equipped with the natural maps ψ_k, $k = 1, 2, \cdots$, the geometric power maps and with the involution τ, defined by

$$\psi_k(\alpha)(\theta) = \alpha(k\theta)$$
$$\tau(\alpha)(\theta) = \alpha(-\theta)$$

with $\alpha \in P^{S^1}$, and $\theta \in S^1$.

The involution τ is the restriction of the action of $O(2)$ to the reflection $\theta \to -\theta$ in S^1. Then $(\Omega^*_X(P^{S^1}), d^*, i^X_*)$ is a mixed CDGA, hence a mixed complex with power maps Ψ_k, τ and involution τ induced from ψ_k and τ.

Suppose $f : P_1 \to P_2$ is a smooth map. It induces a smooth equivariant map $f^{S^1} : P_1^{S^1} \to P_2^{S^1}$ whose restriction to the fixed point set is exactly f. If f is a homotopy equivalence then so is f^{S^1}.

Introduce the notation

$$hH^*(P) := H^*(P^{S^1}),$$
$$cH^*(P) := H^*_{S^1}(\tilde{P^{S^1}}),$$
$$sH^*(P) :=^- H^*_{S^1}(\tilde{P^{S^1}}).$$

The assignments $P \rightsquigarrow hH^*(P)$, $P \rightsquigarrow cH^*(P)$, $P \rightsquigarrow sH^*(P)$ are functors [8] with the property that $hH^*(f), cH^*(f), sH^*(f)$ are isomorphism if f is a homotopy equivalence, hence they are all homotopy functors . They are related by the commutative diagram below. This diagram is the same as diagram (Fig 3) applied to $\tilde{M} = \tilde{P^{S^1}}$ with the specifications provided by Proposition 3.

$$\begin{tikzcd}
\cdots \arrow[r] & cH^{r-2}(P) \arrow[r, "S^{r-2}"] \arrow[d, "id"] & cH^r(P) \arrow[r, "J^r"] \arrow[d, "\mathbb{I}^r"] & hH^r(P) \arrow[r, "B^r"] \arrow[d] & cH^{r-1}(P) \arrow[r] \arrow[d, "id"] & \cdots \\
\cdots \arrow[r] & cH^{r-2}(P) \arrow[r, "\mathbb{I}^{r-2}"] \arrow[d, "S"] & K^r(P) \arrow[r, "\mathbb{J}^r"] & sH^r(P) \arrow[r, "\mathbb{B}^r"] & cH^{r-1}(P) \arrow[r] & \cdots \\
& cH^r(P) \arrow[r] \arrow[ur, "\mathbb{I}^r"] & \varinjlim cH^{r+2k}(P) \arrow[u] & & &
\end{tikzcd}$$

Fig 6

where $\varinjlim\ cH^{r+2k}(P) = \varinjlim\{\cdots \to cH^{r+2k}(P) \xrightarrow{S} cH^{r+2k+2}(P) \to \cdots\}$.

The linear map

$$cH^*(P) = H^*_{S^1}(\tilde{P^{S^1}}) \xrightarrow{\mathbb{I}^r} K^*$$

factors through $\varinjlim\ cH^{r+2k}(P)$ which depends only on the fundamental group of P. Indeed, it is shown in [**5**] that if $P(1)$ [9] is a smooth manifold (possibly of infinite dimension) which has the homotopy type of $K(\pi, 1)$ and $p(1) : P \to P(1)$ is smooth map inducing an isomorphism for the fundamental group then $\varinjlim\ H^{r+2k}_{S^1}(\tilde{P^{S^1}}) \to \varinjlim\ H^{r+2k}_{S^1}(\tilde{P(1)^{S^1}})$ is an isomorphism. cf [**5**]. Then if one denotes by $\overline{cH^*}(M) := \mathrm{coker}(cH^*(M) \to cH^*(\mathrm{pt}))$ and $\overline{K^*}(M) := \mathrm{coker}(K^*(M) \to K^*(\mathrm{pt}))$[10] one obtains

[8] the notations hH^*, cH^* are motivated by the Hochschild resp. cyclic homology interpretation of these functors, while sH^* is abbreviation from string cohomology.

[9] the notation for the first stage Postnikov term of P.

[10] clearly $K^r(\mathrm{pt}) = H^r_{S^1}(\mathrm{pt}) = \kappa$, resp. 0, if r is even, resp. odd.

THEOREM 2. *If P is a 1-connected smooth manifold then we have the following short exact sequence:*

$$0 \to \overline{K}^r(P) \otimes \kappa \xrightarrow{\mathbb{J}^r} sH^r(P) \xrightarrow{\mathbb{B}^r} \overline{c}H^{r-1}(P) \to 0$$

where $\kappa = \mathbb{R}$ or $\mathbb{C}$.

OBSERVATION 2. *The vector space $K^*(P)$ can be identified via the Chern character with the Atiyah–Hirzebruch (complex) $K-$theory tensored with the field $\kappa = \mathbb{R}$ or $\mathbb{C}$, depending on what sort of differential forms one considers (real or complex valued). When P is 1-connected $\overline{c}H^*(P)$ identifies with $\mathrm{Hom}(\tilde{\mathrm{A}}_*(\mathrm{P}), \mathrm{k})$, where $\tilde{A}_*(P)$ denotes the reduced Waldhausen algebraic $K-$theory*[11]*, cf* [**4**]*. From this perspective sH^* unifies topological (Atiyah–Hirzebruch) $K-$theory and Waldhausen algebraic $K-$theory.*

OBSERVATION 3. *In view of the definition of ${}^-H^*_\beta(C^*, \delta^*, \beta_*)$, cf. section 2, observe that $sH^*(P)$ is represented by infinite sequences* [12]*, rather than eventually finite sequences of invariant differential forms on P^{S^1}. If instead of "infinite sequences" we would have considered "eventually finite sequences" the outcome would have been different for infinite dimensional manifolds. The difference between "infinite sequences" and "eventually finite sequences" exists only for infinite dimensional manifolds, which P^{S^1} always is.*

The power maps ψ_k induce the endomorphisms $h\Psi_k$, $c\Psi_k$ $s\Psi_k$ and $K\Psi_k$ on hH^*, cH^*, sH^*, and K^*.

In general only the $K\Psi_k$ are easy to describe. Precisely, if r is even then $K^r = \prod_{i\geq 0} H^{2i}(P)$ and if r is odd then $K^r = \prod_{i\geq 0} H^{2i+1}(P)$, and in both cases $K\Psi_k = \prod_{i\geq 0} k^{i-r}\mathrm{Id}$.

The symmetric part with respect to the involution $c\Psi_{-1}$, i.e. the eigenspaces corresponding to the eigenvalues $+1$ identifies with $H^*_{O(2)}(P^{S^1})$, the equivariant cohomology for the canonical $O(2)-$action.

However, if P is 1-connected, in view of the section 6, one can describe both the eigenvalues and the eigenspaces of the power maps $h\Psi_k$ and $c\Psi_k$ and then of $s\Psi_k$. We have:

THEOREM 3. *Let P be a 1-connected manifold.*

1. All eigenvalues of the endomorphisms $h\Psi_k$ and $c\Psi_k$ are $k^r, r = 0, 1, 2 \cdots$, and the eigenspaces corresponding to k^r are independent of k provided $k \geq 2$.

2. Denote these eigenspaces by $hH^*(M)(r)$ and $cH^*(M)(r)$. Then
$hH^*(0) = H^*(X;\kappa), cH^*(0) = H^{*+1}(X;\kappa)$, and
$hH^r(p) = cH^r(p) = 0, p \geq r + 1$.

3. If $\sum_i \dim \pi_i(P) \otimes \kappa < \infty$, κ the field of real or complex numbers and $\sum_i \dim(H^i(P)) < \infty$ then for any $r \geq 0$ one has

$$\sum_i \dim hH^i(P)(r) < \infty, \quad \sum_i \dim cH^i(P)(r) < \infty.$$

[11] often referred to as $A-$ theory.

[12] $sH^*(P)$ is the cohomology of the cochain complex $({}^-C^*, {}^-D^*)$ with ${}^-C^r = \prod_{k\geq 0} \Omega^{r+2k}_{inv}(P^{S^1})$ and ${}^-D^r(\cdots, \omega_{r+2}, \omega_r) = (\cdots, (i^X\omega_{r+2} + d\omega_r))$, cf. section 2.

If P is "formal" in the sense of rational homotopy theory [13] (a projective complex algebraic variety, or more generally a closed Kähler manifold is formal, cf [**13**]) then the Euler Poincaré characteristic

$$\chi^h(\lambda) := \sum_{i,r} \dim hH^i(r)\lambda^r$$

and

$$\chi^c(\lambda) := \sum_{i,r} \dim cH^i(r)\lambda^r$$

can be explicitly calculated in terms of the numbers $\dim H^i(P)$, cf [B]. The explicit formulae are quite complicated. They require the results of P.Hanlon [**15**] about the eigenspaces of Adams operations in Hochschild and cyclic homology as well as the identification of $hH^*(P)$ resp. $cH^*(P)$ with the Hochschild, resp. cyclic, homology of the graded algebra $H^*(P)$. These are not discussed in this paper but the reader can consult [**8**] and [**6**] for precise statements.

The functor $\overline{sH}^r(P)$ is of particular interest in geometric topology. In the case P is 1-connected it calculates in some ranges the homotopy groups of the (homotopy) quotient space of homotopy equivalences by the group of diffeomorphisms [**3**], [**6**].

6. The free loop construction for CDGA

The "free loop " construction associates to a free connected CDGA, $(\Lambda[V], d_V)$ a mixed CDGA, $(\Lambda[V \oplus \overline{V}], \delta_V, i^V)$, endowed with power maps Ψ_k and involution τ defined as follows.

(i) Let $\overline{V} = \oplus_{i\geq 0}\overline{V}^i$ with $\overline{V}^i := V^{i+1}$ and let $\Lambda[V \oplus \overline{V}]$ be the commutative graded algebra generated by $V \oplus \overline{V}$.

(ii) Let $i^V : \Lambda[V\oplus\overline{V}] \to \Lambda[V\oplus\overline{V}]$be the unique internal differential (of degree -1) which extends $i^V(v) = \overline{v}$ and $i^V(\overline{v}) = 0$.

(iii) Let $\delta_V : \Lambda[V \oplus \overline{V}] \to \Lambda[V \oplus \overline{V}]$ be the unique external differential (of degree $+1$) which extends $\delta_V(v) = d(v)$ and $\delta(\overline{v}) = -i^V(d(v))$.

(iv) Let $\Psi_k : (\Lambda[V\oplus\overline{V}], \delta_V) \to (\Lambda[V\oplus\overline{V}], \delta_V), k = -1, 1, 2, \cdots$ be the unique morphisms of CDGA which extends $\Psi_k(v) = v, \Psi_k(\overline{v}) = k\overline{v}$. We put $\tau := \Psi_{-1}$. The maps Ψ_k $k \geq 1$ are called the power maps and τ the canonical involution. One has

$$\begin{aligned}\Psi_k \cdot \Psi_r =&\Psi_{kr}\\ \Psi_k \cdot i_V =&ki_V \cdot \Psi_k\end{aligned}$$

(v) Let $\Lambda^+[V\oplus\overline{V}]$ be the ideal of $\Lambda[V\oplus\overline{V}]$ generated by $V\oplus\overline{V}$ or the kernel of the augmentation which vanishes on $V \oplus \overline{V}$.

Note that :

Observation 4. *1.* $(\mathrm{Im(i^V)}, \delta_\mathrm{V}, 0)$ *is a mixed subcomplex of* $(\Lambda^+[V\oplus\overline{V}], \delta_V, i^V) \subset (\Lambda[V \oplus \overline{V}], \delta_V, i^V)$

2. Ψ_k, $k = -1, 1, 2, \cdots$ *leave* $(\Lambda^+[V\oplus\overline{V}], \delta_V, i^V)$ *and* $(\mathrm{Im(i^V)}, \delta)$ *invariant and have* k^r $r = 0, 1, 2, \cdots$ *as eigenvalues. These are all the eigenvalues.*

[13] i.e. for each connected component of P, the de Rham algebra and the cohomology algebra equipped with the differential 0 are homotopy equivalent, cf section 6.

For $k \geq 2$ the eigenspace of $\Psi_r : \Lambda[V \oplus \overline{V}] \to \Lambda[V \oplus \overline{V}]$ corresponding to the eigenvalue k^r is exactly $\Lambda[V] \otimes \overline{V}^{\otimes r}$, resp. $\Lambda^+[V \oplus \overline{V}] \cap \Lambda^[V] \otimes \overline{V}^{\otimes r}$, resp. $\mathrm{Im(i^V)(r) = Im(i^V) \cap \Lambda[V] \otimes \overline{V}^{\otimes r}}$, and hence independent of k. Each such eigenspace is δ_V-invariant.

3. The mixed complex $(\Lambda^+[V \oplus \overline{V}], \delta_V, i^V)$ is i^V-acyclic,

4. We have the decomposition

$$(\Lambda[V \oplus \overline{V}], \delta_V) = \bigoplus_{r \geq 1} (\Lambda[V] \otimes \overline{V}^{\otimes r}, \delta_V)$$

and the analogous decomposition for $(\Lambda^+[V \oplus \overline{V}], \delta_V)$ and $(\mathrm{Im(i^V)}, \delta_\mathrm{V})$, referred to from now on as the weight decomposition.

Consider the complex $(\Lambda[V] \otimes \overline{V}^{\otimes r}, \delta_V)$ and the filtration provided by $\Lambda[V] \otimes F_p(\overline{V}^{\otimes r})$ with $F_p(\overline{V}^{\otimes r})$ the span of elements in $\overline{V}^{\otimes r}$ of total degree $\leq p$.

For a graded vector space $W = \oplus_i W^i$ denote $\dim W = \sum \dim W^i$.

Observation 5. *1. $(\Lambda[V] \otimes F_p(\overline{V}^{\otimes r}), \delta_V)$ is a subcomplex of $(\Lambda[V] \otimes \overline{V}^{\otimes r}, \delta_V)$.*

2.If $(\Lambda[V], d_V)$ is minimal and 1– connected then, by Observation 1, $\delta(F_p(\overline{V}^{\otimes r})) \subset \Lambda[V] \otimes F_{p-1}(\overline{V}^{\otimes r})$ and then

$$(\Lambda[V] \otimes F_p(\overline{V}^{\otimes r}) / \Lambda[V] \otimes F_{p-1}(\overline{V}^{\otimes r}) \ , \delta_V) = (\Lambda[V], d_V) \otimes F_p(\overline{V}^{\otimes r}) / F_{p-1}(\overline{V}^{\otimes r}).$$

2. $\sum_p \dim(F_p(\overline{V}^{\otimes r}) / F_{p-1}(\overline{V}^{\otimes r})) = \dim(\overline{V}^{\otimes r}) = (\dim V)^r$.

If $f : (\Lambda[V], d_V) \to \Lambda(W, d_W)$ is a morphism of CDGAs then it induces $\tilde{f} : (\Lambda[V \oplus \overline{V}], \delta_V, i^V) \to (\Lambda[W \oplus \overline{W}], \delta_W, i^W)$ which intertwines Ψ'_ks and then preserves the weight decompositions.

We introduce the the notation $HH^*, CH^* PH^*$

$$\begin{aligned} HH^*(\Lambda[V], d_V) :=& H^*(\Lambda[V \oplus \overline{V}], \delta_V) \\ CH^*(\Lambda[V], d_V) :=& {}^+H^*_{i^V}(\Lambda[V \oplus \overline{V}], \delta_V, i^V) \\ PH^*(\Lambda[V], d_V) :=& PH^*_{i^V}(\Lambda[V \oplus \overline{V}], \delta_V, i^V) \end{aligned}$$

and for a morphism f denote by $HH(f), CH(f), PH(f)$ the linear maps induced by $\tilde{f}$. The assignments HH^*, CH^*, PH^* provide functors from the category of fc–CDGA's to graded vector spaces. They come equipped with the operations $H\Psi_k, C\Psi_k$, etc. induced from Ψ_k. Since for f a quasi–isomorphism $HH^*(f), CH^*(f), PH^*(f)$ are isomorphisms, these functors, as shown in section 3, extend to the category of c–CDGA's. We have the following result.

Theorem 4. *Let $(\mathcal{A}, d_\mathcal{A})$ be a connected CDGA.*

1. All the eigenvalues of the endomorphisms $H\Psi_k$ and $C\Psi_k$ are $k^r, r = 0, 1, 2, \ldots$ and their eigenspaces are independent of k provided $k \geq 2$. One denotes them by $HH(\mathcal{A}, d_\mathcal{A})(r)$, and $CH(\mathcal{A}, d_\mathcal{A})(r)$.

2.

$$\begin{aligned} HH^*(\mathcal{A}, d_\mathcal{A})(0) =& H^*(\mathcal{A}, d_\mathcal{A}), \\ CH^*(\mathcal{A}, d_\mathcal{A})(0) =& H^{*+1}(\mathcal{A}, d_\mathcal{A}) \\ HH^r(\mathcal{A}, d_\mathcal{A})(p) =& CH^r(\mathcal{A}, d_\mathcal{A})(p) = 0, \quad p \geq r+1 \end{aligned}$$

3.Suppose $(\mathcal{A}, d_{\mathcal{A}})$ *is 1-connected with minimal model* $(\Lambda[V], d_V)$. *If* $\sum_i \dim V^i < \infty$ *and* $\sum_i \dim H^i(\mathcal{A}, d_{\mathcal{A}}) < \infty$ *then for any* $r \geq 0$ *one has*

$$\sum_i \dim HH^i(\mathcal{A}, d_{\mathcal{A}})(r) < \infty, \quad \sum_i \dim CH^i(\mathcal{A}, d_{\mathcal{A}})(r) < \infty.$$

PROOF. It suffices to check the statements for $(\mathcal{A}, d_{\mathcal{A}}) = (\Lambda[V], d_V)$ minimal. Items (1) and (2) are immediate consequences of Observation 4.

Item 3) follows from Observation 5. Indeed for a fixed r one has

$$\sum_i \dim H^i(\Lambda[V] \otimes \overline{V}^{\otimes r}, \delta_V)) \leq$$
$$\sum_{i,p} \dim H^i(\Lambda[V] \otimes F_p(\overline{V}^{\otimes r}/\Lambda[V] \otimes F_{p-1}(\overline{V}^{\otimes r}),\ \delta_V) =$$
$$(\dim V)^r \cdot \sum_i \dim H^i(\Lambda[V], d_V)$$

□

In addition to $\chi(\mathcal{A}, d_{\mathcal{A}}) := \sum (-1)^i \dim H^i(\mathcal{A}, d_{\mathcal{A}})$ one can consider
$\chi^H(\mathcal{A}, d_{\mathcal{A}})(r) := \sum (-1)^i \dim HH^i(\mathcal{A}, d\mathcal{A})(r)$ and
$\chi^C(\mathcal{A}, d_{\mathcal{A}})(r) := \sum (-1)^i \dim CH^i(\mathcal{A}, d_{\mathcal{A}})(r)$,
and then the power series in λ,

$$\chi^H(\mathcal{A}, d_{\mathcal{A}})(\lambda) := \sum \chi^H(\mathcal{A}, d_{\mathcal{A}})(r)\lambda^r, \quad \chi^C(\mathcal{A}, d\mathcal{A})(\lambda) := \sum \chi^C(\mathcal{A}, d\mathcal{A})(r)\lambda^r.$$

Theorem 4 (3) implies that for $(\mathcal{A}, d_{\mathcal{A}})$ 1-connected with $\sum_i \dim V^i < \infty$ and $\sum_i \dim H^i(\mathcal{A}, d_{\mathcal{A}}) < \infty$ the partial Euler–Poincaré characteristics $\chi^H(\mathcal{A}, d_{\mathcal{A}})(r)$ and $\chi^C(\mathcal{A}, d_{\mathcal{A}})(r)$ and therefore the power series $\chi^H(\mathcal{A}, d_{\mathcal{A}})(\lambda)$ and $\chi^C(\mathcal{A}, d_{\mathcal{A}})(\lambda)$ are well defined. The results of Hanlon [**15**] permit one to calculate explicitly $\chi^H(\lambda)$ and $\chi^C(\lambda)$ in terms of $\dim H^i(\mathcal{A}, d_A)$ if $(\mathcal{A}, d_{\mathcal{A}})$ is 1-connected and formal, i.e. there exists a quasi–isomorphism $(\Lambda[V], d) \to (H^*(\Lambda[V], d), 0)$, $(\Lambda[V], d)$ to a minimal model of $(\mathcal{A}, d_{\mathcal{A}})$.

We want to define an algebraic analogue of the functor sH^* on the category of cCDGA's. Recall that for a morphism $f^* : (C_1^*, d_1^*) \to (C_2^*, d_2^*)$ the "mapping cone" Cone(f*) is the cochain complex with components $C_f^* = C_2^* \oplus C_1^{*+1}$ and with $d_f^* = \begin{pmatrix} d_2^* & f^{*+1} \\ 0 & -d_1^{*+1} \end{pmatrix}$.

Notice that, when f^* is injective, the morphism $\mathrm{Cone(f^*)} \to \mathrm{C_2^*/f^*(C_1^*)}$ defined by the composition $C_2^* \oplus C_1^{*+1} \to C_2^* \to C_2^*/f^*(C_1^*)$ is a quasi–isomorphism.

We will consider the composition

$$\underline{I}^{*-2} :{}^+ \mathcal{C}^{*-2}_{i^V}(\Lambda[V \oplus \overline{V}], \delta_V, i^V) \overset{I^{*-2}}{\to} \mathbb{P}\mathcal{C}^*(\Lambda[V \oplus \overline{V}], \delta_V, i^V) \overset{\mathbb{P}^*(p)}{\to} \mathbb{P}\mathcal{C}^r(\Lambda[V], d_V, 0)$$

with the first arrow provided by the natural transformation $I^{*-2} :{}^+ C_\beta^{*-2} \to \mathbb{P}C^*$ described in section 2 applied to the mixed complex $(\Lambda[V \oplus \overline{V}], \delta_V, i^V)$ and the second induced by the projection on the zero weight component of $(\Lambda[V \oplus \overline{V}], \delta_V, i^V)$.

The mapping cone $\mathrm{Cone}(\underline{I}^{*-2})$, is functorial when regarded on the category of fc–CDGA's. Define

$$SH^*(\Lambda[V], d_V) := H^*(\mathrm{Cone}(\underline{I}^*)).$$

The assignment $(\Lambda[V], d_V) \rightsquigarrow SH^*(\Lambda[V], d_V)$ is a homotopy functor.

Consider the commutative diagrams

$$\begin{tikzcd}
{}^{+}\mathcal{C}^{*-2}_{i^V}(\Lambda[V\oplus\overline{V}],\delta_V,i^V) \arrow[r, "i^{*-2}"] \arrow[d, "id"] & {}^{+}\mathcal{C}^{*}_{i^V}(\Lambda[V\oplus\overline{V}],\delta_V,i^V) \arrow[r] \arrow[d, "\underline{I}^*"] & \mathrm{Cone}(i^{*-2}) \arrow[d] \\
{}^{+}\mathcal{C}^{*-2}_{i^V}(\Lambda[V\oplus\overline{V}],\delta_V,i^V) \arrow[r, "\underline{I}^{*-2}"] & \mathbb{P}\mathcal{C}^*(\Lambda[V],d_V,0) \arrow[r] & \mathrm{Cone}(\underline{\mathrm{I}}^{*-2}).
\end{tikzcd}$$

and

$$\begin{tikzcd}
{}^{+}\mathcal{C}^{*-2}_{i^V}(\Lambda[V\oplus\overline{V}],\delta_V,i^V) \arrow[r, "i^{*-2}"] \arrow[d, "id"] & {}^{+}\mathcal{C}^{*}_{i^V}(\Lambda[V\oplus\overline{V}],\delta,i^V) \arrow[r] \arrow[d, "id"] & \mathrm{Cone}(\mathrm{i}^{*-2}) \arrow[d] \\
{}^{+}\mathcal{C}^{*-2}_{i^V}(\Lambda[V\oplus\overline{V}],\delta_V,i^V) \arrow[r, "i^{*-2}"] & {}^{+}\mathcal{C}^{*}_{i^V}(\Lambda[V\oplus\overline{V}],\delta_V,i^V) \arrow[r] & C^*(\Lambda[V\oplus\overline{V}],\delta_V)
\end{tikzcd}$$

with the last vertical arrow in the second diagram a quasi–isomorphism as noted above.

The long exact sequence induced by passing to cohomology in the first diagram combined with the identifications implied by the second diagram lead to

$$\begin{tikzcd}
\cdots \arrow[r] & CH^{r-2} \arrow[r, "S^{r-2}"] \arrow[d, "id"] & CH^r \arrow[r, "J^r"] \arrow[d, "T^r"] & HH^r \arrow[r, "B^r"] \arrow[d] & CH^{r-1} \arrow[r] \arrow[d, "id"] & \cdots \\
\cdots \arrow[r] & CH^{r-2} \arrow[r, "\mathbb{I}^{r-2}"] \arrow[d, "S^{r-2}"] & K^r \arrow[r, "\mathbb{J}^r"] & SH^{r+2} \arrow[r, "\mathbb{B}^r"] & CH^{r-1} \arrow[r] & \cdots \\
& CH^r \arrow[r] \arrow[ur, "\mathbb{I}^r"] & PH^r \arrow[u] & & &
\end{tikzcd}$$

Fig 7

with $PH^r = \varinjlim CH^{r+2k}$ and $K^r := K^r(\Lambda[V],d_V)$ given by $=\prod_k H^{2k}(\Lambda[V],d_V)$ resp. $\prod_k H^{2k+1}(\Lambda[V],d_V)$if r is even, resp. odd. It is immediate that Theorem 2 remains true for sH^*, cH^* replaced by SH^*, CH^* as follows easily from the diagram Fig 7. The diagram Fig 7 should be compared with diagram Fig 6. This explains why $SH^*(\Lambda[V],d^V)$ will be regarded as the algebraic analogue of $sH^*(P)$.

It is natural to ask if the functors HH^*, CH^*, SH^* applied to $(\Omega^*(P),d^*)$ calculate hH^*, cH^*, sH^* applied to P and the diagram Fig 7 identifies to the diagram Fig 6. The answer is in general no, but is yes if P 1-connected .

The minimal model theory, discussed in the next section, permits one to identify $HH^*(\Omega^*(P),d^*), CH^*(\Omega^*(P),d^*)$ with $hH^*(P), CH^*(P)$ and then $SH^*(\Omega^*(P),d^*)$ with $SH^*(P)$ and actually diagram Fig 6 with diagram Fig 7 when P is 1-connected.

7. Minimal models and the proof of Theorem 3

Observe that if $(\mathcal{A}^*,d^*,\beta_*)$ is a mixed CDGA equipped with the power maps and involution $\Psi_k, k=-1,1,2,\cdots$, then the diagram

$$\begin{tikzcd}
{}^{+}H^*_\beta(\mathcal{A}^*,d^*,\beta_*) \arrow[r] \arrow[d, "{}^{+}\underline{\Psi}_k"] & H^*(\mathcal{A}^*,d^*) \arrow[d, "{}^{+}\underline{\Psi}_k"] \\
{}^{+}H^*_\beta(\mathcal{A}^*,d^*,\beta_*) \arrow[r] & H^*(\mathcal{A}^*,d^*)
\end{tikzcd}$$

can be derived by passing to cohomology in the commutative diagram of CDGA's.

$$\begin{tikzcd}
(\mathcal{A}^*\otimes\Lambda[u],\mathcal{D}[u]) \arrow[r] \arrow[d, "\Psi_k[u]"] & (\mathcal{A}^*,d^*) \arrow[d, "\Psi_k"] \\
(\mathcal{A}^*\otimes\Lambda[u],\mathcal{D}[u]) \arrow[r] & (\mathcal{A}^*,d^*)
\end{tikzcd}$$

where $\Lambda[u]$ is the free commutative graded algebra generated by the symbol u of degree 2, $\mathcal{D}[u](a\otimes u^r)=d(a)\otimes u^r+\beta(a)\otimes u^{r+1}$ and $\Psi[u](a\otimes u^r)=1/k^r\Psi(a)\otimes u^r$.

For P 1-connected and $(\Lambda([V],d_V)$ a minimal model of $(\Omega^*(P),d^*)$ we want to establish the existence of the homotopy commutative diagram

$$\begin{tikzcd}
 & A \arrow[rr] \arrow[dd, "\Psi_k^P[u]"' near start] & & B \arrow[dd, "\Psi_k^P"] \\
C \arrow[ur, "\tilde\theta"] \arrow[rr, crossing over] \arrow[dd, "\Psi_k[u]"'] & & D \arrow[ur, "\theta"] & \\
 & A \arrow[rr] & & B \\
C \arrow[ur, "\tilde\theta"] \arrow[rr] & & D \arrow[ur, "\theta"] \arrow[from=uu, crossing over, "\Psi_k"' near end] &
\end{tikzcd}$$

where

$$\begin{aligned}
A =&(\Omega_X(P^{S^1})\otimes\Lambda[u],\mathcal{D}[u])\\
B =&(\Omega_X(P^{S^1}),d)\\
C =&(\Lambda[V\oplus\overline{V}]\otimes\Lambda[u],\delta[u])\\
D =&(\Lambda[V\oplus\overline{V}],\delta)
\end{aligned}$$

with

$$\begin{aligned}
\mathcal{D}[u](\omega\otimes u^r) =&d(\omega)\otimes u^r+i^X(\omega)\otimes u^{r+1}\\
\delta[u](a\otimes u^r) =&\delta(a)\otimes u^r+i^V(\omega)\otimes u^{r+1}.
\end{aligned}$$

The existence of the quasi–isomorphism θ was established in [**20**]. The existence of the quasi–isomorphism $\tilde\theta$ and the homotopy commutativity of the top square was established in [**9**] and the homotopy commutativity of the side squares was verified in [**8**]. The right side square, resp. left side square, in this diagram provides identifications of $HH^*(\Lambda[V],d_V)$ with $hH^*(P)$, resp. of $CH^*(\Lambda[V],d_V)$ with $cH^*(P)$. These identifications are compatible with all natural transformations defined above and with the endomorphisms induced by the algebraic resp. geometric power maps. In particular one derives Theorem 3 from Theorem 4. It is tedious but straightforward to derive, under the hypothesis of 1– connectivity for P, the identification of the diagram Fig 6 for P and the diagram Fig 7 for $(\Omega^*(P),d^*)$.

References

[1] M.Atiyah and R.Bott The moment map and equivariant cohomology , Topology **23**, (1986), 1-28

[2] D. Burghelea The rational homotopy groups of Diff(M) and Homeo(M) in stability range LNM , Springer -Verlag **763** 1978, 604-626

[3] D. Burghelea The free loop space Part 1– Algebraic topology Contemporary Math. **96** 1989 59-85
[4] D. Burghelea Cyclic homology and k-theory of spaces I. Contemporary Math. **55** 1986 89-115
[5] D. Burghelea A localization theorem for functional S^1-spaces Math. Ann. **282** 1988 513-27
[6] D. Burghelea Free loop spaces, power maps and K-theory Contemporary of math. **1**99 1996 35–58
[7] D. Burghelea and Z.Fiedorowicz Cyclic homology and k-theory of spaces II. Topology **25** 1986 303-317
[8] D. Burghelea, Z.Fiedorowicz and W.Gajda Adams operations in Hochschild and Cyclic Homology of the de Rham Algebra and Free Loop Spaces K-Theory **4** 1991 269-87
[9] D. Burghelea and M.Vigué- Poirrier Cyclic homology of commutative algebras I LNM , Springer Verlag **1318** 1985 51-72
[10] D. Burghelea and M.Vigué- Poirrier A model for cyclic homology and algebraic K-theory of spaces J.Differential Geom. **22** 1985 243-53
[11] M. Chas and D.Sullivan String Topology arxiv; math 9911159 **1**999
[12] T. Goodwillie Cyclic homology, derivations and the free loop space Topology **24** 1985 187-215
[13] P.Deligne, Ph.Griffiths, J.Morgan and D.Sullivan Real Homotopy Theory of Kähler Manifolds Inventiones Math. **29** 1975 245-274
[14] S.Halperin Lectures on minimal models Mem.Soc.Math.France **n**o 9/10 1983
[15] P. Hanlon The action of S_n on the components of the Hodge decomposition of Hochschild homology Michigan Math. J **37** 1990 105-24
[16] J.D.S.Jones Cyclic homology and equivariant cohomology Invent.Math. **87** 1987 403-423
[17] J.D.S.Jones and S.B.Petrack The fixed point theorem in equivariant cohomology preprint Warwick **1**987
[18] D.Lehmann Theorie homotopique des formes differentielles Asterisque, S.M.F **45** 1977
[19] J.L.Loday Cyclic Homology, (book) Grundlehren der mathematischen Wissenchaften, Springer Verlag **301** 1991
[20] M.Vigué-Poirrier and D.Sullivan The homology theory of the closed geodesics problem J.Differential Geom. **11** 1976 653-44

Dept. of Mathematics, The Ohio State University, 231 West Avenue, Columbus, OH 43210, USA.
E-mail address: burghele@mps.ohio-state.edu

Clay Mathematics Proceedings
Volume 11, 2010

Vinberg Algebras, Lie Groups and Combinatorics

Pierre Cartier

Dedicated to Alain Connes on the occasion of his 60th birthday, to witness our long lasting friendship

ABSTRACT. Vinberg algebras are usually called pre-Lie algebras and were introduced long ago by Gerstenhaber. We propose to follow a different route by motivating these algebras by problems coming from differential geometry, and first studied in depth by Vinberg. We shall recall how the Lie bracket of vector fields can be obtained by skewsymmetrizing a more fundamental product. We shall then develop a combinatorial method for the higher order differential operators, quite similar to the procedure used in studying Runge–Kutta methods. We shall then move to nilpotent (or pronilpotent) Lie groups. In the last part of these lectures, I shall apply the previous methods in the renormalization theory of quantum fields (à la Connes–Kreimer).

1. Introduction

Vinberg (or pre-Lie) algebras have become important new tools in combinatorics and differential geometry. They generate a special class of Lie algebras. Our purpose in these notes is to describe them in some detail, and to apply them in the method of renormalization theory introduced by Alain Connes and Dirk Kreimer. These authors have introduced a Hopf algebra, here we consider a simpler algebraic tool, the Vinberg algebras. Such a Vinberg algebra gives rise to a Lie algebra; hence to a Lie group (or rather inverse limit of Lie groups). This provides an alternative route to the results of Connes and Kreimer.

2. Vinberg (pre-Lie) algebras

2.1. The basic concept. Associative algebras and Lie algebras have been with us for a while. Historically, there have been many attempts to define other types of algebras. But the efforts were not systematic, or the right viewpoint was

2010 *Mathematics Subject Classification.* Primary 17A30; Secondary 16T30, 18D50, 81Q30, 81T15.

This paper is based on a short course I gave at Luminy, as part of the meeting "Operads 2005" (4-6 April 2005). I thank Jean-Louis Loday for inviting me to give this course, as well as Jose Gracia-Bondía who wrote a first version of these notes. Thanks also to David Ellwood for his continuing interest.

not reached, and only the theory of Jordan algebras was developed to a reasonable extent. Lie remarked that for commutators $[ab] := ab - ba$ the Jacobi identity,

$$J(a,b,c) := [a[bc]] + [b[ca]] + [c[ab]] = 0,$$

pertaining to the definition of Lie algebras, is implied by associativity. The computation is trivial, but it is useful for us to recall it. For a, b, c elements of any algebra:

$$\begin{aligned} [a[bc]] &= a(bc) - a(cb) - (bc)a + (cb)a, \\ [b[ca]] &= b(ca) - b(ac) - (ca)b + (ac)b, \\ [c[ab]] &= c(ab) - c(ba) - (ab)c + (ba)c. \end{aligned} \tag{1}$$

Define the *associator* $A(a,b,c)$ of three elements a, b, c as

$$A(a,b,c) = a(bc) - (ab)c.$$

An algebra is associative iff $A(a,b,c)$ vanishes identically. Then (1) tells us clearly that

$$J(a,b,c) = \text{total skewsymmetrization of } A(a,b,c).$$

Therefore, for J to vanish, it is not necessary that A vanish in turn; it is enough that

$$A(a,b,c) - A(b,a,c) = a(bc) - (ab)c - b(ac) + (ba)c = 0. \tag{2}$$

We call this the *Vinberg identity*, and algebras with this property will be called *Vinberg algebras*. They were introduced with the name *pre-Lie algebras* by Gerstenhaber in 1962, and around the same epoch by Vinberg in relation with problems in differential geometry. A more general definition of a pre-Lie algebra would be that of an algebra with a product such that the corresponding commutators define a Lie algebra. In the previous definition, a Vinberg algebra is one in which $A(a,b,c)$ is symmetric in a, b; hence a more precise terminology would be *left-symmetric Vinberg algebra*. Similarly, if $A(a,b,c)$ is symmetric in b, c, we get a *right-symmetric Vinberg algebra*. For us, a pre-Lie algebra shall be a right-symmetric Vinberg algebra.

Let A be a (left-symmetric) Vinberg algebra, with product ab. According to the previous explanations, we define a Lie algebra A_{Lie} with bracket

$$[ab] = ab - ba\,.$$

For each a in A, let L_a be the linear operator $b \mapsto ab$ of left multiplication by a in A. Vinberg's identity can be written as

$$L_{[ab]} = L_a\,L_b - L_b\,L_a\,,$$

hence the operators L_a provide a representation of the Lie algebra A_{Lie} in the space A, the so-called *half-adjoint representation*. We have a similar property for the right-symmetric Vinberg algebras, and the operators $R_a : b \mapsto -ab$.

2.2. Examples.

2.2.1. Consider vector fields written in local coordinates $(x^1, \dots, x^n) \equiv (x^\alpha) \equiv x$, with $x \in \mathbb{R}^n$. To a vector function $X^\alpha(x)$ one can associate the Lie derivative $\mathcal{L}_X$, that is, the differential operator defined by

$$\mathcal{L}_X f = \sum_{\alpha=1}^{n} X^\alpha \frac{\partial f}{\partial x^\alpha} =: X^\alpha \, \partial_\alpha f;$$

we of course use Einstein's notation in this Einstein year![1] The little miracle is that

$$[\mathcal{L}_X \mathcal{L}_Y] = \mathcal{L}_X \mathcal{L}_Y - \mathcal{L}_Y \mathcal{L}_X$$

is again a first order differential operator, hence of the form $\mathcal{L}_{[X,Y]}$; and then the Jacobi identity for the Lie bracket $[X, Y]$ comes for free from associativity of the algebra of differential operators.

However, we can do things differently. We define $D_X Y$ by

$$(D_X Y)^\beta := X^\alpha \, \partial_\alpha Y^\beta. \tag{3}$$

This definition is not intrinsic, in the sense of not being consistent under general changes of coordinates. But allow us to go on. Suppose now we define a bracket $[X, Y]$ by

$$[X, Y] := D_X Y - D_Y X \tag{4}$$

and use also notations $X \rightarrow Y$ and $X * Y$ for $D_X Y$. Soon it will be apparent that the grafting notation $X \rightarrow Y$ for the product of X and Y given by $D_X Y$ is very pertinent. Now we check that $X \rightarrow Y$ satisfies a Vinberg identity:

$$\begin{aligned} X \rightarrow (Y \rightarrow Z) - (X \rightarrow Y) \rightarrow Z &= D_X D_Y Z - D_{D_X Y} Z \\ &= X^\alpha \, \partial_\alpha (Y^\beta \, \partial_\beta Z^\bullet) - (D_X Y)^\beta \, \partial_\beta Z^\bullet \\ &= X^\alpha \, \partial_\alpha (Y^\beta \, \partial_\beta Z^\bullet) - X^\alpha \, \partial_\alpha Y^\beta \, \partial_\beta Z^\bullet \\ &= X^\alpha Y^\beta \, \partial_\alpha \partial_\beta Z^\bullet. \end{aligned}$$

Because $\partial_\alpha \partial_\beta Z^\bullet$ is symmetric in α, β, we then see that $A(X, Y, Z) = A(Y, X, Z)$ with the operation $\rightarrow$, and this is all we need to establish the Jacobi identity for the Lie bracket. We shall eventually see (section 2.2) that this calculation can be given an intrinsic meaning, after all.

A calculation similar to the previous one establishes that $\mathcal{L}_X \mathcal{L}_Y f - \mathcal{L}_{D_X Y} f$ is symmetric in X, Y, that is,

$$\mathcal{L}_X \mathcal{L}_Y f - \mathcal{L}_Y \mathcal{L}_X f = \mathcal{L}_{D_X Y} f - \mathcal{L}_{D_Y X} f = \mathcal{L}_{[X,Y]} f$$

for an arbitrary smooth function f. Hence the two definitions of the bracket $[X, Y]$ agree!

2.2.2. Next consider polynomial vector fields; that is, $X(x)$ is a function from $\mathbb{R}^n$ to $\mathbb{R}^n$ such that

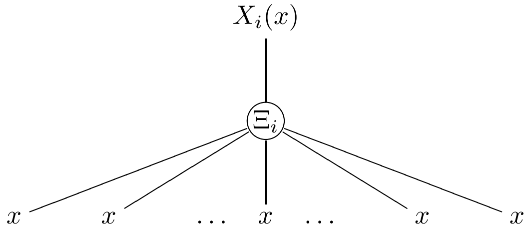
$$X = X_0 + X_1 + X_2 + \cdots$$

where X_i is a (vector-valued) homogeneous polynomial of degree i in n variables. It is well known that any of these is of the form

$$X_i(x) = \Xi_i(x, \dots, x),$$

[1] This was written in 2005!

for $\Xi_i(y_1, \ldots, y_i)$ a uniquely defined *symmetric* multilinear function. We represent the last identity by means of a graph:

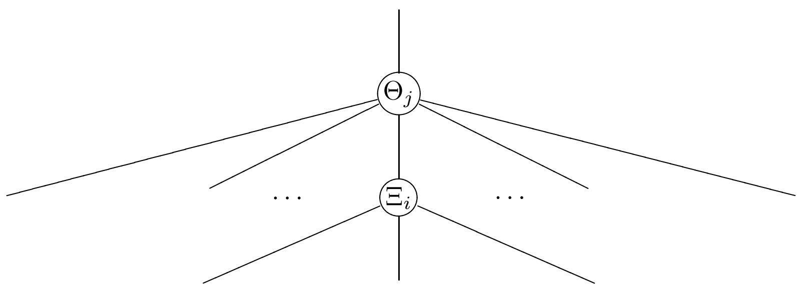


Here we have a rooted tree with the root on top and with unordered leaves.

Now we reconsider $D_X Y$ with X homogeneous of degree i and Y homogeneous of degree j. Let the symmetric function Θ_j with j entries correspond to Y as Ξ_i to X above. Then $Z = D_X Y$ has degree $i + j - 1$. Precisely, the Leibniz rule says that Z is obtained by considering the substitution for $\Xi_i(x, \ldots, x)$ of each variable argument in the symmetric function Θ_j, and summing over all the terms obtained. We define, for r between 1 and j,

(5)

$$\Theta_j \mid_r \Xi_i(x_1, \ldots, x_{i+j-1}) = \Theta_j(x_1, \ldots, x_{r-1}, \Xi_i(x_r, \ldots, x_{i+r-1}), x_{i+r}, \ldots, x_{i+j-1}),$$



and the sum

$$\Theta_j \circ \Xi_i = \sum_{r=1}^{j} \Theta_j \mid_r \Xi_i, \tag{6}$$

which is a symmetric function in $i+j-1$ entries (also written as $\Xi_i \rightarrow \Theta_j$). Thereby Z is given as a sum of insertions or graftings:

$$Z(x) = \Theta_j \circ \Xi_i(x, \ldots, x).$$

We may look at the Vinberg property in the light of this graphical representation. Consider in turn $\Xi \rightarrow (\Theta \rightarrow \Lambda)$ in relation with $(\Xi \rightarrow \Theta) \rightarrow \Lambda$: when grafting Ξ on $\Theta \rightarrow \Lambda$, we can choose to do it on a Θ-part or on a Λ-part. Now, the insertions on a Θ-part are totally cancelled in $\Xi \rightarrow (\Theta \rightarrow \Lambda) - (\Xi \rightarrow \Theta) \rightarrow \Lambda$, and there remain the insertions on Λ-parts. But the latter are *symmetric* in Ξ, Θ: here $\partial_\alpha \partial_\beta = \partial_\beta \partial_\alpha$ is the rule of symmetry of the insertions! Thus the Vinberg property holds.

2.2.3. We look now at Poisson brackets, in two variables for simplicity of notation:

$$\{f, g\} = \partial_p f \, \partial_q g - \partial_p g \, \partial_q f.$$

Suppose we define a 'star' product by

$$f \star g := \partial_p f \, \partial_q g,$$

so

$$\{f, g\} = [f \star g].$$

Then

$$(f \star g) \star h = \partial_p(f \star g)\, \partial_q h = \partial^2_{pp} f\, \partial_q g\, \partial_q h + \partial_p f\, \partial^2_{pq} g\, \partial_q h$$

has two terms. The Jacobi identity certainly holds; it contains 24 summands. But $\star$ is a counterexample for the Vinberg property, since the six expressions obtained by permutation of f, g, h in $A(f, g, h)$ are distinct in general.

2.2.4. We now come to Lie groups. Let e be the neutral element of one of such G. Consider local coordinates $x \equiv (x^1, \ldots, x^n)$ on G with the property $x^i(e) = 0$. In general in those coordinates, the group product rule will have the form

$$z = F(x; y) = \sum_{p \geq 0,\, q \geq 0} F_{p,q}(x; y) \quad \text{if} \quad z = x \cdot y. \tag{7}$$

We trust that the reader will be able to distinguish when we refer to abstract elements of the group and when to their coordinates, in our notation.

Here $F_{p,q}(x^1, \ldots, x^n; y^1, \ldots, y^n)$ is a polynomial in $2n$ variables, homogeneous of degree p in x, q in y; moreover $F_{0,0} = 0$ and $F_{p,0}(x; y) = x$, $F_{0,p}(x; y) = y$ for any $p \geq 1$. Let $\mathfrak{g}$ denote the tangent (Lie) algebra of G. Consider

$$(x, y) := xyx^{-1}y^{-1};$$

then, in coordinates,

$$(x, y) = G(x; y) = \sum_{p \geq 0,\, q \geq 0} G_{p,q}(x; y), \tag{8}$$

where $G_{0,0} = G_{1,0} = G_{0,1} = 0$ and $G_{1,1}(x; y) = F_{1,1}(x; y) - F_{1,1}(y; x)$ is a bilinear function. If we make the identification with tangent vectors at the identity, then $G_{1,1}$ defines the bracket in $\mathfrak{g}$, hence must satisfy the Jacobi identity. But we want a calculational explanation for this fact. The foregoing is obviously related to Poincaré's bilinear groups. Given an (associative, finite-dimensional, unital) algebra A, its units form a group $A^\times$ with $e = 1_A$. For x, y small enough, $1 + x$ and $1 + y$ lie in $A^\times$ with $(1 + x)(1 + y) = 1 + (x + y + xy)$ (hence the name "bilinear"), and the commutator $(1 + x, 1 + y)$ coincides, up to third order terms, and shifting the origin from 1 to 0, with

$$(1 + x)(1 + y) - (1 + y)(1 + x) = xy - yx.$$

Hence, the Lie algebra corresponding to the Lie group $A^\times$ is the vector space A endowed with the bracket $[ab] = ab - ba$.

More generally, write $x * y$ for $F_{1,1}(x; y)$ and using polarization expand the product $x \cdot y$ as follows:

$$F(x; y) = x + y + x * y + L(x, x, y) + M(x, y, y) + O_4(x, y),$$

where L and M are trilinear, satisfying

$$L(x, y, z) = L(y, x, z)$$

$$M(x, y, z) = M(x, z, y),$$

and O_4 contains no term of total degree < 4. By associativity of the group law, we obtain

$$F(x; F(y; z)) - F(F(x; y); z) = 0.$$

Keeping only the terms trilinear in x, y, z, we obtain the identity

$$A(x,y,z) = 2L(x,y,z) - 2M(x,y,z)$$

for the associator $A(x,y,z) = x*(y*z) - (x*y)*z$. From the symmetry properties of $L(x,y,z)$ and $M(x,y,z)$, it follows that the skew-symmetrization of $A(x,y,z)$ is 0, hence the bracket $[xy] = x*y - y*x = G_{1,1}(x;y)$ on $\mathbb{R}^d$ satisfies Jacobi identity.

If $F(x;y) - x$ is linear in y, then $M = 0$, hence the operation $x*y$ satisfies the left-symmetric Vinberg identity: the Lie algebra of G comes from a left-symmetric Vinberg algebra. Similarly, if $F(x;y) - y$ is linear in x, then the product $x*y$ defines on $\mathbb{R}^d$ a right-symmetric Vinberg algebra. More on this topic in section 2.3.

2.3. The Gerstenhaber approach and noncommutative polynomials. Gerstenhaber arrived at the concept of *pre-Lie algebra* when working on Hochschild cohomology. Let A be an associative algebra, and consider cochains:

$$c_p : A^{\otimes p} \to A; \qquad d_q : A^{\otimes q} \to A.$$

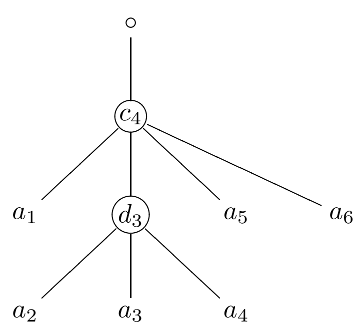
Then, for $i = 1, \ldots, p$, we form the cochain with $p+q-1$ arguments given by:

(9)
$$c_p \mid_i d_q(a_1, \ldots, a_{p+q-1}) = c_p(a_1, \ldots, a_{i-1}, d_q(a_i, \ldots, a_{i+q-1}), a_{i+q}, \ldots, a_{p+q-1});$$

and we set

$$c_p \circ d_q := \sum_{i=1}^{p} c_p \mid_i d_q. \tag{10}$$

Here, for instance, is $c_4 \mid_2 d_3$:



The small miracle is that $[c_p\, d_q] := c_p \circ d_q - d_q \circ c_p$ satisfies the Jacobi identity! (This, by the way, is nowadays known as the Gerstenhaber bracket, and plays an important role for instance in the theory of deformations.) We are going to see that the Jacobi identity holds because $c_p \circ d_q$ happens to satisfy the right-symmetric Vinberg property. Indeed, we can represent graphically the cochains as we did with the symmetric functions in example 1.2.2; now d_q is grafted on c_p in all possible ways, and everything works the same, with the only difference that the leaves now have a natural order. But this does not affect the conclusion that $c_p \circ (d_q \circ e_r) - (c_p \circ d_q) \circ e_r$ is symmetric in d_q, e_r. Hence the Jacobi identity holds.

We have moreover gleaned an interpretation of the cochains as noncommutative vector fields.

3. Some good reasons to study Vinberg algebras

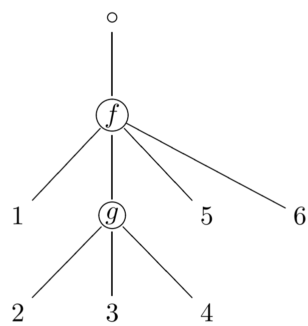
3.1. Operads. Let V be a vector space over a suitable field k of characteristic zero, and let us denote $E_V(1) := \operatorname{End} V$ and, for $n \geq 2$,

$$E_V(n) := \operatorname{Hom}(V^{\otimes n}, V); \qquad E_V = \bigoplus_{n\geq 1} E_V(n).$$

Consider then $f_n \in E_V(n)$, $g_p \in E_V(p)$ and, somewhat similarly to the previous section, define, for $i = 1, \ldots, n$, an element of $E_V(n + p - 1)$ by

$$f_n \mid_i g_p = f_n \circ (\mathrm{id}_V^{\otimes i-1} \otimes g_p \otimes \mathrm{id}_V^{\otimes n-i}). \tag{11}$$

This we represent by trees with half-edges (those branches that do not connect blobs) and edges (those that do connect blobs to effect the insertions). For example,



To each rooted planar tree does correspond an operation of this kind. In these operations there are numerous compatibility conditions of the associativity type in that

$$(f_m \mid_i g_n) \mid_j h_p = f_m \mid_i (g_n \mid_{j-i+1} h_p) \quad \text{when} \quad \begin{cases} 1 \leq i \leq m, \\ i \leq j \leq i + n - 1; \end{cases} \tag{12}$$

and of the commutativity type:

$$(f_m \mid_i g_n) \mid_j h_p = (f_m \mid_{j-n+1} h_p) \mid_i g_n \quad \text{when} \quad \begin{cases} 1 \leq i \leq m, \\ i + n \leq j \leq n + m - 1 \end{cases} \tag{13}$$

which is a kind of locality principle in the insertions. This construction gives rise to an *operad* P, that we regard as a collection of vector spaces $\mathsf{P}(n)$ indexed by the positive integers, together with bilinear maps:

$$\mathsf{P}(n) \otimes \mathsf{P}(p) \to \mathsf{P}(n + p - 1) : f \otimes g \mapsto f \mid_i g, \quad \text{for each} \quad i = 1, \ldots, n$$

satisfying the above-mentioned properties. In the standard definition, the $\mathsf{P}(n)$ are $k[S_n]$-modules, but we shall not employ that yet. In the example there are moreover maps

$$\xi_n : E_V(n) \otimes V^{\otimes n} \to V,$$

with obvious associativity and unity properties. In general, given the operad P, a P-algebra is a vector space A with maps

$$\mathsf{P}(n) \otimes A^{\otimes n} \to A$$

with the analogous properties. This gives rise to an 'operadic map' $\mathsf{P} \to E_A$.

What is the relation to Vinberg algebras? Take $f \circ g = \sum_{i=1}^{n} f \mid_i g$ and define θ_g by $\theta_g(f) = f \circ g$, for $f, g \in \mathsf{P}$. Then, although $\circ$ is not associative, $\{\theta_g\}_{g \in E}$ is a Lie algebra of operators acting on P, since the bracket $[hg] = g \circ h - h \circ g$ satisfies

$$[\theta_h \; \theta_g] := \theta_{[hg]} \tag{14}$$

by a reasoning like that of the previous section. This is the "half-adjoint representation" of Gerstenhaber. Identity (14) means that $\mathsf{P} = \bigoplus_{n>0} \mathsf{P}(n)$ is a right-symmetric Vinberg algebra.

In general, given a category of algebras, we can associate to it an operad describing the "natural" operations that can be defined on it; if, for instance, we consider the category of associative algebras, all the trees of the same size in the construction above will determine the same operation in the corresponding operad. Thus there are different classes of operads, according to the basic properties of defining operations. Some are represented in the following table.

P	Operations	Relations
As	xy	$(xy)z = x(yz)$
Com	$xy = yx$	$(xy)z = x(yz)$
Lie	$[xy] = -[yx]$	Jacobi
Mag	xy	no relation
Vinb	xy	$x(yz) - (xy)z = y(xz) - (yx)z$
Zinb	xy	$(xy)z = x(yz) + x(zy)$
2-as	$x \cdot y$, $x * y$	both associative
Dend	$x \prec y$, $x \succ y$	
	$x * y = x \prec y + y \succ x$	(see below)

Dialgebras, that is, vector spaces with two multiplications, can be considered as well. A particularly interesting case of dialgebras are the dendriform dialgebras of Loday. They can be obtained as follows. Let $(D, *)$ be an associative algebra. Assume that D is a bimodule over itself, $D \equiv {}_D D_D$. We write $\succ$ and $\prec$ for the left and right actions, respectively. Assume moreover that, for all $a, b \in D$,

$$a * b = a \prec b + a \succ b. \tag{15}$$

Then by definition D is a dendriform dialgebra. In detail, the dendriform properties are

$$\begin{aligned}
&a \prec (b * c) = a \prec (b \succ c + b \prec c) = (a \prec b) \prec c;\\
&(a * b) \succ c = (a \succ b + a \prec b) \succ c = a \succ (b \succ c);\\
&(a \succ b) \prec c = a \succ (b \prec c).
\end{aligned}$$

Conversely, these last relations (on the right hand sides) are enough to establish associativity of $*$ defined by the equality in (15). Without changing the underlying linear structure, D gives rise not only to an associative algebra, but also to a Vinberg algebra and, in two different ways, to the same Lie algebra. For that, consider the following:

$$x \rightarrow y := x \succ y - y \prec x. \tag{16}$$

This defines a Vinberg algebra structure. Moreover,

$$x \to y - y \to x = x \succ y - y \prec x - y \succ x + x \prec y = x * y - y * x,$$

so the corresponding Lie algebra structures coincide. We have then the following quadrilateral of functors between categories of algebras, with the same underlying vector structure:

$$\begin{tikzcd} \mathsf{Dend\text{-}alg} \arrow[r, "\to"] \arrow[d, "*"'] & \mathsf{Vinb\text{-}alg} \arrow[d, "{[\cdot\,\cdot]}"] \\ \mathsf{As\text{-}alg} \arrow[r, "{[\cdot\,\cdot]}"] & \mathsf{Lie\text{-}alg} \end{tikzcd}$$

To conclude this part, we remark on the affinity between the notion of operad and the functor of restriction of scalars. Let $\mathcal{A}$ and $\mathcal{B}$ be rings with unit, and let $\phi : \mathcal{A} \to \mathcal{B}$ be a unital ring homomorphism. Then any (say) right $\mathcal{B}$-module $\mathcal{F}$ becomes a right $\mathcal{A}$-module by defining

$$t \cdot a := t\,\phi(a) \quad \text{for} \quad t \in \mathcal{F},\ a \in \mathcal{A}.$$

If $\mathcal{G}$ is another right $\mathcal{B}$-module and $\psi \in \mathrm{Hom}_{\mathcal{B}}(\mathcal{F}, \mathcal{G})$, then $\psi(t \cdot a) = \psi(t\phi(a)) = \psi(t)\,\phi(a) = \psi(t) \cdot a$, so ψ can also be regarded as a member of $\mathrm{Hom}_{\mathcal{A}}(\mathcal{F}, \mathcal{G})$. In this way, ϕ defines a functor R_ϕ from the category of right $\mathcal{B}$-modules to the category of right $\mathcal{A}$-modules. This functor is called *restriction of scalars* in regard of the case in which ϕ is the inclusion map of a subring $\mathcal{A}$ into a larger ring $\mathcal{B}$, for obvious reasons.

Let us interpret in a similar way the previous functors between categories of algebras, for instance going from associative algebras to Lie algebras. For the operads As, the n-th term $\mathsf{As}(n)$ has a basis consisting of the operations $(x_1, \dots, x_n) \mapsto x_{\sigma(1)} \dots x_{\sigma(n)}$ for the $n!$ permutations σ of $\{1, \dots, n\}$. Similarly, $\mathsf{Lie}(n)$ is the multilinear part in the free Lie algebra with generators $x_1, \dots, x_n$. There is a map ϕ_n from $\mathsf{Lie}(n)$ to $\mathsf{As}(n)$ interpreting iterated brackets by means of the rule $[ab] = ab - ba$. For instance, $\mathsf{Lie}(3)$ has a basis $[x_1[x_2x_3]]$, $[x_2[x_1x_3]]$ and $[x_1[x_2x_3]]$ for instance is mapped by ϕ_3 to $x_1x_2x_3 - x_1x_3x_2 - x_2x_3x_1 + x_3x_2x_1$. An associative algebra is a vector space on which the elements of $\mathsf{As} = \bigoplus_{n\geq 1} \mathsf{As}(n)$ give operations. By means of the operadic map $\phi = \bigoplus_{n\geq 1} \phi_n$ from Lie to As, we interpret the natural operations in Lie as natural operations for associative algebras, hence the previous functor from associative algebras to Lie algebras.

3.2. More on Vinberg algebras in differential geometry. Let $M = M^d$ denote a manifold with $\dim M = d$, and let TM be its tangent bundle. We define a *linear connection* on M as an $\mathbb{R}$-bilinear operation ∇ that, given two vector fields $X, Y \in \Gamma(M, TM)$, produces a new vector field $\nabla_X Y$ with the properties

$$\nabla_{fX} Y = f\nabla_X Y \quad (C^\infty\text{-linearity in } X); \tag{17}$$

$$\nabla_X(fY) = (\mathcal{L}_X f)Y + f\nabla_X Y \quad \text{(Leibniz rule)}. \tag{18}$$

Here $\mathcal{L}_X$ denotes the Lie derivative with respect to the vector field X. Now define the *torsion* T and *curvature* R of ∇ respectively by

$$T(X,Y) := \nabla_X Y - \nabla_Y X - [X,Y]; \tag{19}$$

$$R(X,Y)Z := \big(\nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]}\big)Z, \tag{20}$$

with $Z \in \Gamma(M, TM)$, too. One checks without too much difficulty C^∞-bilinearity of T,

$$T(fX, Y) = T(X, fY) = fT(X, Y). \tag{21}$$

This means that the map $T : \Gamma(M, TM)^{\otimes 2} \to \Gamma(M, TM)$ descends to a map $T_xM \otimes T_xM \to T_xM$, for all $x \in M$. Similarly for $R(\cdot, \cdot)Z$. This allows the definition of the torsion and curvature as tensors, familiar from Riemannian geometry. Suppose now $R = 0$, $T = 0$. That is:

$$[X, Y] := \nabla_X Y - \nabla_Y X;$$

and thus

$$\nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z = \nabla_{[X,Y]} Z = \nabla_{\nabla_X Y - \nabla_Y X} Z.$$

We can see the Vinberg relation here; in fact this is how he came by it, when dealing with simply transitive affine actions of groups on $\mathbb{R}^d$.

The *covariant derivative map* $\nabla X : Y \mapsto \nabla_Y X$ is a bundle map from TM to TM. In general we envisage $\mathrm{Hom}(TM^{\otimes p}, TM^{\otimes q}) =: \mathcal{T}_{p,q}$, where the maps considered are fibrewise linear. The sections of these bundles are called *tensor fields*. We have $\mathcal{T}_{0,1} \simeq TM$; $\mathcal{T}_{1,1} \simeq \mathrm{End}\, TM$, and so on. We can look at ∇ as a map

$$\nabla : \Gamma(M, \mathcal{T}_{0,1}) \ni X \mapsto \nabla X \in \Gamma(M, \mathcal{T}_{1,1}),$$

with the property $\nabla(fX) = df \otimes X + f\nabla X$. This map can be extended to all tensor fields:

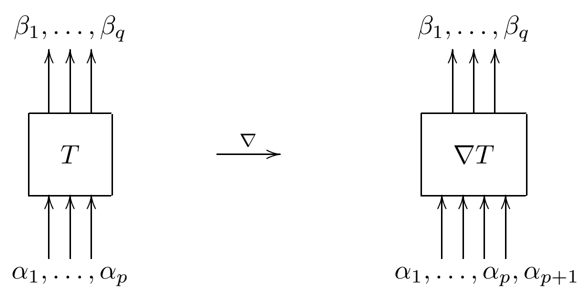
$$\nabla : \Gamma(M, \mathcal{T}_{p,q}) \ni T \mapsto \nabla T \in \Gamma(M, \mathcal{T}_{p+1,q}),$$

by

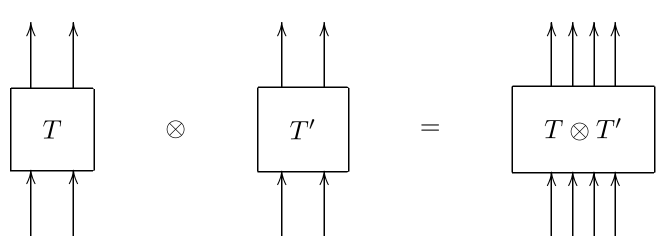
$$\nabla(T \otimes T') = \nabla T \otimes T' + T \otimes \nabla T'. \tag{22}$$

We have then $\nabla f = df$, for f a section of $\mathcal{T}_{0,0}$.

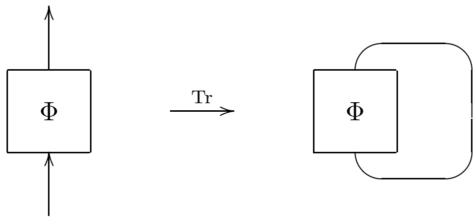
In the index notation, we go from $T^{\beta_1,\ldots,\beta_q}_{\alpha_1,\ldots,\alpha_p}$ to $T^{\beta_1,\ldots,\beta_q}_{\alpha_1,\ldots,\alpha_p;\alpha_{p+1}}$. We can think of T as a 'box' that takes the 'input' $(\alpha_1, \ldots, \alpha_p)$ and transforms it into the 'output' $(\beta_1, \ldots, \beta_q)$. See this in the next figures, as the action of ∇:



and the tensor product as a 'glueing' or juxtaposing operation:



Another operation is *contraction*: given $\Phi \in \operatorname{End} TM$, we can consider $\operatorname{Tr}\Phi = \sum \Phi^\alpha_\alpha$. This is represented by 'sewing':



The graphical representation makes the point, often underlined by Penrose, that this apparently coordinate-dependent definition, and the tensor index notation in general, are actually intrinsic.

Let us now assume that on M we have given a connection with vanishing torsion and curvature (so the manifold is "locally flat"). Then there exist local coordinates $(x^1, \dots, x^d)$ with $\nabla^2 x^\alpha = 0$, such that a smooth function f on M, with vanishing $\nabla^2 f$ in $\mathcal{T}_{2,0}$, is affine:

$$f = c_0 + c_1 x^1 + \cdots + c_d x^d$$

with real constants $c_0, c_1, \dots, c_d$. Moreover, $\nabla_X Y = D_X Y$ in those coordinates; and so we come to see that

$$[X, Y] := D_X Y - D_Y X$$

is a special form of an intrinsic object! Under the present hypothesis, we can further consider the maps

$$\nabla^m X : \operatorname{Sym}^m TM \to TM,$$

where $\nabla^m X$ is of the form $t^\beta_{\alpha_1,\dots,\alpha_m}$, with symmetry in the indices; symmetry comes from torsion-freeness, of course. If we put

$$\{X \mid Y_1, \dots, Y_m\} := \nabla^m X(Y_1, \dots, Y_m), \tag{23}$$

we find the properties

(24)

$$\{X \mid Y_1, \dots, Y_{m-1}, Y_m\} = \nabla_{Y_m}\{X \mid Y_1, \dots, Y_{m-1}\} - \sum_{i=1}^{m-1}\{X \mid Y_1, \dots, \nabla_{Y_m} Y_i, \dots, Y_{m-1}\},$$

characteristic of the 'brace' operation in Vinberg algebras.

3.3. Lie groups with affine coordinates. The developments in the previous section enable us to recast the results of the end of section 1.2.4 in a more invariant way. Let us start from a Lie algebra $\mathfrak{g}$ and a Lie group G with Lie algebra $\mathfrak{g}$. We identify $\mathfrak{g}$ with the space $T_e G$ of vectors tangent to the group manifold G at the unit element e of G. Changing slightly the notation, *a pre-Lie algebra structure on* $\mathfrak{g}$*, compatible with the given Lie bracket, is a bilinear operation* $(X, Y) \mapsto \{X \mid Y\}$ *in* $\mathfrak{g}$ *satisfying the following rules:*

$$\{X \mid Y\} - \{Y \mid X\} = [X, Y]\,, \tag{25}$$

$$\{X \mid Y_1, Y_2\} = \{X \mid Y_2, Y_1\}\,. \tag{26}$$

The definition of $\{X \mid Y_1, Y_2\}$ is the special case $m = 2$ in the inductive definition of the *braces*

(27)
$$\{X \mid Y_1, \ldots, Y_m\} = \{\{X \mid Y_1, \ldots, Y_{m-1}\} \mid Y_m\} - \sum_{i=1}^{m-1} \{X \mid Y_1, \ldots, \{Y_i \mid Y_n\}, \ldots, Y_{m-1}\}$$

similar to formula (24).

A vector field X on the Lie group G is called *right-invariant* if it is invariant under the right translations $g \mapsto gg_0$ acting on the manifold G. There is a Similar definition for tensor fields or connections. For any X in $\mathfrak{g}$, there exists a unique right-invariant vector field, whose value at e is X, to be denoted X^r. If $(X_1, \ldots, X_n)$ is a basis of $\mathfrak{g}$ (over the real field $\mathbb{R}$), then every vector field X on G can be written as

$$X = f_1 X_1^r + \cdots + f_n X_n^r \,, \tag{28}$$

with a unique set of smooth functions $f_1, \ldots, f_n$ on G.

There exists a unique connection ∇^r on G such that the right-invariant vector fields are characterized by the equation $\nabla^r X = 0$. More explicitly, if X is given as in (28), one has

$$\nabla_Y^r X = \mathcal{L}_Y f_1 \cdot X_1^r + \cdots + \mathcal{L}_Y f_n \cdot X_n^r \tag{29}$$

or, more compactly $\nabla^r X = \sum\limits_{i=1}^{n} df_i \otimes X_i^r$. An arbitrary right-invariant connection ∇ on G is of the form

$$\nabla_Y X = \nabla_Y^r X + A(X, Y) \,, \tag{30}$$

where A is a right-invariant tensor field of type $\mathcal{T}_{2,1}$. Since A is right-invariant, there exists a bilinear operation $(X, Y) \mapsto \{X \mid Y\}$ on $\mathfrak{g}$ such that

$$A(X^r, Y^r) = \{X \mid Y\}^r \,, \tag{31}$$

hence

$$\nabla_{Y^r} X^r = \{X \mid Y\}^r \tag{32}$$

for X, Y in $\mathfrak{g}$. According to the well-known formula

$$[X^r, Y^r] = -[X, Y]^r \,, \tag{33}$$

the torsion T and the curvature R of the connection ∇ are right-invariant tensor fields, whose values at the unit e of G are given by

$$T_e(X, Y) = [X, Y] - \{X \mid Y\} + \{Y \mid X\} \tag{34}$$

$$R_e(X, Y) \cdot Z = \{Z \mid Y, X\} - \{Z \mid X, Y\} + \{Z \mid T_e(X, Y)\} \tag{35}$$

for X, Y, Z in $\mathfrak{g}$. Comparing these formulas with (25) and (26) we can conclude:

The formula (32) *establishes a bijection between the right-invariant connections on G, with vanishing torsion and curvature, and the pre-Lie algebra operations on $\mathfrak{g}$, generating via* (25) *the given Lie bracket $\mathfrak{g}$.*

We finish by describing the development map $c : G \to \mathfrak{g}$ associated to a connection ∇ as before. Assume that G is *simply connected.* Define the *sheaf* $\mathcal{F}$ on G such that, for any open set U in G, $\mathcal{F}(U)$ consists of the smooth functions f on U

such that $\nabla^2 f = 0$. More explicitly, using local coordinates x^α, with corresponding partial derivatives $\partial_\alpha = \partial/\partial x^\alpha$, and the coefficients of the connection ∇ given by

$$\nabla_\alpha \, \partial_\beta = \Gamma^\gamma_{\alpha\beta} \, \partial_\gamma \,, \tag{36}$$

the equation $\nabla^2 f = 0$ reads $\partial_\alpha \, \partial_\beta f = \Gamma^\gamma_{\alpha\beta} \, \partial_\gamma f$. Each point g_0 of G belongs to the domain U_0 of a system of local coordinates x^α, such that $\nabla^2 x^\alpha = 0$. For every connected open set U contained in U_0, $\mathcal{F}(U)$ consists of the functions of the form $c_0 + \sum_{\alpha=1}^{n} c_\alpha x^\alpha$, with constants $c_0, c_1, \ldots, c_n$. That is, the sheaf $\mathcal{F}$ is locally constant. Since G is assumed to be simply connected, the sheaf $\mathcal{F}$ is globally constant. It means that taking $g_0 = e$, there exist globally defined smooth functions $x^1, \ldots, x^n$ on G, vanishing at e, satisfying the equations $\nabla^2 x^\alpha = 0$, which induce a coordinate system around e.

We reformulate this as follows: *there exists a smooth map $c : G \to \mathfrak{g}$, such that $c(e) = 0$, $\nabla^2 c = 0$, and the derivative $d_e \, c$ at the origin is the identity map of $T_e \, G$ onto $\mathfrak{g}$.* We call c the *development map.*

Using the development map, we can describe the group law in G by the explicit formula

$$c(x \cdot y) = c(y) + \sum_{m \geq 0} \frac{1}{m!} \, \{c(x) \mid \underbrace{c(y), \ldots, c(y)}_{m}\} \tag{37}$$

involving the braces, and valid in a neighbourhood of e in G. Otherwise stated, the development map c defines a local chart $c : U \to \mathfrak{g}$ of G around e, such that the product in this chart is given by

$$X \cdot Y = Y + \sum_{m \geq 0} \frac{1}{m!} \, \{X \mid Y^m\} \,. \tag{38}$$

We can introduce a norm in $\mathfrak{g}$ such that

$$\|\{X \mid Y\}\| \leq \|X\| \cdot \|Y\| \,. \tag{39}$$

An easy induction gives the estimate

$$\|\{X \mid Y_1, \ldots, Y_m\}\| \leq m! \|X\| \cdot \|Y_1\| \ldots \|Y_m\| \,; \tag{40}$$

hence the series (38) converges for $\|Y\| < 1$ and all X.

Remarks. 1) In a Lie group like G, there is a locally defined *logarithm* $\log : U \to \mathfrak{g}$ and the multiplication in G is described by the *Campbell-Baker-Hausdorff formula*

$$\log(x \cdot y) = \mathrm{CH}(\log x, \log y) \,; \tag{41}$$

the right-side can be calculated from the Lie bracket in $\mathfrak{g}$, in a very complicated way. Formula (38) is much simpler.

2) Comparing formulas (7) and (38), we may identify $F_{1,1}(x; y)$ with $\{x \mid y\}$ and hence the commutator

$$G_{1,1}(x; y) = F_{1,1}(x; y) - F_{1,1}(y; x)$$

with $\{x \mid y\} - \{y \mid x\}$. This is in agreement with formula (25).

4. The Connes–Kreimer paradigm

4.1. Graded Vinberg algebras.

4.1.1. *General theory.* In view of the applications, we work in this section with algebras over the field $\mathbb{C}$ of complex numbers. The case of the field $\mathbb{R}$ of real numbers requires very small changes. By replacing Lie groups with algebraic groups, one could cover the case of an arbitrary ground field of characteristic 0.

We consider an algebra A which can be a Lie algebra, or a Vinberg algebra (left- or right-symmetric). We say A is *graded* if it is a direct sum $A = \bigoplus_{n\geq 1} A_n$, where each subspace A_n is finite-dimensional, and the multiplication satisfies the rule:

(G) $\qquad a_p \cdot a_q$ *belongs to* A_{p+q} *for* a_p *in* A_p, a_q *in* A_q.

The subspace $\bigoplus_{n>p} A_n$ is an ideal in A, and the quotient algebra $A(p) := A/\bigoplus_{n>p} A_n$ (isomorphic to $\bigoplus_{1\leq n\leq p} A_n$) is finite-dimensional. We call $A(p)$ the *truncation* of A.

We assume now that A is a pre-Lie algebra, with product $\{X \mid Y\}$, defining the Lie bracket

$$[X, Y] = \{X \mid Y\} - \{Y \mid X\}.$$

The corresponding Lie algebra is graded and denoted by $\mathfrak{g}$. To the truncated pre-Lie algebra $A(p)$ corresponds the truncated Lie algebra $\mathfrak{g}(p)$. It follows from (G) that any iterated Lie bracket in $\mathfrak{g}(p)$ with $n > p$ factors vanishes, hence $\mathfrak{g}(p)$ is a finite-dimensional nilpotent Lie algebra. By well-known results, if $G(p)$ is the simply connected complex Lie group with Lie algebra $\mathfrak{g}(p)$, the exponential map is an isomorphism of complex manifolds of $\mathfrak{g}(p)$ onto $G(p)$, hence there exists a globally defined logarithm map $\log : G(p) \to \mathfrak{g}(p)$.

Introducing as before the braces $\{X \mid Y_1, \ldots, Y_m\}$ in the pre-Lie algebra A, hence in its truncations $A(p)$, it follows from (G) that $\{X \mid Y_1, \ldots, Y_m\}$ is of degree $d + e_1 + \cdots + e_m$ if $X \in A_d$, $Y_i \in A_{e_i}$; since $d \geq 1$, $e_i \geq 1$, one obtains $d+e_1+\cdots+e_m \geq m+1$, hence the braces $\{X \mid Y_1, \ldots, Y_m\}$ vanish in the truncated pre-Lie algebra $A(p)$ for any m with $m > p - 1$. According to formula (38), we define a (non-linear) multiplication in $A(p)$ by the rule

$$X \cdot Y = Y + \sum_{m=0}^{p-1} \{X \mid Y^m\}/m!\,. \tag{42}$$

It can be shown, using the properties of braces, that this is a group law on $A(p) = \mathfrak{g}(p)$. Hence $A(p)$ is a simply connected complex Lie group with Lie algebra $\mathfrak{g}(p)$. From the uniqueness of simply connected Lie groups, it follows that *the development map $c : G(p) \to \mathfrak{g}(p)$ is an isomorphism of $G(p)$ with the group defined by the multiplication* (42) *in* $\mathfrak{g}(p)$.

We can go to the limit. The truncated pre-Lie algebras $A(p)$, the corresponding truncated Lie algebras $\mathfrak{g}(p)$ as well as the groups $G(p)$ form inverse systems. The inverse limit $\hat{A} = \varprojlim A(p)$ is the direct product $\prod_{m\geq 1} A_m$, whose elements can be represented as series $a_1 + a_2 + \cdots$ with a_i in A_i. The product in $\hat{G} = \varprojlim \hat{G}(p)$ is given by $X \cdot Y = Y + \sum_{m\geq 0} \{X \mid Y^m\}/m!$: in this series, any homogeneous component is obtained using finitely many operations.

4.1.2. *An example.* Suppose P is an operad with $\mathsf{P}(1) = 0$ and each $\mathsf{P}(n)$ of finite dimension. In section 2.1, we introduced a composition $f \circ g = \sum_{i=1}^{n} f \mid_i g$, for f in $\mathsf{P}(n)$, g in $\mathsf{P}(p)$, with values in $\mathsf{P}(n+p-1)$. Shifting degrees, we define $A(n) = \mathsf{P}(n+1)$ for $n \geq 1$ and the properties of the product $f \circ g$ show that $A = \bigoplus_{n \geq 1} A_n$ is a graded pre-Lie algebra. We can, as before, associate to it a Lie algebra, and an inverse limit of complex nilpotent Lie groups.

Take for instance the operad Com corresponding to the category of commutative and associative algebras. For each $n \geq 1$, the space $A(n) = \mathsf{Com}(n+1)$ is one-dimensional, generated by an element z_n corresponding to the product $x_1 \ldots x_{n+1}$ in a commutative algebra. The composition is given for $n \geq 1$, $p \geq 1$ by

$$z_n \mid_i z_p = z_{n+p} \quad (1 \leq i \leq n+1)\,, \tag{43}$$

hence the pre-Lie bracket

$$\{z_n \mid z_p\} = (n+1)\, z_{n+p} \tag{44}$$

and the Lie bracket

$$[z_n, z_p] = (n-p)\, z_{n+p} \tag{45}$$

(Witt algebra).

The elements in $\hat{A} = \prod_{n \geq 1} A_n$ can be identified with complex formal power series of the form $\sum_{n \geq 1} c_n\, t^{n+1} = c(t)$ with the brace

$$\{c(t) \mid d(t)\} = c'(t)\, d(t) \tag{46}$$

(where $c'(t)$ is the derivative of $c(t)$). The corresponding Lie bracket is given by

$$[c(t), d(t)] = c'(t)\, d(t) - d'(t)\, c(t)\,. \tag{47}$$

The map $c(t) \mapsto -c(t)\frac{d}{dt}$ gives an isomorphism of this Lie algebra with the Lie algebra of formal vector fields $\sum_{n \geq 1} c_n\, t^{n+1} \frac{d}{dt}$, and z_n corresponds to $-t^{n+1}\frac{d}{dt}$. For the group $\hat{G}$, it is the group of formal power series

$$f(t) = t + u_1\, t^2 + u_2\, t^3 + \cdots \tag{48}$$

with composition

$$(f \circ g)(t) = f(g(t))\,. \tag{49}$$

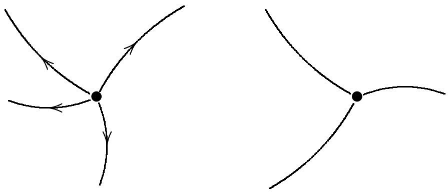
The explicit form of this multiplication rests on the Faà di Bruno formula – which preserves its validity for analytic functions on the appropriate domains [**1**].

Exercise. Show that the braces are given by

$$\{c \mid d_1, \ldots, d_m\} = c^{(m)} d_1 \cdots d_m$$

where $c^{(m)}$ is the m-th derivative of c. Derive the value of $\{z_n \mid z_{p_1}, \ldots, z_{p_m}\}$.

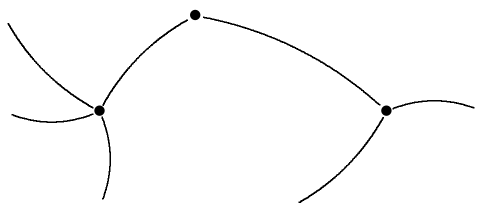
4.2. Graph combinatorics in physics. We begin by some Feynman-graphology. Feynman diagrams are constructed out of sets of vertices $v \in V$ and sets of rays R_v originating in each vertex. This gives rise to 'stars' or 'corollas'.



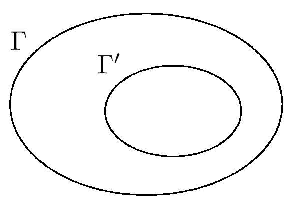
We sew them by choosing sets of disjoint pairs in $R := \bigsqcup_v R_v$ and involutive maps $s : R \to R$ that interchange elements in a chosen pair

$$s(a) = b; \quad s(b) = a.$$

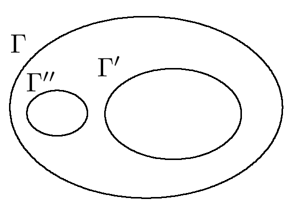
The pair (a, b) becomes an *edge*.



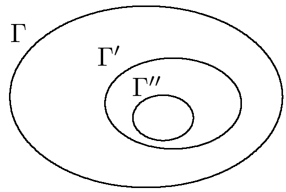
Graphical insertion of middle points to keep track of edges is useful in some circumstances. Rays that do not become edges may be called legs or half-edges. On the graphs we define a grafting operation, that slightly generalizes the grafting of trees. We graft a connected diagram Γ' into another connected diagram Γ by choosing a vertex $v \in \Gamma$ and establishing a bijection φ between R_v and the legs of Γ'. Grafting in the middle of edges is accepted. Obviously one needs the cardinality of those sets to coincide to proceed; but the order implied in the bijection is also important. The result is called $\Gamma \circ_{v,\varphi} \Gamma'$.



As before, the insertion of Γ'' into $\Gamma \circ_{v,\varphi} \Gamma'$ can be done like this:



which has the commutativity property, or like this:



which has the associativity property.

Then one sums (in the sense of formal linear combinations) over the vertices and possible bijections:

$$\Gamma' \to \Gamma := \Gamma \circ \Gamma' := \sum_{v,\varphi} \Gamma \circ_{v,\varphi} \Gamma'.$$

Some of the graphs we sum over might be isomorphic; we disregard 'symmetry factors'. *This grafting product* $\Gamma' \to \Gamma$ *satisfies the (left-symmetric) Vinberg identity.*

The set of graphs is endowed with a weight function, given by the number n of external legs, and a grading, given by the number L of loops. Grafting does not change the weight of a diagram[2], whereas the grade of $\Gamma' \to \Gamma$ is $L(\Gamma) + L(\Gamma')$. If π denotes a Feynman rule (an appropriate linear map of the space of linear combinations of graphs into the complex numbers), then the operator $\theta_{\Gamma'}$ is defined by

$$\theta_{\Gamma'}\pi(\Gamma) = \pi(\Gamma' \to \Gamma).$$

In view of the preceding considerations,

$$[\theta_{\Gamma_1}\, \theta_{\Gamma_2}] := \theta_{\Gamma_2 \to \Gamma_1} - \theta_{\Gamma_1 \to \Gamma_2},$$

just like in (14), is a Lie bracket. Consider now truncated spaces[3] of graphs $W_n(L)$ with a fixed number n of external legs and bounded loop grading. We can regard W_n as an algebraic unipotent group $\lim_{L\to\infty} W_n(L)$, an inverse limit of complex nilpotent Lie groups. The situation is similar to the case of Diff with its standard affine coordinates.

The important point is that *the Lie algebra corresponding to the group* W_n (or its truncation $W_n(L)$) *is a vector space whose basis consists of diagrams, with bracket* $[\Gamma, \Gamma'] = \Gamma' \to \Gamma - \Gamma \to \Gamma'$ *derived from the Vinberg bracket defined above.*

4.3. Renormalization scheme according to Connes and Kreimer. Let us fix the weight n, that is, the number of external legs. The simply connected nilpotent Lie group $W_n(L)$, for every L, can be realized, according to well-known results, as an algebraic subgroup of the group $T_p(\mathbb{C})$ of complex triangular matrices of the form

$$\begin{pmatrix} 1 & & * \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

(with p rows and p columns). Via dimensional regularization (see section 3.4), one defines a Feynman rule associating to every Feynman graph Γ a formal Laurent series in a variable ε. From this Feynman rule, we construct, for fixed n and L, an

[2]That is, the weight of $\Gamma' \to \Gamma$ is the weight of Γ.

[3]If we don't allow vertices of valence 2, then the spaces $W_n(L)$ are finite-dimensional, and W_n becomes a graded pre-Lie algebra in the sense of section 3.1.

element $g(\varepsilon)$ of the group $W_n(L)$. We can interpret it as a matrix $g(\varepsilon) = (g_{ij}(\varepsilon))$ in the triangular group $T_p(\mathbb{C})$: each entry $g_{ij}(\varepsilon)$ belongs to the ring $\mathbb{C}\,[[\varepsilon, \varepsilon^{-1}]]$ of formal Laurent series.

When a diagram is divergent, but has no subdivergences, the corresponding amplitude is of the form

$$g(\epsilon) = \begin{pmatrix} 1 & a(\epsilon) \\ 0 & 1 \end{pmatrix}.$$

Here

$$a(\epsilon) = a_-(\epsilon^{-1}) + a_+(\epsilon),$$

where a_- is a polynomial and a_+ is a series, and the ambiguity in their definition is solved by deciding $a_-(0) = 0$. Then the famous 'Birkhoff decomposition' by Connes and Kreimer [**5**] is simply given in this case by

$$g(\epsilon) = \begin{pmatrix} 1 & a_-(\epsilon^{-1}) \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & a_+(\epsilon) \\ 0 & 1 \end{pmatrix} =: g_-(\epsilon^{-1})\, g_+(\epsilon).$$

Suppose now the diagram is divergent overall, and has a subdivergence; then the corresponding amplitude is typically of the form

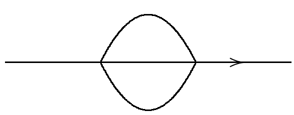
$$g(\epsilon) = \begin{pmatrix} 1 & a(\epsilon) & b(\epsilon) \\ 0 & 1 & c(\epsilon) \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & a_-(\epsilon^{-1}) + a_+(\epsilon) & b_-(\epsilon^{-1}) + b_+(\epsilon) \\ 0 & 1 & c_-(\epsilon^{-1}) + c_+(\epsilon) \\ 0 & 0 & 1 \end{pmatrix}.$$

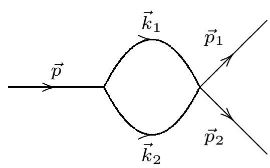
We are going to factorize this in the form $g(\epsilon) = g_-(\epsilon^{-1})g_+(\epsilon)$ again. We obtain

$$\begin{aligned} g(\epsilon) &= \begin{pmatrix} 1 & a_-(\epsilon^{-1}) & b_-(\epsilon^{-1}) - (a_-c_+)_-(\epsilon^{-1}) \\ 0 & 1 & c_-(\epsilon^{-1}) \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & a_+(\epsilon) & b_+(\epsilon) - (a_-c_+)_+(\epsilon) \\ 0 & 1 & c_+(\epsilon) \\ 0 & 0 & 1 \end{pmatrix} \\ &=: g_-(\epsilon^{-1})g_+(\epsilon). \end{aligned}$$

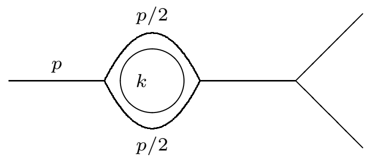
That is, one proceeds subdiagonal by subdiagonal, effecting the previous renormalization of the subdivergence —this Kurusch Ebrahimi-Fard would do by use of abstract Rota–Baxter operator properties.

4.4. Analytical considerations. Consider now a diagram Γ with n external legs (corresponding to high-energy reactions involving n particles). The figure shows two examples, respectively with $n = L = 2$; and with $n = 3$ and $L = 1$.





In the figure on the right, we must have $\vec{p} = \vec{p}_1 + \vec{p}_2$, by momentum conservation. In fact, one considers $p \in M_4$, the relativistic four-momentum, to include conservation of energy. (By the way, possible symmetries here, like the one implemented by change of the time orientation $p \mapsto -p$ correspond to deep conservation principles in physics.) The result of the calculation of the scattering amplitude defined by the diagram will depend on p, obviously. In the diagram



we put $k_1 = p/2 + k$, $k_2 = p/2 - k$; to it corresponds the integral

$$I(p) = \int \frac{d^4k}{((p/2+k)^2 + m^2)((p/2-k)^2 + m^2)}. \tag{50}$$

By means of analytic continuation we have gone over to "Euclidean" integrals; so the squares in this formula have their ordinary meaning. By power counting, it is clear that $I(p)$ diverges. We can make a cutoff $|k| < \Lambda$ (physically this is justified, as we have no information on ultra-high frequencies) and study the asymptotic behaviour of the integral with Λ. This method is however awkward for diagrams with subdivergences. Following Schwinger, Feynman, Symanzik, Nambu, Nakanishi and many others since the 50's and 60's, we can pass to the *effective action* integral, mathematically given by the identity

$$\frac{1}{p^2 + m^2} = \int_0^\infty d\alpha\, e^{-\alpha(p^2+m^2)}. \tag{51}$$

If we now substitute these integrals for the fractions in (50), after doing the easy Gaussian integral, we are left with an integral of the form

$$\int_0^\infty \int_0^\infty d\alpha_1\, d\alpha_2\, \frac{e^{-(\alpha_1+\alpha_2)m^2}\, e^{-\frac{\alpha_1\alpha_2}{2(\alpha_1+\alpha_2)}p^2}}{(\alpha_1+\alpha_2)^2}. \tag{52}$$

Whereas Λ had dimensions of energy, the α's behave as the square of a minimal length. In order to avoid the divergence of this integral near zero, one can regularize by choosing $\alpha_1, \alpha_2 \geq \epsilon$; a harmless alternative is to cut the corner out, $\alpha_1 + \alpha_2 \geq \epsilon$. In the case of having three denominators, we would get numerators of the type $\alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3$ in the exponentials; and so on. These integrals can be attacked by blow-up methods. Let us consider the relatively simple related integral:

$$\int_0^1 \int_0^1 \frac{d\alpha_1\, d\alpha_2}{1 - \alpha_1\alpha_2}.$$

This is finite, as the limit of the integral with the corner cut out (yielding the Stasheff polyhedron $\mathcal{P}_4$ [**2**]) exists and is equal to $\zeta(2) = \pi^2/6$ —by the way, as an exercise the reader is challenged to find an elementary proof of this last fact, discovered by Euler. One can practice a blow-up at the point $\alpha_1 = \alpha_2 = 1$ with the introduction of a new coordinate $\alpha_3 := \frac{1-\alpha_1}{1-\alpha_1\alpha_2}$, to obtain a surface in the real three-space as domain of integration; then the singularity disappears.

We make a final comment on dimensional renormalization. This is a form of analytic regularization, in turn developed in the last century by Hadamard, Riesz, Gelfand, Bernstein,... and is concerned with integrals of the form

$$\int_{\mathcal{D}} \frac{\varphi(x)}{[P(x)]^s}\, dx$$

with P a polynomial and s a complex parameter. The denominator in (52) that comes from the Gaussian integral is in fact $(\alpha_1 + \alpha_2)^{d/2}$, where d is the dimension. The idea is to make $d = 4 - \epsilon$ a complex variable. Feynman amplitudes, always for a fixed number of external legs, become integrals $I(\Gamma, p, \epsilon)$, that give rise to Laurent expansions in ϵ. The elements of the Connes–Kreimer group of diffeographisms that 'kill' the divergences are of the form

$$g_-(\epsilon^{-1}) = \exp\left(\frac{\beta_{-1}}{\epsilon} + \frac{\beta_{-2}}{\epsilon^2} + \cdots\right),$$

where the β_{-i} live in the Lie algebra of that group. This is related to the notion of the motivic Galois group, investigated at present by Alain Connes, Matilde Marcolli and myself.

References

[1] S. G. Krantz and H. R. Parks, *A Primer of Real Analytic Functions*, Birkhäuser, Boston, 1992.

[2] P. Cartier, "Combinatorics of trees", in *Surveys in Modern Mathematics*, V. Prasolov and Y. Ilyashenko, eds., Cambridge University Press, Cambridge, 2005; pp. 274–282.

[3] J.-L. Loday and M. Ronco, Combinatorial Hopf algebras, 2008, preprint ArXiv:math:08100435, in this volume

[4] M. Gerstenhaber, The cohomology structure of an associative ring. Ann. of Math. (2) 78 1963 267–288.

[5] A. Connes and D. Kreimer, "Renormalization in the quantum field theory and the Riemann problem I. The Hopf algebra structure of graphs and the main theorem", *Comm. Math. Phys.* **210** (2000), pp. 249-273.

Institut des Hautes Études Scientifiques, Bures-sur-Yvette, France

Clay Mathematics Proceedings
Volume 11, 2010

Noncommutative Geometry as the Key to Unlock the Secrets of Space-Time

Ali H. Chamseddine

Abstract. I give a summary of the progress made on using the elegant construction of Alain Connes' noncommutative geometry to explore the nature of space-time at very high energies. In particular I show that by making very few natural and weak assumptions about the structure of the noncommutative space, one can deduce the structure of all fundamental interactions at low energies.

1. Introduction

This article is dedicated to Alain Connes on the occasion of his 60th birthday. I came to know Alain well during my first visit to IHES in 1996. I was immediately overwhelmed with his brilliance and the inflow of his ideas, and within a short time started to collaborate with him on the interface of noncommutative geometry, his invention, and the ideas of unification in theoretical physics. This collaboration has been very fruitful, and we have come to appreciate the mysterious links between geometry and physics. Many problems remain, but I am optimistic that the challenge of finding a quantum theory of gravity using the geometric tools that Alain developed is within reach. At the personal level, I discovered that Alain is a very warm person, full of life, and has a fantastic sense of humor. I am proud of his friendship.

What I will present here is a summary of a forthcoming long article in collaboration with Alain, which hopefully will appear in the near future, where a self-contained exposition of the methods of noncommutative geometry applied to particle physics is explained in a language accessible to physicists [**1**]. A good part of this forthcoming article will elaborate and build on the results that were first obtained with the crucial input of the collaboration with Marcolli [**2**]. In addition, the introduction present in a recent paper [**3**] can be used to help introduce the reader to the general philosophy of our program. Our aim is to provide enough material to help students and young researchers who wish to learn about this promising direction of research.

2010 *Mathematics Subject Classification.* Primary 81T75; Secondary 58B34, 58J42, 81R60, 81V22.

This is supported in part by the Arab Fund for Social and Economic Development and by NSF grant Phys-0653300.

The laws of physics at low energies are well encoded by the action functional which is the sum of the Einstein-Hilbert action and that of the standard model. These two parts have different properties, the first being dependent on the geometry of the underlying manifold (M, g) where g is the metric, while the other is governed by internal symmetries of a gauge group G which can be well described using the language of vector bundles. The underlying symmetries are also different. General relativity is governed by diffeomorphism invariance (outer automorphisms of (M, g)) while gauge symmetries are based on local gauge invariance (inner automorphisms). Thus the natural group of invariance is the semi-direct product

$$G = U \rtimes \mathrm{Diff}\,(M)$$

where

$$U = C^{\infty}\left(M, U(1) \times SU(2) \times SU(3)\right).$$

It is possible to trace back the failure of finding a unified theory of all interactions including quantum gravity to the difference between these two kinds of symmetries. In addition, there are many unanswered questions within the established formulation of the standard model. For example, the following questions have no compelling answer: Why is the gauge group specifically given by $U(1)\times SU(2)\times SU(3)$? Why do the fermions occupy the particular representations that they do? Why are there three families and why are there 16 fundamental fermions per family? What is the theoretical origin of the Higgs mechanism and the spontaneous breakdown of gauge symmetries? What is the Higgs mass and how does one explain all the fermionic masses? These are only a few of the questions that have to be answered by the ultimate unified theory of all interactions. We shall attempt to answer some of these questions taking as a starting point the following observations. At energies well below the Planck scale

$$M_P = \sqrt{\frac{1}{8\pi G}} \equiv \frac{1}{\kappa} = 2.43 \times 10^{18}\ \text{Gev}$$

gravity can be safely considered as a classical theory. But as energies approach the Planck scale one expects the quantum nature of space-time to reveal itself, and for the Einstein-Hilbert action to become an approximation of some deformed theory. In addition the other three forces must be unified with gravity in such a way that all interactions will correspond to one underlying symmetry. One thus would expect that the nature of space-time, and thus of geometry, would change at Planckian energies in such a way that at lower energies, one recovers the above picture of diffeomorphisms and internal gauge symmetries. It is not realistic to guess the exact properties of space-time at Planckian energies and to make directly an extrapolation of 17 orders of magnitude from our present energies. We are therefore led to take an indirect approach where we search for the hidden structure in the functional of gravity coupled to the standard model at present energies. To do this we shall make a basic conjecture which we will take as a starting point:

Conjecture 1. *At some energy level, space-time is the product of a continuous four-dimensional manifold times a discrete space F.*

The aim then is to find supporting evidence for this conjecture. Once this is done the next step would be to find the true geometry at Planckian energies, for which this product in turn is a limit.

This is the minimal extension where no new extra dimensions are assumed. The task now is to determine with minimal input the properties of the discrete space F, and construct the associated physical theory. Remarkably, we will show that this information will allow us to determine the hidden structure of space-time, and answer some, but not all (so far) of the questions posed above.

2. *A Brief Summary of Alain Connes NCG*

The basic idea is based on physics. The modern way of measuring distances is spectral. The unit of distance is taken as the wavelength of atomic spectra. To adapt this geometrically the notion of real variable which one takes as a function f on a set X where $f : X \to \mathbb{R}$ has to be replaced. This is now taken to be a self-adjoint operator in a Hilbert space as in quantum mechanics. The space X is described by the algebra $\mathcal{A}$ of coordinates which is represented as operators in a fixed Hilbert space $\mathcal{H}$. The geometry of the noncommutative space is determined in terms of the spectral data $(\mathcal{A}, \mathcal{H}, \mathcal{D}, J, \gamma)$. A real, even spectral triple is defined by [**4**], [**5**]

- $\mathcal{A}$ is an associative algebra with unit 1 and involution $*$.
- $\mathcal{H}$ is a complex Hilbert space carrying a faithful representation π of the algebra.
- D is a self-adjoint operator on $\mathcal{H}$ with the resolvent $(D-\lambda)^{-1}, \lambda \notin \mathbb{R}$ of D compact.
- J is an anti–unitary operator on $\mathcal{H}$, which is a real structure (charge conjugation.)
- γ is a unitary operator on $\mathcal{H}$, the chirality.

We require the following axioms to hold:

- $J^2 = \varepsilon$, ($\varepsilon = 1$ in dimension zero and $\varepsilon = -1$ in 4 dimensions).
- $[a, b^o] = 0$ for all $a, b \in \mathcal{A}$, where $b^o = Jb^*J^{-1}$. This is the zeroth order condition. This is needed to define the right action on elements of $\mathcal{H}$: $\zeta b = b^o \zeta$, and is the statement that left action and right action commute.
- $DJ = \varepsilon' JD, \quad J\gamma = \varepsilon'' \gamma J, \quad D\gamma = -\gamma D$ where $\varepsilon, \varepsilon', \varepsilon'' \in \{-1, 1\}$. The reality conditions resemble the conditions governing the existence of Majorana (real) fermions.
- $[[D, a], b^o] = 0$ for all $a, b \in \mathcal{A}$. This is the first order condition.
- $\gamma^2 = 1$ and $[\gamma, a] = 0$ for all $a \in \mathcal{A}$. These properties allow the decomposition $\mathcal{H} = \mathcal{H}_L \oplus \mathcal{H}_R$.
- $\mathcal{H}$ is endowed with the $\mathcal{A}$-bimodule structure $a\zeta b = ab^o\zeta$.
- The notion of dimension is governed by growth of eigenvalues, and may be fractal or complex.
- We stress that we are considering spaces with Euclidean signature.

$\mathcal{A}$ has a well-defined unitary group

$$\mathcal{U} = \{u \in \mathcal{A}; \quad u\,u^* = u^*u = 1\}.$$

The natural adjoint action of $\mathcal{U}$ on $\mathcal{H}$ is given by $\zeta \to u\zeta u^* = u\,J\,u\,J^*\zeta \quad \forall \zeta \in \mathcal{H}$. Then

$$\langle \zeta, D\zeta \rangle$$

is not invariant under the above transformation:

$$(u\,J\,u\,J^*)\,D\,(u\,J\,u\,J^*)^* = D + u\,[D, u^*] + J\,(u\,[D, u^*])\,J^*.$$

However, the action $\langle \zeta, D_A \zeta \rangle$ is invariant, where

$$D_A = D + A + \varepsilon' JAJ^{-1}, \quad A = \sum_i a^i \left[D, b^i\right]$$

and $A = A^*$ is self-adjoint. This is similar to the appearance of the interaction term for the photon with electrons

$$i\overline{\psi}\gamma^\mu \partial_\mu \psi \to i\overline{\psi}\gamma^\mu \left(\partial_\mu + ieA_\mu\right)\psi$$

to maintain invariance under the variations $\psi \to e^{ie\alpha(x)}\psi$.

The properties listed above of the anti-linear isometry $J : \mathcal{H} \to \mathcal{H}$ are characteristic of a real structure of KO-dimension $n \in \mathbb{Z}/8$ on the spectral triple $(\mathcal{A}, \mathcal{H}, D)$. The numbers $\varepsilon, \varepsilon', \varepsilon'' \in \{-1, 1\}$ are a function of n mod 8 given by

n	0	1	2	3	4	5	6	7
ε	1	1	-1	-1	-1	-1	1	1
ε'	1	-1	1	1	1	-1	1	1
ε''	1		-1		1		-1	

We take the algebra $\mathcal{A}$ to be given by a tensor product which geometrically corresponds to a product space. The spectral geometry of $\mathcal{A}$ is given by the product rule $\mathcal{A} = C^\infty(M) \otimes \mathcal{A}_F$ where the algebra $\mathcal{A}_F$ is finite dimensional, and

$$\mathcal{H} = L^2(M, S) \otimes \mathcal{H}_F, \quad D = D_M \otimes 1 + \gamma_5 \otimes D_F,$$

where $L^2(M, S)$ is the Hilbert space of L^2 spinors, and D_M is the Dirac operator of the Levi-Civita spin connection on M, $D_M = \gamma^\mu (\partial_\mu + \omega_\mu)$. The Hilbert space $\mathcal{H}_F$ is taken to include the physical fermions. The chirality operator is $\gamma = \gamma_5 \otimes \gamma_F$ and the reality operator is $J = C \otimes J_F$, where C is the charge conjugation matrix.

In order to avoid the fermion doubling problem where the fermions $\zeta, \zeta^c, \zeta^*, \zeta^{c*}$, $\zeta \in \mathcal{H}$, should not be all independent, it was shown that the finite dimensional space must be taken to be of K-theoretic dimension 6 [**6**], [**7**], where in this case $(\varepsilon, \varepsilon', \varepsilon") = (1, 1, -1)$ (so as to impose the condition $J\zeta = \zeta$). This makes the total K-theoretic dimension of the noncommutative space to be 10 and would allow one to impose the reality (Majorana) condition and the Weyl condition simultaneously in the Minkowskian continued form, a situation very familiar in ten-dimensional supersymmetry. In the Euclidean version, the use of J in the fermionic action would give for the chiral fermions in the path integral a Pfaffian instead of a determinant [**6**], and will thus cut down the fermionic degrees of freedom by a factor of 2. In other words, in order to have the fermionic sector free of the fermionic doubling problem we must make the choice

$$J_F^2 = 1, \qquad J_F D_F = D_F J_F, \qquad J_F \gamma_F = -\gamma_F J_F.$$

In what follows we will restrict our attention to determination of the finite algebra, and will omit the subscript F.

3. Classification of Finite Noncommutative Spaces

There are two main constraints on the algebra from the axioms of noncommutative geometry. We first look for involutive algebras $\mathcal{A}$ of operators in $\mathcal{H}$ such that,

$$[a, b^o] = 0\,, \quad \forall\, a, b \in \mathcal{A}\,,$$

where for any operator a in $\mathcal{H}$, $a^o = Ja^*J^{-1}$. This is called the zeroth order condition. We shall assume that the representations of $\mathcal{A}$ and J in $\mathcal{H}$ are *irreducible.*

The classification of the irreducible triplets $(\mathcal{A}, \mathcal{H}, J)$ leads to the following theorem [**8**], [**9**]:

THEOREM 2. The center $Z(\mathcal{A}_{\mathbb{C}})$ is $\mathbb{C}$ or $\mathbb{C}\oplus\mathbb{C}$.

If the center $Z(\mathcal{A}_{\mathbb{C}})$ is $\mathbb{C}$ then we can state the following theorem:

THEOREM 3. *Let $\mathcal{H}$ be a Hilbert space of dimension n. Then an irreducible solution with $Z(\mathcal{A}_{\mathbb{C}}) = \mathbb{C}$ exists iff $n = k^2$ is a square. It is given by $A_{\mathbb{C}} = M_k(\mathbb{C})$ acting by left multiplication on itself and the anti-linear involution*

$$J(x) = x^*, \quad \forall x \in M_k(\mathbb{C}).$$

For $\mathcal{A}_{\mathbb{C}} = M_k(\mathbb{C})$ we have $\mathcal{A} = M_k(\mathbb{C})$, $M_k(\mathbb{R})$ or $M_a(\mathbb{H})$ for even $k = 2a$, where $\mathbb{H}$ is the field of quaternions [**10**]. These correspond respectively to the unitary, orthogonal and symplectic case.

If the center $Z(\mathcal{A}_{\mathbb{C}})$ is $\mathbb{C}\oplus\mathbb{C}$ then we can state the theorem:

THEOREM 4. *Let $\mathcal{H}$ be a Hilbert space of dimension n. Then an irreducible solution with $Z(\mathcal{A}_{\mathbb{C}}) = \mathbb{C}\oplus\mathbb{C}$ exists iff $n = 2k^2$ is twice a square. It is given by $A_{\mathbb{C}} = M_k(\mathbb{C}) \oplus M_k(\mathbb{C})$ acting by left multiplication on itself and anti-linear involution*

$$J(x, y) = (y^*, x^*), \quad \forall x, y \in M_k(\mathbb{C}).$$

With each of the $M_k(\mathbb{C})$ in $\mathcal{A}_{\mathbb{C}}$ we can have the three possibilities $M_k(\mathbb{C})$, $M_k(\mathbb{R})$, or $M_a(\mathbb{H})$, where $k = 2a$. At this point we make the *hypothesis* that we are in the "symplectic–unitary" case, thus restricting the algebra $\mathcal{A}$ to the form

$$\mathcal{A} = M_a(\mathbb{H}) \oplus M_k(\mathbb{C}), \qquad k = 2a.$$

The dimension of the Hilbert space is $n = 2k^2$; however, because of the reality condition, these correspond to k^2 fundamental fermions , where $k = 2a$ is an even integer. The first possible value for k is 2 corresponding to a Hilbert space of four fermions and an algebra $\mathcal{A} = \mathbb{H} \oplus M_2(\mathbb{C})$. This is ruled out because it does not allow one to impose a grading on the algebra. It is also ruled out by the existence of quarks. The next possible value for k is 4, thus predicting the number of fermions to be 16.

In the $Z(\mathcal{A}_{\mathbb{C}}) = \mathbb{C}$ case, one can show that it is not possible to have the finite space to be of K-theoretic dimension 6 consistent with the relation $J\gamma = -\gamma J$ [**8**]. We therefore can proceed directly to the second case.

One then takes the grading γ of $\mathcal{H}$ so that the K-theoretic dimension of the finite space is 6 and this is consistent with the condition $J\gamma = -\gamma J$. It is given by

$$\gamma(\zeta, \eta) = (\gamma\zeta, -\gamma\eta).$$

This grading breaks the algebra $\mathcal{A} = M_2(\mathbb{H}) \oplus M_4(\mathbb{C})$, which is non-trivially graded only for the $M_2(\mathbb{H})$ component, to its even part:

$$\mathcal{A}^{\text{ev}} = \mathbb{H}_R \oplus \mathbb{H}_L \oplus M_4(\mathbb{C}).$$

The Dirac operator must connect the two pieces non-trivially, and therefore must satisfy

$$[D, Z(\mathcal{A})] \neq \{0\}.$$

The physical meaning of this constraint is to allow some of the fermions to acquire Majorana masses, realizing the see-saw mechanism, and thus connecting the fermions to their conjugates.

We have to look for subalgebras $\mathcal{A}_F \subset \mathcal{A}^{\mathrm{ev}}$, the even part of the algebra $\mathcal{A}$, for which $[[D,a],b^o]=0, \quad \forall\, a,b \in \mathcal{A}_F$. We can state the theorem:

THEOREM 5. Up to automorphisms of A^{ev}, there exists a unique involutive subalgebra $A_F \subset A^{\mathrm{ev}}$ of maximal dimension admitting off-diagonal Dirac operators

$$\begin{aligned}\mathcal{A}_F &= \left\{ \left(\lambda \oplus \overline{\lambda}\right) \oplus q,\ \lambda \oplus m \,|\, \lambda \in \mathbb{C}, \quad q \in \mathbb{H}, \quad m \in M_3\left(\mathbb{C}\right) \right\} \\ &\subset \mathbb{H} \oplus \mathbb{H} \oplus M_4\left(\mathbb{C}\right).\end{aligned}$$

It is isomorphic to $\mathbb{C} \oplus \mathbb{H} \oplus M_3\left(\mathbb{C}\right)$.

4. Tensor Notation

It is helpful to write the results obtained about the standard model using tensor notation. The Dirac action must take the form

$$\Psi_M^* D_M^N \Psi_N$$

where $\Psi_M = \begin{pmatrix} \psi_A \\ \psi_{A'} \end{pmatrix}$ and we have denoted $\psi_{A'} = \psi_A^c$, the conjugate spinor. We start with the algebra

$$\mathcal{A} = M_4\left(\mathbb{C}\right) \oplus M_4\left(\mathbb{C}\right)$$

and denote the spinors by $\psi_A = \psi_{\alpha I}$, $A = \alpha I, \quad \alpha = 1, \cdots, 4, \quad I = 1, \cdots, 4$, and thus $D_A^B = D_{\alpha I}^{\beta J}$. The Dirac operator takes the form

$$D = \begin{pmatrix} D_A^B & D_A^{B'} \\ D_{A'}^B & D_{A'}^{B'} \end{pmatrix},$$

where $A' = \alpha' I', \quad \alpha' = 1', \cdots, 4', \quad I' = 1', \cdots, 4'$, and $D_{A'}^{B'} = \overline{D}_A^B$, $D_{A'}^B = \overline{D}_A^{B'}$ and overbar denotes complex conjugation.

Elements of the algebra $\mathcal{A}$ are matrices a_M^N of the special form:

$$a = \begin{pmatrix} X_\alpha^\beta \delta_I^J & 0 \\ 0 & \delta_{\alpha'}^{\beta'} Y_{I'}^{J'} \end{pmatrix},$$

where X_α^β is an element of the first $M_4\left(\mathbb{C}\right)$ and $Y_{I'}^{J'}$ is an element of the second $M_4\left(\mathbb{C}\right)$. The reality operator J is defined by

$$J = \begin{pmatrix} 0 & \delta_\alpha^{\beta'} \delta_I^{J'} \\ \delta_{\alpha'}^\beta \delta_{I'}^J & 0 \end{pmatrix} \times \text{complex conjugation.}$$

In this representation we deduce that a^o takes the form

$$a^o = J a^* J^{-1} = \begin{pmatrix} \delta_\alpha^\beta \widetilde{Y}_I^J & 0 \\ 0 & \widetilde{X}_{\alpha'}^{\beta'} \delta_{I'}^{J'}, \end{pmatrix},$$

where$^\sim$ denotes transposition. It is trivial to verify that $[a, b^o] = 0$.

The order one condition is

$$\left[\left[D,a\right], b^o\right] = 0$$

If we write

$$b^o = \begin{pmatrix} \delta_\alpha^\beta W_I^J & 0 \\ 0 & Z_{\alpha'}^{\beta'} \delta_{I'}^{J'} \end{pmatrix},$$

then the commutator $[[D,a], b^o]$ is given by

$$\begin{pmatrix} [[D,X], W]_A^B & ((DY - XD)Z - W(DY - XD))_A^{B'} \\ ((DX - YD)W - Z(DX - YD))_{A'}^B & [[D,Y], Z]_{A'}^{B'} \end{pmatrix}$$

the vanishing of which implies the equations:

$$\left(D_{\alpha I}^{\gamma K} X_\gamma^\beta - X_\alpha^\gamma D_{\gamma I}^{\beta K}\right) W_K^J - W_I^K \left(D_{\alpha K}^{\gamma J} X_\gamma^\beta - X_\alpha^\gamma D_{\gamma K}^{\beta J}\right) = 0$$
$$\left(D_{\alpha I}^{\gamma' K'} Y_{K'}^{J'} - X_\alpha^\gamma D_{\gamma I}^{\gamma' K}\right) Z_{\gamma'}^{\beta'} - W_I^K \left(D_{\alpha K}^{\beta' K'} Y_{K'}^{J'} - X_\alpha^\gamma D_{\gamma K}^{\beta' J'}\right) = 0.$$

We have shown [**8**], [**2**], that the only non-zero solution of the second equation is

$$D_{\alpha I}^{\beta' K'} = \delta_\alpha^{\dot{1}} \delta_{\dot{1}'}^{\beta'} \delta_I^1 \delta_{1'}^{K'} k^{*\nu_R}$$

which means that there can be only *one non-zero single entry* in the off-diagonal 16×16 matrix $D_A^{B'}$, and this implies that

$$D_{\alpha I}^{\beta J} = D_{\alpha(l)}^\beta \delta_I^1 \delta_1^J + D_{\alpha(q)}^\beta \delta_I^i \delta_j^J \delta_i^j$$
$$Y_{I'}^{J'} = \delta_{I'}^{1'} \delta_{1'}^{J'} Y_{1'}^{1'} + \delta_{I'}^{i'} \delta_{j'}^{J'} Y_{i'}^{j'}$$
$$X_{\dot{1}}^{\dot{1}} = Y_{1'}^{1'}, \; X_{\dot{1}}^\alpha = 0, \quad \alpha \neq \dot{1},$$

where we have split the index $I = 1, i$, and $I' = 1', i'$. From the property of commutation of the grading operator

$$g_\alpha^\beta = \begin{pmatrix} 1_2 & 0 \\ 0 & -1_2 \end{pmatrix}$$
$$[g, a] = 0 \quad a \in M_4(\mathbb{C}),$$

the algebra $M_4(\mathbb{C})$ reduces to $M_2(\mathbb{C})_R \oplus M_2(\mathbb{C})_L$. We further impose the condition of symplectic isometry on $M_2(\mathbb{C})_R \oplus M_2(\mathbb{C})_L$,

$$\sigma_2 \otimes 1_2 (\overline{a}) \sigma_2 \otimes 1_2 = a,$$

which reduces it to $\mathbb{H}_R \oplus \mathbb{H}_L$. We will be using the notation

$$\alpha = \dot{1}, \dot{2}, a \;\text{ where } \xi_{\dot{1},\dot{2}} \in \mathbb{H}_R, \, \xi_a \in \mathbb{H}_L.$$

Together with the above condition this implies that

$$X_\alpha^\beta = \delta_\alpha^{\dot{1}} \delta_{\dot{1}}^\beta X_{\dot{1}}^{\dot{1}} + \delta_\alpha^{\dot{2}} \delta_{\dot{2}}^\beta \overline{X}_{\dot{1}}^{\dot{1}} + \delta_\alpha^a \delta_b^\beta X_a^b$$

and the algebra $\mathbb{H}_R \oplus \mathbb{H}_L \oplus M_4(\mathbb{C})$ reduces to

$$\mathbb{C} \oplus \mathbb{H} \oplus M_3(\mathbb{C})$$

because $X_{\dot{1}}^{\dot{1}} = Y_{1'}^{1'}$. Expanding the Dirac action we get

$$\psi_A^* D_A^B \psi_B + \psi_{\dot{1}'1'}^* D_{\dot{1}'1'}^B \psi_B + \psi_A^* D_A^{\dot{1}'1'} \psi_{\dot{1}'1'} + \psi_{A'}^* D_{A'}^{B'} \psi_{B'}$$

The spinors can thus be denoted by

$$\begin{aligned}\psi_A &= \psi_{\alpha I} = (\psi_{\alpha 1}, \psi_{\alpha i}) \\ &= \left(\psi_{\dot{1}1}, \psi_{\dot{2}1}, \psi_{a1}, \psi_{\dot{1}i}, \psi_{\dot{2}i}, \psi_{ai}\right) \\ &\equiv (\nu_R, e_R, l_a, u_{Ri}, d_{Ri}, q_{ai})\,,\end{aligned}$$

where $l_a = (\nu_L, e_L)$ and $q_{ai} = (u_{Li}, d_{Li})$. The component $\psi_{\dot{1}'1'} = \psi^c_{\dot{1}1} = \nu^c_R$, which implies that the Dirac action

$$\psi^*_A D^B_A \psi_B + \nu^{*c}_R k^{*\nu_R} \nu_R + \text{c.c}$$

has only a mixing term for the right-handed neutrinos.

Having determined the structure of the Dirac operator of the discrete space, we can form the Dirac operator of the product space of this discrete space times a four-dimensional Riemannian manifold:

$$D = D_M \otimes 1 + \gamma_5 \otimes D_F.$$

As D_F is a 32×32 matrix tensored with the 3×3 matrices of generation space and with the Clifford algebra, D is 384×384 matrix.

To take inner automorphisms into account, we have to evaluate the Dirac operator

$$D_A = D + A + JAJ^{-1},$$

where

$$A = \sum a\,[D, b]\,.$$

In particular

$$A^B_A = \sum a^C_A \left(D^D_C b^B_D - b^D_C D^B_D\right).$$

Note there are no mixing terms like $D^{D'}_C b^B_{D'}$ because b is block diagonal.

Evaluating all components of the full Dirac operator D_M^N , quoting only the result as the full derivation will be given in a forthcoming paper [**1**], we obtain:

$$
\begin{aligned}
(D)_{\dot{1}1}^{\dot{1}1} &= \gamma^\mu \otimes D_\mu \otimes 1_3, \quad D_\mu = \partial_\mu + \frac{1}{4}\omega_\mu^{cd}\,(e)\,\gamma_{cd}, \quad 1_3 = \text{generations} \\
(D)_{\dot{1}1}^{a1} &= \gamma_5 \otimes k^{*\nu} \otimes \epsilon^{ab} H_b \qquad k^\nu = 3\times 3 \text{ neutrino mixing matrix} \\
(D)_{\dot{2}1}^{\dot{2}1} &= \gamma^\mu \otimes (D_\mu + i g_1 B_\mu) \otimes 1_3 \\
(D)_{\dot{2}1}^{a1} &= \gamma_5 \otimes k^{*e} \otimes \overline{H}^a \\
(D)_{a1}^{\dot{1}1} &= \gamma_5 \otimes k^{\nu} \otimes \epsilon_{ab}\overline{H}^b \\
(D)_{a1}^{\dot{2}1} &= \gamma_5 \otimes k^{e} \otimes H_a \\
(D)_{a1}^{b1} &= \gamma^\mu \otimes \left(\left(D_\mu + \frac{i}{2} g_1 B_\mu\right)\delta_a^b - \frac{i}{2} g_2 W_\mu^\alpha \left(\sigma^\alpha\right)_a^b\right) \otimes 1_3, \qquad \sigma^\alpha = \text{Pauli} \\
(D)_{1i}^{1j} &= \gamma^\mu \otimes \left(\left(D_\mu - \frac{2i}{3} g_1 B_\mu\right)\delta_i^j - \frac{i}{2} g_3 V_\mu^m \left(\lambda^m\right)_i^j\right) \otimes 1_3, \qquad \lambda^i = \text{Gell-Mann} \\
(D)_{\dot{1}i}^{aj} &= \gamma_5 \otimes k^{*u} \otimes \epsilon^{ab} H_b \delta_i^j \\
(D)_{\dot{2}i}^{\dot{2}j} &= \gamma^\mu \otimes \left(\left(D_\mu + \frac{i}{3} g_1 B_\mu\right)\delta_i^j - \frac{i}{2} g_3 V_\mu^m \left(\lambda^m\right)_i^j\right) \otimes 1_3 \\
(D)_{\dot{2}i}^{aj} &= \gamma_5 \otimes k^{*d} \otimes \overline{H}^a \delta_i^j \\
(D)_{ai}^{bj} &= \gamma^\mu \otimes \left(\left(D_\mu - \frac{i}{6} g_1 B_\mu\right)\delta_a^b \delta_i^j - \frac{i}{2} g_2 W_\mu^\alpha \left(\sigma^\alpha\right)_a^b \delta_i^j - \frac{i}{2} g_3 V_\mu^m \left(\lambda^m\right)_i^j \delta_a^b\right) \otimes 1_3 \\
(D)_{ai}^{\dot{1}j} &= \gamma_5 \otimes k^{u} \otimes \epsilon_{ab}\overline{H}^b \delta_i^j \\
(D)_{ai}^{\dot{2}j} &= \gamma_5 \otimes k^{d} \otimes H_a \delta_i^j \\
(D)_{\dot{1}1}^{\dot{1}'1'} &= \gamma_5 \otimes k^{*\nu_R}\sigma \qquad \text{generate scale } M_R \text{ by } \langle\sigma\rangle = M_R \\
(D)_{\dot{1}'1'}^{\dot{1}1} &= \gamma_5 \otimes k^{\nu_R}\sigma \\
D_{A'}^{B'} &= \overline{D}_A^B, \qquad D_{A'}^{B} = \overline{D}_A^{B'}, \qquad D_A^{B'} = \overline{D}_{A'}^{B}
\end{aligned}
$$

where B_μ, W_μ^α and V_μ^m are the $U(1)$, $SU(2)$ and $SU(3)$ gauge fields, and H is a complex doublet scalar field and σ is a singlet real scalar field. We have assumed that the unitary algebra $\mathcal{U}(\mathcal{A})$ is restricted to $\mathcal{SU}(\mathcal{A})$ to eliminate a superfluous

$U(1)$ gauge field. Pictorially, the matrix D_M^N has the structure:

$$
\begin{array}{c|cccccc}
 & \dot{1}1 & \dot{2}1 & a1 & \dot{1}i & \dot{2}i & ai \\
 & v_R & e_R & l_a & u_{iR} & d_{iR} & q_{iL} \\
\hline
\dot{1}1 & (D)^{\dot{1}1}_{\dot{1}1} & 0 & (D)^{a1}_{\dot{1}1} & 0 & 0 & 0 \\
\dot{2}1 & 0 & (D)^{\dot{2}1}_{\dot{2}1} & (D)^{a1}_{\dot{2}1} & 0 & 0 & 0 \\
b1 & (D)^{\dot{1}1}_{b1} & (D)^{\dot{2}1}_{b1} & (D)^{b1}_{a1} & 0 & 0 & 0 \\
\dot{1}j & 0 & 0 & 0 & (D)^{\dot{1}i}_{\dot{1}j} & 0 & (D)^{ai}_{\dot{1}j} \\
\dot{2}j & 0 & 0 & 0 & 0 & (D)^{\dot{2}i}_{\dot{2}j} & (D)^{ai}_{\dot{2}j} \\
bj & 0 & 0 & 0 & (D)^{\dot{1}i}_{bj} & (D)^{\dot{2}i}_{bj} & (D)^{ai}_{bj}
\end{array}
$$

Needless to say the term $\psi_M^* D_M^N \psi_N$ contains all the fermionic terms and their interactions in the standard model.

5. *The Spectral Action Principle*

There is a shift of point of view in NCG similar to Fourier transform, where the usual emphasis on the points $x \in M$ of a geometric space is now replaced by the spectrum Σ of the operator D. The existence of Riemannian manifolds which are isospectral but not isometric shows that the following hypothesis is stronger than the usual diffeomorphism invariance of the action of general relativity

The physical action depends only on Σ

This is the spectral action principle [**11**]. The spectrum is a geometric invariant and replaces diffeomorphism invariance. We now apply this basic principle to the noncommutative geometry defined by the spectrum of the standard model to show that the dynamics of all interactions, including gravity is given by the spectral action

$$\text{Trace } f\left(\frac{D_A}{\Lambda}\right) + \frac{1}{2}\left\langle J\Psi, D_A\Psi\right\rangle,$$

where f is a positive function, Λ a cut-off scale needed to make $\frac{D_A}{\Lambda}$ dimensionless, and Ψ is a Grassmann variable which represents fermions.

In the case of the cut-off function, f only plays a role through its moments f_0, f_2, f_4, where

$$f_k = \int_0^\infty f(v) v^{k-1} dv, \quad \text{for } k > 0, \quad , \quad f_0 = f(0).$$

These will serve as three free parameters in the model. In this case the action $S_\Lambda[D_A]$ is the number of eigenvalues λ of D_A counted with their multiplicities such that $|\lambda| \leq \Lambda$.

To illustrate how this comes about, expand the function f in terms of its Laplace transform

$$
\begin{aligned}
\mathrm{Trace} f(P) &= \sum_s f_s \mathrm{Trace}\left(P^{-s}\right) \\
\mathrm{Trace}\left(P^{-s}\right) &= \frac{1}{\Gamma(s)} \int_0^\infty t^{s-1} \mathrm{Trace}\left(e^{-tP}\right) dt \qquad \mathrm{Re}(s) \geq 0 \\
\mathrm{Trace}\left(e^{-tP}\right) &\simeq \sum_{n \geq 0} t^{\frac{n-m}{d}} \int_M a_n(x, P)\, dv(x),
\end{aligned}
$$

where $m = 4$ is the dimension of the manifold M and $d = 2$ is the order of the elliptic operator D^2. Gilkey gives generic formulas for the Seeley-DeWitt coefficients $a_n(x, P)$ for a large class of differential operators P [**12**]. The details are explained in preceding papers [**11**], [**2**] or using the tensorial notation, in a forthcoming paper [**1**].

The bosonic part of the spectral action gives an action that unifies gravity with $SU(2) \times U(1) \times SU(3)$ Yang-Mills gauge theory, with a Higgs doublet H and spontaneous symmetry breaking and a real scalar field σ. It is given by [**11**], [**2**]

$$
\begin{aligned}
S = & \frac{48}{\pi^2} f_4 \Lambda^4 \int d^4x \sqrt{g} \\
& - \frac{4}{\pi^2} f_2 \Lambda^2 \int d^4x \sqrt{g} \left(R + \frac{1}{2} a \overline{H} H + \frac{1}{4} c \right) \\
& + \frac{1}{2\pi^2} f_0 \int d^4x \sqrt{g} \left[\frac{1}{30} \left(-18 C^2_{\mu\nu\rho\sigma} + 11 R^* R^* \right) \right. \\
& \qquad + \frac{5}{3} g_1^2 B^2_{\mu\nu} + g_2^2 \left(W^\alpha_{\mu\nu} \right)^2 + g_3^2 \left(V^m_{\mu\nu} \right)^2 \\
& \qquad + \frac{1}{6} a R H_a \overline{H}^a + b \left(\overline{H} H \right)^2 + a \left| \nabla_\mu H_a \right|^2 \\
& \qquad \left. + 2e \overline{H} H\, \sigma^2 + \frac{1}{2} d\, \sigma^4 + \frac{1}{12} c R \sigma^2 + \frac{1}{2} c \left(\partial_\mu \sigma \right)^2 \right] \\
& + f_{-2} \Lambda^{-2} a_6 + \cdots
\end{aligned}
$$

This can be rearranged, after normalizing the kinetic energies and ignoring the σ field which only plays a role in cosmology, to the form:

$$
\begin{aligned}
S = & \int \left(\frac{1}{2\kappa_0^2} R + \alpha_0 C_{\mu\nu\rho\sigma} C^{\mu\nu\rho\sigma} + \gamma_0 + \tau_0 R^* R^* \right. \\
+ & \ \frac{1}{4} G^i_{\mu\nu} G^{\mu\nu i} + \frac{1}{4} F^\alpha_{\mu\nu} F^{\mu\nu\alpha} + \frac{1}{4} B_{\mu\nu} B^{\mu\nu} \\
+ & \left. \frac{1}{2} |D_\mu \mathbf{H}|^2 - \mu_0^2 |\mathbf{H}|^2 - \xi_0 R |\mathbf{H}|^2 + \lambda_0 |\mathbf{H}|^4 \right) \sqrt{g}\, d^4x,
\end{aligned}
$$

where

$$\frac{1}{\kappa_0^2} = \Lambda^2 \, \frac{96\, f_2 \, - f_0\, c}{12\, \pi^2}, \qquad \mu_0^2 = \Lambda^2 \, \left(2\, \frac{f_2}{f_0} - \frac{e}{a} \right)$$

$$\alpha_0 = -\frac{3\, f_0}{10\, \pi^2}, \qquad \tau_0 = \frac{11\, f_0}{60\, \pi^2}, \qquad \lambda_0 = \frac{\pi^2}{2\, f_0} \frac{b}{a^2}$$

$$\gamma_0 = \, \Lambda^4 \frac{1}{\pi^2} (48\, f_4 \, - f_2 \, c + \frac{1}{4} f_0 d), \qquad \xi_0 = \frac{1}{12}.$$

The parameters a, b, c, d, e are all dimensionless and related to the Yukawa couplings that give the fermionic masses after the spontaneous breaking of symmetry:

$$a = \mathrm{Tr}\left(k^{e*} k^e + k^{\nu *} k^\nu + 3 k^{u*} k^u + 3 k^{d*} k^d \right)$$

$$b = \mathrm{Tr}\left(\left(k^{e*} k^e \right)^2 + \left(k^{\nu *} k^\nu \right)^2 + 3 \left(k^{u*} k^u \right)^2 + 3 \left(k^{d*} k^d \right)^2 \right)$$

$$c = \mathrm{Tr}\left(k_R^* k_R \right), \quad d = \mathrm{Tr}\left((k_R^* k_R)^2 \right), \quad e = \mathrm{Tr}\left(k_R^* k_R k^{\nu *} k^\nu \right).$$

6. *Predictions of Spectral Action for Standard Model*

We shall first perform our analysis by assuming that the function f is well approximated by the cut-off function, thus allowing us to ignore higher order terms. We will determine to what extent such an approximation could be made, and its effects on the predictions. The normalization of the kinetic terms imposes a relation between the coupling constants g_1, g_2, g_3 and the coefficient f_0, of the form

$$\frac{g_3^2\, f_0}{2\pi^2} = \frac{1}{4}, \qquad g_3^2 = g_2^2 = \frac{5}{3}\, g_1^2 \, .$$

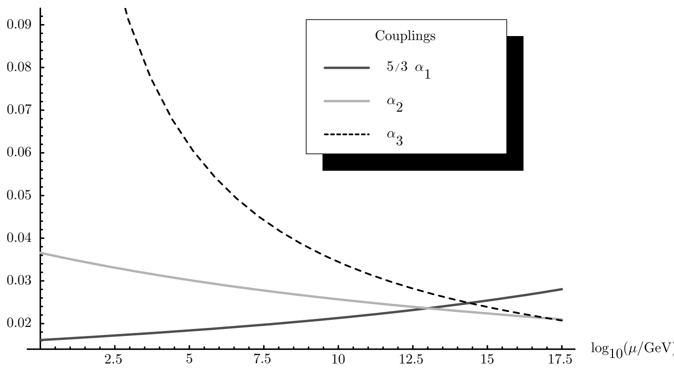
This gives the relation $\sin^2 \theta_W = \frac{3}{8}$, a value also obtained in $SU(5)$ and $SO(10)$ grand unified theories. The three moments of the function f_0, f_2 and f_4 can be used to specify the initial conditions on the gauge couplings, the Newton constant and the cosmological constant. We deduce that the geometrical picture is valid at high energies, and the spectral action must be considered in the Wilsonian approach, where all coupling constants are energy dependent and follow the renormalization group equations. For example, The fine structure constant α_{em} is given by

$$\alpha_{\mathrm{em}} = \, \sin(\theta_w)^2 \, \alpha_2 \, , \quad \alpha_i = \frac{g_i^2}{4\pi}.$$

Its infrared value is $\sim 1/137.036$ but it is running as a function of the energy and increases to the value $\alpha_{\mathrm{em}}(M_Z) = 1/128.09$ already at the energy $M_Z \sim 91.188$ Gev.

Assuming the "big desert" hypothesis, the running of the three couplings α_i is known. With 1-loop corrections only, it is given by [**13**]

$$\beta_{g_i} = (4\pi)^{-2} \, b_i \, g_i^3, \qquad \text{with} \quad b = (\frac{41}{6}, -\frac{19}{6}, -7),$$



so that

$$\begin{aligned}\alpha_1^{-1}(\Lambda) &= \alpha_1^{-1}(M_Z) - \frac{41}{12\pi}\log\frac{\Lambda}{M_Z}\\ \alpha_2^{-1}(\Lambda) &= \alpha_2^{-1}(M_Z) + \frac{19}{12\pi}\log\frac{\Lambda}{M_Z}\\ \alpha_3^{-1}(\Lambda) &= \alpha_3^{-1}(M_Z) + \frac{42}{12\pi}\log\frac{\Lambda}{M_Z},\end{aligned}$$

where M_Z is the mass of the Z^0 vector boson.

In fact, if one considers the actual experimental values

$$g_1(M_Z) = 0.3575, \quad g_2(M_Z) = 0.6514, \quad g_3(M_Z) = 1.221,$$

one obtains the values

$$\alpha_1(M_Z) = 0.0101, \quad \alpha_2(M_Z) = 0.0337, \quad \alpha_3(M_Z) = 0.1186.$$

The graphs of the running of the three constants α_i do not meet exactly, hence do not specify a unique unification energy.

Next we study the running of the Higgs quartic coupling λ [**14**]:

$$\frac{d\lambda}{dt} = \lambda\gamma + \frac{1}{8\pi^2}(12\lambda^2 + B),$$

where

$$\begin{aligned}\gamma &= \frac{1}{16\pi^2}(12k_t^2 - 9g_2^2 - 3g_1^2)\\ B &= \frac{3}{16}(3g_2^4 + 2g_1^2\, g_2^2 + g_1^4) - 3\,k_t^4\,.\end{aligned}$$

The Higgs mass is then given by

$$m_H^2 = 8\lambda\,\frac{M^2}{g^2}\,, \quad m_H = \sqrt{2\lambda}\,\frac{2M}{g}.$$

The numerical solution to these equations with the boundary value $\lambda_0 = 0.356$ at $\Lambda = 10^{17}$ Gev gives $\lambda(M_Z) \sim 0.241$ and a Higgs mass of the order of 170 Gev. This value will receive substantial corrections when gravitational loop effects are taken into account, and these will also effect the running of the gauge couplings.

The mass of the top quark is governed by the top quark Yukawa coupling k_t and is given by the equation

$$m_{\text{top}}(t) = \frac{1}{\sqrt{2}} \frac{2M}{g} k_t = \frac{1}{\sqrt{2}} v\, k_t,$$

where $v = \frac{2M}{g}$ is the vacuum expectation value of the Higgs field. There is a relation between the Yukawa and the gauge couplings which emerges as a consequence of normalizing the Higgs interactions. This relation is a consequence of the fact that all fermions get their masses by coupling to the same Higgs through interactions of the form

$$kH\overline{\psi}\psi.$$

After normalizing the kinetic energies of the Higgs field through the redefinition $H \to \frac{\pi}{\sqrt{af_0}} H$, the mass terms take the form

$$\frac{\pi}{\sqrt{f_0}} \frac{k}{\sqrt{a}} H\overline{\psi}\psi.$$

Using the identity $\sum_i \left(\frac{k_i}{\sqrt{a}}\right)^2 = 1$ gives a relation among the fermion masses and W mass [**2**]

$$\sum_{\text{generations}} m_e^2 + m_\nu^2 + 3m_d^2 + 3m_u^2 = 8M_W^2.$$

The value of g at a unification scale of 10^{17} Gev is ~ 0.517. Thus, neglecting the τ neutrino Yukawa coupling, we get the simplified relation

$$k_t = \frac{2}{\sqrt{3}} g \sim 0.597\ .$$

The numerical integration of the differential equation yields an acceptable value for the top quark mass of 179 Gev [**2**]. One reason why the resulting top quark mass is acceptable while the Higgs mass is not is because the latter is dependent on the cut-off function.

The fact that the coupling constants do not meet is giving us information about the nature of the function f used in the spectral action. Our results were obtained under the assumption that the function f is the cut-off function for which all coefficients of the higher order terms in the asymptotic expansion vanish. These coefficients are given by derivatives of the function evaluated at zero. We can infer from these results, especially from the near meeting of the coupling constants, the good approximate values for $\sin^2\theta$ and the top quark mass, that the function f is well approximated by the cut-off function, but deviates slightly from it. What is needed then is for the Taylor coefficients of the function to be very small but not zero.

To prove that this is indeed the case we compute the gauge and Higgs contributions to the next order, i.e. a_6, in the asymptotic expansion. It is enough to look only at the non-gravitational terms [**1**]:

$$
\begin{aligned}
&-\frac{f'(0)}{12\pi^2\Lambda^2}\left[c_1\overline{H}H\left(\frac{1}{4}g_2^2\left(W^\alpha_{\mu\nu}\right)^2\right)+c_2\overline{H}H\left(g_3^2\left(V^m_{\mu\nu}\right)^2\right)\right.\\
&\qquad+c_3\overline{H}\sigma^\alpha H\left(\frac{1}{2}g_1g_2B_{\mu\nu}W^\alpha_{\mu\nu}\right)+c_4\left(\overline{H}H\right)^3+c_5\left(\overline{H}H\right)^2\sigma^2\\
&\qquad+c_6\left(\left(\overline{H}\nabla_\mu H\right)^2+\left(\nabla_\mu\overline{H}H\right)^2\right)+c_7\left(\nabla_\mu\nabla_\nu\overline{H}\right)\left(\nabla_\mu\nabla_\nu H\right)\\
&\qquad+c_8\left(\overline{H}H\left|\nabla_\mu H\right|^2+\left|\overline{H}\nabla_\mu H\right|^2\right)+c_9\left|\nabla_\mu\left(H\sigma\right)\right|^2+c_{10}\left|\epsilon^{ab}H_a\nabla_\mu H_b\right|^2\\
&\qquad\left.+c_{11}\nabla_\mu\overline{H}\nabla_\nu H\left(\frac{3}{2}ig_1B_{\mu\nu}\right)+c_{12}\nabla_\mu\overline{H}\sigma^\alpha\nabla_\nu H\left(\frac{3}{2}ig_2W^\alpha_{\mu\nu}\right)\right]
\end{aligned}
$$

where the coefficients $c_1,\cdots,c_{12}$ depend only on the Yukawa couplings. The exact expression will be given in reference [**1**]. This clearly shows that the kinetic terms of the gauge fields get modified, and are all multiplied with the coefficients $f_{-2}=f'(0)$. The remarkable thing is that if we rescale the Higgs field by

$$H=\varphi\frac{\Lambda}{|k^t|},$$

assuming the top quark mass dominates the other fermion masses, then the potential will depend on Λ through an overall scale and the $|k^t|$ dependence drops out:

$$V=\frac{3\Lambda^4}{\pi^2}\left(-2f_2\overline{\varphi}\varphi+\frac{1}{2}f_0\left(\overline{\varphi}\varphi\right)^2+\frac{1}{3}f_{-2}\left(\overline{\varphi}\varphi\right)^3+\cdots\right).$$

Now since φ is a dimensionless doublet field, the vev

$$\langle\varphi\rangle=v\begin{pmatrix}0\\1\end{pmatrix},$$

will have a numerical value that depends only on the coefficients f_2, f_0

$$v_0^2=\frac{f_0}{2f_{-2}},$$

and will be perturbed very slightly by the higher coefficients $f_{-2},f_{-4},\ldots$, provided they decrease very rapidly. Looking at the minimum of the potential with the three terms above we have

$$v^2=\frac{f_0}{2f_{-2}}\left(-1+\sqrt{1+8\frac{f_2f_{-2}}{f_0^2}}\right).$$

Thus the condition that the higher order term in the potential perturb the minimum v_0 slightly requires the condition

$$f_{-2}\ll\frac{f_0^2}{8f_2},$$

so that

$$v^2\simeq v_0^2\left(1-4\frac{f_2f_{-2}}{f_0^2}\right).$$

We can get a rough estimate of the coefficients f_0 and f_2 at unification scale by setting

$$\frac{4f_2\Lambda^2}{\pi^2}=\frac{1}{2\kappa^2},\qquad\kappa=4.2\times10^{-19}\mathrm{Gev}^{-1}$$

which implies that

$$f_2 \simeq \left(\frac{\pi^2}{8}\right)\left(\frac{1}{\kappa\Lambda}\right)^2.$$

Thus, if Λ is of the order of M_{Planck} then $f_2 \sim 1$ while if $\Lambda \sim 10^{17}$ then $f_2 \sim 10^2$. We also have

$$\frac{f_0 g_3^2}{2\pi^2} = \frac{1}{4},$$

thus

$$f_0 = \frac{\pi}{8\alpha_s} \sim 20, \qquad \alpha_s = \frac{g_3^2}{4\pi}$$

at unification scale. Therefore we must have

$$f_{-2} \ll \frac{10^2}{f_2}$$

and this can be anywhere between 10^2 and 10^{-2} depending on whether Λ is at the Planck mass or two orders less.

We can now speculate on the form of the function $F(D^2) = f(D)$. This function must have rapidly decreasing Taylor coefficients (these are $F_0 = F(0)$, $F_{-2} = -F'(0)$, $F_{-4} = F''(0)\cdots$) while the Mellin coefficients F_2, F_4 should behave independently. The cut-off function can be approximated by the sequence $F_{\{N\}}(x)$

$$F_{\{N\}}(x) = A\left(1 + x + \frac{1}{2!}x^2 + \cdots + \frac{1}{N!}x^N\right)e^{-x}$$

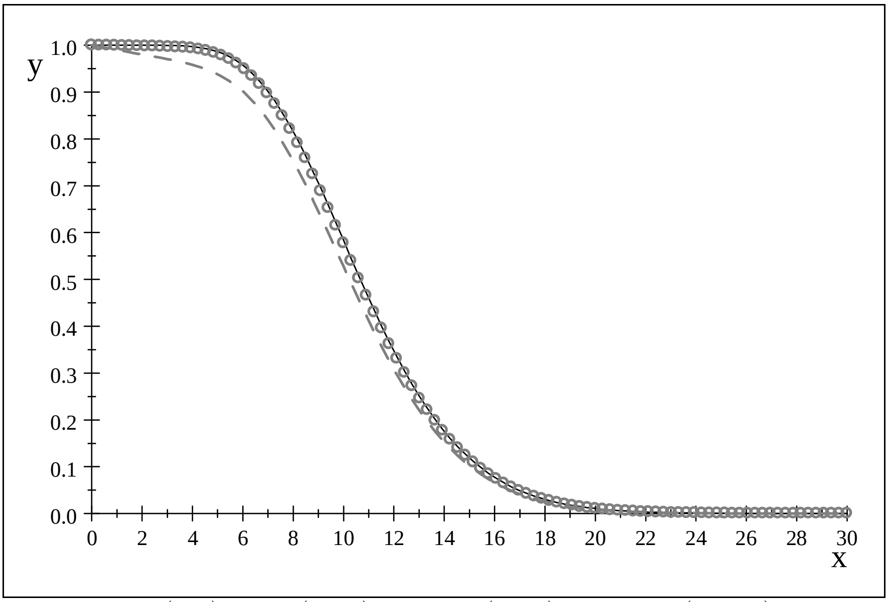
where

$$A \sim 20.$$

This function has the property that the first N coefficients in the Taylor expansion vanish, and is thus a very good approximation to a cut-off function. A slightly perturbed form of this function is given by

$$F_{\{N\}}(x,\epsilon) = e^{-\epsilon x}F_{\{N\}}(x)$$

where $\epsilon \leq \pm 10^{-2}$. In this case, we have $f_{-2} = A\epsilon$, $f_{-4} = A\epsilon^2$. To get a feeling for this function we can plot $F_{\{10\}}(x,\epsilon)$



$F_{\{10\}}(x,\epsilon)$ $\epsilon = 0$ (solid) $\epsilon = 0.01$ (dash) $\epsilon = 0.001$ (circles)

This shows that ϵ should be at least of order 10^{-2} to 10^{-3} in order not to disturb the cut-off function much, in the region where the scale is comparable to Λ. As seen from the plot, the function $F_N(x,\epsilon)$ is indistinguishable from $F_N(x)$ for $\epsilon \sim 10^{-3}$. From this we deduce that higher order terms in the heat kernel expansion will be suppressed by the Taylor coefficients of the function, and the perturbation can be trusted to within one order from the Planck scale. This property will ensure that the initial conditions on the RG equations for the gauge coupling constant get modified. To see this, we have, to lowest order, the modification to the gauge kinetic energies [**1**]:

$$\begin{aligned}
&\frac{f_{-2}v_0^2}{12\pi^2}\left[\left(\frac{17}{12}g_1^2B_{\mu\nu}^2\right)+\left(\frac{3}{4}g_2^2\left(W_{\mu\nu}^{\alpha}\right)^2\right)+\left(g_3^2\left(V_{\mu\nu}^{m}\right)^2\right)+\frac{9i}{4}g_1g_2^2B_{\mu\nu}W_\mu^+W_\nu^-\right.\\
&+\frac{1}{2}g_1g_2B_{\mu\nu}W_{\mu\nu}^3-\frac{3}{2}v^2\left(g_1B_\mu-g_2W_\mu^3\right)^2+6g_2^2W_\mu^+W_\mu^-+6v^2\left(g_1B_\mu-g_2W_\mu^3\right)^2\\
&+\frac{3}{2}g_2^2\left|\partial_\mu W_\nu^--\frac{i}{2}\left(g_1B_\mu-g_2W_\mu^3\right)W_\nu^--\frac{i}{2}W_\mu^-\left(g_1B_\nu-g_2W_\nu^3\right)\right|^2\\
&\left.+\frac{3}{4}\left|\partial_\mu\left(g_1B_\nu-g_2W_\nu^3\right)-ig_2^2W_\mu^+W_\nu^--\frac{i}{2}\left(g_1B_\mu-g_2W_\mu^3\right)\left(g_1B_\nu-g_2W_\nu^3\right)\right|^2\right]
\end{aligned}$$

It remains to show that this form, for some value of f_{-2}, can provide a mechanism for the unification of the three gauge couplings at some energy not far from the Planck scale. Similarly, the contributions to the Higgs potential are expected to modify the prediction of the Higgs mass [**15**]. The analysis of the running of the gauge coupling constants and the Higgs mass, taking these higher order terms into account, is presently under study. We hope to report on this in the near future.

7. Spectral Action for Noncommutative Spaces with Boundary

In the Hamiltonian quantization of gravity it is essential to include boundary terms in the action as this allows one to define consistently the momentum conjugate to the metric. This makes it necessary to modify the Einstein-Hilbert action by adding to it a surface integral term so that the variation of the action is well defined. The reason for this is that the curvature scalar R contains second derivatives of the metric, which are removed after integrating by parts to obtain an action which is quadratic in first derivatives of the metric. To see this note that the curvature $R \sim \partial\Gamma + \Gamma\Gamma$ where $\Gamma \sim g^{-1}\partial g$ has two parts; one part is of second order in derivatives of the form $g^{-1}\partial^2 g$ and the second part is the square of derivative terms of the form $\partial g \partial g$. To define the conjugate momenta in the Hamiltonian formalism, it is necessary to integrate by parts the term $g^{-1}\partial^2 g$ and change it to the form $\partial g \partial g$. These surface terms, which turned out to be very important, are canceled by modifying the Euclidean action to

$$I = -\frac{1}{16\pi}\int\limits_{M} d^4x\sqrt{g}R - \frac{1}{8\pi}\int\limits_{\partial M} d^3x\sqrt{h}K,$$

where ∂M is the boundary of M, h_{ab} is the induced metric on ∂M and K is the trace of the second fundamental form on ∂M. Notice that there is a relative factor of 2 between the two terms, and that the boundary term has to be completely fixed. This is a delicate fine tuning and is not determined by any symmetry, but only by the consistency requirement. There is no known symmetry that predicts this combination and it is always added by hand [**16**]. In contrast we can compute the spectral action for manifolds with boundary. The Hermiticity of the Dirac operator

$$(\psi|\, D\psi) = (D\psi|\, \psi),$$

is satisfied provided that $\pi_-\psi|_{\partial M} = 0$, where $\pi_- = \frac{1}{2}(1-\chi)$ is a projection operator on ∂M with $\chi^2 = 1$. To compute the spectral action for manifolds with boundary we have to specify the condition $\pi_- D\psi|_{\partial M} = 0$. The result of the computation gives the remarkable result that the Gibbons-Hawking boundary term is generated without any fine tuning [**17**]. Adding matter interactions does not alter the relative sign and coefficients of these two terms, even when higher orders are included. The Dirac operator for a product space such as that of the standard model must now be taken to be of the form

$$D = D_1 \otimes \gamma_F + 1 \otimes D_F,$$

instead of

$$D = D_1 \otimes 1 + \gamma_5 \otimes D_F,$$

because γ_5 does not anticommute with D_1 on ∂M.

8. Dilaton and the dynamical generation of scale

Replacing the cutoff scale Λ in the spectral action and replacing $f(\frac{D^2}{\Lambda^2})$ by $f(P)$ where $P = e^{-\phi}D^2e^{-\phi}$ modifies the spectral action with dilaton dependence to the form [**18**]

$$\mathrm{Tr}\ F(P) \simeq \sum_{n=0}^{6} f_{4-n}\int d^4x\sqrt{g}e^{(4-n)\phi}a_n\left(x, D^2\right).$$

One can then show that the dilaton dependence almost disappears from the action if one rescales the fields according to

$$\begin{aligned} G_{\mu\nu} &= e^{2\phi} g_{\mu\nu} \\ H' &= e^{-\phi} H \\ \psi' &= e^{-\frac{3}{2}\phi}\psi. \end{aligned}$$

With this rescaling one finds the result that the spectral action

$$I\left(g_{\mu\nu}, H, \psi, \phi\right) = I\left(G_{\mu\nu}, H', \psi', \phi = 0\right) + \frac{24 f_2}{\pi^2} \int d^4x \sqrt{G} G^{\mu\nu} \partial_\mu \phi \partial_\nu \phi$$

is scale invariant (independent of the dilaton field) except for the kinetic energy of the dilaton field ϕ. The dilaton field has no potential at the classical level. It acquires a Coleman-Weinberg potential [**19**] through quantum corrections, and thus a vev and a very small mass. [**20**]. The Higgs sector in this case becomes identical with the Randall-Sundrum model [**21**]. In that model there are two branes in a five-dimensional space, one located at $x_5 = 0$ representing the invisible sector, and another located at $x_5 = \pi r_c$, the visible sector. The physical masses are set by the symmetry breaking scale $v = v_0 e^{-k r_c \pi}$ so that $m = m_0 e^{-k r_c \pi}$. If the bare symmetry breaking scale is taken at $m_0 \sim 10^{19}$ Gev, then by taking $k r_c \pi = 10$ one gets the low-energy mass scale $m \sim 10^2$ Gev. It is not surprising that the Randall-Sundrum scenario is naturally incorporated in the noncommutative geometric model [**22**], [**23**], because intuitively one can think of the discrete space as providing the different right-handed and left-handed brane sectors.

9. Speculations on the Structure of the Noncommutative Space and Quantum Gravity

The small deviation from experimental results of the predictions of the standard model derived from the spectral action can have the following interpretation. This is an indication that the basic assumption we made about space-time as a product of a continuous four-dimensional manifold times a discrete space breaks down at energies just below the unification (Planck) scale. This will lead us to postulate that, at Planckian energies, the structure of space-time becomes noncommutative in a nontrivial way, which will change in an intrinsic way the particle spectrum. On the other hand, the encouraging results we obtained about the unique prediction of the spectrum of the standard model, the determination of the gauge group and the particle representations, can be taken as a guide that the true geometry should reproduce at lower energies, the product structure we assumed. The starting point is to look for a noncommutative space whose KO-dimension is ten (mod 8) and whose metric dimension is dictated by the growth of eigenvalues of the Dirac operator to be four. A good starting point would be to mesh in a smooth manner the four-dimensional manifold with the discrete space $M_2(\mathbb{H}) \oplus M_4(\mathbb{C})$. The appearance of 4×4 matrices and their relation to a four-dimensional space-time may not be accidental. In particular, we can define the four-dimensional manifold through the following data. The C^*-algebra is generated by $M_2(\mathbb{H})$ and a projection $e = e^2 = e^*$

such that [**24**]

$$\left\langle e-\frac{1}{2}\right\rangle =0$$

$$\left\langle \left(e-\frac{1}{2}\right)[D,e]^{2n}\right\rangle =\left\{\begin{array}{ll} 0, & n=0,1\\ \gamma, & n=2\end{array}\right\},$$

where γ is the chirality operator satisfying

$$\gamma^2=\gamma,\qquad \gamma=\gamma^*,\qquad \gamma e=e\gamma,\qquad D\gamma=-\gamma D$$

The constraint on e forces it to be of the form

$$e=\left(\begin{array}{cccc} \frac{1}{2}+t & 0 & \alpha & \beta\\ 0 & \frac{1}{2}+t & -\beta^* & \alpha^*\\ \alpha^* & -\beta & \frac{1}{2}-t & 0\\ \beta^* & \alpha & 0 & \frac{1}{2}-t\end{array}\right)$$

where t,α,α^*,β and β^* all commute and satisfy the relation

$$t^2+|\alpha|^2+|\beta|^2=\frac{1}{4}.$$

One can then check that $\mathcal{A}=C\left(S^4\right)$. The differential constraints are then satisfied by any Riemannian structure with a given volume form on S^4. This space can be deformed by considering the algebra to be generated by $M_4\left(\mathbb{C}\right)$ and e, where [**25**]

$$e=\left(\begin{array}{cc} q_{11} & q_{12}\\ q_{21} & q_{22}\end{array}\right)$$

and each q is a 2×2 matrix of the form

$$q=\left(\begin{array}{cc} \alpha & \beta\\ -\lambda\beta & \alpha^*\end{array}\right)$$

In this case the projection constraints imply

$$e=\left(\begin{array}{cccc} \frac{1}{2}+t & 0 & \alpha & \beta\\ 0 & \frac{1}{2}+t & -\lambda\beta^* & \alpha^*\\ \alpha^* & -\overline{\lambda}\beta & \frac{1}{2}-t & 0\\ \beta^* & \alpha & 0 & \frac{1}{2}-t\end{array}\right)$$

satisfying

$$\alpha\alpha^*=\alpha^*\alpha,\quad \beta\beta^*=\beta^*\beta,\quad \alpha\beta=\lambda\beta\alpha,\quad \alpha^*\beta=\overline{\lambda}\beta\alpha$$

giving rise to deformed S^4.

The idea now is to define the noncommutative space by marrying the concept of generating a manifold as instantonic solution of a set of equations, and to blend these with the finite space. For further details see [**1**].

10. Conclusions

We summarize the main assumptions made:

- Space-time is a product of a continuous four-dimensional manifold times a finite space.
- One of the algebras $M_4\left(\mathbb{C}\right)$ is subject to symplectic symmetry reducing it to $M_2\left(\mathbb{H}\right)$.
- The commutator of the Dirac operator with the center of the algebra is non-trivial, $[D,Z\left(\mathcal{A}\right)]\neq 0$.

- The unitary algebra $\mathcal{U}(\mathcal{A})$ is restricted to $\mathcal{SU}(\mathcal{A})$.

These give rise to the following results:

- The number of fundamental fermions is 16.
- The algebra of the finite space is $\mathbb{C} \oplus \mathbb{H} \oplus M_3(\mathbb{C})$.
- The correct representations of the fermions with respect to $SU(3) \times SU(2) \times U(1)$ are derived.
- The Higgs doublet appears as part of the inner fluctuations of the metric, and the spontaneous symmetry breaking mechanism appears naturally with the negative mass term without any tuning.
- Mass of the top quark of around 179 Gev.
- See-saw mechanism to give very light left-handed neutrinos.

The following problems are encountered:

- The unification of the gauge couplings with each other and with the Newton constant do not meet at one point which is expected to be one order below the Planck scale.
- The naive prediction of the mass of the Higgs field is around 170 Gev. This however, depends on the value of the gauge couplings at the unification scale and the form of the spectral function and higher order corrections.
- No new particles besides those of the Standard Model. This will be problematic if new physics is observed at LHC.
- No explanation of the number of generations.
- No constraints on the values of the Yukawa couplings which are the non-zero entries in the Dirac operator of the finite space.

From these results we can deduce the following:

- It is necessary to include the higher order corrections to the spectral action using a convergent series for the heat kernel expansion. This step is now done, and shows clearly that the corrections cannot be ignored if the spectral function deviates even slightly from the cut-off function. What remains to be done is to input these corrections into the RG equations and prove that this mechanism does produce gauge couplings unification, and thus will enable us to get an accurate prediction for the Higgs mass.
- To get an insight into the problem of quantum gravity, it is essential to find the noncommutative space whose limit is the product $M_4 \times F$. We speculated that this could be done by adopting the strategy of generating a continuous manifold through instantonic solutions of algebraic and differential constraints. This step has to be elaborated on and we must construct in detail the structure of such a space, to study its properties at the Planck scale and to show that the usual space-time can be recovered from the geometry of a non-trivial noncommutative space.
- The results obtained so far are very encouraging and we hope to report on future positive developments.

References

[1] A. Chamseddine and A. Connes, *Noncommutative geometry as a framework to unify all fundamental interactions including gravity, Part I,* Fortschr. Phys. **58** (2010) 553-600. Part II in preparation.

[2] A. Chamseddine, A. Connes, M. Marcolli, *Gravity and the standard model with neutrino mixing*, Adv. Theor. Math. **11** (2007) 991-1090.

[3] A. Chamseddine and A. Connes, *The uncanny precision of the spectral action*, Comm. Math. Phys. **293** (2010) 867-897.

[4] A. Connes, "Noncommutative Geometry", Academic Press, N.Y. 1994.

[5] A. Connes, *Gravity coupled with matter and the foundation of noncommutative geometry,* Comm. Math. Phys. **182** (1996) 155-176.

[6] A. Connes, *Noncommutative geometry and the standard model with neutrino mixing,* JHEP **0611:081** (2006)

[7] John Barrett, *The Lorentzian version of the noncommutative geometry model of particle physics*, *J. Math. Phys.* **48:** 012303 (2007).

[8] A. Chamseddine and A. Connes, *Why the standard model*, Jour. Geom. Phys. **58** (2008) 38-47.

[9] A. Chamseddine and A. Connes, *Conceptual explanation for the algebra in the noncommutative approach to the standard model*, Phys. Rev. Lett. **99** (2007) 191601.

[10] N. Bourbaki, *Lie groups and Lie algebras,* Chapters 7-9, Appendix II, page 385, Springer, Berlin 2000.

[11] A. Chamseddine and A. Connes, *The spectral action principle*, Comm. Math. Phys. **186** (1997), 731–750.

[12] P. Gilkey, *Invariance theory, the heat equation and the Atiyah-Singer index theorem,* Wilmington, Publish or Perish, 1984.

[13] For a review see G. Ross, *Grand unified theories,* Benjamin-Cummings, 1984.

[14] For a review see M. Sher, *Electroweak Higgs potential and vacuum stability,* Phys. Rep. **179** (1977) 273.

[15] G. Isidori, V. Rychkov, A. Sturmia and N. Tetradis, *Gravitational corrections to the standard model vacuum decay,* Phys. Rev. D**77** (2008) 025034.

[16] G. Gibbons and S. Hawking, *Action integrals and partition functions in quantum gravity,* Phys. Rev. D**15** (1977) 2752.

[17] A. Chamseddine and A. Connes, *Quantum gravity boundary terms from the spectral action of noncommutative space,* Phys. Rev. Lett. **99** (2007) 071302.

[18] A. Chamseddine and A. Connes, *Scale invariance in the spectral action,* Jour. Math. Phys. **47** (2006) 063504.

[19] S. Coleman and E. Weinberg, *Radiative corrections as the origin of spontaneous symmetry breaking*, *Phys. Rev.* **D7,** (1973) 1888.

[20] W. Buchmüller and C. Busch, *Symmetry breaking and mass bounds in the standard model with hidden scale invariance*, *Nucl. Phys.* **B 349** (1991) 71.

[21] L. Randall and R. Sundrum, *An alternative to compactification*, *Phys. Rev. Lett.* **83** (1999) 3370.

[22] A. Chamseddine and J. Fröhlich, *Constraints on the Higgs masses and top quark masses from the effective potential and noncommutative geometry*, *Phys. Lett.* **B 314** (1993) 308.

[23] F. Lizzi, G. Mangano and G. Miele, *Another alternative to compactification: noncommutative geometry and the Randall-Sundrum models, Mod. Phys. Lett.* **A16** (2001) 1.

[24] A. Connes, *A short survey of noncommutative geometry, J. Math. Phys.* **41** (2000) 3832.

[25] A. Connes and G. Landi, *Noncommutative manifolds: The instanton algebra and isospectral deformations,* Comm. Math. Phys. **221** (2001) 141.

Physics Department, American University of Beirut, Lebanon

Clay Mathematics Proceedings
Volume 11, 2010

The Regular C*-Algebra of an Integral Domain

Joachim Cuntz and Xin Li

Dedicated to Alain Connes on the occasion of his 60th birthday

ABSTRACT. To each integral domain R with finite quotients we associate a purely infinite simple C*-algebra in a very natural way. Its stabilization can be identified with the crossed product of the algebra of continuous functions on the "finite adèle space" corresponding to R by the action of the $ax+b$-group over the quotient field $Q(R)$. We study the relationship to generalized Bost-Connes systems and deduce for them a description as universal C*-algebras with the help of our construction.

CONTENTS

1. Introduction

In [**Cun1**], the first named author introduced C*-algebras $\mathcal{Q}_{\mathbb{Z}}$ and $\mathcal{Q}_{\mathbb{N}}$ associated with the ring of integers $\mathbb{Z}$ or also with the semiring $\mathbb{N}$, respectively, and which can be obtained from the natural actions of $\mathbb{Z}$ and $\mathbb{N}$, by multiplication and addition, on the Hilbert spaces $\ell^2(\mathbb{Z})$ and $\ell^2(\mathbb{N})$.

This was originally motivated by the well known construction by Bost and Connes [**BoCo**] who had introduced a C*-dynamical system $(\mathcal{C}_{\mathbb{Q}}, \sigma_t)$ and studied its KMS states.

2010 *Mathematics Subject Classification.* Primary: 46L55; Secondary: 11R04, 11M55, 11R56, 46L05.

Research supported by the Deutsche Forschungsgemeinschaft (SFB 478).

The Bost-Connes algebra $\mathcal{C}_{\mathbb{Q}}$ is naturally embedded into $\mathcal{Q}_{\mathbb{N}}$. The difference between the two algebras lies in the fact that $\mathcal{Q}_{\mathbb{N}}$ contains, besides the operators induced by multiplication in $\mathbb{N}$ also those corresponding to addition.

A main result in [**Cun1**] was the proof that the algebras $\mathcal{Q}_{\mathbb{Z}}$ and $\mathcal{Q}_{\mathbb{N}}$ are simple purely infinite and, after stabilization, can also be described as crossed products of the algebra of functions on the finite adèle space $\mathbb{A}_f$ over $\mathbb{Q}$ by the $ax+b$-groups over $\mathbb{Q}$ or $\mathbb{Q}^+$. This leads in particular to a simple presentation, by generators and relations, of the C*-algebras generated by the "left regular representations" of $\mathbb{Z}$ and $\mathbb{N}$.

In the present paper we extend the construction of [**Cun1**] to an arbitrary commutative ring R without zero divisors (an integral domain) subject to a finiteness condition which is typically satisfied by the integral domains considered in number theory (rings of integers in algebraic number fields or polynomial rings over finite fields). We denote the associated C*-algebra by $\mathfrak{A}[R]$. We generalize the result from [**Cun1**] by showing that $\mathfrak{A}[R]$ and its stabilization $\mathfrak{A}(R)$ are purely infinite simple and that $\mathfrak{A}(R)$ can be represented as a crossed product of the algebra of functions on the "finite adèle space", corresponding to the profinite completion $\hat{R}$ of R, by the action of the $ax+b$-group over the quotient field $Q(R)$. At the same time we streamline and improve the arguments given in [**Cun1**] in the case $R=\mathbb{Z}$.

We also show that the higher dimensional analogues of the Bost-Connes system, studied in [**CMR**] for imaginary quadratic number fields and in [**LLN**] for arbitrary number fields, embed into the C*-algebra $\mathfrak{A}[R]$ if the number field affords at most one real place and the class number is one. We use this to deduce a description of the algebras considered in [**CMR**], [**LLN**] (for number fields with at most one real place and class number one) in terms of generators and relations.

The main result of [**LLN**] was to construct and classify all KMS$_\beta$ states of generalized Bost-Connes systems which were originally introduced by Ha and Paugam in [**HaPa**]. Now, our description of the algebras arising in these C*-dynamical systems can be used to construct all extremal KMS$_\beta$ states in a very natural way (in complete analogy to the original case of $\mathbb{Q}$ treated in [**BoCo**]).

2. Universal C*-algebras

Throughout this article, R will denote an integral domain with the following properties:

1. the set of units R^* in R does not equal $R^\times := R\setminus\{0\}$ (so we exclude fields)
2. for each $m\in R^\times$ the ideal (m) generated by m in R is of finite index in R.

We will always think of R as a subring of its quotient field $Q(R)$.

Now, let us introduce our C*-algebras $\mathfrak{A}[R]$ in a universal way in terms of generators and relations. Later on, we will see more concrete models for $\mathfrak{A}[R]$.

Definition 2.1. Let $\mathfrak{A}[R]$ be the universal C*-algebra generated by isometries $\{s_m\colon m\in R^\times\}$ and unitaries $\{u^n\colon n\in R\}$ with the relations

(i) $s_k s_m = s_{km}$
(ii) $u^l u^n = u^{l+n}$
(iii) $s_m u^n = u^{mn} s_m$
(iv) $\sum_{n+(m)\in R/(m)} u^n e_m u^{-n} = 1$

for all $k,m\in R^\times$, $l,n\in R$, where $e_m = s_m s_m^*$ is the final projection corresponding to s_m.

The sum is taken over all cosets $n+(m)$ in $R/(m)$ and $u^n e_m u^{-n}$ is independent of the choice of n. This follows from (i) , (ii) and (iii) (once they are valid).

$\mathfrak{A}[R]$ exists as the generators must have norm 1. To show that this universal C*-algebra is not trivial, it suffices to give an explicit nontrivial representation of these generators and relations on a Hilbert space. For this purpose, we consider the "left regular representation" on the Hilbert space $\ell^2(R)$ given by the operators

$$\begin{aligned} S_m(\xi_r) &:= \xi_{rm} \\ U^n(\xi_r) &:= \xi_{r+n}. \end{aligned}$$

We immediately check the relations: (iii) reflects distributivity and (iv) holds since $U^n S_m S_m^* U^{-n}$ is the projection onto $\overline{\text{span}}\,(\{\xi_r\colon\, r \in n+(m)\})$ and R is the disjoint union of the cosets $\{n+(m)\colon\, n \in R\}$. Hence, the universal property provides a nontrivial representation via $s_m \mapsto S_m$, $u^n \mapsto U^n$.

In analogy to the case of groups, one can think of

$$\mathfrak{A}_r[R] := C^*\left(\left\{S_m\colon\, m \in R^\times\right\} \cup \{U^n\colon\, n \in R\}\right) \subseteq \mathcal{L}(\ell^2(R))$$

as the reduced C*-algebra associated to R.

Moreover, we define $P_R := R \rtimes R^\times$ where $R^\times$ acts on R via multiplication. P_R is called the $ax+b$-semigroup over R. This semigroup describes all affine transformations of R and thereby incorporates the ring structure of R. P_R is not a group as R is not a field. Furthermore, P_R can be realized as the subsemigroup $\{\left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right) \,:\, a \in R^\times\ ,\, b \in R\}$ in $M_2(R)$. For us, it is important to note that we have a natural representation of P_R in $\mathfrak{A}[R]$ given by $\left(\begin{smallmatrix} a & b \\ 0 & 1 \end{smallmatrix}\right) \longmapsto u^b s_a$.

3. The Inner Structure

In order to see that $\mathfrak{A}[R]$ is simple and purely infinite, we proceed similarly as in [**Cun1**]. This means that we construct a faithful conditional expectation out of certain group actions and describe this expectation with the help of appropriate projections (actually, this idea already appears in [**Cun2**]).

3.1. Preparations. We begin with some immediate consequences of the characteristic relations defining $\mathfrak{A}[R]$. First of all, the projections $u^n e_m u^{-n}$, $u^l e_m u^{-l}$ are orthogonal if $n+(m) \neq l+(m)$ because of (iv) . Denote by P the set of all these projections, $P = \{u^n e_m u^{-n} \,:\, m \in R^\times\ ,\, n \in R\}$. We have the following

Lemma 3.1. *The formula*

$$e_m = \sum_{n+(k) \in R/(k)} u^{mn} e_{mk} u^{-mn}$$

is valid for all k, $m \in R^\times$. Furthermore, the projections in P commute and span(P) is multiplicatively closed.

Proof. This follows by

$$\begin{aligned} e_m &= s_m 1 s_m^* \\ &= s_m\Big(\sum_{n+(k) \in R/(k)} u^n e_k u^{-n}\Big) s_m^* \\ &= \sum_{n+(k) \in R/(k)} u^{mn} e_{mk} u^{-mn}. \end{aligned}$$

Given two projections $u^n e_m u^{-n}$, $u^l e_k u^{-l}$, we can use the formula above to write both projections as sums of conjugates of $e_{mk} = e_{km}$. Hence it follows that they commute and that their product is in span(P). □

As span(P) is obviously a subspace closed under involution, we get that $C^*(P) = \overline{\text{span}}(P)$ is a commutative C*-subalgebra of $\mathfrak{A}[R]$. We denote it by $\mathfrak{D}[R]$ and investigate its structure later on.

Now we present the "standard form" of elements in the canonical dense subalgebra of $\mathfrak{A}[R]$.

Lemma 3.2. *Define* $S := \left\{ s_m^* u^n f u^{-n'} s_{m'} : m,\, m' \in R^\times \; ; n, n' \in R \; ; f \in P \right\}$. *span*$(S)$ *is the *-algebra in* $\mathfrak{A}[R]$ *generated by* $\{s_m : m \in R^\times\} \cup \{u^n : n \in R\}$.

Proof. Since S contains the generators and is a subset of the smallest *-algebra containing them, we just have to prove that span(S) is closed under multiplication (as it obviously is an involutive subspace). This follows from the following calculation:

$$\begin{aligned}
& s_m^* u^n f u^{-n'} s_{m'} \cdot s_k^* u^l e u^{-l'} s_{k'} \\
= \;& s_m^* u^n f u^{-n'} s_{m'} s_{m'}^* s_k^* s_{m'} u^l e u^{-l'} s_{k'} \\
= \;& s_m^* u^{n-n'} \underbrace{u^{n'} f u^{-n'}}_{=:\tilde{f}} e_{m'} s_k^* s_{m'} \underbrace{u^l e u^{-l}}_{=:\tilde{e}} u^{l-l'} s_{k'} \\
= \;& s_m^* u^{n-n'} s_k^* s_k \tilde{f} s_k^* s_k e_{m'} s_k^* s_{m'} \tilde{e} s_{m'}^* s_{m'} u^{l-l'} s_{k'} \\
= \;& s_{km}^* u^{kn-kn'} \underbrace{s_k \tilde{f} s_k^*}_{\in P} \underbrace{s_k e_{m'} s_k^*}_{\in P} \underbrace{s_{m'} \tilde{e} s_{m'}^*}_{\in P} u^{lm'-l'm'} s_{k'm'}.
\end{aligned}$$

As span(P) is closed under multiplication, we conclude that the same holds for span(S). □

3.2. A Faithful Conditional Expectation.

Proposition 3.3. There is a faithful conditional expectation

$$\Theta\colon\, \mathfrak{A}[R] \longrightarrow \mathfrak{D}[R]$$

characterized by

$$\Theta(s_m^* u^n f u^{-n'} s_{m'}) = \delta_{m,m'} \delta_{n,n'} s_m^* u^n f u^{-n} s_m$$

for all m, $m' \in R^\times$; n, $n' \in R$; $f \in P$.

Proof. Θ will be constructed as the composition of two faithful conditional expectations

$$\begin{aligned}
&\Theta_s\colon\, \mathfrak{A}[R] \longrightarrow C^* \left(\left\{ e_m\colon\, m \in R^\times \right\} \cup \{u^n\colon\, n \in R\} \right) \\
&\Theta_u\colon\, \Theta_s(\mathfrak{A}[R]) \longrightarrow \mathfrak{D}[R]
\end{aligned}$$

both arising from group actions on $\mathfrak{A}[R]$ or $\Theta_s(\mathfrak{A}[R])$ respectively.

1. Construction of Θ_s:

Consider the Pontrjagin dual group $\hat{G}$ of the discrete multiplicative group $G := (Q(R)^\times, \cdot)$ in the quotient field of R. To each character ϕ in $\hat{G}$ we assign the automorphism $\alpha_\phi \in \text{Aut}(\mathfrak{A}[R])$ given by $\alpha_\phi(s_m) = \phi(m) s_m$, $\alpha_\phi(u^n) = u^n$ for all

$m \in R^\times$, $n \in R$. The existence of α_ϕ is guaranteed by the universal property of $\mathfrak{A}[R]$. In this way, we get a group homomorphism

$$\begin{aligned} \hat{G} &\longrightarrow \mathrm{Aut}(\mathfrak{A}[R]) \\ \phi &\longmapsto \alpha_\phi \end{aligned}$$

which is continuous for the point-norm topology.

It is known that Θ_s defined by

$$\Theta_s(x) = \int_{\hat{G}} \alpha_\phi(x) d\mu(\phi)$$

is a faithful conditional expectation from $\mathfrak{A}[R]$ onto the fixed-point algebra $\mathfrak{A}[R]^{\hat{G}}$, where μ is the normalized Haar measure on the compact group $\hat{G}$ (see [**Bla**], II.6.10.4 (v)).

It will be useful to determine $\mathfrak{A}[R]^{\hat{G}}$ more precisely. In order to do so let us calculate

$$\begin{aligned} & \Theta_s(s_m^* u^n f u^{-n'} s_{m'}) \\ = & \int_{\hat{G}} \alpha_\phi(s_m^* u^n f u^{-n'} s_{m'}) d\mu(\phi) \\ = & \left(\int_{\hat{G}} \phi(m^{-1}m') d\mu(\phi) \right) s_m^* u^n f u^{-n'} s_{m'} \\ = & \delta_{m,m'} s_m^* u^n f u^{-n'} s_{m'}. \end{aligned}$$

Thus, $\mathfrak{A}[R]^{\hat{G}} = \Theta_s(\mathfrak{A}[R]) = \overline{\mathrm{span}}(\left\{ s_m^* u^n f u^{-n'} s_m \colon m \in R^\times;\, n,\, n' \in R;\, f \in P \right\}$ as $\mathfrak{A}[R] = \overline{\mathrm{span}}(\left\{ s_m^* u^n f u^{-n'} s_{m'} \colon m,\, m' \in R^\times;\, n,\, n' \in R;\, f \in P \right\}$ by Lemma 3.2. But we can do even better, claiming

$$\mathfrak{A}[R]^{\hat{G}} = \overline{\mathrm{span}}(\left\{ u^n e_m u^{-n'} \colon m \in R^\times;\, n,\, n' \in R; \right\}),$$

because we have

$$\begin{aligned} & s_m^* u^n e_k u^{-n'} s_m \\ = & s_m^* e_m u^n e_k u^{-n} u^{n-n'} s_m \\ = & s_m^* \sum_{l+(k) \in R/(k)} u^{lm} e_{km} u^{-lm} u^n \sum_{i+(m) \in R/(m)} u^{ik} e_{km} u^{-ik} u^{-n} u^{n-n'} s_m \\ = & s_m^* \sum_a u^{am} e_{km} u^{-am} u^{n-n'} s_m \\ = & \sum_a u^a e_k u^{-a} s_m^* u^{n-n'} s_m. \end{aligned}$$

where the sums are taken over appropriate indices a (this being justified by Lemma 3.1).

Additionally,

$$\begin{aligned} s_m^* u^{n-n'} s_m &= s_m^* e_m u^{n-n'} e_m s_m \\ = & \begin{cases} 0 \text{ if } n - n' \notin (m) \\ u^{m^{-1}(n-n')} \text{ if } n - n' \in (m) \end{cases} \end{aligned}$$

so that each $s_m^* u^n f u^{-n'} s_m$ lies in $\overline{\mathrm{span}}(\left\{u^n e_m u^{-n'}\colon m \in R^\times;\ n,\ n' \in R\right\})$. This implies $\mathfrak{A}[R]^{\hat{G}} = C^*(\{e_m\colon m \in R^\times\} \cup \{u^n\colon n \in R\})$.

2. Construction of Θ_u:

Defining $H := (R, +)$, each $\chi \in \hat{H}$ gives an automorphism $\beta_\chi \in \mathrm{Aut}(\Theta_s(\mathfrak{A}[R]))$ with the properties $\beta_\chi(e_m) = e_m$ and $\beta_\chi(u^n) = \chi(n)u^n$. To see the existence of β_χ, we fix $m \in R^\times$ and consider $C^*(\{u^n\colon n \in R\}, e_m)$.

Lemma 3.4. *$C^*(\{u^n\colon n \in R\}, e_m)$ is the universal C^*-algebra generated by unitaries $\{v^n\colon n \in R\}$ and one projection f_m such that*

$$\begin{aligned} v^n v^{n'} &= v^{n+n'} \\ \sum_{n+(m) \in R/(m)} v^n f_m v^{-n} &= 1, \end{aligned}$$

the latter relation implicitly including $v^{lm} f_m = f_m v^{lm}$ for all $l \in R$.

Proof. The universal C*-algebra given by these generators and relations can be faithfully represented on a (necessarily infinite-dimensional) Hilbert space. Then, we can identify this algebra with $M_p(C^*(\{v^n\colon n \in R\}))$ where $p := \#[R/(m)]$. The isomorphism is provided by the p pairwise orthogonal projections $v^n f_m v^{-n}$ each being equivalent to 1 (where $\{n + (m)\} = R/(m)$). Now the same argument shows $C^*(\{u^n\colon n \in R\}, e_m) \cong M_p(C^*(\{u^n\colon n \in R\}))$. Thus it remains to show that

$$C^*(\{v^n\colon n \in R\}) \longrightarrow C^*(\{u^n\colon n \in R\}); v^n \longmapsto u^n$$

is an isomorphism. This follows by the following observations:

For each n, $\mathrm{Sp}(u^n)$ is maximal, meaning that it is $\mathbb{T}$ if $\mathrm{char}(R) = 0$ and $\{\zeta \in \mathbb{T}\colon \zeta^p = 1\}$ if $\mathrm{char}(R) = p$ (in this case we have $(u^n)^p = 1$ for all $n \in R$). This follows from the "left regular representation" of $\mathfrak{A}[R]$ discussed above. Therefore, $\mathrm{Sp}(v^n) = \mathrm{Sp}(u^n)$ for all $n \in R$.

Given $n_1, ..., n_i \in R$, we have $C^*(\{v^{n_1}, \dots, v^{n_i}\}) \cong C^*(\{u^{n_1}, \dots, u^{n_i}\})$. To see this, we can assume that the n_1, ..., n_i are linearly independent over the prime ring of R, so that we get $\mathrm{Spec}(C^*(\{u^{n_1}, \dots, u^{n_i}\}) \cong \mathrm{Sp}(u^{n_1}) \times ... \times \mathrm{Sp}(u^{n_i})$ which can be identified with $\mathrm{Sp}(v^{n_1}) \times ... \times \mathrm{Sp}(v^{n_i}) \cong \mathrm{Spec}(C^*(\{v^{n_1}, \dots, v^{n_i}\})$.

Now the claim follows by taking the inductive limit of the isomorphisms obtained via the identification of these spectra, and we again get an isomorphism sending v^n to u^n. □

This Lemma yields automorphisms $\beta_{\chi,m} \in \mathrm{Aut}(C^*(\{u^n\colon n \in R\}, e_m))$ given by $\beta_{\chi,m}(e_m) = e_m$, $\beta_{\chi,m}(u^n) = \chi(n)u^n$. Now, since $\beta_{\chi,km}|_{C^*(\bigcup\{u^n\colon n \in R\}, e_m)} = \beta_{\chi,m}$, β_χ can be constructed as the inductive limit of the $\beta_{\chi,m}$. Here, we use $C^*(\{e_m\colon m \in R^\times\} \cup \{u^n\colon n \in R\}) = \varinjlim C^*(\{u^n\colon n \in R\}, e_m)$ with the inclusions $C^*(\{u^n\colon n \in R\}, e_m) \hookrightarrow C^*(\{u^n\colon n \in R\}, e_{km})$ given by Lemma 3.1.

Clearly, $\hat{H}$ acts on $\Theta_s(\mathfrak{A}[R])$ via $\chi \longmapsto \beta_\chi$, which is again continuous for the point-norm topology. So we can proceed just as before defining

$$\Theta_u(y) = \int_{\hat{H}} \beta_\chi(y) d\mu(\chi),$$

and an analogous calculation shows $\Theta_u(u^n e_m u^{-n'}) = \delta_{n,n'} u^n e_m u^{-n}$. Hence it follows that $(\Theta_s(\mathfrak{A}[R]))^{\hat{H}} = \left(\mathfrak{A}[R]^{\hat{G}}\right)^{\hat{H}} = \mathfrak{D}[R]$.

As mentioned at the beginning, we set $\Theta := \Theta_u \circ \Theta_s$ which obviously yields a faithful conditional expectation with the property

$$\Theta(s_m^* u^n f u^{-n'} s_{m'}) = \Theta_u(\delta_{m,m'} s_m^* u^n f u^{-n'} s_m) = \delta_{m,m'}\delta_{n,n'} s_m^* u^n f u^{-n} s_m$$

□

In the following we want to give an alternative description of Θ with the help of sufficiently small projections. Let y be in $\mathrm{span}(S)$, which means

$$y = \sum_{m,m',n,n',f} a_{(m,m',n,n',f)} s_m^* u^n f u^{-n'} s_{m'}.$$

In this sum, there are only finitely many projections lying in P which appear with nontrivial coefficients. Write them as sums of mutually orthogonal projections $u^{n_1} e_M u^{-n_1}, \dots , u^{n_N} e_M u^{-n_N}$.

PROPOSITION 3.5. There are N pairwise orthogonal projections f_i in P such that

I.

$$\begin{aligned} \Phi : C^*(\{u^{n_1} e_M u^{-n_1}, \dots , u^{n_N} e_M u^{-n_N}\}) &\longrightarrow C^*(\{f_1, \dots , f_N\}) \\ z &\longmapsto \sum_{i=1}^{N} f_i z f_i \end{aligned}$$

is an isomorphism.

II. $\Phi(\Theta(y)) = \sum_{i=1}^{N} f_i y f_i$

PROOF. We will find appropriate ν_i and μ so that $f_i := u^{\nu_i} e_\mu u^{-\nu_i}$ satisfies I. and II.

As a first step, the conditions

$$\begin{aligned} &\nu_i + (M) = n_i + (M) \text{ for all } 1 \le i \le N \\ &\mu \in (M) \end{aligned}$$

enforce mutual orthogonality and imply I. as we have for $\lambda = M^{-1}\mu$ (in R by the second condition)

$$\begin{aligned} & f_i u^{n_j} e_M u^{-n_j} f_i \\ = \;& f_i \sum_{l+(\lambda)\in R/(\lambda)} u^{n_j + lM} e_\mu u^{-n_j - lM} f_i \\ = \;& \delta_{i,j} f_i \end{aligned}$$

because

$$\begin{aligned} & f_i u^{n_j + lM} e_\mu u^{-n_j - lM} \neq 0 \text{ for some } l \in R \\ \Leftrightarrow \;& \nu_i + (\mu) = n_j + lM + (\mu) \text{ for some } l \in R \\ \Leftrightarrow \;& \nu_i + (M) = n_j + (M) (\text{since } \mu \in (M)) \\ \Leftrightarrow \;& i = j (\text{as } \nu_i + (M) = n_i + (M) \neq n_j + (M) \text{ for all } i \neq j). \end{aligned}$$

Therefore, Φ maps $u^{n_i} e_M u^{-n_i}$ to f_i and is thus an isomorphism.

To find sufficient conditions on ν_i and μ for II., let us consider those summands in y with $a_{(m,m',n,n',f)} \neq 0$ and $\delta_{m,m'}\delta_{n,n'} = 0$. Call the corresponding indices (m, m', n, n', f) critical; there are only finitely many of them. As Θ maps such summands to 0, we have to ensure that $f_i s_m^* u^n f u^{-n'} s_{m'} f_i = 0$ for those critical indices.

We have

$$\begin{aligned} & f_i s_m^* u^n f u^{-n'} s_{m'} f_i \\ = \quad & s_m^* u^n \left(u^{-n} s_m f_i s_m^* u^n\right) f \left(u^{-n'} s_{m'} f_i s_{m'}^* u^{n'}\right) u^{-n'} s_{m'} \\ = \quad & s_m^* u^n (u^{m\nu_i - n} e_{m\mu} u^{-m\nu_i + n} u^{m'\nu_i - n'} e_{m'\mu} u^{-m'\nu_i + n'}) f u^{-n'} s_{m'} \end{aligned}$$

and the term in brackets can be described as

$$\begin{aligned} & u^{m\nu_i - n} e_{m\mu} u^{-m\nu_i + n} u^{m'\nu_i - n'} e_{m'\mu} u^{-m'\nu_i + n'} \\ = \quad & \sum_{a+(m')\in R/(m')} u^{-n+m\nu_i+am\mu} e_{mm'\mu} u^{n-m\nu_i-am\mu} \\ \cdot \quad & \sum_{b+(m)\in R/(m)} u^{-n'+m'\nu_i+bm'\mu} e_{mm'\mu} u^{n-m'\nu_i-bm'\mu} \end{aligned}$$

Now we see that the projections in these two sums are pairwise orthogonal if

$$\begin{aligned} & -n + m\nu_i + am\mu + (mm'\mu) \neq -n' + m'\nu_i + bm'\mu + (mm'\mu) \\ \Leftrightarrow \quad & n - n' + \nu_i(m' - m) + (bm' - am)\mu \notin (mm'\mu) \text{ for all } a, b \text{ in } R. \end{aligned}$$

This can be enforced by the even stronger condition

$$n - n' + \nu_i(m' - m) \notin (\mu),$$

which we have to satisfy for each critical index simultaneously.

On the whole, the projections f_i satisfy I. and II. if ν_i and μ have the three properties

- $\nu_i + (M) = n_i + (M)$ for all $1 \leq i \leq N$
- $\mu \in (M)$
- $n - n' + \nu_i(m' - m) \notin (\mu)$ for all critical indices.

One could, for example, choose ν_i such that $\nu_i + (M) = n_i + (M)$ for all $1 \leq i \leq N$ and $n - n' + \nu_i(m' - m) \neq 0$ for all critical indices. This can be simultaneously done as there are infinitely many possibilities for the ν_i to satisfy the first condition, while the second one only excludes finitely many (namely $-(m' - m)^{-1}(n - n')$ for all critical indices with $m \neq m'$, otherwise this condition is automatically valid as $\delta_{m,m'}\delta_{n,n'} = 0$). Then just take an element $r \in R^\times$ which is not invertible and set

$$\mu := rM \prod [n - n' + \nu_i(m' - m)] \in R^\times$$

where the product is taken over all critical indices. It is immediate that this choice of μ enforces the second and third condition. □

3.3. Purely Infinite Simple C*-algebras. With the help of these ingredients it is now possible to prove the following result:

THEOREM 3.6. *$\mathfrak{A}[R]$ is simple and purely infinite, i.e. for all $0 \neq x \in \mathfrak{A}[R]$ there are a, b in $\mathfrak{A}[R]$ with $axb = 1$.*

PROOF. Consider first a positive, nontrivial element x in $\mathfrak{A}[R]$. Recall that we have constructed a faithful conditional expectation Φ in Proposition 3.3. As $\Theta(x) \neq 0$ we can assume $\|\Theta(x)\| = 1$. Since $\operatorname{span}(S)$ is dense in $\mathfrak{A}[R]$ (compare

Lemma 3.2) we can find $y \in \operatorname{span}(S)_+$ with $\|x - y\| < \frac{1}{2}$, $\|\Theta(y)\| = 1$. Proposition 3.5 gives pairwise orthogonal projections f_i and Φ depending on y with

$$\Phi(\Theta(y)) = \sum_{i=1}^{N} f_i y f_i = \sum_{i=1}^{N} \lambda_i f_i$$

for some nonnegative λ_i, as we know that $\Phi(\Theta(y))$ lies in $C^*(\{f_1, \dots , f_N\})$ and that $\Theta(y)$ is positive. Since Φ is isometric, we have $1 = \|\Phi(\Theta(y))\|$. Thus, there must be an index j with $\lambda_j = 1$, as $\|\sum_{i=1}^{N} \lambda_i f_i\| = \sup_{1 \leq i \leq N} \lambda_i$. Consider the isometry $s := u^{\nu_j} s_\mu$. It has the properties $ss^* = f_j$ and $s^* f_j s = s^* ss^* s = 1$, so that

$$s^* y s = s^* f_j s s^* y s s^* f_j s = s^* f_j^2 y f_j^2 s = s^* f_j s = 1. \tag{3.1}$$

Therefore, we conclude that

$$\|s^* x s - 1\| = \|s^*(x - y)s\| < \frac{1}{2}.$$

This implies that $s^* x s$ is invertible in $\mathfrak{A}[R]$.

Set $a := (s^* x s)^{-1} s^*$ and $b := s$; this gives $axb = (s^* x s)^{-1} s^* x s = 1$ as claimed.

Given an arbitrary nontrivial element x, we get by the same argument as above, applied to $x^* x$, elements a' and b' with $a' x^* x b' = 1$. Then, we can set $a := a' x^*$ and $b := b'$. □

Remark 3.7. An immediate consequence is the fact that every nontrivial C*-algebra generated by unitaries and isometries satisfying the characteristic relations is canonically isomorphic to $\mathfrak{A}[R]$.

As a special case of this observation, we get $\mathfrak{A}_r[R] \cong \mathfrak{A}[R]$.

4. Representation as a Crossed Product

This section is about representing $\mathfrak{A}[R]$ as a crossed product involving some kind of a generalized finite adèle ring and the $ax + b$-group $P_{Q(R)}$.

4.1. Ring-theoretical Constructions. We start with some ring-theoretical constructions. Set

$$\hat{R} = \varprojlim \{R/(m); p_{lm,m}\}$$

where $p_{lm,m}$: $R/(lm) \longrightarrow R/(m)$ is the canonical projection. This is the profinite completion of R.

A concrete description would be

$$\hat{R} = \left\{ (r_m)_m \in \prod_{m \in R^\times} R/(m) : p_{lm,m}(r_{lm}) = r_m \right\}$$

with the induced topology of the product $\prod_m R/(m)$, each finite Ring $R/(m)$ carrying the discrete topology. $\hat{R}$ is a compact ring with addition and multiplication defined componentwise. Furthermore, we have the diagonal embedding

$$R \hookrightarrow \hat{R}; r \longmapsto (r)_m$$

and we will identify R with a subring of $\hat{R}$ via this embedding.

Moreover, for $l \in R^\times$ we have the canonical projection $\hat{R} \twoheadrightarrow R/(l)$. Its kernel equals $l\hat{R}$ as those elements are mapped to 0, while an element $(r_m)_m$ in the kernel can be written as $l \cdot (l^{-1} r_{lm})_m \in l\hat{R}$. Therefore, we get an isomorphism $\hat{R}/l\hat{R} \cong R/(l)$.

As a next step, set

$$\mathscr{R} := \varinjlim \{\mathscr{R}_m;\ \phi_{m,lm}\}$$

where $\mathscr{R}_m = \hat{R}$ for all $m \in R^\times$ and $\phi_{m,lm}$ is multiplication by l.

An explicit picture for $\mathscr{R}$ is

$$\coprod_{m \in R^\times} \mathscr{R}_m / \sim$$

where $x_l \sim y_m \Leftrightarrow m x_l = l y_m$ for $x_l \in \mathscr{R}_l$, $y_m \in \mathscr{R}_m$. Denote by p the canonical projection $\coprod_{m \in R^\times} \mathscr{R}_m \twoheadrightarrow \mathscr{R}$ and by ι_m the embedding

$$\hat{R} \to \mathscr{R}_m \to \mathscr{R}; x \longmapsto p(x).$$

$\mathscr{R}$ is a locally compact ring via

$$\iota_m(x) + \iota_l(y) = \iota_{lm}(lx + my), \iota_m(x) \cdot \iota_l(y) = \iota_{lm}(xy).$$

Again, we identify $\hat{R}$ with a subring of $\mathscr{R}$ via ι_1.

An immediate observation is the fact that $\iota_m(\hat{R})$ is compact and open in $\mathscr{R}$. Compactness is clear as ι_m is continuous and $\hat{R}$ is compact. Furthermore,

$$\begin{aligned} & \mathscr{R}_l \cap p^{-1}(\iota_m(\hat{R})) \\ = \ & \{x_l \in \mathscr{R}_l \colon x_l \sim y_m \text{ for some } y_m \in \mathscr{R}_m\} \\ = \ & \phi_{l,lm}^{-1}(l\hat{R}) \end{aligned}$$

and $l\hat{R}$ is open in $\hat{R}$ because $\hat{R} \backslash l\hat{R} = \bigcup_{r+(l) \neq (l)} r + l\hat{R}$ using the isomorphism $\hat{R}/l\hat{R} \cong R/(l)$. Therefore, $\hat{R} \backslash l\hat{R}$ is a finite union of compact sets, thus closed.

4.2. Description of the Algebra. With these preparations, we can establish connections with the C*-algebra $\mathfrak{A}[R]$.

Observation 4.1. $\mathfrak{D}[R] \cong C(\hat{R})$ via $u^n e_m u^{-n} \longmapsto p_{m\hat{R}+n}$, where $p_{m\hat{R}+n}$ denotes the characteristic function on the compact and open subset $m\hat{R} + n \subseteq \hat{R}$.

Proof. Consider $D_m = C^*(\{u^n e_m u^{-n} \colon n \in R/(m)\})$ together with the inclusions $D_m \hookrightarrow D_{lm}$. $\mathfrak{D}[R]$ can be described as the inductive limit of this system. Furthermore, $\mathrm{Spec}(D_m) \cong R/(m)$ as the projections $u^n e_m u^{-n}$ are mutually orthogonal. Moreover,

$$\mathrm{Spec}(D_{lm}) \longrightarrow \mathrm{Spec}(D_m); \chi \longmapsto \chi|_{D_m}$$

corresponds to

$$p_{lm,m} \colon R/(lm) \longrightarrow R/(m); r + (lm) \longmapsto r + (m)$$

via this identification. Therefore, we have $\mathrm{Spec}(D) \cong \varprojlim \{R/(m); p_{lm,m}\} = \hat{R}$. Thus we get the isomorphism

$$\alpha \colon \mathfrak{D}[R] \longrightarrow C(\hat{R}); u^n e_m u^{-n} \longmapsto p_{m\hat{R}+n}.$$

□

Definition 4.2. The stabilization of $\mathfrak{A}[R]$, denoted by $\mathfrak{A}(R)$, is defined as the inductive limit of the system $\{\mathfrak{A}(R)_m;\ \varphi_{m,lm}\}$ where we take $\mathfrak{A}(R)_m = \mathfrak{A}[R]$ and $\varphi_{m,lm} \colon \mathfrak{A}[R] \longrightarrow \mathfrak{A}[R]$ given by $x \longmapsto s_l x s_l^*$.

Furthermore, we set $\mathfrak{D}(R) = \varinjlim \{\mathfrak{D}(R)_m; \varphi_{m,lm}\}$ with $\mathfrak{D}(R)_m = \mathfrak{D}[R]$ and $\varphi_{m,lm}$ just defined as above. $\mathfrak{D}(R)$ can obviously be identified with a C*-subalgebra of $\mathfrak{A}(R)$.

OBSERVATION 4.3. We have $\mathfrak{D}(R) \cong C_0(\mathscr{R})$.

PROOF. The maps $\varphi_{m,lm}$, conjugated by α (see Observation 4.1), give maps

$$\psi_{m,lm} := \alpha \circ \varphi_{m,lm} \circ \alpha^{-1} \colon C(\hat{R}) \longrightarrow C(\hat{R})$$

where $\psi_{m,lm}(f)(x) = f(l^{-1}x)p_{l\hat{R}}(x)$. This follows from the calculation

$$\begin{aligned} & \psi_{m,lm} \circ \alpha(u^n e_m u^{-n})(x) \\ = \; & \psi_{m,lm}(p_{m\hat{R}+n})(x) \\ = \; & p_{m\hat{R}+n}(l^{-1}x)p_{l\hat{R}}(x) \\ = \; & p_{lm\hat{R}+ln}(x) \\ = \; & \alpha(u^{ln}e_{lm}u^{-ln})(x) \\ = \; & \alpha \circ \varphi_{m,lm}(u^n e_m u^{-n})(x). \end{aligned}$$

This yields an isomorphism $\mathfrak{D}(R) \longrightarrow \varinjlim \left\{ C(\hat{R}); \psi_{m,lm} \right\}$.

Additionally, we consider homomorphisms

$$\kappa_k \colon C(\hat{R}) \longrightarrow C_0(\mathscr{R}); f \longmapsto f \circ \iota_k^{-1} \cdot p_{\iota_k(\hat{R})}.$$

They satisfy $\kappa_{lm} \circ \varphi_{m,lm} = \kappa_m$ because

$$\begin{aligned} & \kappa_{lm} \circ \varphi_{m,lm}(f)(x) \\ = \; & \varphi_{m,lm}(f)(\iota_{lm}^{-1}(x))p_{\iota_{lm}(\hat{R})}(x) \\ = \; & f(l^{-1}\iota_{lm}^{-1}(x)) \cdot p_{\iota_{lm}(l\hat{R})}(x) \\ = \; & f(\iota_m^{-1}(x))p_{\iota_m}(x) \\ = \; & \kappa_m(f)(x). \end{aligned}$$

Hence these homomorphisms give rise to a homomorphism

$$\varinjlim \left\{ C(\hat{R}); \psi_{m,lm} \right\} \longrightarrow C_0(\mathscr{R})$$

which is injective as each κ_k is injective because of $\kappa_k(f) \circ \iota_k = f$ and surjective because of $\mathscr{R} = \bigcup_{m \in R^\times} \iota_m(\hat{R})$ and Stone-Weierstrass. □

Finally, we come to the already mentioned picture of $\mathfrak{A}(R)$.

THEOREM 4.4. *$\mathfrak{A}(R)$ is isomorphic to $C_0(\mathscr{R}) \rtimes P_{Q(R)}$ where the $ax+b$-group acts on $\mathscr{R}$ via affine transformations.*

PROOF. The first step is the observation that we have a canonical isomorphism

$$p_{\hat{R}}(C_0(\mathscr{R}) \rtimes P_{Q(R)})p_{\hat{R}} \cong \mathfrak{A}[R]$$

denoted by β:

To this end, consider $u^n := V_{\left(\begin{smallmatrix} 1 & n \\ 0 & 1 \end{smallmatrix}\right)} p_{\hat{R}}$ and $s_m := V_{\left(\begin{smallmatrix} m & 0 \\ 0 & 1 \end{smallmatrix}\right)} p_{\hat{R}}$. Here, $V_{\left(\begin{smallmatrix} 1 & n \\ 0 & 1 \end{smallmatrix}\right)}$ and $V_{\left(\begin{smallmatrix} m & 0 \\ 0 & 1 \end{smallmatrix}\right)}$ denote the unitaries in $\mathcal{M}(C_0(\mathscr{R}) \rtimes P_{Q(R)})$ implementing the action of $P_{Q(R)}$. u^n and s_m are unitaries and isometries in $p_{\hat{R}}(C_0(\mathscr{R}) \rtimes P_{Q(R)})p_{\hat{R}}$ satisfying the characteristic relations of $\mathfrak{A}[R]$. Furthermore, $u^n s_m s_m^* u^{-n} = p_{m\hat{R}+n}$ so that

$C^*(\{u^n s_m s_m^* u^{-n}\colon m,\, m' \in R^\times;\, n \in R\})$ is a closed C*-subalgebra of $C(\hat{R})$ separating points and thus equal to $C(\hat{R}) = p_{\hat{R}} C_0(\mathscr{R})$ by Stone-Weierstrass. Hence it follows that $p_{\hat{R}}(C_0(\mathscr{R}) \rtimes P_{Q(R)}) p_{\hat{R}}$ is the C*-algebra generated by the u^n and s_m and thus isomorphic to $\mathfrak{A}[R]$ by Remark 3.7.

Secondly, define

$$\tilde{\varphi}_{m,lm}\colon p_{\hat{R}}(C_0(\mathscr{R}) \rtimes P_{Q(R)}) p_{\hat{R}} \longrightarrow p_{\hat{R}}(C_0(\mathscr{R}) \rtimes P_{Q(R)}) p_{\hat{R}}$$

to be conjugation by $V_{\left(\begin{smallmatrix} l & 0 \\ 0 & 1 \end{smallmatrix}\right)}$. It is clear that we have $\beta \circ \varphi_{m,lm} \circ \beta^{-1} = \tilde{\varphi}_{m,lm}$, thus an isomorphism

$$\mathfrak{A}(R) \longrightarrow \varinjlim \left\{ p_{\hat{R}}(C_0(\mathscr{R}) \rtimes P_{Q(R)}) p_{\hat{R}}; \tilde{\varphi}_{m,lm} \right\}.$$

Moreover, set

$$\begin{aligned} \lambda_k\colon p_{\hat{R}}(C_0(\mathscr{R}) \rtimes P_{Q(R)}) p_{\hat{R}} &\longrightarrow C_0(\mathscr{R}) \rtimes P_{Q(R)} \\ z &\longmapsto V^*_{\left(\begin{smallmatrix} k & 0 \\ 0 & 1 \end{smallmatrix}\right)} z V_{\left(\begin{smallmatrix} k & 0 \\ 0 & 1 \end{smallmatrix}\right)}. \end{aligned}$$

As

$$\begin{aligned} & \lambda_{lm} \circ \tilde{\varphi}_{m,lm}(z) \\ = \; & V^*_{\left(\begin{smallmatrix} lm & 0 \\ 0 & 1 \end{smallmatrix}\right)} V_{\left(\begin{smallmatrix} l & 0 \\ 0 & 1 \end{smallmatrix}\right)} z V^*_{\left(\begin{smallmatrix} l & 0 \\ 0 & 1 \end{smallmatrix}\right)} V_{\left(\begin{smallmatrix} lm & 0 \\ 0 & 1 \end{smallmatrix}\right)} \\ = \; & V^*_{\left(\begin{smallmatrix} m & 0 \\ 0 & 1 \end{smallmatrix}\right)} z V_{\left(\begin{smallmatrix} m & 0 \\ 0 & 1 \end{smallmatrix}\right)} = \lambda_m(z), \end{aligned}$$

this gives a homomorphism

$$\lambda\colon\ \varinjlim \left\{ p_{\hat{R}}(C_0(\mathscr{R}) \rtimes P_{Q(R)}) p_{\hat{R}};\ \tilde{\varphi}_{m,lm} \right\} \longrightarrow C_0(\mathscr{R}) \rtimes P_{Q(R)}$$

which is injective as this is the case for each λ_m, and it is surjective as $\lambda_m(p_{\hat{R}}) = p_{\iota_m(\hat{R})}$ is an approximate unit for $C_0(\mathscr{R}) \rtimes P_{Q(R)}$. □

Remark 4.5. Combining this result with the preceding remark, we see that

$$\mathfrak{A}_r[R] \cong \mathfrak{A}[R] \cong p_{\hat{R}}(C_0(\mathscr{R}) \rtimes P_{Q(R)}) p_{\hat{R}}$$

Therefore, $p_{\hat{R}}(C_0(\mathscr{R}) \rtimes P_{Q(R)}) p_{\hat{R}}$ can be faithfully represented on $\ell^2(R)$ in a very natural way.

Remark 4.6. We call $\mathfrak{A}(R)$ the stabilization because $\mathfrak{A}(R) \cong \mathcal{K} \otimes \mathfrak{A}[R]$. This comes from the observation that $\mathfrak{A}[R]$ is isomorphic to $M_L(\mathfrak{A}[R])$ with regard to the L pairwise orthogonal projections $\{u^n e_l u^{-n} : n \in R\}$, where $L = \#[R/(l)]$. And under this identification, conjugation with s_l (which is $\varphi_{m,lm}$) corresponds to the inclusion of $\mathfrak{A}[R]$ into the upper left corner of $M_L(\mathfrak{A}[R])$.

In other words, using the theory of crossed products by semigroups, we can also say that $\mathfrak{A}[R] \cong C(\hat{R}) \rtimes P_R$ and that the dynamical system corresponding to $\mathfrak{A}(R)$ is just the associated minimal dilation system (see [**Lac**]).

Remark 4.7. Having the classification programme for C*-algebras in mind, one should note that each of the algebras $\mathfrak{A}[R]$ is nuclear, as $P_{Q(R)} \cong Q(R) \rtimes Q(R)^\times$ is always amenable because it is solvable.

5. Links to Algebraic Number Theory

The typical examples we have in mind are the ring of integers in an algebraic number field and polynomial rings with coefficients in a finite field. These are exactly the objects of interest in algebraic number theory.

Let $R = \mathfrak{o}$ be such a ring (the conditions from the beginning are satisfied). First of all, we have in this case $\hat{\mathfrak{o}} \cong \prod \mathfrak{o}_\nu$, where the product is taken over all the finite places ν over $K = Q(\mathfrak{o})$. If $\mathfrak{o}$ has positive characteristic (i.e. $\mathfrak{o}$ sits in a finite extension of $\mathbb{F}_p(T)$), we call the place corresponding to T^{-1} infinite. Here, $\hat{\mathfrak{o}}$ is to be understood in the sense of the previous section, and $\mathfrak{o}_\nu$ is the completion of $\mathfrak{o}$ with respect to ν.

Furthermore, we have $\mathscr{R} \cong \mathbb{A}_f$, which is the finite adèle ring.

These identifications can be proven as follows:

$$
\begin{aligned}
& \hat{\mathfrak{o}} = \varprojlim{}_m \{\mathfrak{o}/(m)\} \\
\cong \; & \varprojlim{}_{\wp_i, n_i} \{\mathfrak{o}/\wp_1^{n_1} \cdots \wp_l^{n_l}\} \\
& (\mathfrak{o} \text{ is a Dedekind ring, thus unique factorization of ideals holds}) \\
\cong \; & \varprojlim{}_{n_\wp} \left\{ \prod_{\wp \in \operatorname{Spec}(\mathfrak{o}) \setminus \{0\}} \mathfrak{o}/\wp^{n_\wp} \right\} \\
& (\text{Chinese remainder theorem}) \\
\cong \; & \prod_{\wp \in \operatorname{Spec}(\mathfrak{o}) \setminus \{(0)\}} \varprojlim{}_n \{\mathfrak{o}/\wp^n\} \\
\cong \; & \prod_{\nu \text{ finite}} \varprojlim{}_n \{\mathfrak{o}_\nu / P_\nu^n\} \\
& (\text{there is a bijection between nontrivial prime ideals and finite places}) \\
\cong \; & \prod_{\nu \text{ finite}} \mathfrak{o}_\nu \\
& (\mathfrak{o} \text{ is a Dedekind ring})
\end{aligned}
$$

where P_ν is the subset of $\mathfrak{o}_\nu$ with valuation strictly smaller than 1.

The second identification comes from

$$\mathscr{R} \cong \varinjlim \hat{\mathfrak{o}} \cong \varinjlim \left\{ \prod \mathfrak{o}_\nu \right\} \cong (\mathfrak{o}^\times)^{-1} \prod \mathfrak{o}_\nu \cong \mathbb{A}_f.$$

The details can be found in [**Wei**] and [**Neu**].

So all in all, we have purely infinite C*-algebras $\mathfrak{A}[\mathfrak{o}] \cong C(\hat{\mathfrak{o}}) \rtimes P_{\mathfrak{o}}$ with stabilization $\mathfrak{A}(\mathfrak{o}) \cong C_0(\mathbb{A}_f) \rtimes P_K$.

6. Relationship to Bost-Connes Systems

As mentioned at the beginning, our investigations are partly motivated by the work of Bost and Connes, who studied a C*-dynamical system for $\mathbb{Q}$ which had several interesting properties: e.g. it revealed connections to explicit class field theory over the rational numbers (see [**BoCo**]). As a next step, Connes, Marcolli and Ramachandran succeeded in constructing a C*-dynamical system for imaginary quadratic number fields and establishing analogous connections to explicit class field theory for these (see [**CMR**]).

In the meantime, there have been several attempts to construct systems with similar properties for arbitrary number fields (see [**CoMa**] for an overview).

Most recently, Laca, Larsen and Neshveyev studied C*-dynamical systems for arbitrary number fields generalizing the systems mentioned above for the case of $\mathbb{Q}$ and imaginary quadratic fields. These dynamical systems have already been considered in [**HaPa**], but Laca, Larsen and Neshveyev managed to classify the corresponding KMS states, which was a key ingredient in setting up connections to class field theory. Still, these results have not led to more insights concerning explicit class field theory.

Our aim in the following section is to embed these generalized Bost-Connes algebras into $\mathfrak{A}$, at least for a certain class of number fields. Viewing these generalized Bost-Connes systems as subalgebras of our C*-algebra $\mathfrak{A}$, it will be possible to deduce for them a description as universal C*-algebras with generators and relations, as Bost and Connes originally did in the case of $\mathbb{Q}$.

First of all, let us very briefly explain the construction in [**LLN**], to set up the notation:

Fix an algebraic number field K and let $\mathfrak{o}$ be its ring of integers.

Denote the ring of finite adèles by $\mathbb{A}_f = \prod'_{\nu \text{ finite}} K_\nu$, where we take the restricted direct product with respect to the ring inclusions $\mathfrak{o}_\nu \subseteq K_\nu$; let $K_\infty = \prod_{\nu \text{ infinite}} K_\nu$ be the product of the completions K_ν over the set of infinite places; then the ring of adèles can be written as $\mathbb{A} = K_\infty \times \mathbb{A}_f$.

Furthermore, the group of idèles is $\mathbb{A}^* = K_\infty^\times \times \prod'_{\nu \text{ finite}} K_\nu^\times$; this time the restricted product is taken with respect to $\mathfrak{o}_\nu^* \subseteq K_\nu^\times$. Let us write $K_{\infty,+}$ for the component of the identity in K_∞.

Moreover, take $\hat{\mathfrak{o}} = \prod_{\nu \text{ finite}} \mathfrak{o}_\nu$ and $\hat{\mathfrak{o}}^* = \prod_{\nu \text{ finite}} \mathfrak{o}_\nu^*$.

We will frequently think of subsets of $\mathbb{A}_f$ as embedded in $\mathbb{A}$, just by filling in zeros at the infinite places (or identities in the multiplicative case). Moreover, the algebraic number field (or subsets in K) can always be thought of as subsets of the adèles (or of the idèles in the multiplicative case) using the diagonal embedding.

Now, for each number field K, Laca, Larsen and Neshveyev construct a topological space (which was originally considered in [**CMR**])

$$X = \mathbb{A}^* / \overline{K^\times K_{\infty,+}^\times} \times_{\hat{\mathfrak{o}}^*} \mathbb{A}_f$$

which is a quotient of $\mathbb{A}^* / \overline{K^\times K_{\infty,+}^\times} \times \mathbb{A}_f$ with respect to the equivalence relation

$$\begin{aligned} & ((x_\nu),(y_\nu)) \sim ((x'_\nu),(y'_\nu)) \\ \Leftrightarrow \quad & \text{there exists } (r_\nu) \in \hat{\mathfrak{o}}^* \text{ with } ((r_\nu)(x_\nu),(r_\nu)^{-1}(y_\nu)) = ((x'_\nu),(y'_\nu)). \end{aligned}$$

For brevity, let us write U for $\overline{K^\times K_{\infty,+}^\times}$. $Y := \mathbb{A}^*/U \times_{\hat{\mathfrak{o}}^*} \hat{\mathfrak{o}}$ is a clopen subset of X. Furthermore, they consider an action of $\mathbb{A}_f^*/\hat{\mathfrak{o}}^*$ on X given by $(z_\nu)((x_\nu),(y_\nu)) = ((z_\nu)^{-1}(x_\nu),(z_\nu)(y_\nu))$. Finally, their C*-algebra is given by

$$\mathcal{A} = 1_Y \left(C_0(X) \rtimes \mathbb{A}_f^*/\hat{\mathfrak{o}}^* \right) 1_Y.$$

At this point, we should note that - presented in this way - this is a purely adelic-idelic way of describing the system, but that these objects have their natural meaning in number theory via certain abstract identifications (mostly provided by class field theory), for instance:

$\mathbb{A}^*/U \cong \mathrm{Gal}(K^{\mathrm{ab}}/K)$, where $\mathrm{Gal}(K^{\mathrm{ab}}/K)$ is the Galois group of the maximal abelian field extension of K, or

$\mathbb{A}_f^*/\hat{\mathfrak{o}}^* \cong J_K$, where J_K is the group of fractional ideals (see [**Wei**], IV 3).

6.1. Comparison of the Adelic-Idelic Constructions. We start the comparison on a purely topological level considering the adelic-idelic constructions. The first aim will be to establish a relationship between $\mathbb{A}^*/U \times_{\hat{\mathfrak{o}}^*} \mathbb{A}_f$ and $\mathbb{A}_f$.

There is a canonical map

$$\begin{aligned} \psi^* : \mathbb{A}_f &\longrightarrow \mathbb{A}^*/U \times_{\hat{\mathfrak{o}}^*} \mathbb{A}_f \\ (y_\nu) &\longmapsto ((1)^\bullet, (y_\nu))^\bullet \end{aligned}$$

which we would like to investigate in detail. By $(\cdot)^\bullet$, we mean the corresponding equivalence classes.

From the definitions, we immediately get

$$\begin{aligned} & \psi^*((y_\nu)) = \psi^*((\tilde{y}_\nu)) \\ \Leftrightarrow\quad & ((1)^\bullet, (y_\nu)) \sim ((1)^\bullet, (\tilde{y}_\nu)) \\ \Leftrightarrow\quad & \text{there exists } (z_\nu) \in \hat{\mathfrak{o}}^* \text{ such that } ((1)^\bullet, (y_\nu)) = ((z_\nu)^\bullet, (z_\nu)^{-1}(\tilde{y}_\nu)) \\ \Leftrightarrow\quad & \text{there exists } (z_\nu) \in \hat{\mathfrak{o}}^* \cap U \text{ with } (y_\nu) = (z_\nu)^{-1}(\tilde{y}_\nu). \end{aligned}$$

Let us calculate $\hat{\mathfrak{o}}^* \cap U$, as this will be needed later on:

LEMMA 6.1.

$$\hat{\mathfrak{o}}^* \cap U = \overline{\mathfrak{o}^* \cap \bigcap_{\nu \text{ real}} \nu^{-1}(\mathbb{R}_{>0})}$$

PROOF. The inclusion "$\subseteq$" holds because we have

$$\mathfrak{o}^* \subseteq \hat{\mathfrak{o}}^* \text{ and } \mathfrak{o}^* \cap \bigcap_{\nu \text{ real}} \nu^{-1}(\mathbb{R}_{>0}) \subseteq K^\times K^\times_{\infty,+}.$$

To get the other inclusion, observe

$$\begin{aligned} & (z_\nu) \in \hat{\mathfrak{o}}^* \cap U \\ \Leftrightarrow\quad & (z_\nu) \in \hat{\mathfrak{o}}^* \text{ and there exists a sequence } (z_\nu^{(n)}) \text{ in } K^\times K^\times_{\infty,+} \text{ with} \\ & (z_\nu) = \lim_{n\to\infty} (z_\nu^{(n)}) \text{ in } \mathbb{A}^*. \end{aligned}$$

By the definition of the topology on $\mathbb{A}^*$, there is a finite set of places P such that

$$\begin{aligned} & (z_\nu) \in \prod_{\nu \in P} K_\nu^\times \times \prod_{\nu \notin P} \mathfrak{o}_\nu^* \\ \Rightarrow\quad & \text{there is } \tilde{N} \in \mathbb{N} \text{ with } (z_\nu^{(n)}) \in \prod_{\nu \in P} K_\nu^\times \times \prod_{\nu \notin P} \mathfrak{o}_\nu^* \text{ for all } n \geq \tilde{N}. \end{aligned}$$

As $\lim_{n\to\infty} (z_\nu^{(n)}) = (z_\nu)$, we conclude that $\lim_{n\to\infty} z_\nu^{(n)} = z_\nu$ for all places ν in $K_\nu^\times$. But, as $z_\nu^{(n)} \in \mathfrak{o}_\nu^*$ for almost all finite places if $n \geq \tilde{N}$ and because $\mathfrak{o}_\nu^*$ is open in $K_\nu^\times$, there must be $N \in \mathbb{N}$ ($N \geq \tilde{N}$) such that:

$$z_\nu^{(n)} \in \mathfrak{o}_\nu^* \text{ for all finite places } \nu \text{ and for all } n \geq N.$$

Also, as $(z_\nu^{(n)})$ lies in $K^\times K_{\infty,+}^\times$, there exists, for each n, $z^{(n)}$ in $K^\times$ such that $z_\nu^{(n)} = z^{(n)}$ at every finite place (K is diagonally embedded). Thus, we have $z^{(n)} = z_\nu^{(n)} \in K^\times \cap \mathfrak{o}_\nu^*$ for all finite places and for all $n \geq N$, which means that $z^{(n)} \in K^\times \cap \bigcap_{\nu \text{ finite}} \mathfrak{o}_\nu^* = \mathfrak{o}^*$ for all $n \geq N$. Hence it follows that

$$(z_\nu^{(n)}) = (z^{(n)}) \in \mathfrak{o}^* \subseteq \hat{\mathfrak{o}}^* \text{ for all } n \geq N.$$

Moreover, as (z_ν) lies in $\hat{\mathfrak{o}}^*$, we must have $z_\nu = 1$ for all infinite places, so that $z_\nu^{(n)} \in \mathbb{R}_{>0}$ for all real places and for sufficiently large n. Using the observation above ($z^{(n)} = z_\nu^{(n)}$), we conclude that $z^{(n)} \in \nu^{-1}(\mathbb{R}_{>0})$ for all real places and for n large enough, and hence

$$(z_\nu) \in \overline{\mathfrak{o}^* \cap \bigcap_{\nu \text{ real}} \nu^{-1}(\mathbb{R}_{>0})}$$

as claimed. □

Let us now consider the quotient space $\mathbb{A}_f / \sim_{\hat{\mathfrak{o}}^* \cap U}$, where

$(y_\nu) \sim_{\hat{\mathfrak{o}}^* \cap U} (\tilde{y}_\nu)$ if and only if there exists $(z_\nu) \in \hat{\mathfrak{o}}^* \cap U$ such that $(y_\nu) = (z_\nu)(\tilde{y}_\nu)$.

Using the universal property of this quotient, we get a continuous and injective map

$$\varphi^* : \mathbb{A}_f / \sim_{\hat{\mathfrak{o}}^* \cap U} \longrightarrow \mathbb{A}^*/U \times_{\hat{\mathfrak{o}}^*} \mathbb{A}_f$$

with $\psi^* = \varphi^* \circ p$, where p is the projection map $\mathbb{A}_f \to \mathbb{A}_f / \sim_{\hat{\mathfrak{o}}^* \cap U}$.

Lemma 6.2. *φ^* is closed.*

Proof. It suffices to show that ψ^* is closed, because of the following: Take $A \subseteq \mathbb{A}_f / \sim_{\hat{\mathfrak{o}}^* \cap U}$ to be an arbitrary closed set. As p is continuous, $p^{-1}(A)$ is closed in $\mathbb{A}_f$. Assuming that ψ^* is closed, $\varphi^*(A) = \varphi^* p p^{-1}(A) = \psi^* p^{-1}(A)$ is closed in $\mathbb{A}^* / \sim_{\hat{\mathfrak{o}}^* \cap U}$.

Now take $A \subseteq \mathbb{A}_f$ to be an arbitrary closed set, and let

$$\pi : \mathbb{A}^*/U \times \mathbb{A}_f \to \mathbb{A}^*/U \times_{\hat{\mathfrak{o}}^*} \mathbb{A}_f$$

be the canonical projection. We have to show that $\psi^*(A)$ is closed, which is equivalent to closedness of $\pi^{-1}\psi^*(A)$. We have

$$\begin{aligned}
&\pi^{-1}\psi^*(A) \\
= \; &\{((a_\nu)^\bullet, (b_\nu)) \in \mathbb{A}^* \times \mathbb{A}_f : \pi((a_\nu)^\bullet, (b_\nu)) \in \psi^*(A)\} \\
= \; &\{((a_\nu)^\bullet, (b_\nu)) \in \mathbb{A}^* \times \mathbb{A}_f : \exists\, (y_\nu) \in A : ((a_\nu)^\bullet, (b_\nu)) \sim ((1)^\bullet, (y_\nu))\} \\
= \; &\big\{((a_\nu)^\bullet, (b_\nu)) : \exists\, (y_\nu) \in A,\, (z_\nu) \in \hat{\mathfrak{o}}^* : (a_\nu)^\bullet = (z_\nu)^\bullet \wedge (b_\nu) = (z_\nu)^{-1}(y_\nu)\big\} \\
= \; &\big\{((z_\nu)^\bullet, (z_\nu)^{-1}(y_\nu)) \in \mathbb{A}^* \times \mathbb{A}_f : (z_\nu) \in \hat{\mathfrak{o}}^*, (y_\nu) \in A\big\}
\end{aligned}$$

Now suppose we have a sequence $((z_\nu^{(n)})^\bullet, (z_\nu^{(n)})^{-1}(y_\nu^{(n)}))$ in $\pi^{-1}\varphi^*(A)$ converging to $((a_\nu)^\bullet, (b_\nu)) \in \mathbb{A}^*/U \times \mathbb{A}_f$. Then we claim: $((a_\nu)^\bullet, (b_\nu)) \in \pi^{-1}\psi^*(A)$.

Indeed: $\hat{\mathfrak{o}}^*$ is compact, therefore there is a convergent subsequence $(z_\nu^{(n_k)})$ with limit $(z_\nu) \in \hat{\mathfrak{o}}^*$. Thus, $(z_\nu)^\bullet = \lim_{k\to\infty} (z_\nu^{(n_k)})^\bullet = \lim_{n\to\infty} (z_\nu^{(n)})^\bullet = (a_\nu)^\bullet$ and $(y_\nu^{(n_k)}) = (z_\nu^{(n_k)})(z_\nu^{(n_k)})^{-1}(y_\nu^{(n_k)}) \underset{k}{\longrightarrow} (z_\nu)(b_\nu)$. Hence, $(y_\nu^{(n_k)})$ is a sequence in A converging in $\mathbb{A}_f$, therefore $(z_\nu)(b_\nu) = \lim_{k\to\infty}(y_\nu^{(n_k)}) =: (y_\nu)$ lies in A. Thus we have $(b_\nu) = (z_\nu)^{-1}(y_\nu)$ and hence $((a_\nu)^\bullet, (b_\nu)) = ((z_\nu)^\bullet, (z_\nu)^{-1}(y_\nu)) \in \pi^{-1}\psi^*(A)$ which proves our claim and therefore the Lemma. □

It remains to investigate under which conditions φ^* is surjective.

LEMMA 6.3. *φ^* is surjective if $h_K = 1$ and there is at most one real (infinite) place of K.*

PROOF. First of all, φ^* is surjective if and only if ψ^* is so. Now, ψ^* is surjective if for any $(a_\nu)^\bullet \in \mathbb{A}^*/U$, $(b_\nu) \in \mathbb{A}_f$ there are $(y_\nu) \in \mathbb{A}_f$, $(z_\nu) \in \hat{\mathfrak{o}}^*$ such that $(a_\nu)^\bullet = (z_\nu)^\bullet$ and $(b_\nu) = (z_\nu)^{-1}(y_\nu)$ (this is equivalent to $((a_\nu)^\bullet, (b_\nu)) \sim ((1)^\bullet, (y_\nu))$). As the first condition is the crucial one (once it holds, the second one can be enforced), ψ^* is surjective if (and only if) $\mathbb{A}^* = \hat{\mathfrak{o}}^* \cdot U$.

Assuming that the number of real places is not bigger than one, we have $K^\times K_\infty^\times \hat{\mathfrak{o}}^* = K^\times K_{\infty,+}^\times \hat{\mathfrak{o}}^*$ because given $(a)(b_\nu)(c_\nu) \in K^\times K_\infty^\times \hat{\mathfrak{o}}^*$ with $b_\nu < 0$ at the real place, we have $(a)(b_\nu)(c_\nu) = (-a)(-b_\nu)(-c_\nu) \in K^\times K_{\infty,+}^\times \hat{\mathfrak{o}}^*$.

Now, $h_K = 1$ implies $1 = \#[J_K/P_K] = \#[I(K)/P(K)] = \#[\mathbb{A}^*/K^\times K_\infty^\times \hat{\mathfrak{o}}^*]$ (see [**Wei**], V 3). Hence it follows that we have $\mathbb{A}^* = K^\times K_\infty^\times \hat{\mathfrak{o}}^* = K^\times K_{\infty,+}^\times \hat{\mathfrak{o}}^*$. This implies that ψ^* is surjective. □

Summarizing our observations to this point, we get

PROPOSITION 6.4. If $h_K = 1$ and there is at most one real place, then the map

$$\begin{aligned} \varphi^* : \mathbb{A}_f/\sim_{\hat{\mathfrak{o}}^* \cap U} &\longrightarrow \mathbb{A}^*/U \times_{\hat{\mathfrak{o}}^*} \mathbb{A}_f \\ (y_\nu)^\bullet &\longmapsto ((1)^\bullet, (y_\nu))^\bullet \end{aligned}$$

is a homoemorphism.

REMARK 6.5. One should note that Lemma 6.3 is not optimal in the sense that φ^* can be surjective even if K has more than one real place. The crucial criterion is whether $\mathfrak{o}^*$ is embedded in $K_\infty^\times$ in such a way that every possible combination of signs (in the real places) can be arranged (compare [**LaFr**], Proposition 4 for a similar problem).

6.2. Comparison of the C*-algebras. As a next step, let us study the corresponding C*-algebras and the crossed products in the situation of the last proposition ($h_K = 1$, at most one real place):

PROPOSITION 6.6. φ^* induces *-isomorphisms

$$C_0(X) \rtimes \mathbb{A}_f^*/\hat{\mathfrak{o}}^* \xrightarrow{\simeq} C_0(\mathbb{A}_f/\sim_{\hat{\mathfrak{o}}^* \cap U}) \rtimes K^\times/\mathfrak{o}^*$$

and

$$\mathcal{A} \xrightarrow{\simeq} 1_{\hat{\mathfrak{o}}/\sim_{\hat{\mathfrak{o}}^* \cap U}} \left(C_0(\mathbb{A}_f/\sim_{\hat{\mathfrak{o}}^* \cap U}) \rtimes K^\times/\mathfrak{o}^*\right) 1_{\hat{\mathfrak{o}}/\sim_{\hat{\mathfrak{o}}^* \cap U}}$$

if there are no real places of K.

If there is a real place, we get *-isomorphisms

$$C_0(X) \rtimes \mathbb{A}_f^*/\hat{\mathfrak{o}}^* \xrightarrow{\simeq} C_0(\mathbb{A}_f/\sim_{\hat{\mathfrak{o}}^* \cap U}) \rtimes K_{>0}^\times/\mathfrak{o}_{>0}^*$$

and

$$\mathcal{A} \xrightarrow{\simeq} 1_{\hat{\mathfrak{o}}/\sim_{\hat{\mathfrak{o}}^* \cap U}} \left(C_0(\mathbb{A}_f/\sim_{\hat{\mathfrak{o}}^* \cap U}) \rtimes K_{>0}^\times/\mathfrak{o}_{>0}^*\right) 1_{\hat{\mathfrak{o}}/\sim_{\hat{\mathfrak{o}}^* \cap U}}.$$

Here we have fixed a real embedding corresponding to the real place and we think of K as a subset of $\mathbb{R}$ via this embedding.

Proof. As a first step, φ^* induces a *-isomorphism

$$C_0(X) \xrightarrow{\simeq} C_0(\mathbb{A}_f/\sim_{\hat{\mathfrak{o}}^*\cap U})$$

(recall $X = \mathbb{A}^*/U \times_{\hat{\mathfrak{o}}^*} \mathbb{A}_f$).

Now, let us assume that K does not have any real places. It remains to prove that the actions of $\mathbb{A}_f^*/\hat{\mathfrak{o}}^*$ and $K^\times/\mathfrak{o}^*$ are compatible. But this follows from the fact that we have an isomorphism (because of $h_K = 1$)

$$\begin{array}{rl} K^\times/\mathfrak{o}^* & \xrightarrow{\simeq} \mathbb{A}_f^*/\hat{\mathfrak{o}}^* \\ \lambda^\bullet & \longmapsto (\lambda)^\bullet \end{array}$$

and the following computation:

$$\begin{array}{rl} & (\lambda)^\bullet \cdot \varphi^*((y_\nu)) \\ = & (\lambda)^\bullet \cdot ((1)^\bullet, (y_\nu))^\bullet \\ = & ((\lambda^{-1})^\bullet, (\lambda)(y_\nu))^\bullet \\ = & ((1)^\bullet, (\lambda)(y_\nu))^\bullet \\ = & \varphi^*(\lambda^\bullet \cdot (y_\nu)^\bullet). \end{array}$$

This gives us the first isomorphism, which we denote by φ.

To get the second one, we have to show $\varphi(1_Y) = 1_{\hat{\mathfrak{o}}/\sim_{\hat{\mathfrak{o}}^*\cap U}}$, which is equivalent to $\varphi^*(\hat{\mathfrak{o}}/\sim_{\hat{\mathfrak{o}}^*\cap U}) = Y$. To see this, let us prove $Y \subseteq \varphi^*(\hat{\mathfrak{o}}/\sim_{\hat{\mathfrak{o}}^*\cap U})$, since the other inclusion is certainly valid. Take any $((a_\nu)^\bullet, (b_\nu))^\bullet \in Y$, as φ^* is surjective we can find $(y_\nu) \in \mathbb{A}_f$ such that $((a_\nu)^\bullet, (b_\nu))^\bullet = ((1)^\bullet, (y_\nu))^\bullet$. Therefore, we can also find (z_ν) in $\hat{\mathfrak{o}}^*$ such that $((z_\nu)^\bullet, (z_\nu)^{-1}(y_\nu)) = ((a_\nu)^\bullet, (b_\nu))$, and thus, $(y_\nu) = (z_\nu)(b_\nu) \in \hat{\mathfrak{o}}$. This means that $((a_\nu)^\bullet, (b_\nu))^\bullet \in \varphi^*(\hat{\mathfrak{o}}/\sim_{\hat{\mathfrak{o}}^*\cap U})$.

This completes the proof for the case of no real places. If K has one real place, the proof will be exactly the same. But one should note that in the computation above, one really needs the restriction to $K^\times_{>0}/\mathfrak{o}^*_{>0}$ because for $\lambda \in K$, $(\lambda)^\bullet \in \mathbb{A}_f^*$ lies in U if and only if λ is positive. □

To get the relationship with our algebras $\mathfrak{A}$, we remark that there are canonical embeddings

$$C_0(\mathbb{A}_f/\sim_{\hat{\mathfrak{o}}^*\cap U}) \rtimes K^\times/\mathfrak{o}^* \hookrightarrow C_0(\mathbb{A}_f) \rtimes K^\times \hookrightarrow C_0(\mathbb{A}_f) \rtimes P_K \cong \mathfrak{A}(\mathfrak{o})$$

if K does not have real places and

$$C_0(\mathbb{A}_f/\sim_{\hat{\mathfrak{o}}^*\cap U}) \rtimes K^\times_{>0}/\mathfrak{o}^*_{>0} \hookrightarrow C_0(\mathbb{A}_f) \rtimes K^\times \hookrightarrow C_0(\mathbb{A}_f) \rtimes P_K \cong \mathfrak{A}(\mathfrak{o})$$

for the case of one real place (see Theorem 4.4 for the description of $\mathfrak{A}(\mathfrak{o})$).

Restricting these embeddings to the generalized Bost-Connes algebra $\mathcal{A}$ (using the *-isomorphisms of Proposition 6.6), we get embeddings $\mathcal{A} \hookrightarrow \mathfrak{A}[\mathfrak{o}]$. At this point, we should note that there is no distinction between reduced or full crossed products as all the groups are amenable. Therefore, we really get embeddings.

Remark 6.7. In the case of purely imaginary number fields of class number one, the observations made in Remark 2.2.(iii) of [**LLN**] are related to our results (compare also the paper [**LaFr**]).

6.3. Representation of the Bost-Connes Algebras. Regarding the generalized Bost-Connes systems as subalgebras of our algebras as above, it is possible to get an alternative description of $\mathcal{A}$ as a universal C*-algebra with generators and relations:

THEOREM 6.8. *Let $h_K = 1$ and assume that K has no real places. In this case, $\mathcal{A}$ is the universal C*-algebra generated by nontrivial projections*

$$f(m,n) \text{ for every } m \in \mathfrak{o}^\times / \sim_{\mathfrak{o}^*} ,\ n \in (\mathfrak{o}/(m)) / \sim_{\mathfrak{o}^*}$$

and isometries

$$s_p \text{ for each } p \in \mathfrak{o}^\times / \sim_{\mathfrak{o}^*}$$

satisfying the relations

- $s_p s_q = s_{pq}$ *for all* $p, q \in \mathfrak{o}^\times / \sim_{\mathfrak{o}^*}$
- $f(1,0) = 1$
- $s_p f(m,n) s_p^* = f(mp, np)$ *for all* $p, m \in \mathfrak{o}^\times/\mathfrak{o}^*$, $n \in (\mathfrak{o}/(m)) / \sim_{\mathfrak{o}^*}$
- $\sum_j f(mp, j) = f(p,k)$ *for all* $m, p \in \mathfrak{o}^\times/\mathfrak{o}^*$, $k \in (\mathfrak{o}/(p)) / \sim_{\mathfrak{o}^*}$

The sum in the last relation is taken over $\pi^{-1}_{mp,m}(k)$, where

$$\pi_{mp,m} : (\mathfrak{o}/(mp)) / \sim_{\mathfrak{o}^*} \longrightarrow (\mathfrak{o}/(p)) / \sim_{\mathfrak{o}^*}$$

is the canonical projection.

If there is one real place, one has to substitute $\mathfrak{o}^\times$ by $\mathfrak{o}^\times_{>0}$ and $\mathfrak{o}^*$ by $\mathfrak{o}^*_{>0}$.

Before we start with the proof, it should be noted that one can think of the projection $f(m,n)$ as $\sum u^l e_m u^{-l}$, where the sum is taken over all classes $l + (m)$ in $\mathfrak{o}/(m)$ which are in the same equivalence class as n with respect to $\sim_{\mathfrak{o}^*}$. This is exactly the way how these elements are embedded into $\mathfrak{A}[\mathfrak{o}]$. Moreover, using the characteristic relations in $\mathfrak{A}[\mathfrak{o}]$, the relations above can be checked in a straightforward manner.

PROOF. Let us prove the theorem in the case of no real places, the other case can be proven in an analogous way.

The first step is to establish a *-isomorphism of the commutative C*-algebras $C(\hat{\mathfrak{o}}/ \sim_{\hat{\mathfrak{o}}^* \cap U})$ and $C^*(\{f(m,n) : m \in \mathfrak{o}^\times / \sim_{\mathfrak{o}^*} ,\ n \in (\mathfrak{o}/(m)) / \sim_{\mathfrak{o}^*}\})$.

We already had the description $\hat{\mathfrak{o}} = \varprojlim \{\mathfrak{o}/(m) \ ; \phi_{lm,m}\}$. It will be convenient to describe $\hat{\mathfrak{o}}^* \cap U = \overline{\mathfrak{o}^*}$ in a similar way using projective limits.

We claim: $\hat{\mathfrak{o}^*} \cap U \cong \varprojlim \{(\mathfrak{o}^* + (m))/(m) \ ; \phi_{lm,m}\}$ where the $\phi_{lm,m}$ are the canonical projections $(\mathfrak{o}^* + (lm))/(lm) \to (\mathfrak{o}^* + (m))/(m)$ as above.

To prove the claim, consider the following continuous embeddings

$$\mathfrak{o}^* \hookrightarrow \varprojlim \{(\mathfrak{o}^* + (m))/(m) \ ; \phi_{lm,m}\} \hookrightarrow \hat{\mathfrak{o}}.$$

Their composition is exactly the diagonal embedding of $\mathfrak{o}^*$ into $\hat{\mathfrak{o}}$.

Moreover, $\varprojlim \{(\mathfrak{o}^* + (m))/(m) \ ; \phi_{lm,m}\}$ (identified with its image in $\hat{\mathfrak{o}}$) is compact and contains $\mathfrak{o}^*$. As it follows from the construction of this projective limit that $\mathfrak{o}^*$ (embedded in the projective limit) is dense, we have proven the claim (compare Lemma 6.1).

Furthermore, we have

$$\begin{aligned} & \hat{\mathfrak{o}}/\sim_{\hat{\mathfrak{o}}^*\cap U} \\ \cong\ & \varprojlim \{\mathfrak{o}/(m)\ ;\ \phi_{lm,m}\}/\sim_{\varprojlim \{(\mathfrak{o}^*+(m))/(m)\ ;\ \phi_{lm,m}\}} \\ \cong\ & \varprojlim \{(\mathfrak{o}/(m))\sim_{\mathfrak{o}^*}\ ;\ \phi_{lm,m}\}. \end{aligned}$$

The first identification has already been proven; for the second one, consider the following maps:

$$\begin{array}{rcl} \varprojlim \{\mathfrak{o}/(m)\}/\sim_{\varprojlim \{(\mathfrak{o}^*+(m))/(m)\}} & \leftrightharpoons & \varprojlim \{(\mathfrak{o}/(m))/\sim_{\mathfrak{o}^*}\} \\ (a_m)^\bullet & \mapsto & (a_m^\bullet) \\ (b_m)^\bullet & \leftarrow\!\shortmid & (b_m^\bullet) \end{array}$$

Existence and continuity of the upper map is given by the universal properties of projective limits and quotient spaces. The lower map is well-defined for the following reason:

Let $(b_m^\bullet) = (c_m^\bullet)$, we have to show $(b_m)^\bullet = (c_m)^\bullet$.

For each $m \in \mathfrak{o}^\times$, there is an $r_m \in \mathfrak{o}^*$ with the property that $b_m + (m) = r_m c_m + (m)$. But the net $((r_m))_m$ has a convergent subnet with limit (s_m) in $\varprojlim \{(\mathfrak{o}^* + (m))/(m)\ ;\ \phi_{lm,m}\}$ as this set is compact. And the choice of the r_m ensures that we have $(b_m) = (s_m)(c_m)$; thus $(b_m) \sim_{\varprojlim \{(\mathfrak{o}^*+(m))/(m)\ ;\ \phi_{lm,m}\}} (c_m)$. Therefore, the lower map exists as well.

Now, the upper map is a bijective continuous map, and the range is Hausdorff, whereas the domain is quasi-compact. Hence it follows that these maps are mutually inverse homoemorphisms.

After this step, we can now use Laca's result on crossed products by semigroups (see [**Lac**]) to conclude the proof:

The universal C*-algebra with the generators and relations as listed above is exactly given by the crossed product

$$C^*(\{f(m,n) : m \in \mathfrak{o}^\times/\sim_{\mathfrak{o}^*},\ n \in (\mathfrak{o}/(m))/\sim_{\mathfrak{o}^*}\}) \rtimes \left(\mathfrak{o}^\times/\sim_{\mathfrak{o}^*}\right)$$

where we take the endomorphisms given by conjugation with s_p. This is valid as both C*-algebras have the same universal properties.

Furthermore, the identification above shows that

$$\begin{aligned} & C^*(\{f(m,n) : m \in \mathfrak{o}^\times/\sim_{\mathfrak{o}^*},\ n \in (\mathfrak{o}/(m))/\sim_{\mathfrak{o}^*}\}) \rtimes \left(\mathfrak{o}^\times/\sim_{\mathfrak{o}^*}\right) \\ \cong\ & C(\hat{\mathfrak{o}}/\sim_{\hat{\mathfrak{o}}^*\cap U}) \rtimes \left(\mathfrak{o}^\times/\sim_{\mathfrak{o}^*}\right) \end{aligned}$$

and the last C*-algebra is isomorphic to

$$1_{\hat{\mathfrak{o}}/\sim_{\hat{\mathfrak{o}}^*\cap U}} \left(C_0(\mathbb{A}_f/\sim_{\hat{\mathfrak{o}}^*\cap U}) \rtimes K^\times/\mathfrak{o}^*\right) 1_{\hat{\mathfrak{o}}/\sim_{\hat{\mathfrak{o}}^*\cap U}}.$$

This follows from the work of Laca, since in Laca's notation, $C_0(\mathbb{A}_f/\sim_{\hat{\mathfrak{o}}^*\cap U})$ together with the action of $K^\times/\mathfrak{o}^*$ is the minimal automorphism dilation corresponding to $C(\hat{\mathfrak{o}}/\sim_{\hat{\mathfrak{o}}^*\cap U}) \rtimes (\mathfrak{o}^\times/\sim_{\mathfrak{o}^*})$ (see [**Lac**]).

Finally, Proposition 6.6 implies

$$\begin{aligned} & 1_{\hat{\mathfrak{o}}/\sim_{\hat{\mathfrak{o}}^*\cap U}} \left(C_0(\mathbb{A}_f/\sim_{\hat{\mathfrak{o}}^*\cap U}) \rtimes K^\times/\mathfrak{o}^*\right) 1_{\hat{\mathfrak{o}}/\sim_{\hat{\mathfrak{o}}^*\cap U}} \\ \cong\ & 1_Y \left(C_0(X) \rtimes \mathbb{A}_f^*/\hat{\mathfrak{o}}^*\right) 1_Y = \mathcal{A}. \end{aligned}$$

$\square$

REMARK 6.9. Again, using $\mathfrak{A}_r[\mathfrak{o}] \cong \mathfrak{A}[\mathfrak{o}]$, this result gives us a faithful (and again rather natural) representation of $\mathcal{A}$ on $\ell^2(\mathfrak{o})$.

As a last point, we explain how to use Theorem 6.8 to construct the extremal KMS_β states of the C*-dynamical system $(\mathcal{A}, \sigma_t)$, where $\sigma_t(s_p) = N(p)^{it} s_p$ and N is the norm of the number field K. σ_t exists because of the universal property of $\mathcal{A}$ (see Theorem 6.8). $(\mathcal{A}, \sigma_t)$ is exactly the system considered in [**LLN**].

We essentially follow the construction in [**BoCo**], THEOREM 25 (a), in the sense that for each element of the Galois group $\mathrm{Gal}(K^{\mathrm{ab}}/K)$, we can construct a representation of $\mathcal{A}$ using its universal property which yields the corresponding KMS_β state.

First of all, we can associate to each $\alpha \in \mathrm{Gal}(K^{\mathrm{ab}}/K)$ the *-representation

$$\pi_\alpha : \mathcal{A} \longrightarrow \mathcal{L}\left(\ell^2(\mathfrak{o}^\times / \sim_{\mathfrak{o}^*})\right)$$

by

$$\pi_\alpha(s_p)\xi_r = \xi_{pr}$$

$$\pi_\alpha(f(m,n))\xi_r = \begin{cases} \xi_r \text{ if } \overline{\alpha}(r + (m)) = n \in (\mathfrak{o}/(m)) / \sim_{\mathfrak{o}^*} \\ 0 \text{ otherwise} \end{cases}$$

Here, $\overline{\alpha}$ is the image of α under the composition

$$\mathrm{Gal}(K^{\mathrm{ab}}/K) \longrightarrow \hat{\mathfrak{o}}^*/\overline{\mathfrak{o}^*} \longrightarrow (\mathfrak{o}/(m))^* / \sim_{\mathfrak{o}^*} .$$

The existence of π_α follows from the universal property of $\mathcal{A}$ described in Theorem 6.8.

Now, let us define $H(\xi_r) = \log(N(r))\xi_r$. With this (unbounded) operator we can construct the following KMS_β state

$$\varphi_{\beta,\alpha}(x) = \zeta(\beta)^{-1} \mathrm{tr}(\pi_\alpha(x) e^{-\beta H})$$

where ζ is the zeta-function of the number field K and $1 < \beta < \infty$.

This observation gives us candidates for the extremal KMS_β states, and this construction follows an alternative, more operator-theoretic route compared to the rather measure-theoretic approach of [**LLN**]. But it is another question whether these $\varphi_{\beta,\alpha}$ are precisely all the extremal KMS_β states of this C*-dynamical system for $1 < \beta < \infty$. This is answered in the affirmative in [**LLN**], Theorem 2.1.(iii). The connection is built by identifying the semigroup of integral ideals, J_K^+, with $\mathfrak{o}^\times / \sim_{\mathfrak{o}^*}$ using our assumption on K that the class number is 1.

References

[Bla] B. BLACKADAR, *Operator Algebras, Theory of C*-Algebras and von Neumann Algebras*, Encyclopaedia of Mathematical Sciences, Vol. 122, Springer-Verlag, Berlin Heidelberg New York, 2006.

[BoCo] J. B. BOST and A. CONNES, *Hecke algebras, Type III Factors and Phase Transitions with Spontaneous Symmetry Breaking in Number Theory*, Selecta Math., New Series, Vol. 1, *3* (1995), 411-457.

[CoMa] A. CONNES and M. MARCOLLI, *From Physics to Number Theory via Noncommutative Geometry, Part I: Quantum Statistical Mechanics of* $\mathbb{Q}$*-lattices* in Frontiers in Number Theory, Physics and Geometry, I, Springer-Verlag, Berlin Heidelberg New York, 2006, 269-350.

[CMR] A. CONNES, M. MARCOLLI and N. RAMACHANDRAN, *KMS states and complex multiplication*, Selecta Math., New Series, *11* (2005), 325-347.

[Cun1] J. Cuntz, *C*-algebras associated with the ax+b-semigroup over* $\mathbb{N}$, Proc. of the Conference on K-Theory and Non-Commutative Geometry, Valladolid 2006, 201-215.

[Cun2] J. Cuntz, *Simple C*-algebras generated by isometries*, Comm. Math. Phys. *57* (1977), 173-185.

[HaPa] E. Ha, F. Paugam, *Bost-Connes-Marcolli systems for Shimura varieties. I. Definitions and formal analytic properties*, IMRP Int. Math. Res. Rap. 2005, no. 5, 237-286.

[Lac] M. Laca, *From endomorphisms to automorphisms and back: dilations and full corners*, J. London Math. Soc. *61* (2000), 893-904.

[LaFr] M. Laca and M. van Frankenhuijsen *Phase transitions on Hecke C*-algebras and class-field theory over* $\mathbb{Q}$, J. reine angew. Math. *595* (2006), 25-53.

[LLN] M. Laca, N. S. Larsen and S. Neshveyev, *On Bost-Connes type systems for number fields*, J. Number Theory *129* (2009), no. 2, 325-338.

[Neu] J. Neukirch, *Algebraische Zahlentheorie*, Springer-Verlag, Berlin Heidelberg New York, 1992.

[Rør] M. Rørdam, *Classification of Nuclear C*-Algebras* in Classification of Nuclear C*-Algebras. Entropy in Operator Algebras, Encyclopaedia of Mathematical Sciences, Vol. 126, Springer-Verlag, Berlin Heidelberg New York, 2002.

[Wei] A. Weil, *Basic number theory* Die Grundlehren der mathematischen Wissenschaften, Band 144, Springer-Verlag, Berlin Heidelberg New York, 1974.

Mathematisches Institut, Westfaelische Wilhelms-Universitaet Muenster, Einsteinstrasse 62, 48149 Muenster, Germany

E-mail address: cuntz@uni-muenster.de

Mathematisches Institut, Westfaelische Wilhelms-Universitaet Muenster, Einsteinstrasse 62, 48149 Muenster, Germany

E-mail address: xinli.math@uni-muenster.de

Clay Mathematics Proceedings
Volume 11, 2010

Noncommutative Coordinate Algebras

Michel Dubois-Violette

Dédié à Alain Connes

Abstract. We discuss the noncommutative generalizations of polynomial algebras which after appropriate completions can be used as coordinate algebras in various noncommutative settings (noncommutative differential geometry, noncommutative algebraic geometry, etc.). These algebras have finite presentations and are completely characterized and classified by their (noncommutative) volume forms.

Il est bon de lire entre les lignes, cela fatigue moins les yeux
Sacha Guitry

Contents

2010 *Mathematics Subject Classification.* Primary: 16S37, Secondary: 16S36, 16W50.

Introduction

The universal skeletons for coordinate algebras in classical geometry (differential geometry, algebraic geometry, etc.) are polynomial algebras. The appropriate function algebras are obtained by completions with respect to the adapted topologies and either by a gluing process or by taking quotient algebras.

Our aim here is to discuss the noncommutative generalizations of polynomial algebras which can be used similarly in various noncommutative settings: noncommutative differential geometry [**13**], [**14**], [**15**], noncommutative algebraic geometry, etc. as well as in applications in physics.

At the very beginning, one has to face the question of what class of algebras we should consider as generalizations of the algebras of polynomial functions on finite-dimensional vector spaces. It seems clear that one must stay within the class of the $\mathbb{N}$-graded algebras which are connected, generated in degree 1 with finite presentations. There is a minimal choice which is the class of quadratic algebras which are Koszul (see below) of finite global dimension, which have polynomial growth and satisfy a version of Poincaré duality refered to as the Gorenstein property in [**1**]. A bigger class is the class of regular algebras in the sense of [**1**], which shall be referred to as the class of AS-regular algebras in the following. We shall consider here a bigger class in that we shall drop the condition of polynomial growth included in the AS-regularity condition. We shall refer to this bigger class of algebras as regular algebras. Although polynomial growth is a very natural condition for noncommutative coordinate algebras (and from the point of view of deformation theory), it turns out that for our analysis we do not need it and that by imposing polynomial growth one eliminates algebras which in spite of the fact that they do not have interpretation of (noncommutative) coordinate algebras are very interesting and are furthermore relevant for physics. Of course, at any stage one can restrict attention to the subclass of algebras with polynomial growth (or which are quadratic, etc.). For global dimensions $D = 2$ and $D = 3$, the regular algebras are N-homogeneous and Koszul. We shall recall what this Koszul property means. This is a very desirable property that one can formulate only for N-homogeneous algebras for the moment (i.e. algebras with relations of degree N). This is why, for global dimensions $D \geq 4$, we shall impose N-homogeneity and Koszulity.

In the following we shall review various concepts and results. We shall in particular give a survey of the results of [**22**], [**23**] in which we shall insist on the conceptual points and omit technical proofs. We shall illustrate the main points by a lot of examples. A central result is that the algebras under consideration are completely specified by multilinear forms on finite-dimensional vector spaces. Given such an algebra, the corresponding multilinear form, which is unique up to a nonvanishing scale factor, plays the role of the (noncommutative) volume form. Furthermore, isomorphic algebras correspond to multilinear forms which are in the same orbit of the corresponding linear group ($GL(g)$ for g generators). The determination of the moduli space of these algebras is of course of mathematical interest in itself. Concerning physics, the classification of these algebras can become of great importance since in a noncommutative geometrical approach to the quantum theory of space and gravitation one should expect the occurrence at some approximation of a superposition of noncommutative geometries.

It is worth noticing that the results of [**23**] have been recently generalized to the quiver case in [**9**]. The correspondence between [**23**] and [**9**] should read : multilinear forms or volumes $\leftrightarrow$ superpotentials.

Finally one should point out that this article is not only a survey but that it also contains new results and concepts.

Let us give some indications on the notation. Throughout the paper $\mathbb{K}$ denotes a field, all vector spaces and algebras are over $\mathbb{K}$, the dual of a vector space E is denoted by E^* and the symbol $\otimes$ denotes the tensor product over $\mathbb{K}$. Without other specifications, an algebra will always be an associative unital algebra. A graded algebra will be a $\mathbb{N}$-graded algebra $\mathcal{A} = \oplus_{n\in\mathbb{N}}\mathcal{A}_n$. Such a graded algebra is said to be connected whenever $\mathcal{A}_0 = \mathbb{K}\mathbb{1}$. Given an (r, s)-matrix A, we denote by A^t its transposed (s, r)-matrix. We use the Einstein summation convention of repeated up down indices in the formulas.

1. Regular algebras

The aim of this section is to make explicit the general class of algebras that we wish to investigate and to set up some notations.

1.1. Graded algebras. The algebras that we shall consider will be connected $\mathbb{N}$-graded algebras which are finitely generated in degree 1 and finitely presented with homogeneous relations of degrees ≥ 2. These algebras are the objects of the category **GrAlg**, the morphisms of this category being the homogeneous algebra homomorphisms of degree 0.

An algebra $\mathcal{A} \in$ **GrAlg** is of the form $\mathcal{A} = A(E, R) = T(E)/[R]$ where $E = \mathcal{A}_1$ is finite-dimensional, $R = \oplus_{n\geq 2}R_n$ is a finite-dimensional graded subspace of $T(E)$ such that (independence)

$$R_n \cap [\oplus_{m<n}R_m] = \{0\}$$

for any n ($R_n = \{0\}$ for $n < 2$) and where $[F]$ denotes for any subset $F \subset T(E)$ the two-sided ideal generated by F. The graded vector space R is the space of independent relations of $\mathcal{A}$.

By choosing a basis $(x^\lambda)_{\lambda\in\{1,\dots,g\}}$ of E and a homogeneous basis $(f_\alpha)_{\alpha\in\{1,\dots,r\}}$ of R one can also write

$$\mathcal{A} = \mathbb{K}\langle x^1,\dots,x^g\rangle/[f_1,\dots,f_r]$$

where $f_\alpha \in E^{\otimes^{N_\alpha}}$, $N_\alpha \geq 2$. Notice that $r(=\dim R)$ is well defined (i.e. it only depends on $\mathcal{A}$).

If R is concentrated in degree N (≥ 2) i.e. if $R\subset E^{\otimes^N}$ then $\mathcal{A}$ will be said to be an N-homogeneous algebra. The N-homogeneous algebras form a full subcategory $\mathbf{H_N Alg}$ of $\mathbf{GrAlg}$, [**3**], [**6**].

1.2. Dimension. Let $\mathcal{A}\in \mathbf{GrAlg}$ so $\mathcal{A} = \mathbb{K}\langle x^1,\dots,x^g\rangle/[f_1,\dots,f_r]$ and one can define $M_{\alpha\lambda}\in E^{\otimes^{N_\alpha-1}}$ by setting $f_\alpha = M_{\alpha\lambda}\otimes x^\lambda \in E^{\otimes^{N_\alpha}}$. Then the presentation of $\mathcal{A}$ by generators and relations is equivalent to the exactness of the sequence of left $\mathcal{A}$-modules [**1**]

$$\mathcal{A}^r \overset{M}{\to} \mathcal{A}^g \overset{x}{\to} \mathcal{A} \overset{\varepsilon}{\to} \mathbb{K}\to 0 \tag{1.1}$$

where M means right multiplication (in $\mathcal{A}$) by the matrix $(M_{\alpha\lambda})$, x means right multiplication by the column (x^λ) and ε is the projection onto $\mathcal{A}_0=\mathbb{K}$. In more intrinsic notation the exact sequence (1.1) reads for $\mathcal{A}=A(E,R)$

$$\mathcal{A}\otimes R\to \mathcal{A}\otimes E \overset{m}{\to} \mathcal{A}\overset{\varepsilon}{\to}\mathbb{K}\to 0 \tag{1.2}$$

where m is the multiplication of $\mathcal{A}$ and the first arrow is as in (1.1). The exact sequence (1.2) corresponding to the presentation of $\mathcal{A}$ extends as a minimal projective resolution

$$\cdots \to \mathcal{E}_n\to\mathcal{E}_{n-1}\to\mathcal{E}_0\to\mathbb{K}\to 0$$

of the left $\mathcal{A}$-module $\mathbb{K}$ which is in fact a free resolution [**11**]

$$\cdots\to\mathcal{A}\otimes E_n\to\mathcal{A}\otimes E_{n-1}\to\cdots\to\mathcal{A}\to\mathbb{K}\to 0 \tag{1.3}$$

and it follows from the very definition of $\mathrm{Ext}_{\mathcal{A}}(\mathbb{K},\mathbb{K})$ that one can make the identifications

$$E^*_n = \mathrm{Ext}^n_{\mathcal{A}}(\mathbb{K},\mathbb{K}) \tag{1.4}$$

which read $R^*=\mathrm{Ext}^2_{\mathcal{A}}(\mathbb{K},\mathbb{K})$ and $E^*=\mathrm{Ext}^1_{\mathcal{A}}(\mathbb{K},\mathbb{K})$ for $n=2$ and $n=1$. The Yoneda algebra $\mathrm{Ext}_{\mathcal{A}}(\mathbb{K},\mathbb{K})$ is the cohomology of a graded differential algebra, from which it follows that it carries a canonical A_∞-structure [**29**], [**33**]. It turns out that one can reconstruct the graded algebra $\mathcal{A}$ from the A_∞-algebra $\mathrm{Ext}_{\mathcal{A}}(\mathbb{K},\mathbb{K})$ [**30**], [**33**]. Thus the A_∞-algebra $\mathrm{Ext}_{\mathcal{A}}(\mathbb{K},\mathbb{K})$ is a natural dual of the graded algebra $\mathcal{A}$. In the case of an N-homogeneous algebra $\mathcal{A}$, there is another natural dual of $\mathcal{A}$ which is its Koszul dual $\mathcal{A}^!$ [**6**] (see below). In the case of a Koszul algebra these two notions are strongly connected and coincide in the quadratic case ($N=2$), [**7**].

The length of the resolution (1.3) is the projective dimension of the left module $\mathbb{K}$. It is classical [**11**], [**2**] that the left global dimension of $\mathcal{A}$ (for $\mathcal{A}\in\mathbf{GrAlg}$) coincides with the projective dimension of $\mathbb{K}$ as left module and that it also coincides with the right global dimension (and with the projective dimension of $\mathbb{K}$ as right module). Furthermore it has been shown recently [**4**] that this dimension also coincides with the Hochschild dimension of $\mathcal{A}$ in homology as well as in cohomology.

So for an algebra $\mathcal{A} \in \mathbf{GrAlg}$ there is a unique definition of the dimension from a homological point of view, which will be referred to as *its global dimension* in the sequel. In the following, we shall only consider algebras in **GrAlg** with finite global dimension.

It is worth noticing here that there is another dimension for $\mathcal{A} \in \mathbf{GrAlg}$, which is the Gelfand-Kirillov dimension, but since in the following polynomial growth plays no role (and therefore will not be assumed) we shall only consider the global dimension for our general analysis.

1.3. Poincaré duality. We now assume that $\mathcal{A} = A(E,R) \in \mathbf{GrAlg}$ is of finite global dimension D. The (free) resolution (1.3) of $\mathbb{K}$ then reads

$$0 \to \mathcal{A} \otimes E_D \to \cdots \to \mathcal{A} \otimes E_1 \to \mathcal{A} \to \mathbb{K} \to 0$$

with $E_1 = \mathcal{A}_1 = E$.

By applying the functor $\mathrm{Hom}_{\mathcal{A}}(\bullet, \mathcal{A})$ to the chain complex

$$0 \to \mathcal{A} \otimes E_D \to \cdots \to \mathcal{A} \otimes E \to \mathcal{A} \to 0$$

of (free) left $\mathcal{A}$-modules, one obtains a cochain complex $\mathcal{E}'$

$$0 \to \mathcal{E}'_0 \to \mathcal{E}'_1 \to \cdots \to \mathcal{E}'_D \to 0$$

of right $\mathcal{A}$-modules. The cohomology of this complex is by definition $\mathrm{Ext}_{\mathcal{A}}(\mathbb{K}, \mathcal{A})$ that is, one has

$$H^n(\mathcal{E}') = \mathrm{Ext}^n_{\mathcal{A}}(\mathbb{K}, \mathcal{A}) \tag{1.5}$$

for any $n \in \mathbb{N}$.

By definition $\mathcal{A}$ is said to be Gorenstein if one has $\mathrm{Ext}^D_{\mathcal{A}}(\mathbb{K}, \mathcal{A}) = \mathbb{K}$ and $\mathrm{Ext}^n_{\mathcal{A}}(\mathbb{K}, \mathcal{A}) = 0$ for $n \neq D$. This means that

$$0 \to \mathcal{E}'_0 \to \cdots \to \mathcal{E}'_D \to \mathbb{K} \to 0$$

is a free resolution of $\mathbb{K}$ as right $\mathcal{A}$-module. This resolution is then a minimal projective resolution of the right $\mathcal{A}$-module $\mathbb{K}$, which implies the isomorphisms

$$E^*_n \simeq E_{D-n}$$

of vector spaces and therefore

$$\dim(E_n) = \dim(E_{D-n}) \tag{1.6}$$

for $0 \le n \le D$.

Thus the Gorenstein property is a variant of the Poincaré duality property.

1.4. Regularity. Let $\mathcal{A} = A(E,R)$ be a graded algebra of **GrAlg**. $\mathcal{A}$ will be said to be *regular* if it is of finite global dimension, $\mathrm{gldim}(\mathcal{A}) = D < \infty$, and is Gorenstein. This definition of regularity is directly inspired from that of [**1**] which will be referred to as AS-regularity, the only difference being that we have dropped the condition of polynomial growth since we do not need it for the analysis in the sequel and since it would eliminate very interesting examples.

This is the class of algebras that we would like to analyse and we shall do it for low global dimensions $D = 2$ and $D = 3$. For higher global dimension, we shall restrict a little the class of algebras that we will consider. In order to understand this let us recall the following result [**7**].

PROPOSITION 1. *Let $\mathcal{A}$ be a regular algebra of global dimension D.*
(i) *If $D = 2$ then $\mathcal{A}$ is quadratic and Koszul.*
(ii) *If $D = 3$ then $\mathcal{A}$ is N-homogeneous with $N \geq 2$ and Koszul.*

Thus, for $D < 4$, regularity implies N-homogeneity (with $N = 2$ for $D = 2$) and Koszulity. We shall explain later what the Koszul property is. This is a very desirable property that one can formulate for the moment only for homogeneous algebras. This is why we shall restrict attention in the following to regular algebras which are N-homogeneous (with $N \geq 2$) and Koszul. In view of the above proposition this is not a restriction for regular algebras of global dimension $D = 2$ and $D = 3$; however one knows examples of regular algebras in global dimension 4 and higher which are not homogeneous.

2. Global dimension $D = 2$

This section is devoted to the description of the regular algebras of global dimension $D = 2$.

2.1. General results. Let us use the notation of the beginning of §1.2; so let $\mathcal{A} = \mathbb{K}\langle x^1, \dots, x^g\rangle/[f_1, \dots, f_r]$ and consider the exact sequence (1.1) corresponding to the presentation of $\mathcal{A}$. The algebra $\mathcal{A}$ has global dimension $D = 2$ if and only if (1.1) extends as an exact sequence

$$0 \to \mathcal{A}^r \stackrel{M}{\to} \mathcal{A}^g \stackrel{x}{\to} \mathcal{A} \stackrel{\varepsilon}{\to} \mathbb{K} \to 0$$

i.e. as a free resolution of $\mathbb{K}$ of length $D = 2$.

Assume now that $D = 2$ and that $\mathcal{A}$ is Gorenstein. Then the Gorenstein property implies that $r = 1$, that degree(M) = degree$(x) = 1$, so $M = (B_{\rho\lambda}x^\rho)$, and that the matrix $(B_{\lambda\mu}) \in M_g(\mathbb{K})$ is invertible. The above free resolution of $\mathbb{K}$ then reads

$$0 \to \mathcal{A} \stackrel{x^tB}{\to} \mathcal{A}^g \stackrel{x}{\to} \mathcal{A} \stackrel{\varepsilon}{\to} \mathbb{K} \to 0 \tag{2.1}$$

with obvious notations.

Conversely, let b be a nondegenerate bilinear form on $\mathbb{K}^g$ with matrix elements $B_{\lambda\mu}$ in the canonical basis and let $\mathcal{A}$ be the (quadratic) algebra generated by g generators x^λ with relation $B_{\lambda\mu}x^\lambda x^\mu = 0$; then $\mathcal{A}$ is Gorenstein of global dimension $D = 2$. One has the following theorem [**23**].

THEOREM 2. *Let b be a nondegenerate bilinear form on $\mathbb{K}^g$ ($g \geq 2$) with components $B_{\mu\nu} = b(e_\mu, e_\nu)$ in the canonical basis (e_λ) of $\mathbb{K}^g$. Then the quadratic algebra $\mathcal{A}$ generated by the elements x^λ ($\lambda \in \{1, \dots, g\}$) with the relation $B_{\mu\nu}x^\mu x^\nu = 0$ is regular of global dimension $D = 2$. Conversely, any regular algebra of global dimension $D = 2$ is of the above kind for some $g \geq 2$ and some nondegenerate bilinear form b on $\mathbb{K}^g$. Furthermore, two such algebras $\mathcal{A}$ and $\mathcal{A}'$ are isomorphic if and only if $g = g'$ and $b' = b \circ L$ for some $L \in GL(g, \mathbb{K})$.*

The last part of this theorem is almost obvious and gives a description of the moduli space of the regular algebras of global dimension $D = 2$.

The right action $b \mapsto b \circ L$ of the linear group on bilinear forms is a particular case of the right action of the linear group $GL(V)$ on multilinear forms on a vector space V defined for an n-linear form w by

$$w \circ L(v_1, \dots, v_n) = w(Lv_1, \dots, Lv_n) \tag{2.2}$$

for any $v_k \in V$, $k \in \{1, \dots, n\}$.

For reasons which will become clear, the algebra $\mathcal{A}$ (regular of global dimension $D = 2$) associated to the nondegenerate bilinear form b on $\mathbb{K}^g$ as in Theorem 2 will be denoted $\mathcal{A}(b, 2)$ in the following.

2.2. Poincaré series and polynomial growth. Let $\mathcal{A}$ be a regular algebra of global dimension $D = 2$. Then the exact sequence (2.1) splits as

$$0 \to \mathcal{A}_{n-2} \overset{x^tB}{\to} \mathcal{A}^g_{n-1} \overset{x}{\to} \mathcal{A}_n \to 0$$

for $n \neq 0$ with of course $\mathcal{A}_0 = \mathbb{K}$ and $\mathcal{A}_n = 0$ for $n < 0$. It follows that the Poincaré series $P_\mathcal{A}(t)$ of $\mathcal{A}$ is given by

$$P_\mathcal{A}(t) = \frac{1}{1 - gt + t^2} \tag{2.3}$$

in view of the Euler-Poincaré formula.

For $g = 2$ one has

$$P_\mathcal{A}(t) = \left(\frac{1}{1-t}\right)^2$$

so $\mathcal{A}$ then has polynomial growth (with GKdim=2) while for $g > 2$ one has

$$P_\mathcal{A}(t) = \frac{1}{(1 - k^{-1}t)(1 - kt)}$$

with

$$k = \frac{1}{2}(g + \sqrt{g^2 - 4}) > 1$$

so $\mathcal{A}$ then has exponential growth.

Let us now discuss the case of the regular algebras of global dimension 2 with $g = 2$ generators i.e. which have polynomial growth. In view of Theorem 2 these algebras are classified by the $GL(2, \mathbb{K})$-orbits of nondegenerate bilinear forms on $\mathbb{K}^2$. Assuming that $\mathbb{K}$ is algebraically closed, it is easy to classify these $GL(2, \mathbb{K})$-orbits of nondegenerate bilinear forms according to the rank **rk** of their symmetric parts [**24**] :

(0) **rk** $= 0$ - there is only one orbit, which is the orbit of the bilinear form $b = \varepsilon$ with matrix of components

$$B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

which corresponds to the relations $x^1x^2 - x^2x^1 = 0$, so $\mathcal{A}$ is isomorphic to the polynomial algebra $\mathbb{K}[x^1, x^2]$,

(1) **rk** $= 1$ - there is only one orbit, which is the orbit of the bilinear form b with matrix of components

$$B = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}$$

which corresponds to the relations $x^1x^2 - x^2x^1 - (x^2)^2 = 0$,

(2) **rk** $= 2$ - there is a 1-parameter family of orbits which are the orbits of the bilinear forms $b = \varepsilon_q$ with matrices of components

$$B = \begin{pmatrix} 0 & -1 \\ q & 0 \end{pmatrix}$$

for $q \in \mathbb{K}$ with $q^2 - q \neq 0$ modulo $q \sim q^{-1}$, which corresponds to the relations $x^1x^2 - qx^2x^1 = 0$.

The case (0) corresponds to the ordinary plane, the case (1) corresponds to the Jordanian plane and the cases (2) correspond to the Manin planes. One thus recovers the usual description of the algebras which are regular in the sense of [**1**] i.e. AS-regular of global dimension 2, [**28**], [**1**].

2.3. Hecke symmetries. Any linear mapping

$$R : \mathbb{K}^g \otimes \mathbb{K}^g \to \mathbb{K}^g \otimes \mathbb{K}^g$$

is characterized by its components $R^{\mu\nu}_{\lambda\rho}$ defined by

$$R(e_\lambda \otimes e_\rho) = R^{\mu\nu}_{\lambda\rho} e_\mu \otimes e_\nu$$

in the canonical basis (e_λ) of $\mathbb{K}^g$.

Let b be a nondegenerate bilinear form on $\mathbb{K}^g$ with components $B_{\lambda\rho} = b(e_\lambda, e_\rho)$ and let $K^{\mu\nu}$ be the components of a bilinear form on the dual vector space of $\mathbb{K}^g$ in the dual basis of (e_λ). Define the endomorphism R of $\mathbb{K}^g \otimes \mathbb{K}^g$ by setting

$$R^{\mu\nu}_{\lambda\rho} = \delta^\mu_\lambda \delta^\nu_\rho + K^{\mu\nu} B_{\lambda\rho} \tag{2.4}$$

for $\mu, \nu, \lambda, \rho \in \{1, \dots, g\}$. Assume now that the above R defined by (2.4) satisfies the Yang-Baxter equation

$$(I \otimes R)(R \otimes I)(I \otimes R) = (R \otimes I)(I \otimes R)(R \otimes I) \tag{2.5}$$

on $(\mathbb{K}^g)^{\otimes^3}$ where I denotes the identity mapping of $\mathbb{K}^g$ onto itself. One verifies that (2.5) is equivalent to

$$\left\{ \begin{array}{l} KBK^tB^t + (1 + \mathrm{tr}(KB^t))\mathbb{1} = 0 \\ K^tB^tKB + (1 + \mathrm{tr}(KB^t))\mathbb{1} = 0 \end{array} \right. \tag{2.6}$$

where K and B are the matrices $(K^{\mu\nu})$ and $(B_{\lambda\rho})$ of $M_g(\mathbb{K})$ and where the product is the matrix product. Equations (2.6) then imply that one has

$$(R - \mathbb{1})(R - (1 + \mathrm{tr}(KB^t))\mathbb{1}) = 0 \tag{2.7}$$

which means that R is a Hecke symmetry in the terminology of [**26**].

Given the nondegenerate bilinear for b, one can always solve (2.6). For instance

$$K = qB^{-1} \tag{2.8}$$

with $q \in \mathbb{K}$ such that

$$q + q^{-1} + \mathrm{tr}(B^{-1}B^t) = 0 \tag{2.9}$$

is a solution of (2.6). The corresponding Hecke symmetries will be called *the standard Hecke symmetries associated with* (the nondegenerate bilinear form) b while more generally the Hecke symmetries associated with the solutions of (2.6) will be said to be *associated with* b. There are generically two standard Hecke symmetries corresponding to the two roots of Equation (2.9).

Notice that (2.6) implies that $K \neq 0$ so if R is a Hecke symmetry associated to b, the defining relation of $\mathcal{A}(b,2)$ namely $B_{\mu\nu}x^\mu x^\nu = 0$ is equivalent to the quadratic relations

$$x^\mu x^\nu = R^{\mu\nu}_{\lambda\rho} x^\lambda x^\rho \tag{2.10}$$

for $\mu, \nu \in \{1, \dots, g\}$.

In the case $g = 2$ with $b = \varepsilon_q$, i.e.

$$B = \begin{pmatrix} 0 & -1 \\ q & 0 \end{pmatrix}$$

with $q \neq 0$ which includes cases (0) and (2) of §2.2, one can take

$$K = \begin{pmatrix} 0 & 1 \\ -p & 0 \end{pmatrix}, \ p \in \mathbb{K}$$

as solution of (2.6). Equation (2.7) then reads

$$(R - \mathbb{1})(R + pq) = 0$$

and for $p = q$, R is a standard Hecke symmetry for $b = \varepsilon_q$. In the classical situation $q = 1$, i.e. for $\mathcal{A} = \mathbb{K}[x^1, x^2]$, the two standard Hecke symmetries coincide and reduce to the flip

$$x \otimes y \mapsto y \otimes x$$

of $\mathbb{K}^2 \otimes \mathbb{K}^2$.

2.4. Actions of quantum groups. There are quantum groups acting on the noncommutative planes corresponding to the regular algebras of global dimension $D = 2$. For the Manin planes corresponding to the $\mathcal{A}(\varepsilon_q, 2)$ these are the quantum groups $SL_q(2)$, $GL_q(2)$ and $GL_{p,q}(2)$ [**34**], [**35**].

For the noncommutative plane corresponding to $\mathcal{A}(b,2)$ where b is a nondegenerate bilinear form on $\mathbb{K}^g$, the generalization of $SL_q(2)$ is the quantum group of the nondegenerate bilinear form b [**24**]. Let us recall the definition of this object. Let $\mathcal{H}(b)$ be the unital associative algebra generated by the g^2 elements u^μ_ν ($\mu, \nu \in \{1, \dots, g\}$) with the relations

$$B_{\lambda\rho} u^\lambda_\mu u^\rho_\nu = B_{\mu\nu}\mathbb{1} \tag{2.11}$$

and

$$B^{\mu\nu} u^\lambda_\mu u^\rho_\nu = B^{\lambda\rho}\mathbb{1} \tag{2.12}$$

where the $B^{\mu\nu}$ are the matrix elements of the inverse matrix B^{-1} of the matrix B of the components $B_{\mu\nu} = b(e_\mu, e_\nu)$ of b. One verifies easily that there is a unique structure of Hopf algebra on $\mathcal{H}(b)$ with coproduct Δ, counit ε and antipode S such that

$$\Delta(u^\mu_\nu) = u^\mu_\rho \otimes u^\rho_\nu \tag{2.13}$$

$$\varepsilon(u^\mu_\nu) = \delta^\mu_\nu \tag{2.14}$$

$$S(u^\mu_\nu) = B^{\mu\lambda} B_{\rho\nu} u^\rho_\lambda \tag{2.15}$$

the product and the unit being the original ones on $\mathcal{H}(b)$.

There is a canonical algebrahomomorphism $\Delta_L : \mathcal{A}(b,2) \to \mathcal{H}(b) \otimes \mathcal{A}(b,2)$ such that

$$\Delta_L(x^\lambda) = u^\lambda_\mu \otimes x^\mu$$

for $\lambda \in \{1, \dots, g\}$. This equips $\mathcal{A}(b,2)$ with a structure of $\mathcal{H}(b)$-comodule. The dual object of $\mathcal{H}(b)$ is the quantum group of the nondegenerate bilinear form b. The analysis of the category of representations of this quantum group has been done in [**8**]. To the coaction Δ_L of $\mathcal{H}(b)$ on $\mathcal{A}(b,2)$ corresponds an action of this quantum group on the noncommutative plane corresponding to $\mathcal{A}(b,2)$.

The (quadratic) homogeneous part of the relations (2.11) and (2.12) reads

$$u^\mu_\alpha u^\nu_\beta R^{\alpha\beta}_{\lambda\rho} = R^{\mu\nu}_{\alpha\beta} u^\alpha_\lambda u^\beta_\rho \tag{2.16}$$

where R is a standard Hecke symmetry of b. In fact (2.16) together with (2.13) and (2.14) define a bialgebra with counit for any R. In the case where R is a standard Hecke symmetry then $B^{\mu\nu} B_{\rho\lambda} u^\lambda_\mu u^\rho_\nu$ is in the center and the Hopf algebra $\mathcal{H}(b)$ corresponding to the quantum group of b is the quotient of the bialgebra by the ideal generated by the element

$$B^{\mu\nu} B_{\rho\lambda} u^\lambda_\mu u^\rho_\nu - g\mathbb{1}$$

of the center. In fact $\mathcal{H}(b)$ is a quotient of a bigger Hopf algebra associated with the homogeneous relations (2.16) which is the generalization of the Hopf algebra corresponding to $GL_q(2)$ in the case $b = \varepsilon_q$, $(g = 2)$. More generally, if R is an arbitrary Hecke symmetry associated with b, there is a Hopf algebra associated with the quadratic relations (2.16) which coacts on $\mathcal{A}(b,2)$ and corresponds to the generalization of $GL_{p,q}(2)$.

3. Global dimension $D = 3$

In this section we shall analyse regular algebras of global dimension $D = 3$ and describe some representative examples. For global dimensions $D \geq 3$ what replaces the bilinear forms of global dimension $D = 2$ (last section) are multilinear forms, so we start this section with a discussion on multilinear forms.

3.1. Multilinear forms. Let V be a vector space with $\dim(V) \geq 2$, Q be an element of the linear group $GL(V)$ and m be an integer with $m \geq 2$. Then an m-linear form w on V (i.e. a linear form on $V^{\otimes^m}$) will be said to be *Q-cyclic* if one has

$$w(X_1, \dots, X_m) = w(QX_m, X_1, \dots, X_{m-1}) \tag{3.1}$$

for any $X_1, \dots, X_m \in V$.

Let w be Q-cyclic; then one has

$$w(X_1, \dots, X_m) = w(QX_k, \dots, QX_m, X_1, \dots, X_{k-1})$$

for $1 \leq k \leq m$ so in particular one has

$$w(X_1, \dots, X_m) = w(QX_1, \dots, QX_m)$$

for any $X_1, \dots, X_m \in V$, which also reads $w = w \circ Q$ and means that w is invariant by Q.

Now let now w be an arbitrary Q-invariant m-linear form on V; then the m-linear form $\pi_Q(w)$ on V defined by

$$\pi_Q(w)(X_1, \dots, X_m) = \frac{1}{m} \sum_{k=1}^{m} w(QX_k, \dots, QX_m, X_1, \dots, X_{k-1})$$

for any $X_1, \dots, X_m \in V$ is Q-cyclic and this defines a projection π_Q of the space of Q-invariant m-linear forms onto the space of Q-cyclic m-linear forms on V. This projection is $GL(V)$-equivariant in the sense that if w is Q-invariant (resp. Q-cyclic) then $w \circ L$ is $L^{-1}QL$-invariant (resp. $L^{-1}QL$-cyclic) for any $L \in \mathrm{GL}(V)$.

The m-linear form w on V will be said to be *preregular* if it satisfies the following conditions (i) and (ii) :
(i) $w(X, X_1, \dots, X_{m-1}) = 0$ for any $X_1, \dots, X_{m-1} \in V$ implies $X = 0$,
(ii) there is a $Q_w \in GL(V)$ such that w is Q_w-cyclic.

Condition (i) implies that Q_w is unique under (ii) and Conditions (ii) and (i) imply that w satisfies the following condition (i') which is stronger than (i) :
(i') $w(X_1, \dots, X_k, X, X_{k+1}, \dots, X_{m-1}) = 0$ for any $X_1, \dots, X_{m-1} \in V$ implies $X = 0$, for any $k \in \{0, \dots, m-1\}$.

An m-linear form w on V satisfying (i') will be said to be 1-*site-nondegenerate*. The set of preregular m-linear forms on V is invariant by the action of $GL(V)$ and one has

$$Q_{w \circ L} = L^{-1} Q_w L \tag{3.2}$$

for any preregular m-linear form w on V.

A bilinear form b on $\mathbb{K}^g$ is preregular if and only if it is nondegenerate; one then has $Q_b = (B^{-1})^t B$ where B is the matrix of components of b.

The condition of preregularity will be involved throughout the paper. We now introduce a stronger condition which is involved specifically in the description of the regular algebras of global dimension $D = 3$. Let N be an integer with $N \geq 2$; then an $(N+1)$-linear form w on V will be said to be 3-*regular* if it is preregular and satisfies the following condition (iii) :
(iii) If L_0 and L_1 are endomorphisms of V satisfying

$$w(L_0X_0, X_1, X_2, \dots, X_N) = w(X_0, L_1X_1, X_2, \dots, X_N)$$

for any $X_0, \dots, X_n \in V$, then $L_0 = L_1 = k\mathbb{1}$ for some $k \in \mathbb{K}$.

The set of 3-regular $(N+1)$-linear forms is also invariant by $GL(V)$.

Condition (iii) is a sort of 2-sites nondegeneracy condition. Consider the stronger condition (iii') :
(iii') $\sum_i w(Y_i, Z_i, X_1, \dots, X_{N-1}) = 0$ for any $X_1, \dots, X_{N-1} \in V$ implies

$$\sum_i Y_i \otimes Z_i = 0.$$

It is clear that (iii') $\Rightarrow$ (iii), however it is a strictly stronger condition. For instance let ε be the completely antisymmetric $(N+1)$-linear form on $\mathbb{K}^{N+1}$ with $\varepsilon(e_0, \dots, e_N) = 1$. Then ε is 3-regular but one has

$$\varepsilon(Y, Z, X_1, \dots, X_{N-1}) + \varepsilon(Z, Y, X_1 \dots, X_{N-1}) = 0$$

identically and this does not imply $Y \otimes Z + Z \otimes Y = 0$.

3.2. General results for $D = 3$. Let w be a preregular $(N+1)$-linear form on $\mathbb{K}^g$ with components $W_{\lambda_0 \dots \lambda_N} = w(e_{\lambda_0}, \dots, e_{\lambda_N})$ in the canonical basis (e_λ) of $\mathbb{K}^g$ and let $\mathcal{A}(w, N)$ be the N-homogeneous algebra generated by the g elements x^λ ($\lambda \in \{1, \dots, g\}$) with the g relations

$$W_{\lambda\lambda_1 \dots \lambda_N} x^{\lambda_1} \cdots x^{\lambda_N} = 0 \tag{3.3}$$

for $\lambda \in \{1, \dots, g\}$. In other words one has $\mathcal{A}(w, N) = A(E, R)$ with $E = \oplus_\lambda \mathbb{K} x^\lambda$ and $R = \sum_\lambda \mathbb{K} W_{\lambda\lambda_1 \dots \lambda_N} x^{\lambda_1} \otimes \cdots \otimes x^{\lambda_N}$. Condition (i) implies that $\dim(R) = g$, that is, that the latter sum is direct and that the relations (3.3) are independent.

Let us now use again the notation of the beginning of §1.2, so let $\mathcal{A} \in$ **GrAlg** with $\mathcal{A} = \mathbb{K}\langle x^1, \dots, x^g \rangle / [f_1, \dots, f_r]$ and consider the exact sequence (1.1) corresponding to the presentation of $\mathcal{A}$. Then $\mathcal{A}$ has global dimension $D = 3$ if and only if (1.1) extends as an exact sequence

$$0 \to \mathcal{A}^s \to \mathcal{A}^r \stackrel{M}{\to} \mathcal{A}^g \stackrel{x}{\to} \mathcal{A} \stackrel{\varepsilon}{\to} \mathbb{K} \to 0$$

i.e. as a free resolution of $\mathbb{K}$ of length $D = 3$. Assume now that $\mathcal{A}$ is regular. Then the Gorenstein property (Poincaré duality) implies immediately that $r = g$, that $s = 1$, that the above resolution reads with an appropriate choice of the relations f_λ

$$0 \to \mathcal{A} \stackrel{x^t}{\to} \mathcal{A}^g \stackrel{M}{\to} \mathcal{A}^g \stackrel{x}{\to} \mathcal{A} \stackrel{\varepsilon}{\to} \mathbb{K} \to 0 \tag{3.4}$$

and that $w = x^\lambda \otimes f_\lambda$ is homogeneous, say of degree $N+1$, and is preregular [**1**]. So $\mathcal{A} = \mathcal{A}(w, N)$ as above. In fact one has the following theorem [**23**].

THEOREM 3. *Let $\mathcal{A}$ be a regular algebra of global dimension $D = 3$. Then $\mathcal{A} = \mathcal{A}(w, N)$ for some $N \geq 2$, some $g \geq 2$ and some 3-regular $(N+1)$-linear form w on $\mathbb{K}^g$.*

The Poincaré series of $\mathcal{A} = \mathcal{A}(w, N)$ as in the above theorem (i.e. $\mathcal{A}$ regular with $D = 3$) is given by

$$P_{\mathcal{A}}(t) = \frac{1}{1 - gt + gt^N - t^{N+1}} \tag{3.5}$$

in view of (3.4).

If one compares this theorem with Theorem 2 for $D = 2$, one sees that there are two missing items : first there is no converse of the statement in Theorem 3 and second there is no characterization of the isomorphism classes. Concerning the first point, it was conjectured in [**23**] that given a 3-regular $(N+1)$-linear form w on $\mathbb{K}^g$ then $\mathcal{A}(w, N)$ is a regular algebra with $D = 3$, but unfortunately this is wrong and we shall give counterexamples (see below). This means that one has to find some slightly stronger condition than 3-regularity for w (for $D = 3$).

Concerning the second point the following result holds (independently of the regularity of the algebras) [**23**].

Proposition 4. *Let w be a 3-regular $(N+1)$-linear form on $\mathbb{K}^g$ and let w' be a 3-regular $(N'+1)$-linear form on $\mathbb{K}^{g'}$. Then $\mathcal{A}(w, N)$ and $\mathcal{A}(w', N')$ are isomorphic if and only if $g' = g$, $N' = N$ and $w' = w \circ L$ for some $L \in GL(g, \mathbb{K})$.*

The conditions $g' = g$ and $N' = N$ are clear but the 3-regularity is genuinely involved in the proof of this proposition (see in [**23**]).

Following [**1**] one deduces from (3.5) that a regular algebra of global dimension $D = 3$ has polynomial growth if and only if $g = 3$ and $N = 2$ or $g = 2$ and $N = 3$; Otherwise it has exponential growth (for $g \geq 2$ and $N \geq 2$).

3.3. Examples and counterexamples. All AS-regular algebras of global dimension $D = 3$ of course give examples and our notations w, M, Q_w come from [**1**]. In fact, the classification of the regular algebras of global dimension $D = 3$ with polynomial growth is based on the possible Jordan decompositions of the corresponding Q_w's. Let us give some representative examples.
(a) *The 3-dimensional Sklyanin algebra* [**37**], [**36**]. This is the algebra $\mathcal{A}$ generated by 3 elements x, y, z with relations

$$\left\{\begin{array}{l} xy - qyx = pz^2 \\ yz - qzy = px^2 \\ zx - qxz = py^2 \end{array}\right. \tag{3.6}$$

where $p, q \in \mathbb{K}$ with $(p, q) \neq (0,0)$ and $(p^3 + 1, q^3 + 1) \neq (0,0)$.

This algebra is AS-regular with $D = 3$. One has $\mathcal{A} = \mathcal{A}(w, 2)$ with

$$\begin{array}{rcl} w & = & x \otimes y \otimes z + y \otimes z \otimes x + z \otimes x \otimes y \\ & - & q(x \otimes z \otimes y + y \otimes x \otimes z + z \otimes y \otimes x) \\ & - & p(x \otimes x \otimes x + y \otimes y \otimes y + z \otimes z \otimes z) \end{array} \tag{3.7}$$

where we have identified the 3-linear form w on $\mathbb{K}^3$ with the corresponding element of $(\mathbb{K}^{3*})^{\otimes^3}$. One verifies that w is 3-regular and one has

$$Q_w = \mathbb{1} \tag{3.8}$$

for the corresponding element of $GL(3, \mathbb{K})$.

(b) *The q-deformed 3-dimensional polynomial algebra*. This is the algebra $\mathcal{A}$ generated by 3 elements x, y, z with relations

$$\left\{ \begin{array}{l} xy = qcyx \\ yz = qazy \\ zx = qbxz \end{array} \right. \tag{3.9}$$

with $q, a, b, c \in \mathbb{K}$, $abc = 1$ and $q \neq 0$. This algebra is AS-regular with $D = 3$ and one has $\mathcal{A} = \mathcal{A}(w, 2)$ with

$$w = bx \otimes y \otimes z + cy \otimes z \otimes x + az \otimes x \otimes y - q(abx \otimes z \otimes y + bcy \otimes x \otimes z + caz \otimes y \otimes x) \tag{3.10}$$

with the same conventions as above. One verifies that w is 3-regular and one has

$$Q_w = \begin{pmatrix} b/c & 0 & 0 \\ 0 & c/a & 0 \\ 0 & 0 & a/b \end{pmatrix} \tag{3.11}$$

(c) *Type E quadratic AS-algebra* [**1**]. This is the algebra $\mathcal{A}$ generated by 3 elements x, y, z with relations

$$\left\{ \begin{array}{l} x^2 + \zeta^{-1} yz + \zeta\ zy = 0 \\ y^2 + \zeta^{-4} zx + \zeta^4 xz = 0 \\ z^2 + \zeta^{-7} xy + \zeta^7 yx = 0 \end{array} \right. \tag{3.12}$$

where $\zeta \in \mathbb{K}$ is a primitive 9th root of 1, $\zeta^9 = 1$.
This algebra is AS-regular with $D = 3$ and $\mathcal{A} = \mathcal{A}(w, 2)$ with

$$\begin{array}{rcl} w & = & x \otimes z \otimes x + y \otimes x \otimes y + z \otimes y \otimes z \\ & + & \zeta\ z \otimes x \otimes x + \zeta^{-1} x \otimes x \otimes z \\ & + & \zeta^4 x \otimes y \otimes y + \zeta^{-4} y \otimes y \otimes x \\ & + & \zeta^7 y \otimes z \otimes z + \zeta^{-7} z \otimes z \otimes y \end{array} \tag{3.13}$$

which defines a 3-regular 3-linear form on $\mathbb{K}^3$.

One has

$$Q_w = \begin{pmatrix} \zeta & 0 & 0 \\ 0 & \zeta^4 & 0 \\ 0 & 0 & \zeta^7 \end{pmatrix} \tag{3.14}$$

for the corresponding element of $GL(3, \mathbb{K})$.

It is worth noticing here that the algebras of Case (a) and Case (b) are deformations of the polynomial algebra $\mathbb{K}[x, y, z]$ while this is not the case here. In fact the algebra with relations (3.12) is quite rigid.

(d) *Counterexample to the converse of Theorem 3.* Let $\mathcal{A}$ be the algebra generated by 3 elements x, y, z with relations

$$\left\{ \begin{array}{l} x^2 + yz = 0 \\ y^2 + zx = 0 \\ xy = 0 \end{array} \right. \tag{3.15}$$

Then $\mathcal{A} = \mathcal{A}(w, 2)$ where the 3-linear form w on $\mathbb{K}^3$ is given by

$$w = x \otimes x \otimes x + y \otimes y \otimes y + x \otimes y \otimes z + y \otimes z \otimes x + z \otimes x \otimes y \tag{3.16}$$

with the same conventions as before. One verifies that w is again 3-regular and one has $Q_w = \mathbb{1}$. However, $\mathcal{A}$ is not regular of global dimension $D = 3$. Indeed the candidate for (3.4) is

$$0 \to \mathcal{A} \xrightarrow{x^t} \mathcal{A}^3 \xrightarrow{M} \mathcal{A}^3 \xrightarrow{x} \mathcal{A} \xrightarrow{\varepsilon} \mathbb{K} \to 0$$

with $x^t = (x, y, z)$ and

$$M = \begin{pmatrix} x & 0 & y \\ z & y & 0 \\ 0 & x & 0 \end{pmatrix}$$

but this complex is *not* exact in second position : One has

$$(yz, 0, 0) \in \mathrm{Ker}(M)$$

while $(yz, 0, 0)$ is not in the image of x^t.
This algebra is discussed in [**1**] and there is a similar one which is cubic with 2 generators.

(e) *The Yang-Mills algebra* [**17**]. The Yang-Mills algebra is the cubic algebra $\mathcal{A}$ generated by g elements ∇_λ ($\lambda \in \{1, \dots, g\}$) with relations

$$g^{\lambda\mu}[\nabla_\lambda, [\nabla_\mu, \nabla_\nu]] = 0 \tag{3.17}$$

for $\nu \in \{1, \dots, g\}$, where the $g^{\lambda\mu}$ are the components of a symmetric nondegenerate bilinear form on $\mathbb{K}^g$. The use here of covariant instead of contravariant notational conventions has a physical origin. This algebra is regular of global dimension $D = 3$. One has $\mathcal{A} = \mathcal{A}(w, 3)$ where w is the 4-linear form on $\mathbb{K}^g$ with components

$$W^{\alpha_1\alpha_2\alpha_3\alpha_4} = g^{\alpha_1\alpha_2}g^{\alpha_3\alpha_4} + g^{\alpha_2\alpha_3}g^{\alpha_4\alpha_1} - 2g^{\alpha_1\alpha_3}g^{\alpha_2\alpha_4} \tag{3.18}$$

for $\alpha_k \in \{1, \dots, g\}$. This 4-linear form on $\mathbb{K}^g$ is 3-regular, in fact it satisfies the strong condition (iii'), and one has $Q_w = \mathbb{1}$.

(f) *The super Yang-Mills algebra* [**19**]. There is a "super" version of the Yang-Mills algebra which is the cubic algebra $\tilde{\mathcal{A}}$ generated by g elements S_λ ($\lambda \in \{1, \dots, g\}$) with relations

$$g^{\lambda\mu}[S_\lambda, [S_\mu, S_\nu]_+] = 0 \tag{3.19}$$

for $\nu \in \{1, \dots, g\}$, where the $g^{\lambda\mu}$ are as above and $[A, B]_+ = AB + BA$.

This algebra is again regular of global dimension 3 and $\tilde{\mathcal{A}} = \mathcal{A}(\tilde{w}, 3)$ where $\tilde{w}$ is the 4-linear form on $\mathbb{K}^g$ with components

$$\tilde{W}^{\alpha_1\alpha_2\alpha_3\alpha_4} = g^{\alpha_2\alpha_3}g^{\alpha_4\alpha_1} - g^{\alpha_1\alpha_2}g^{\alpha_3\alpha_4} \tag{3.20}$$

for $\alpha_k \in \{1, \dots, g\}$. This $\tilde{w}$ is 3-regular (and satisfies (iii')) and $Q_{\tilde{w}} = -\mathbb{1}$. Notice that the equations (3.19) are equivalent to

$$[S_\lambda, g^{\mu\nu}S_\mu S_\nu] = 0 \tag{3.21}$$

i.e. to the fact that $g^{\mu\nu}S_\mu S_\nu$ is central.

Before leaving this section, it is worth noticing that the Yang-Mills algebra is by its very definition the universal enveloping algebra of a graded Lie algebra. In the case $g = 2$ this is an AS-regular algebra considered in [**1**] which is the universal enveloping algebra of the graded 3-dimensional Lie algebra with basis (∇_1, ∇_2) in degree 1 and C in degree 2 with Lie bracket defined by

$$[\nabla_1, \nabla_2] = C, [\nabla_1, C] = 0, [\nabla_2, C] = 0.$$

In the case $g > 2$, the Yang-Mills algebra has exponential growth.

Similar considerations apply to the super Yang-Mills where the above Lie algebra is replaced by a super Lie algebra.

4. Homogeneous algebras

The aim of this section is to describe properties of N-homogeneous algebras and to introduce and discuss the Koszul property [**3**], [**6**].

4.1. Koszul duality. Let $\mathcal{A} \in \mathbf{H}_N\mathbf{Alg}$ be an N-homogeneous algebra, that is, $\mathcal{A} = A(E, R)$ with $R \subset E^{\otimes^N}$. One defines the (Koszul) *dual* $\mathcal{A}^!$ of $\mathcal{A}$ to be the N-homogeneous algebra

$$\mathcal{A}^! = \mathcal{A}(E^*, R^\perp) \tag{4.1}$$

where $R^\perp \subset E^{*\otimes^N} = (E^{\otimes^N})^*$ is the annihilator of R, i.e. the subspace

$$R^\perp = \{\omega \in (E^{\otimes^N})^* | \omega(x) = 0,\ \forall x \in R\}$$

of $(E^{\otimes^N})^*$ identified with $E^{*\otimes^N}$ (there a canonical identification of $(E^{\otimes^N})^*$ with $E^{*\otimes^N}$ since E is finite-dimensional). One has canonically

$$(\mathcal{A}^!)^! = \mathcal{A} \tag{4.2}$$

and to any morphism $f : \mathcal{A} \to \mathcal{A}' = A(E', R')$ of $\mathbf{H}_N\mathbf{Alg}$ corresponds a morphism $f^! : \mathcal{A}'^! \to \mathcal{A}^!$ which is induced by the transpose of the restriction $f \restriction E : E \to E'$ of f to E. The correspondence $(\mathcal{A} \mapsto \mathcal{A}^!, f \mapsto f^!)$ defines a contravariant involutive functor $((f^!)^! = f)$.

4.2. The Koszul N-complex $K(\mathcal{A})$. .
Let $\mathcal{A} = A(E, R)$ be an N-homogeneous algebra with dual $\mathcal{A}^! = \oplus_n \mathcal{A}^!_n$ and consider the dual vector spaces $\mathcal{A}^{!*}_n$ of the $\mathcal{A}^!_n$. One has

$$\begin{cases} \mathcal{A}^{!*}_n = E^{\otimes^n} \quad \text{for} \quad n < N \\ \mathcal{A}^{!*}_n = \cap_{r+s=n-N} E^{\otimes^r} \otimes R \otimes E^{\otimes^s} \quad \text{for} \quad n \geq N \end{cases} \tag{4.3}$$

so that for any $n \in \mathbb{N}$ one has $\mathcal{A}^{!*}_n \subset E^{\otimes^n}$. Let us then define the sequence of homomorphisms of (free) left $\mathcal{A}$-modules

$$\cdots \xrightarrow{d} \mathcal{A} \otimes \mathcal{A}^{!*}_{n+1} \xrightarrow{d} \mathcal{A} \otimes \mathcal{A}^{!*}_n \xrightarrow{d} \cdots \xrightarrow{d} \mathcal{A} \to 0 \tag{4.4}$$

where $d : \mathcal{A} \otimes \mathcal{A}^{!*}_{n+1} \to \mathcal{A} \otimes \mathcal{A}^{!*}_n$ is induced by the map

$$a \otimes (e_0 \otimes e_1 \otimes \cdots \otimes e_n) \mapsto ae_0 \otimes (e_1 \otimes \cdots \otimes e_n)$$

of $\mathcal{A} \otimes E^{\otimes^{n-1}}$ into $\mathcal{A} \otimes E^{\otimes^n}$. Then one has

$$d^N = 0 \tag{4.5}$$

since $\mathcal{A}^{!*}_n \subset R \otimes E^{\otimes^{n-N}}$ for $n \geq N$. Thus (4.4) defines an N-complex which will be refered to as *the Koszul N-complex of* $\mathcal{A}$ and denoted by $K(\mathcal{A})$.

As for any N-complex [**21**] one obtains from $K(\mathcal{A})$ a family $C_{p,r}(K(\mathcal{A}))$ of ordinary complexes, called *the contractions of* $K(\mathcal{A})$, by putting together alternatively p and $N-p$ arrows d of $K(\mathcal{A})$. The complex $C_{p,r}(K(\mathcal{A}))$ is defined as

$$\cdots \stackrel{d^{N-p}}{\to} \mathcal{A}\otimes\mathcal{A}^{!*}_{Nk+r} \stackrel{d^p}{\to} \mathcal{A}\otimes\mathcal{A}^{!*}_{Nk-p+r} \stackrel{d^{N-p}}{\to} \mathcal{A}\otimes\mathcal{A}^{!*}_{N(k-1)+r} \stackrel{d^p}{\to}\cdots \tag{4.6}$$

for $0\leq r<p\leq N-1$ [**6**] (one verifies that all such complexes are exhausted by these couples (p,r)). For the homology of these complexes one has the following result [**6**].

PROPOSITION 5. *Let $\mathcal{A}=A(E,R)$ be an N-homogeneous algebra with $N\geq 3$. Assume that (p,r) is distinct from $(N-1,0)$ and that $C_{p,r}(K(\mathcal{A}))$ is exact at degree $k=1$. Then $R=0$ or $R=E^{\otimes^N}$.*

Except for $C_{N-1,0}(K(\mathcal{A}))$, a nontrivial acyclicity for $C_{p,r}(K(\mathcal{A}))$ leads to the trivial algebras $\mathcal{A}=T(E)$ or $\mathcal{A}=T(E)^!$.

4.3. Koszul complexes and Koszul property. The last proposition points out the complex $C_{N-1,0}(K(\mathcal{A}))$, which will be denoted by $\mathcal{K}(\mathcal{A},\mathbb{K})$ and referred to as *the Koszul complex of* $\mathcal{A}$. It coincides with the Koszul complex originality introduced in [**3**] without mention to the N-complex $K(\mathcal{A})$. Of course for a quadratic algebra $\mathcal{A}$, i.e. for $N=2$, one has $K(\mathcal{A})=\mathcal{K}(\mathcal{A},\mathbb{K})$ and this coincides with the definition of [**38**] (see also [**34**], [**35**]).

An N-homogeneous algebra $\mathcal{A}$ will be said to be a *Koszul algebra* whenever its Koszul complex $\mathcal{K}(\mathcal{A},\mathbb{K})$ is acyclic in positive degrees, (i.e. $H_n(\mathcal{K}(\mathcal{A},\mathbb{K}))=0$ for $n\geq 1$). This is the generalization given in [**3**] of the definition of [**38**] to N-homogeneous algebras. There are very good reasons explained in [**3**] for this generalization. We content ourselves here with observing that, among the contractions of $K(\mathcal{A})$, the Koszul complex $\mathcal{K}(\mathcal{A},\mathbb{K})$ is distinguished by the fact that it terminates as a projective resolution of $\mathbb{K}$. Indeed, the presentation of $\mathcal{A}=A(E,R)$ is equivalent to the exactness of the sequence

$$\mathcal{A}\otimes R \stackrel{d^{N-1}}{\to} \mathcal{A}\otimes E \stackrel{d}{\to} \mathcal{A} \stackrel{\varepsilon}{\to} \mathbb{K}\to 0$$

as observed before and, on the other hand one has $\mathcal{A}^{!*}_1=E$ and $\mathcal{A}^{!*}_N=R$ so $\mathcal{K}(\mathcal{A},\mathbb{K})$ terminates as

$$\cdots \stackrel{d}{\to} \mathcal{A}\otimes R \stackrel{d^{N-1}}{\to} \mathcal{A}\otimes E \stackrel{d}{\to} \mathcal{A}\to 0.$$

Thus if $\mathcal{A}$ is a Koszul algebra, one has a free resolution of $\mathbb{K}$ which is then in fact a minimal projective resolution of the trivial left $\mathcal{A}$-module $\mathbb{K}$ given by

$$\mathcal{K}(\mathcal{A},\mathbb{K}) \stackrel{\varepsilon}{\to} \mathbb{K}\to 0 \tag{4.7}$$

which is referred to as the Koszul resolution of (the left $\mathcal{A}$-module) $\mathbb{K}$.

Notice that if $\mathcal{A}$ is a regular algebra of global dimension 2 (resp. 3) then (2.1) (resp. (3.4)) are the Koszul resolutions of $\mathbb{K}$ (with a slight abuse of language) so that $\mathcal{A}$ is then a Koszul algebra as announced in Proposition 1. One has the following result [**25**].

PROPOSITION 6. *Let $\mathcal{A}$ be a Koszul N-homogeneous algebra. One has*

$$P_{\mathcal{A}}(t)Q_{\mathcal{A}}(t)=1$$

where the series $Q_{\mathcal{A}}(t)$ is defined by

$$Q_{\mathcal{A}}(t) = \sum_{n\in\mathbb{N}} (\dim(\mathcal{A}^!_{Nn})t^{Nn} - \dim(\mathcal{A}^!_{Nn+1})t^{Nn+1})$$

and where $P_{\mathcal{A}}(t) = \sum_n \dim(\mathcal{A}_n)t^n$ is the Poincaré series of $\mathcal{A}$.

In fact the Koszul N-complex splits into sub-N-complexes for the total degree

$$K(\mathcal{A}) = \oplus K^{(n)}(\mathcal{A})$$

which induces a splitting of the Koszul complex into finite-dimensional subcomplexes

$$\mathcal{K}(\mathcal{A}, \mathbb{K}) = \oplus \mathcal{K}^{(n)}(\mathcal{A}, \mathbb{K})$$

and the proposition follows from the Euler-Poincaré formula applied to each component.

Notice that in the quadratic case, one has $Q_{\mathcal{A}}(t) = P_{\mathcal{A}^!}(-t)$.

If $\mathcal{A}$ is a Koszul N-homogeneous algebra, one has clearly

$$\mathcal{A}^!_{Nn} \simeq \mathrm{Ext}^{2n}_{\mathcal{A}}(\mathbb{K},\mathbb{K}), \mathcal{A}^!_{Nn+1} \simeq \mathrm{Ext}^{2n+1}_{\mathcal{A}}(\mathbb{K},\mathbb{K}) \tag{4.8}$$

and therefore by setting

$$Y_{\mathcal{A}}(t) = \sum_{n\in\mathbb{N}} (\dim(\mathrm{Ext}^{2n}_{\mathcal{A}}(\mathbb{K},\mathbb{K}))t^{Nn} - \dim(\mathrm{Ext}^{2n+1}_{\mathcal{A}}(\mathbb{K},\mathbb{K}))t^{Nn+1}) \tag{4.9}$$

one has $P_{\mathcal{A}}(t)Y_{\mathcal{A}}(t) = 1$. In [**31**] it is shown that, conversely, if $\mathcal{A}$ is an N-homogeneous algebra such that one has

$$P_{\mathcal{A}}(t)Y_{\mathcal{A}}(t) = 1 \tag{4.10}$$

then $\mathcal{A}$ is Koszul. This gives an interesting numerical criterion for Koszulity which has to be compared with the fact that there are N-homogeneous algebras $\mathcal{A}$ satisfying $P_{\mathcal{A}}(t)Q_{\mathcal{A}}(t) = 1$ which are not Koszul (of course then (4.8) does not hold).

In (4.4) the factors $\mathcal{A}$ are considered as left $\mathcal{A}$-modules. By considering $\mathcal{A}$ as a right $\mathcal{A}$-module and by exchanging the factors, one obtains an N-complex $\tilde{K}(\mathcal{A})$ of right $\mathcal{A}$-modules.

$$\cdots \stackrel{\tilde{d}}{\rightarrow} \mathcal{A}^{!*}_{n+1} \otimes \mathcal{A} \stackrel{\tilde{d}}{\rightarrow} \mathcal{A}^{!*}_{n} \otimes \mathcal{A} \stackrel{\tilde{d}}{\rightarrow} \cdots \stackrel{\tilde{d}}{\rightarrow} \mathcal{A} \rightarrow 0 \tag{4.11}$$

where $\tilde{d} : \mathcal{A}^{!*}_{n+1} \otimes \mathcal{A} \to \mathcal{A}^{!*}_{n} \otimes \mathcal{A}$ is induced by the mapping $(e_1 \otimes \cdots \otimes e_{n+1}) \otimes a \mapsto (e_1 \otimes \cdots \otimes e_n) \otimes e_{n+1}a$ of $E^{\otimes^{n+1}} \otimes \mathcal{A}$ into $E^{\otimes^n} \otimes \mathcal{A}$. The fact that $\tilde{d}^N = 0$ follows from $\mathcal{A}^{!*}_N \subset E^{\otimes^{n-N}} \otimes R$ for $n \geq N$. Let us consider the sequences (L, R)

$$\cdots \stackrel{d_L,d_R}{\rightarrow} \mathcal{A} \otimes \mathcal{A}^{!*}_{n+1} \otimes \mathcal{A} \stackrel{d_L,d_R}{\rightarrow} \mathcal{A} \otimes \mathcal{A}^{!*}_{n} \otimes \mathcal{A} \stackrel{d_L,d_R}{\rightarrow} \cdots \stackrel{d_L,d_R}{\rightarrow} \mathcal{A} \otimes \mathcal{A} \rightarrow 0 \tag{4.12}$$

where $d_L = d \otimes I$ and $d_R = I \otimes \tilde{d}$, I being the identity mapping of $\mathcal{A}$ onto itself. One has $d_L^N = d_R^N = 0$ and d_L and d_R are homomorphisms of $(\mathcal{A},\mathcal{A})$-bimodules, i.e. of left $\mathcal{A} \otimes \mathcal{A}^{opp}$-modules. The two N-differentials d_L and d_R commute, so one has

$$(d_L - d_R)\left(\sum_{p=0}^{N-1} d_L^p d_R^{N-p-1}\right) = \left(\sum_{p=0}^{N-1} d_L^p d_R^{N-p-1}\right)(d_L - d_R) = d_L^N - d_R^N = 0.$$

It follows that one defines a complex of free $\mathcal{A}\otimes\mathcal{A}^{opp}$-modules $\mathcal{K}(\mathcal{A},\mathcal{A})$ by setting

$$\left\{\begin{array}{l}\mathcal{K}_{2m}(\mathcal{A},\mathcal{A})=\mathcal{A}\otimes\mathcal{A}^{!*}_{Nm}\otimes\mathcal{A}\\ \\ \mathcal{K}_{2m+1}(\mathcal{A},\mathcal{A})=\mathcal{A}\otimes\mathcal{A}^{!*}_{Nm+1}\otimes\mathcal{A}\end{array}\right.\tag{4.13}$$

with differential δ' defined by

$$\left\{\begin{array}{l}\delta'=d_L-d_R:\mathcal{K}_{2m+1}(\mathcal{A},\mathcal{A})\to\mathcal{K}_{2m}(\mathcal{A},\mathcal{A})\\ \\ \delta'=\sum^{N-1}_{p=0}d^p_Ld^{N-p-1}_R:\mathcal{K}_{2(m+1)}(\mathcal{A},\mathcal{A})\to\mathcal{K}_{2m+1}(\mathcal{A},\mathcal{A})\end{array}\right.\tag{4.14}$$

which will be referred to as *the bimodule Koszul complex of* $\mathcal{A}$.

It turns out that $\mathcal{K}(\mathcal{A},\mathcal{A})$ is acyclic in positive degrees if and only if $\mathcal{K}(\mathcal{A},\mathbb{K})$ is acyclic in positive degrees, that is, if and only if $\mathcal{A}$ is a Koszul algebra. On the other hand one has the obvious exact sequence of bimodules

$$\mathcal{A}\otimes E\otimes\mathcal{A}\stackrel{\delta'}{\to}\mathcal{A}\otimes\mathcal{A}\stackrel{m}{\to}\mathcal{A}\to 0$$

where m denotes the product of $\mathcal{A}$. This means that $H_0(\mathcal{K}(\mathcal{A},\mathcal{A}))=\mathcal{A}$ and therefore whenever $\mathcal{A}$ is Koszul one has a free resolution

$$\mathcal{K}(\mathcal{A},\mathcal{A})\stackrel{m}{\to}\mathcal{A}\to 0$$

of the left $\mathcal{A}\otimes\mathcal{A}^{opp}$-module $\mathcal{A}$ which is a minimal projective resolution of $\mathcal{A}$ and will be referred to as the Koszul resolution of $\mathcal{A}$.

4.4. Small complex and Poincaré duality for Koszul algebras. Let $\mathcal{A}$ be an N-homogeneous Koszul algebra and let $\mathcal{M}$ be a $(\mathcal{A},\mathcal{A})$-bimodule considered as a right $\mathcal{A}\otimes\mathcal{A}^{opp}$-module. Then by interpreting the Hochschild homology $H(\mathcal{A},\mathcal{M})$ of $\mathcal{A}$ with values in $\mathcal{M}$ as $\mathrm{Tor}^{\mathcal{A}\otimes\mathcal{A}^{opp}}(\mathcal{M},\mathcal{A})$ [**12**], one sees that the homology of the complex $\mathcal{M}\otimes_{\mathcal{A}\otimes\mathcal{A}^{opp}}\mathcal{K}(\mathcal{A},\mathcal{A})$ is the $\mathcal{M}$-valued Hochschild homology of $\mathcal{A}$. We shall refer to this latter complex as the *small Hochschild complex* of the Koszul algebra $\mathcal{A}$ with coefficients in $\mathcal{M}$ and denote it by $\mathcal{S}(\mathcal{A},\mathcal{M})$. It reads

$$\cdots\stackrel{\delta}{\to}\mathcal{M}\otimes\mathcal{A}^{!*}_{N(m+1)}\stackrel{\delta}{\to}\mathcal{M}\otimes\mathcal{A}^{!*}_{Nm+1}\stackrel{\delta}{\to}\mathcal{M}\otimes\mathcal{A}^{!*}_{Nm}\stackrel{\delta}{\to}\cdots\tag{4.15}$$

where δ is obtained from δ' by applying the factors d_L to the right of $\mathcal{M}$ and the factors d_R to the left of $\mathcal{M}$.

By construction the lengths of the complexes $\mathcal{K}(\mathcal{A},\mathbb{K})$ and $\mathcal{K}(\mathcal{A},\mathcal{A})$ coincide. Assume that $\mathcal{A}$ is a Koszul algebra; then this implies that the projective dimension of the trivial $\mathcal{A}$-module $\mathbb{K}$ coincides with the Hochschild dimension of $\mathcal{A}$, which is a particular case of the general result of [**4**].

The Koszul complex $\mathcal{K}(\mathcal{A},\mathbb{K})$ is a chain complex since its differential is of degree -1; the same is true for $\mathcal{K}(\mathcal{A},\mathcal{A})$. By applying the functor $\mathrm{Hom}_{\mathcal{A}}(\bullet,\mathcal{A})$ to the chain complex of free left $\mathcal{A}$-modules $\mathcal{K}(\mathcal{A},\mathbb{K})$ one obtains the cochain complex $\mathcal{L}(\mathcal{A},\mathbb{K})$ of free right $\mathcal{A}$-modules

$$0\to\mathcal{L}^0(\mathcal{A},\mathbb{K})\to\cdots\to\mathcal{L}^n(\mathcal{A},\mathbb{K})\to\cdots$$

where $\mathcal{L}^n(\mathcal{A},\mathbb{K})=\mathrm{Hom}_{\mathcal{A}}(\mathcal{K}_n(\mathcal{A},\mathbb{K}),\mathcal{A})$. Assume that $\mathcal{A}$ is Koszul of global dimension D. Then $\mathcal{L}^n(\mathcal{A},\mathbb{K})=0$ for $n>D$ and $\mathcal{A}$ is Gorenstein if and only if $H^n(\mathcal{L}(\mathcal{A},\mathbb{K}))=0$ for $n<D$ and $H^D(\mathcal{L}(\mathcal{A},\mathbb{K}))=\mathbb{K}$. When $\mathcal{A}$ is Koszul of global

dimension D and Gorenstein, this implies a precise form of the Poincaré duality between the Hochschild homology and the Hochschild cohomology of $\mathcal{A}$, [**7**], [**41**], [**42**]. In the case of a regular algebra $\mathcal{A} = \mathcal{A}(w, N)$ of global dimension 3, it reads for an $\mathcal{A}$-bimodule $\mathcal{M}$

$$H_n(\mathcal{A}, \mathcal{M}) = H^{3-n}(\mathcal{A}, \mathcal{M}) \tag{4.16}$$

for $0 \leq n \leq 3$ when $Q_w = \mathbb{1}$, (when $Q_w \neq \mathbb{1}$ it induces an automorphism σ_w of $\mathcal{A}$ and one has to twist by σ_w the left multiplication of $\mathcal{M}$ by $\mathcal{A}$ on the right-hand side of (4.16)).

The complex $\mathcal{L}(\mathcal{A}, \mathbb{K})$ is also a contraction of a natural N-complex $L(\mathcal{A})$. This N-complex $L(\mathcal{A})$ is the cochain N-complex of free right $\mathcal{A}$-modules obtained by applying the functor $\mathrm{Hom}_{\mathcal{A}}(\bullet, \mathcal{A})$ to the Koszul N-complex $K(\mathcal{A})$ (which is a chain N-complex of free left $\mathcal{A}$-modules). The right $\mathcal{A}$-module $L^n(\mathcal{A})$ identifies canonically with $\mathcal{A}^!_n \otimes \mathcal{A}$ while the N-differential of $L(\mathcal{A})$ is then the left multiplication by $x^*_\lambda \otimes x^\lambda$ in $\mathcal{A}^! \otimes \mathcal{A}$, where (x^*_λ) is the dual basis of (x^λ) $(E = \oplus_\lambda \mathbb{K}x^\lambda)$. One has $\mathcal{L}(\mathcal{A}, \mathbb{K}) = C_{1,0}(L(\mathcal{A}))$, i.e. $\mathcal{L}^0(\mathcal{A}, \mathbb{K}) = \mathcal{A} = L^0(\mathcal{A}), \mathcal{L}^1(\mathcal{A}, \mathbb{K}) = L^N(\mathcal{A})$, etc.

4.5. Examples of Koszul algebras. All regular algebras of global dimensions $D = 2$ and $D = 3$ are Koszul so in particular the examples of regular algebras of Sections 2 and 3 are examples of Koszul algebras. We shall describe regular Koszul algebras of higher global dimension D in Section 5. Let us give here some examples of Koszul algebras which are not generically regular.

(a) *Koszul duals of quadratic algebras*. It is well known and not hard to show that if $\mathcal{A}$ is a quadratic algebra, then its Koszul dual $\mathcal{A}^!$ is Koszul if and only if $\mathcal{A}$ is Koszul. Even if $\mathcal{A}$ is regular, $\mathcal{A}^!$ is generically not regular.

For instance the exterior algebra $\bigwedge \mathbb{K}^g$ is the Koszul dual of the algebra of polynomial functions on $\mathbb{K}^g$ which is regular and Koszul of global dimension g; however $\bigwedge \mathbb{K}^g$ is not of finite global dimension.

It is worth noticing here that if $\mathcal{A}$ is an N-homogeneous algebra with $N > 2$, then the Koszulity of $\mathcal{A}$ does not imply the Koszulity of its Koszul dual $\mathcal{A}^!$ (this is due to the jumps in degrees in the Koszul resolution). For instance the Koszul dual $\mathcal{A}^!$ of the Yang-Mills algebra $\mathcal{A}$ (§3.3, example (e)) is such that $P_{\mathcal{A}^!}(t)Q_{\mathcal{A}^!}(t) \neq 1$ (by direct computation) so it is not Koszul in view of Proposition 6.

(b) *Degenerate bilinear form* [**5**]. In the following b is a bilinear form on $\mathbb{K}^g$ with $g \geq 2$, $B = (B_{\mu\nu})$ is the matrix of components $B_{\mu\nu} = b(e_\mu, e_\nu)$ of b in the canonical basis of $\mathbb{K}^g$ and $\mathcal{A} = \mathcal{A}(b, 2)$ is the quadratic algebra generated by g elements x^λ with the relation

$$B_{\mu\nu} x^\mu x^\nu = 0$$

i.e. we generalize the notation of Section 2 to cases where b can be degenerate. In [**5**] one finds the following results (Propositions 5.4 and 5.5 in [**5**]) which contains Theorem 2.

PROPOSITION 7. *Assume that $b \neq 0$, then $\mathcal{A} = \mathcal{A}(b, 2)$ has the following properties :*

1) $\mathcal{A}$ is Koszul,
2) $\mathcal{A}$ has global dimension $D = 2$ except in the case where b is symmetric of rank 1 in which case $D = \infty$,
3) $\mathcal{A}$ is Gorenstein if and only if b is nondegenerate.

Thus for b degenerate one has a lot of examples of Koszul algebras which are not regular. In [**5**] there is a similar statement for N-homogeneous algebras with one relation ($r = 1$) which although slightly more involved permits the construction of examples (see, e.g. Example (c) in the next section §5.3).

(c) *The self-duality algebra* [**17**]. In the case $g = 4$ and $g^{\lambda\mu} = \delta^{\lambda\mu}$, the Yang-Mills algebra (Example (e) in § 3.3) admits the 2 nontrivial quotients $\mathcal{A}^{(+)}$ and $\mathcal{A}^{(-)}$, where $\mathcal{A}^{(\varepsilon)}$ ($\varepsilon = \pm$) is the quadratic algebra generated by the 4 elements ∇_λ ($\lambda \in \{1, 2, 3, 4\}$) with relations

$$[\nabla_4, \nabla_k] = \varepsilon[\nabla_\ell, \nabla_m] \tag{4.17}$$

for any cyclic permutation (k, ℓ, m) of (1,2,3). Let us fix $\varepsilon = +$ and call $\mathcal{A}^{(+)}$ the self-duality algebra (the study of $\mathcal{A}^{(-)}$ is similar). In [**17**] it was shown that this algebra is Koszul of global dimension $D = 2$ and that the Koszul resolution reads

$$0 \to (\mathcal{A}^{(+)})^3 \to (\mathcal{A}^{(+)})^4 \to \mathcal{A}^{(+)} \xrightarrow{\varepsilon} \mathbb{K} \to 0 \tag{4.18}$$

from which it follows that

$$P_{\mathcal{A}^{(+)}}(t) = \frac{1}{(1-t)(1-3t)} \tag{4.19}$$

so $\mathcal{A}^{(+)}$ has exponential growth and is not Gorenstein.

It follows from the definition that $\mathcal{A}^{(+)}$ is the universal enveloping algebra of a Lie algebra which is the semi-direct product of the free Lie algebra $L(\nabla_1, \nabla_2, \nabla_3)$ by the derivation δ given by

$$\delta(\nabla_k) = [\nabla_\ell, \nabla_m] \tag{4.20}$$

for any cyclic permutation (k, ℓ, m) of (1,2,3). Formula (4.19) as well as all the above properties of $\mathcal{A}^{(+)}$ also follow directly from this structure.

(d) *The super self-duality algebra* [**19**]. In a similar way as in the last example, for $g = 4$ and $g^{\lambda\mu} = \delta^{\lambda\mu}$, the super Yang-Mills algebra (Example (f) in § 3.3) admits the 2 nontrivial quotients $\tilde{\mathcal{A}}^{(+)}$ and $\tilde{\mathcal{A}}^{(-)}$, where $\tilde{\mathcal{A}}^{(\varepsilon)}$ ($\varepsilon = \pm$) is the quadratic algebra generated by the 4 elements S_λ ($\lambda \in \{1, 2, 3, 4\}$) with relations

$$i[S_4, S_k]_+ = \varepsilon[S_\ell, S_m] \tag{4.21}$$

for any cyclic permutation (k, ℓ, m) of (1,2,3). Let us fix $\varepsilon = +$ and call $\tilde{\mathcal{A}}^{(+)}$ the super self-duality algebra. This algebra is again a Koszul algebra of global dimension 2 which is not Gorenstein and has Poincaré series given by

$$P_{\tilde{\mathcal{A}}^{(+)}}(t) = \frac{1}{(1-t)(1-3t)} \tag{4.22}$$

and so has also exponential growth. This algebra has direct relations with the 4-dimensional Sklyanin algebra (see in [**19**]).

5. Arbitrary global dimension D

In the previous sections, we have seen that the regular algebras of global dimensions $D = 2$ and $D = 3$ are N-homogeneous (with $N = 2$ for $D = 2$) and Koszul. This very desirable property permits one to write explicit canonical resolutions. On the other hand one can formulate for the moment this Koszul property only for N-homogeneous algebras. This is why in this section we shall restrict attention to Koszul homogeneous algebras and our aim is then to formulate the generalization of Theorem 3 for arbitrary global dimension D. Notice however that for global dimensions $D \geq 4$, regularity does not imply N-homogeneity. It is worth mentioning here that for $D = 4$ the AS-regular algebras, i.e. the regular algebras with polynomial growth, have been recently classified [**32**].

We shall need a class of N-homogeneous algebras associated with preregular multilinear forms that we now describe.

5.1. Homogeneous algebras associated to multilinear forms. In this subsection m and N are integers with $m \geq N \geq 2$ and w is a preregular m-linear form on $\mathbb{K}^g$ ($g \geq 2$) with components $W_{\lambda_1 \dots \lambda_m} = w(e_{\lambda_1}, \dots, e_{\lambda_m})$ in the canonical basis (e_λ) of $\mathbb{K}^g$. Let $\mathcal{A} = \mathcal{A}(w, N)$ be the N-homogeneous algebra generated by the elements x^λ ($\lambda \in \{1, \dots, g\}$) with relation

$$W_{\lambda_1 \dots \lambda_{m-N} \mu_1 \dots \mu_N} x^{\mu_1} \cdots x^{\mu_N} = 0 \tag{5.1}$$

for $\lambda_k \in \{1, \dots, g\}$. Thus one has $\mathcal{A} = A(E, R)$ with $E = \oplus_\lambda \mathbb{K} x^\lambda$ and

$$R = \sum_{\lambda_k} \mathbb{K} W_{\lambda_1 \dots \lambda_{m-N} \mu_1 \dots \mu_N} x^{\mu_1} \otimes \cdots \otimes x^{\mu_N} \subset E^{\otimes^N}.$$

Notice that this generalizes the definitions of Section 2 (which is the case $m = N = 2$) and Section 3 (which is the case $m = N + 1$).

Let us define the subspaces $\mathcal{W}_n \subset E^{\otimes^n}$ for $m \geq n \geq 0$ by

$$\begin{cases} \mathcal{W}_n = E^{\otimes^n} \quad \text{for} \quad N - 1 \geq n \geq 0 \\ \mathcal{W}_n = \sum_{\lambda_k} \mathbb{K} W_{\lambda_1 \dots \lambda_{m-n} \mu_1 \dots \mu_n} x^{\mu_1} \otimes \cdots \otimes x^{\mu_n} \quad \text{for} \quad m \geq n \geq N \end{cases} \tag{5.2}$$

so in particular $\mathcal{W}_1 = E$ and $\mathcal{W}_N = R$. The twisted cyclicity of w (property (ii) of §3.1) and (4.3) imply the following proposition.

PROPOSITION 8. *The sequence*

$$0 \to \mathcal{A} \otimes \mathcal{W}_m \xrightarrow{d} \mathcal{A} \otimes \mathcal{W}_{m-1} \xrightarrow{d} \cdots \xrightarrow{d} \mathcal{A} \to 0 \tag{5.3}$$

is a sub-N-complex of the Koszul N-complex $K(\mathcal{A})$ of $\mathcal{A}$.

In fact one has $\mathcal{W}_n \subset \mathcal{A}^{!*}_n$ and $d(\mathcal{A} \otimes \mathcal{W}_{n+1}) \subset \mathcal{A} \otimes \mathcal{W}_n$. In particular one has $\mathcal{W}_m = \mathbb{K} w \subset \mathcal{A}^{!*}_m$ so w is a linear form on $\mathcal{A}^!_m$. We then define the linear form ω_w on the algebra $\mathcal{A}^!$ by setting

$$\omega_w = w \circ p_m \tag{5.4}$$

where $p_m : \mathcal{A}^! \to \mathcal{A}^!_m$ is the canonical projection onto the degree m component. With $E = \oplus_\lambda \mathbb{K} x^\lambda$, w is canonically an m-linear form on E^* and Q_w an element of $GL(E^*)$. With these identifications one has the following theorem [**23**].

THEOREM 9. *The element Q_w of $GL(E^*)$ induces an automorphism σ_w of the N-homogeneous algebra $\mathcal{A}^! = A(E^*, R^\perp)$ and one has*

$$\omega_w(xy) = \omega_w(\sigma_w(y)x) \tag{5.5}$$

for any $x, y \in \mathcal{A}^!$. The subset of $\mathcal{A}^!$

$$\mathcal{I} = \{y \in \mathcal{A}^! | \omega_w(xy) = 0, \quad \forall x \in \mathcal{A}^!\}$$

is a two-sided ideal of $\mathcal{A}^!$ and the quotient algebra $\mathcal{F}(w, N) = \mathcal{A}^!/\mathcal{I}$ equipped with the linear form induced by ω_w is a graded Frobenius algebra.

To prove this theorem, one first verifies by using the Q_w-invariance of w that one has $Q_w^{\otimes^N} R^\perp \subset R^\perp$, which implies the existence of σ_w. Then (5.5) is just a translation of the Q_w-cyclicity of w. By definition $\mathcal{I}$ is a left ideal and (5.5) implies that it is also a right ideal. The quotient $\mathcal{F} = \mathcal{A}^!/\mathcal{I}$ is a finite-dimensional graded algebra and the pairing induced by $(x, y) \mapsto \omega_w(xy)$ is nondegenerate and is a Frobenius pairing on $\mathcal{F}$.

COROLLARY 10. *Considered as an element of $GL(E)$, the transpose $Q_w^t = Q^w$ of Q_w induces an automorphism σ^w of the N-homogeneous algebra $\mathcal{A} = A(E, R)$.*

Let us end this subsection by noting that, at this level of generality and for $N = 2$ (i.e. in the quadratic case), the multilinear form w induces a (twisted) noncommutative m-form for $\mathcal{A}$. For this let ${}^w\mathcal{A}$ be the $(\mathcal{A}, \mathcal{A})$-bimodule which coincides with $\mathcal{A}$ as right $\mathcal{A}$-module and is such that the structure of left $\mathcal{A}$-module is given by the left multiplication by $(-1)^{(m-1)n}(\sigma^w)^{-1}(a)$ for $a \in \mathcal{A}_n$. One has the following result [**23**].

PROPOSITION 11. *In the case $N = 2$, that is, for $\mathcal{A} = \mathcal{A}(w, 2)$, $\mathbb{1} \otimes w$ is canonically a nontrivial ${}^w\mathcal{A}$-valued Hochschild m-cycle on $\mathcal{A}$.*

In this statement, $\mathbb{1}$ is interpreted as an element of ${}^w\mathcal{A}$ while $w \in E^{\otimes^m}$ is interpreted as an element of $\mathcal{A}^{\otimes^m}$ ($E = \mathcal{A}_1 \subset \mathcal{A}$) so that $\mathbb{1} \otimes w$ is a ${}^w\mathcal{A}$-valued Hochschild m-chain.

5.2. General results for Koszul-Gorenstein algebras. For the N-homogeneous algebras which are Koszul of finite global dimension D and which are Gorenstein (a particular class of regular algebras if $D \geq 4$), one has the following theorem [**23**].

THEOREM 12. *Let $\mathcal{A}$ be an N-homogeneous algebra which is Koszul of finite global dimension D and Gorenstein. Then $\mathcal{A} = \mathcal{A}(w, N)$ for some preregular m-linear form on $\mathbb{K}^g$ for some g. If $N \geq 3$ then $m = Np+1$ and $D = 2p+1$ for some $p \geq 1$ while for $N = 2$ one has $m = D$.*

For the proof we refer to [**23**].

Under the assumptions of Theorem 12 the Koszul resolution of the trivial left $\mathcal{A}$-module $\mathbb{K}$ reads

$$0 \to \mathcal{A} \otimes \mathcal{W}_m \stackrel{d}{\to} \mathcal{A} \otimes \mathcal{W}_{m-1} \stackrel{d^{N-1}}{\to} \cdots \stackrel{d}{\to} \mathcal{A} \otimes \mathcal{W}_N \stackrel{d^{N-1}}{\to} \mathcal{A} \otimes E \stackrel{d}{\to} \mathcal{A} \stackrel{\varepsilon}{\to} \mathbb{K} \to 0$$

or, by setting

$$\begin{cases} \nu_N(2k) = Nk \\ \nu_N(2k+1) = Nk+1 \end{cases} \tag{5.6}$$

for $k \in \mathbb{N}$,

$$0 \to \mathcal{A}\otimes\mathcal{W}_{\nu_N(D)} \xrightarrow{d'} \cdots \xrightarrow{d'} \mathcal{A}\otimes\mathcal{W}_{\nu_N(k)} \xrightarrow{d'} \mathcal{A}\otimes\mathcal{W}_{\nu_N(k-1)} \to \cdots \xrightarrow{d'} \mathcal{A} \xrightarrow{\varepsilon} \mathbb{K} \to 0 \tag{5.7}$$

where d' is defined by

$$\begin{cases} d' = d^{N-1} : \mathcal{A}\otimes\mathcal{W}_{\nu_N(2k)} \to \mathcal{A}\otimes\mathcal{W}_{\nu_N(2k-1)} \\ d' = d : \mathcal{A}\otimes\mathcal{W}_{\nu_N(2k+1)} \to \mathcal{A}\otimes\mathcal{W}_{\nu_N(2k)} \end{cases} \tag{5.8}$$

for $k \in \mathbb{N}$.

Notice that one has

$$\dim(\mathcal{W}_{\nu_N(k)}) = \dim(\mathcal{W}_{\nu_N(D-k)}) \tag{5.9}$$

for $0 \leq k \leq D$. In particular $\mathcal{A}\otimes\mathcal{W}_{\nu_N(D)} = \mathcal{A}\otimes w$ so one sees that $\mathbb{1}\otimes w$ is the generator of the top module of the Koszul resolution, which again corresponds to the interpretation of $\mathbb{1}\otimes w$ as a volume form.

It is worth noticing here that it has been already shown in [**10**] that the quadratic algebras which are Koszul and regular are determined by multilinear form (D-linear for global dimension D) which correspond to volume forms in this noncommutative setting.

Let us come back to a more general situation. Assume that D and N are given integers with $D \geq 2$ and $N \geq 2$ and that $N = 2$ whenever D is an even integer. Then let w be a preregular m-linear form on $\mathbb{K}^g$ with $m = D$ for $N = 2$ and $m = Np+1$ for $D = 2p+1$ and consider the N-homogeneous algebra $\mathcal{A} = \mathcal{A}(w, N)$. The complex

$$0 \to \mathcal{A}\otimes\mathcal{W}_{\nu_N(D)} \xrightarrow{d'} \cdots \xrightarrow{d'} \mathcal{A}\otimes\mathcal{W}_{\nu_N(k)} \xrightarrow{d'} \cdots \xrightarrow{d'} \mathcal{A} \to 0 \tag{5.10}$$

is still well defined, with ν_N as in (5.6) and d' as in (5.8), and is a subcomplex of the Koszul complex $\mathcal{K}(\mathcal{A}, \mathbb{K})$ of $\mathcal{A}$ in view of Proposition 8. It is clear that if this complex is acyclic in positive degree, it coincides with the Koszul complex of $\mathcal{A}$ and that $\mathcal{A}$ is then Koszul of global dimension D and Gorenstein. Thus, as remarked in [**9**] one has the following result, which gives a sort of converse of Theorem 12.

PROPOSITION 13. *Let $\mathcal{A} = \mathcal{A}(w, N)$ be as above; then $\mathcal{A}$ is Koszul of global dimension D and Gorenstein if and only if the complex* (5.10) *is acyclic in positive degrees.*

A weaker assumption on the complex (5.10) is to assume that it coincides with the Koszul complex. In the case where $D = 3$, one has the following proposition [**23**].

PROPOSITION 14. *Let w be a preregular $(N+1)$-linear form on $\mathbb{K}^g$ and let $\mathcal{A} = \mathcal{A}(w, N)$; then the following conditions are equivalent:*

(a) $\mathcal{A}^{!*}_{N+1} = \mathbb{K}w$.

(b) *The complex* (5.10) *coincides with the Koszul complex* $\mathcal{K}(\mathcal{A}, \mathbb{K})$ *of* $\mathcal{A}$.

(c) *w is 3-regular.*

Let us consider $\mathcal{A} = \mathcal{A}(w, N)$ with Koszul dual $\mathcal{A}^! = \oplus_n \mathcal{A}^!_n$ and let us define the graded algebra

$$\mathcal{A}' = \mathcal{A}'(w, N) = \oplus_n \mathcal{A}'_n \tag{5.11}$$

to be $\mathcal{A}^!$ for $N = 2$ and to be defined for $N > 2$ by

$$\mathcal{A}'_n = \mathcal{A}^!_{\nu_N(n)} \tag{5.12}$$

for $n \in \mathbb{N}$ with product $(x, y) \mapsto x \bullet y$ defined by

$$x \bullet y = \pi(xy) \tag{5.13}$$

where $\pi : \mathcal{A}^! \to \mathcal{A}'$ is the canonical projection of $\mathcal{A}^!$ onto $\mathcal{A}' = \oplus_n \mathcal{A}^!_{\nu_N(n)} \subset \mathcal{A}^!$ defined by setting $\pi(\mathcal{A}^!_k) = 0$ whenever k is not in $\nu_N(\mathbb{N})$. Thus this product is defined for two homogeneous elements x and y by

$$xy = 0$$

whenever x and y are both of odd degree and

$$xy = \text{product in } \mathcal{A}^!$$

otherwise. It is clear that this product is associative. One has the following result.

THEOREM 15. *Assume that D, N and w are as above, that is, $N = 2$ for D even and w is a preregular m-linear form on $\mathbb{K}^g$ with $m = D$ for $N = 2$ and $m = Np+1$ for $D = 2p + 1$. Then the following conditions are equivalent.*

(a) *$\mathcal{A}'(w, N)$ equipped with the linear form induced by ω_w is a Frobenius algebra.*

(b) *The complex* (5.10) *coincides with the Koszul complex $\mathcal{K}(\mathcal{A}, \mathbb{K})$ of $\mathcal{A}(w, N)$.*

Proof. The proof of this proposition is almost tautological since conditions (a) and (b) are both equivalent to $\mathcal{W}_{\nu_N(n)} = \mathcal{A}^{!*}_{\nu_N(n)} = \mathcal{A}'^{\,*}_n$ for $n \in \mathbb{N}$. $\square$

This is of course directly inspired by [**7**] and implies Theorem 1.2 of [**7**] since when $\mathcal{A}(w, N)$ is Koszul one has $\mathcal{A}^!_{\nu_N(n)} = \mathrm{Ext}^n_{\mathcal{A}}(\mathbb{K}, \mathbb{K})$ and the product of $\mathcal{A}'(w, N)$ is essentially the Yoneda product ([**7**], Proposition 3.1). Let us recall this Theorem 1.2 of [**7**], which is an important result.

THEOREM 16. *Let $\mathcal{A}$ be an N-homogeneous algebra which is Koszul of finite global dimension. Then $\mathcal{A}$ is Gorenstein if and only if the Yoneda algebra $E(\mathcal{A}) = Ext_{\mathcal{A}}(\mathbb{K}, \mathbb{K})$ is Frobenius.*

As pointed out before this follows from Theorem 12 and Theorem 15 by using the fact that one has $\mathcal{A}' = E(\mathcal{A})$ whenever $\mathcal{A}$ is Koszul.

Remarks.

1) One sees that, with D, N and w as in Theorem 15, one has two natural Frobenius algebras associated with $\mathcal{A}(w, N)$. The first one is the algebra $\mathcal{F}(w, N) = \mathcal{A}^!/\mathcal{I}$ of Theorem 9, the other one is the algebra $\mathcal{F}'(w, N) = \mathcal{A}'/\mathcal{I}'$, where

$$\mathcal{I}' = \{y \in \mathcal{A}' | \omega_w(x \bullet y) = 0, \quad \forall x \in \mathcal{A}'\}$$

is a two-sided ideal since σ_w induces an automorphism of $\mathcal{A}'$ satisfying $\omega_w(x \bullet y) = \omega_w(\sigma_w(y) \bullet x)$. These two Frobenius algebras coincide for $N = 2$ but are different

for $N > 2$.
2) D, N and w being as in Theorem 15, it is tempting in view of Proposition 14 to say that w is *D-regular* whenever the equivalent conditions (a) and (b) are satisfied. In fact Condition (a) contains several nondegeneracy conditions. This notion involves both D and N as above.

5.3. Examples. Of course one has already all the examples of Section 3. Let us give two quadratic examples and a class of N-homogeneous examples.

(a) *The extended 4-dimensional Sklyanin algebra* [**16**], [**18**], [**20**]. In connection with a problem of K-homology, the following quadratic algebra $\mathcal{A}_{\mathbf{u}}$ has been introduced in [**16**] and analyzed in detail in [**18**], [**20**]. The algebra $\mathcal{A}_{\mathbf{u}}$ is the quadratic algebra generated by 4 elements x^λ ($\lambda \in \{0,1,2,3\}$) with relations

$$\cos(\varphi_0 - \varphi_k)[x^0, x^k] = i\sin(\varphi_\ell - \varphi_m)[x^\ell, x^m]_+ \tag{5.14}$$

$$\cos(\varphi_\ell - \varphi_m)[x^\ell, x^m] = i\sin(\varphi_0 - \varphi_k)[x^0, x^k]_+ \tag{5.15}$$

for any cyclic permutation (k, ℓ, m) of (1,2,3). The parameter $\mathbf{u}$ is the element $\mathbf{u} = \left(e^{i(\varphi_1-\varphi_0)}, e^{i(\varphi_2-\varphi_0)}, e^{i(\varphi_3-\varphi_0)}\right)$ of T^3. Thus there are a priori 3 scalar parameters $\varphi_1 - \varphi_0, \varphi_2 - \varphi_0$ and $\varphi_3 - \varphi_0$. However, for generic values of these parameters one can show that $\mathcal{A}_{\mathbf{u}}$ only depends on two scalar parameters and that then by an appropriate linear change of generators it reduces to the 4-dimensional Sklyanin algebra introduced in [**39**] and studied in [**40**] from the point of view of general regularity.

The algebra $\mathcal{A}_{\mathbf{u}}$ is Koszul of global dimension $D = 4$ and is Gorenstein whenever none of the 6 relations (5.14), (5.15) becomes trivial and one then has the nontrivial Hochschild cycle (in $Z(\mathcal{A}, \mathcal{A})$)

$$\begin{aligned} w = \mathrm{ch}_{\frac{3}{2}}(U_{\mathbf{u}}) \quad &= \quad -\textstyle\sum_{\alpha,\beta,\gamma,\delta} \varepsilon_{\alpha\beta\gamma\delta} \cos(\varphi_\alpha - \varphi_\beta + \varphi_\gamma - \varphi_\delta) x^\alpha \otimes x^\beta \otimes x^\gamma \otimes x^\delta \\ &+ \quad i \textstyle\sum_{\mu,\nu} \sin(2(\varphi_\mu - \varphi_\nu)) x^\mu \otimes x^\nu \otimes x^\mu \otimes x^\nu \end{aligned}$$

which defines a 4-linear form on $\mathbb{K}^4$ which is preregular with

$$Q_w = -\mathbb{1}$$

i.e. w is graded-cyclic. One verifies that one then has $\mathcal{A}_{\mathbf{u}} = \mathcal{A}(w, 2)$ and that $\mathbb{1} \otimes w$ is a Hochschild 4-cycle, i.e. $\mathbb{1} \otimes w \in Z_4(\mathcal{A}, \mathcal{A})$.

(b) *The q-deformed D-dimensional polynomial algebra*. This is the algebra $\mathcal{A}$ generated by D elements x^λ ($\lambda \in \{1, \dots, D\}$) with relations

$$x^\mu x^\nu = q^{\mu\nu} x^\nu x^\mu \tag{5.16}$$

for $\mu, \nu \in \{1, \dots, D\}$ where the $q^{\mu\nu} \in \mathbb{K}$ satisfy

$$q^{\mu\nu} q^{\nu\mu} = 1, \quad q^{\lambda\lambda} = 1 \tag{5.17}$$

for any $\lambda, \mu, \nu \in \{1, \dots, D\}$.

This algebra is Koszul of global dimension D and Gorenstein. One has $\mathcal{A} = \mathcal{A}(w, 2)$ with

$$w = \sum_{\pi \in \mathfrak{S}_D} \chi(\pi) x^{\pi(1)} \otimes \cdots \otimes x^{\pi(D)} \tag{5.18}$$

where $\mathfrak{S}_D$ is the group of permutations of $\{1, \dots, D\}$ and where $\chi : \mathfrak{S}_D \to \mathbb{K}$ is given by $\chi(\pi) = \prod_{(\mu\nu)}(-q^{\mu\nu})$ with $\Pi_{(\mu\nu)}$ corresponding to the standard embedding

$$\mathfrak{S}_D \hookrightarrow \{\prod_{(\mu\nu)} b^{\mu\nu}, \mu < \nu\} \subset \mathfrak{B}_D$$

of $\mathfrak{S}_D$ into the group of braids $\mathfrak{B}_D$.

One then has

$$(Q_w)^\mu_\nu = \left(\prod_{\lambda \neq \mu} (-q^{\lambda\mu})\right) \delta^\mu_\nu \tag{5.19}$$

for the matrix element of the corresponding $Q_w \in GL(D, \mathbb{K})$.

(c) *Precommutative examples* [**3**],[**7**]. Let the integers g and N be such that $g \geq N \geq 2$ and let ε be the completely antisymmetric g-linear form on $\mathbb{K}^g$ with $\varepsilon(e_1, \dots, e_g) = 1$. Consider the N-homogeneous algebra $\mathcal{A} = \mathcal{A}(\varepsilon, N)$, i.e. the algebra generated by g elements x^λ ($\lambda \in \{1, \dots, g\}$) with the relations

$$\varepsilon_{\lambda_1 \dots \lambda_{g-N}\ \mu_1 \dots \mu_N} x^{\mu_1} \cdots x^{\mu_N} = 0$$

where $\varepsilon_{\lambda_1 \dots \lambda_g} = \varepsilon(e_{\lambda_1}, \dots, e_{\lambda_g})$. It is clear that ε is preregular with

$$Q_\varepsilon = (-1)^{g-1} \mathbb{1}$$

as associated element of $GL(g, \mathbb{K})$.

It was shown in [**3**] where this algebra was introduced that $\mathcal{A}(\varepsilon, N)$ is a Koszul algebra of finite global dimension and it was shown in [**7**] that it is Gorenstein if and only if either $N = 2$ or $N > 2$ and $g = Np + 1$ for some integer $p \geq 1$. For $N = 2$ this reduces to the algebra polynomial functions on $\mathbb{K}^g$ while for $N > 2$ and $g = Np + 1$ this is a regular algebra of global dimension $D = 2p + 1$. In the latter case, the ideal $\mathcal{I}$ of Theorem 9 is generated by the quadratic elements $\alpha\beta + \beta\alpha$ of $\mathcal{A}(\varepsilon, N)^!$ so that the quotient Frobenius algebra $\mathcal{F}(\varepsilon, N) = \mathcal{A}^!/\mathcal{I}$ reduces to the exterior algebra $\bigwedge \mathbb{K}^g$, which is precisely the Koszul dual algebra of the quadratic algebra of polynomial functions on $\mathbb{K}^g$. Thus by this process one recovers the quadratic relations implying the original N-homogeneous ones.

Notice that for $N > 2$ the algebra $\mathcal{A}(\varepsilon, N)$ has exponential growth [**3**]. In [**27**] a twisted version of this example associated with a Hecke symmetry was introduced and analyzed with similar results. This paper [**27**] even contains a super version of these examples. See also [**26**] (and [**43**]) for the quadratic case associated with a Hecke symmetry.

Remark.
In contrast to the previous example for $N > 2$, in the cases of the Yang-Mills algebra and the super Yang-Mills algebra the ideal $\mathcal{I}$ of Theorem 9 vanishes, that is the

Koszul duals are then Frobenius. The reason is that in these cases the 3-regular multilinear forms (4-linear) w given respectively by (3.18) and (3.20) satisfy the stronger condition (iii') of §3.1.

5.4. Classical limit versus infinitesimal preregularity. We now consider perturbations of the algebra $\mathbb{K}[x^1,\dots,x^g]$ of polynomial functions on $\mathbb{K}^g$. More precisely, one has $\mathbb{K}[x^1,\dots,x^g]=\mathcal{A}(\varepsilon,2)$ where ε is the g-linear form on $\mathbb{K}^g$ which is completely antisymmetric with $\varepsilon_{1\ 2\dots g}=1$, where $\varepsilon_{\lambda_1\dots\lambda_g}=\varepsilon(e_{\lambda_1},\dots,e_{\lambda_g})$ are the components of ε in the canonical basis (e_λ) of $\mathbb{K}^g$. Let w_t be a 1-parameter family of preregular g-linear forms on $\mathbb{K}^g$ with $w_0=\varepsilon$ and let us investigate what happens formally at first order in t. One writes

$$\left\{\begin{array}{l} w_t=\varepsilon+t\dot{w}+o(t^2)\\ Q_{w_t}=(-1)^{g-1}\mathbb{1}+t\dot{Q}+o(t^2)\end{array}\right. \tag{5.20}$$

and the first order Q_{w_t}-cyclicity reads

$$\dot{W}_{\lambda_1\dots\lambda_g}=\dot{Q}^\lambda_{\lambda_g}\varepsilon_{\lambda\lambda_1\dots\lambda_{g-1}}+(-1)^{g-1}\dot{W}_{\lambda_g\lambda_1\dots\lambda_{g-1}} \tag{5.21}$$

with $\dot{W}_{\lambda_1\dots\lambda_g}=\dot{w}(e_{\lambda_1},\dots,e_{\lambda_g})$. This equation implies

$$\mathrm{tr}(\dot{Q})=\dot{Q}^\lambda_\lambda=0 \tag{5.22}$$

which suggests $\det(Q_{w_t})=1$ for a finite version. So a natural question is the following : Does a quadratic AS-regular algebra $\mathcal{A}(w,2)$ is such that $\det(Q_w)=1$? By looking at Example (c) of §3.3, one can see that the answer is no. Notice however that the quadratic AS-algebra of type E is isolated.

References

[1] M. Artin and W.F. Schelter. Graded algebras of global dimension 3. *Adv. Math.*, 66:171–216, 1987.

[2] M. Artin, J. Tate, and M. Van den Bergh. Some algebras associated to automorphisms of elliptic curves. The Grothendieck Festschrift. vol. I. *Prog. Math.*, 86:33–85, 1990.

[3] R. Berger. Koszulity for nonquadratic algebras. *J. Algebra*, 239:705–734, 2001.

[4] R. Berger. Dimension de Hochschild des algèbres graduées. *C.R. Acad.Sci. Paris, Ser. I*, 341:597–600, 2005.

[5] R. Berger. Gerasimov's theorem and N-Koszul algebras. *J. London Soc.*, 79:631–648, 2009.

[6] R. Berger, M. Dubois-Violette, and M. Wambst. Homogeneous algebras. *J. Algebra*, 261:172–185, 2003.

[7] R. Berger and N. Marconnet. Koszul and Gorenstein properties for homogeneous algebras. *Algebras and Representation Theory*, 9:67–97, 2006.

[8] J. Bichon. The representation category of the quantum group of a non-degenerate bilinear form. *Comm. Algebra*, 31:4831–4851, 2003.

[9] R. Bocklandt, T. Schedler, and M. Wemyss. Superpotentials and higher order derivations. *J. Pure Appl. Algebra*, 214:1501–1522, 2010.

[10] A.I. Bondal and A.E. Polishchuk. Homological properties of associative algebras: The method of helices. *Russian Acad. Sci. Izv. Math.*, 42:219–260, 1994.

[11] H. Cartan. Homologie et cohomologie d'une algèbre graduée. *Séminaire Henri Cartan*, 11(2):1–20, 1958.

[12] H. Cartan and S. Eilenberg. *Homological algebra*. Princeton University Press, 1973.

[13] A. Connes. C^* algèbres et géométrie différentielle. *C.R. Acad. Sci. Paris, Série A*, 290:599–604, 1980.

[14] A. Connes. Non-commutative differential geometry. *Publ. IHES*, 62:257–360, 1986.

[15] A. Connes. *Non-commutative geometry*. Academic Press, 1994.

[16] A. Connes and M. Dubois-Violette. Noncommutative finite-dimensional manifolds. I.Spherical manifolds and related examples. *Commun. Math. Phys.*, 230:539–579, 2002.

[17] A. Connes and M. Dubois-Violette. Yang-Mills algebra. *Lett. Math. Phys.*, 61:149–158, 2002.

[18] A. Connes and M. Dubois-Violette. Moduli space and structure of noncommutative 3-spheres. *Lett. Math. Phys.*, 66:91–121, 2003.

[19] A. Connes and M. Dubois-Violette. Yang-Mills and some related algebras. In *Rigorous Quantum Field Theory*, volume 251 of *Progr. Math.*, pages 65–78. Birkhaüser, 2007.

[20] A. Connes and M. Dubois-Violette. Noncommutative finite-dimensional manifolds. II. Moduli space and structure of noncommutative 3-spheres. *Commun. Math. Phys.*, 281(1):23–127, 2008.

[21] M. Dubois-Violette. $d^N = 0$: Generalized homology. *K-Theory*, 14:371–404, 1998.

[22] M. Dubois-Violette. Graded algebras and multilinear forms. *C. R. Acad. Sci. Paris, Ser. I*, 341:719–724, 2005.

[23] M. Dubois-Violette. Multilinear forms and graded algebras. *J. Algebra*, 317:198–225, 2007.

[24] M. Dubois-Violette and G. Launer. The quantum group of a non-degenerate bilinear form. *Phys. Lett.*, 245B:175–177, 1990.

[25] M. Dubois-Violette and T. Popov. Homogeneous algebras, statistics and combinatorics. *Lett. Math. Phys.*, 61:159–170, 2002.

[26] D.I. Gurevich. Algebraic aspects of the quantum Yang-Baxter equation. *Algebra i Analiz (Transl. in Leningrad Math. J. 2 (1991) 801-828)*, 2:119–148, 1990.

[27] P.H. Hai, B. Kriegk, and M. Lorenz. N-homogeneous superalgebras. *J. Noncommut. Geom.*, 2:1–51, 2008.

[28] R.S. Irving. Prime ideals of Ore extensions. *J. Algebra*, 58:399–423, 1979.

[29] T.V. Kadeishvili. On the theory of homology of fiber spaces. *Russ. Math. Surv.*, 35:231–238, 1980.

[30] B. Keller. Introduction to A_∞ algebras and modules. *Homology Homotopy Appl.*, 3:1–35, 2001.

[31] B. Kriegk. Un critère numérique pour la propriété de Koszul. *C.R. Acad. Sci. Paris, Série I*, 344:545–548, 2007.

[32] D.M. Lu, J.H. Palmieri, Q.S. Wu, and J.J. Zhang. Regular algebras of dimension 4 and their A_∞-Ext-algebras. *Duke Math. J.*, 137:537–584, 2007.

[33] D.M. Lu, J.H. Palmieri, Q.S. Wu, and J.J. Zhang. A_∞ structure on Ext-algebras. *J. Pure Appl. Algebra*, 213:2017–2037, 2009.

[34] Yu. I. Manin. Some remarks on Koszul algebras and quantum groups. *Ann. Inst. Fourier, Grenoble*, 37:191–205, 1987.

[35] Yu. I. Manin. *Quantum groups and non-commutative geometry.* CRM Université de Montréal, 1988.

[36] A.V. Odesskii. Elliptic algebras. *Russ. Math. Surv.*, 57:414–442, 2002.

[37] A.V. Odesskii and B.L. Feigin. Sklyanin elliptic algebras. *Func. Anal. Appl.*, 23:207–214, 1989.

[38] S.B. Priddy. Koszul resolutions. *Trans. Amer. Math. Soc.*, 152:39–60, 1970.

[39] E.K. Sklyanin. Some algebraic structures connected with the Yang-Baxter equation. *Func. Anal. Appl.*, 16:263–270, 1982.

[40] S.P. Smith and J.T. Stafford. Regularity of the four dimensional Sklyanin algebra. *Compos. Math.*, 83:259–289, 1992.

[41] M. Van den Bergh. A relation between Hochschild homology and cohomology for Gorenstein rings. *Proc. Amer. Math. Soc.*, 126:1345–1348, 1998.

[42] M. Van den Bergh. Erratum. *Proc. Amer. Math. Soc.*, 130:2809–2810, 2002.

[43] M. Wambst. Complexes de Koszul quantiques. *Ann. Inst. Fourier, Grenoble*, 43:1083–1156, 1993.

Laboratoire de Physique Théorique, UMR 8627, Université Paris XI, Bâtiment 210, F-91 405 Orsay Cedex

E-mail address: `Michel.Dubois-Violette@u-psud.fr`

Clay Mathematics Proceedings
Volume 11, 2010

Random Matrices, Free Probability, Planar Algebras and Subfactors

A. Guionnet, V. F. R. Jones, and D. Shlyakhtenko

Dedicated to Alain Connes.

Abstract. Using a family of graded algebra structures on a planar algebra and a family of traces coming from random matrix theory, we obtain a tower of non-commutative probability spaces, naturally associated to a given planar algebra. The associated von Neumann algebras are II_1 factors whose inclusions realize the given planar algebra as a system of higher relative commutants. We thus give an alternative proof of a result of Popa that every planar algebra can be realized by a subfactor.

Contents

2010 *Mathematics Subject Classification.* Primary: 46L54; Secondary: 15B52, 46L37.

The first author was supported by the Miller Institute for Basic Research in Science, University of California Berkeley.

The second author was supported by NSF grant DMS-0401734.

The third author was supported by NSF grant DMS-0555680.

1. Introduction

It has been apparent for quite some time that there exists a strong connection between subfactors, large random matrices and free probability theory. Perhaps the most clear instance of this connection is that all three theories have an underlying *planar* structure. For example, the standard invariant of a subfactor (i.e., the system of higher relative commutants) is in a natural way a planar algebra [**Jon99**]. Traces of polynomials in random matrices naturally count certain planar objects ([**tH74, BIPZ78, GMS06, MS06, CMŚS07, MŚS07, Zvo97**]). Finally, the combinatorics of free probability theory is intimately tied to that of non-crossing (i.e., planar) partitions [**Spe94**]. Furthermore, techniques from some of these subjects have proved useful for applications to others. For example, there are many connections between work in free probability theory and certain computations in the paper [**BJ97**]. Random matrices and free probability theory were used to construct subfactors [**Răd94, SU02, Pop95, PS03b**]. More recently, Mingo and Speicher and Guionnet and Maurel-Segala [**Gui06, GMS06, MS06, MŚS07, CMŚS07**] have found combinatorial expressions, involving planar diagrams, for the large-N asymptotics of moments of polynomials in certain random matrices.

In this paper we exploit for the first time the *graded algebra* coming from a planar algebra P to obtain a subfactor $N \subset M$ whose standard invariant is P. The essential ingredient is a *trace* on the graded algebra coming from free probability/random matrices, whose use in this context was inspired by [**Gui06, GMS06, GS08**], which promises to be a source of further developments in this direction.

We take the point of view that all of the three subjects mentioned above are intimately related to the notion of a planar algebra. Specifically, the underlying idea is that a planar algebra, endowed with its graded multiplication $\wedge_0$ and trace Tr_0 is a natural replacement for the ring of polynomials occurring in both free probability theory [**VDN92**] and the theory of random matrices with a potential [**Gui06, GMS06, GS08**].

To be more precise, a *subfactor planar algebra* (SPA) P will be a graded vector space $P = (P_n, n > 0, P_0^{\pm})$ which is an algebra over the planar operad of [**Jon01, Jon99, Jon00**] and satisfies certain dimension and positivity conditions outlined in §2. Every extremal finite index subfactor has an SPA as its *standard invariant*.

Given an SPA P, we define the sequence Gr_kP, $k = 0, 1, 2, \dots$ of complex $*$-algebras with $Gr_kP = \oplus_{n \geq k} P_n$ ($P_{0,+} \oplus \bigoplus_{n\geq 1} P_n$ if $k = 0$) and multiplications $\wedge_k : P_n \times P_m \to P_{n+m-k}$ given by tangles as in §2. On each Gr_kP we define a trace $Tr_k : Gr_kP \to \mathbb{C}$ using the sum of all Temperley-Lieb tangles. The trace Tr_0 comes directly from Wick's Theorem applied to large N limit of a certain Gaussian matrix model using Wishart matrices, defined in §3. But once calculated, this trace can be defined entirely in terms of planar algebras.

A rather important special case is when the matrix models may be taken as p independent Hermitian matrices. Then the algebra Gr_0P is the even degree subalgebra of $\mathbb{C}\langle\{X\}\rangle$, the non-commutative polynomials in p self-adjoint variables $\{X\}$ (with $\#\{X\} = p$). The trace Tr_0 is then the one discovered by Voiculescu in the context of his free probability theory [**VDN92, Voi85, Voi91**]. It can be realized as the vacuum expectation value on the full Fock space on a real p-dimensional vector space with basis $\{X\}$, by the representation of $\mathbb{C}\langle\{X\}\rangle$ which sends X to $\ell_X + \ell_X^*$, ℓ_X being the left creation operator of X (see [**Voi85, VDN92**]). The higher multiplications $\wedge_k$ are then given (on monomials of even degree $\geq 2k$) by

$$(X_1X_2\cdots X_r) \wedge_k (Y_1Y_2\cdots Y_s) = (\prod_{i=1}^{k} \delta_{X_{r-i+1},Y_i})X_1\cdots X_{r-k}Y_{k+1}\cdots Y_s.$$

Our main result is, with notation as above and an SPA P, of index parameter δ,

THEOREM. *(i) For each k, tr_k is a faithful tracial state on Gr_kP and the GNS completion of Gr_kP is a II_1 factor M_k as long as $\delta > 1$;*

(ii) There are unital inclusions $Gr_kP \subset Gr_{k+1}P$ which extend to $M_k \subset M_{k+1}$ and projections $\mathbf{e}_k \in Gr_{k+1}P$, such that $(M_{k+1}, \mathbf{e}_k)$ is the tower of basic constructions for the subfactor $M_0 \subset M_1$;

(iii) The relative commutants $M_0' \cap M_k$ are canonically identified with the vector spaces P_k and this identification is a homomorphism of planar $*$-algebras.

This theorem gives a new proof of the breakthrough result of Popa [**Pop95**], showing that any subfactor planar algebra P can indeed be realized by the system of higher relative commutants of a II_1 subfactor.

The key ingredient in the proofs will be representations of the algebras Gr_kP on Fock spaces. In order to define these we will suppose that P is given as a planar subalgebra of the full planar algebra P^Γ of some bipartite graph $\Gamma = \Gamma_+ \amalg \Gamma_-$ as in [**Jon00**]. This is always possible — one may for instance take Γ to be the principal graph of P. A basis of P^Γ is formed by loops on Γ starting and ending in Γ_+ but we will define a slightly different planar algebra structure from that of [**Jon00**], better adapted to graded multiplication.

The Fock space will then be spanned (orthogonally) by paths of varying lengths on Γ, ending in Γ_+. It is naturally $\mathbb{Z}/2\mathbb{Z}$-graded. Note that Γ may be infinite so we need to make a choice of Perron-Frobenius eigenvector and eigenvalue for the adjacency matrix of Γ. There will not necessarily be a Markov trace on P^Γ so we work instead with the center-valued trace. This will restrict to a Markov trace on P.

As in the theory of graph C^*-algebras [**Rae05**], each edge e of Γ defines an operator ℓ_e (of grading 1) on the Fock space, creating an edge on a path. A loop of edges $e_1 \cdots e_{2p}$ in P^Γ is then represented by the product $c(e_1)c(e_2)\cdots c(e_{2p})$ where

$c(e_i)$ is a version of $a_i\ell(e_i) + a_i^{-1}\ell(e_i^*)$ according to the parity of i, where the factors a_i are determined by the Perron-Frobenius eigenvector. We also make use of the fact that the Fock space of loops and the resulting II_1 factor can be embedded into a type III factor canonically associated to the graph and its Perron-Frobenius eigenvector using a free version of the second quantization procedure.

1.1. Notations. To aid the reader, we list here some notation used in the paper.

- Bipartite graph (§2.4): Γ; vertices: Γ; even/odd vertices: $\Gamma_\pm$; Edges: E; positively/negatively oriented edges: $E_\pm$; edges starting at v: $\Gamma_+(v)$; edges ending at v: $\Gamma_-(v)$. All loops of length k starting at even/odd vertex: $L_k^\pm$; all loops starting at a positive/negative vertex: $L^\pm$.
- Planar algebra (Def. 4): P; k-box space: $P_{k,\pm}$; positive/negative part: $P_\pm$; $P_k = P_{k,+}$. Planar algebra of a graph (§2.4): P^Γ, $P^\Gamma_{k,\pm}$ etc. Subfactor planar algebra: Def. 5.
- Graded multiplications: $\wedge_k$ (Def. 7), the algebra Gr_kP. Trace on Gr_kP: Tr_k (Def. 8).

2. On planar algebras

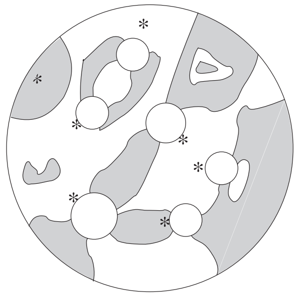
2.1. Definition. We begin with a definition of planar algebra which will be recognizably equivalent to other definitions [**Jon99**] and suited to the purposes of this paper.

Definition 1. (Planar k-tangles.) A planar k-tangle will consist of a smoothly embedded disc D $(= D_0)$ in $\mathbb{R}^2$ minus the interiors of a finite (possibly empty) set of disjoint smoothly embedded discs $D_1, D_2, \ldots, D_n$ in the interior of D. Each disc D_i, $i \geq 0$, will have an even number $2k_i \geq 0$ of marked points on its boundary (with $k = k_0$). Inside D and outside $D_1, D_2, \ldots, D_n$ there is also a finite set of disjoint smoothly embedded curves called strings which are either closed curves or whose boundaries are marked points of the D_i's. Each marked point is a boundary point of some string, and the strings meet the boundaries of the discs transversally, only in the marked points. The connected components of the complement of the strings in $\overset{\circ}{D} \setminus \bigcup_{i=1}^{n} D_i$ are called regions. Those parts of the boundaries of the discs between adjacent marked points (and the whole boundary if there are no marked points) will be called intervals. The regions of the tangle will be shaded black and white so that two regions whose boundaries intersect are shaded differently. (Such a shading is always possible, since there is an even number of marked points.) The shading will be considered to extend to the intervals which are part of the boundary of a region. Finally, to each disc in a tangle there is a distinguished interval on its boundary (which may be shaded black or white).

Definition 2. The set of internal discs of a tangle T will be denoted by $\mathcal{D}_T$.

Remark 1. Observe that diffeomorphisms of $\mathbb{R}^2$ act on planar tangles in the obvious way. In particular if Φ is a diffeomorphism it induces a map $\Phi : \mathcal{D}_T \to \mathcal{D}_{\Phi(T)}$

We will often have to draw pictures of tangles. To indicate the distinguished interval on the boundary of a disc we will place a *, near to that disc, in the region whose boundary contains the distinguished interval. An example of a 4-tangle illustrating all the above ingredients is given below:



We will often use pictures with a given number of strings to illustrate a situation where the number of strings is arbitrary. We hope this will not lead to misinterpretation. Similarly, if the shading is implicit or both possible shadings are intended we will suppress the shading.

Given planar k and k'-tangles T and S respectively, we say they are composable if

(1) The outside boundary of S is equal to the boundary of one of the inside discs of T, where equality means that the marked points are the same, the shadings of the intervals are the same and the distinguished intervals are the same. And,
(2) The union of the strings of S and those of T are smooth curves.

Definition 3. If T and S are composable we define the composition $T \circ S$ to be the union $T \cup S$. The strings of $T \circ S$ are the unions of the strings of T and S.

Since the shadings of T and S agree on their common boundary curve, it is easy to see that $T \cup S$ is a planar k-tangle. This composition operation is often called "gluing" as one may think of S as being glued inside T.

We will now define a notion of planar algebra. Axioms can be subtracted to obtain more general objects but for convenience in this paper the term "planar algebra" will imply all the properties.

Before giving the formal definition we recall the notion of the Cartesian product of vector spaces over an index set $\mathcal{I}$, $\times_{i\in\mathcal{I}} V_i$. This is the set of functions f from $\mathcal{I}$ to the union of the V_i with $f(i) \in V_i$. Vector space operations are pointwise. Multilinearity is defined in the obvious way, and one converts multilinearity into linearity in the usual way to obtain $\otimes_{i\in\mathcal{I}} V_i$, the tensor product indexed by $\mathcal{I}$. A Cartesian product over the empty set will mean the scalars.

Definition 4. A (unital) planar algebra P will be a family of $\mathbb{Z}/2\mathbb{Z}$-graded vector spaces indexed by the set $\{\mathbb{N} \cup \{0\}\}$, where $P_{k,\pm}$ will denote the $\pm$ graded space indexed by k. To each planar tangle T there will be a multilinear map

$$Z_T : \times_{D\in\mathcal{D}_T} P_D \to P_{D_0}$$

where P_D is the vector space indexed by half the number of marked boundary points of D and graded by $+$ if the distinguished interval of D is shaded white and $-$ if it is shaded black.

The maps Z_T are subject to the following two requirements:

(1) (Isotopy invariance) If φ is an orientation preserving diffeomorphism of $\mathbb{R}^2$ then

$$Z_T = Z_{\varphi(T)}$$

where the sets of internal discs of T and $\varphi(T)$ are identified using φ.

(2) (Naturality)

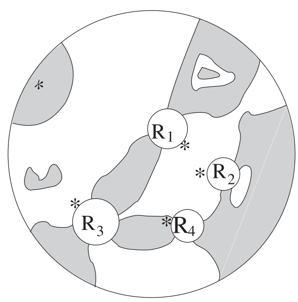
$$Z_{T \circ S} = Z_T \circ Z_S$$

where the right hand side of the equation is defined as follows: first observe that $\mathcal{D}_{T\circ S}$ is naturally identified with $(\mathcal{D}_T \backslash \{D'\}) \cup \mathcal{D}_S$, where D' is the disc of T containing S. Thus, given a function f on $\mathcal{D}_{T\circ S}$ to the appropriate vector spaces, we may define a function $\tilde{f}$ on $\mathcal{D}_T$ by

$$\tilde{f}(D) = \begin{cases} f(D) & \text{if } D \neq D' \\ Z_S(f|_{\mathcal{D}_S}) & \text{if } D = D' \end{cases}$$

Finally, the formula $Z_T \circ Z_S(f) = Z_T(\tilde{f})$ defines the right hand side.

The natural notation for $Z_T(f)$ is to write in the $\{f(D),\ D \in \mathcal{D}_T\}$ into $\mathbb{D}$. This is just like the notation "$y(x_1, \ldots, x_n)$" for a function of several variables, where the x_i are the $f(D)$, and the internal discs correspond to the spaces between the commas. (We also call the internal disks "input discs"). Thus if R_1 and R_2 are in $P_{2,+}$, R_3 is in $P_{2,-}$ and R_4 is in $P_{3,+}$ then the following picture is an element of $P_{4,-}$:

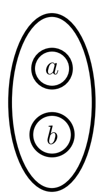


The vector spaces $P_{n,\pm}$ will possess a conjugate linear involution $*$, $x \to x^*$ with the compatibility requirement:

$$Z_T(f^*) = Z_{\Phi(T)}(f \circ \Phi)^*$$

whenever Φ is an orientation reversing diffeomorphism.

Observe that $P_{0,\pm}$ become *unital commutative $*$-algebras* under the multiplication operation (with either shading):

$ab =$ 

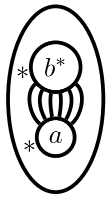
DEFINITION 5. A subfactor planar algebra P will be a planar algebra satisfying the following four conditions:

(a) $\dim(P_{n,\pm}) < \infty$ for all $(n, \pm)$

(b) $\dim(P_{0,\pm}) = 1$

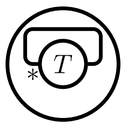
Condition (ii) allows us to canonically identify $P_{0,\pm}$ with $\mathbb{C}$ as *-algebras, 1 being Z (a 0-tangle with nothing in it).

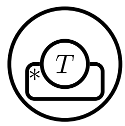
This further allows us to define a sesquilinear form on each $P_{n,\pm}$ by

$\langle a, b\rangle =$ 

where the outside region is shaded according to $\pm$.

(3) The form $\langle\ ,\ \rangle$ is positive definite.
(4) $Z_{T_1} = Z_{T_2}$ where T_1 and T_2 are the following two 0-tangles:

$T_1 =$ 

$T_2 =$ 

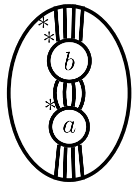
The last condition is topologically natural and corresponds to extremality of the subfactor ([**PP86**],[**Pop94**, 1.2.5]). This condition means that the partition function of a fully-labeled zero-tangle (when considered without its boundary disc) is actually well-defined for that zero-tangle on the sphere S^2 obtained by adding a point at ∞ to $\mathbb{R}^2$. It is natural then to suppress the outer disc of a 0-tangle in pictures.

Remark 2. Once $P_{0,\pm}$ have been identified with the scalars there is a canonical scalar δ associated with a subfactor planar algebra with the property that the multilinear map associated to any tangle containing a closed string is equal to δ times the multilinear map of the same tangle with the closed string removed. By positivity $\delta > 0$ and it is well known that in fact the possible values of δ form the set $\{4\cos^2 \pi/n : n = 3, 4, 5, \ldots\} \cup [4, \infty)$ [**Jon83**].

Remark 3. Since $\delta \neq 0$ it is clear that all the spaces $P_{n,-}$ are redundant and subfactor planar algebra could be axiomatized in terms of $P_{n,+}$. For this reason we will use in what follows P_n to denote $P_{n,+}$ (even in the non-subfactor case).

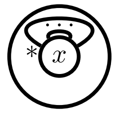
Remark 4. In the development of planar algebras the following structures played a major role:

(1) Multiplication: Each $P_{n,\pm}$ is a *-algebra with the involution defined above and the multiplications

$ab =$ 

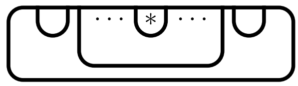
There are two choices of shadings which give in general non-isomorphic algebra structures. (We shall refer to this multiplication sometimes as the "usual" multiplication on $P_{n,\pm}$).

(2) Trace: Each $P_{n,\pm}$ is equipped with a linear map $Tr : P_{n,\pm} \to P_{0,\pm}$ which is given by

$Tr(x) =$ 

(3) The Temperley-Lieb tangles: each tangle consisting of an outside box all of whose $2k$ boundary points are connected by (non-crossing) strings inside of the box determines an element of $P_{k,\pm}$, pairing depending on the shading of the region containing $*$. The set of such tangles is denoted by $TL(k)$.

(4) The Jones projections: $\mathbf{e}_k \in P_{k,+}$ is given by the Temperley-Lieb tangle (having $2k$ boundary points):

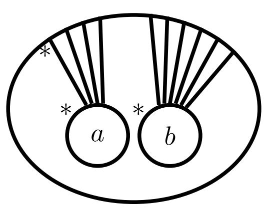
$\mathbf{e}_k =$  .

Remark 5. (Rectangles). It is sometimes very convenient to use rectangles rather than circles for the input and output discs. Strictly speaking this is not allowed since the boundaries are supposed to be smooth. But nothing will happen at the corners of the rectangles so one may simply interpret a picture of a rectangle as one with smoothed corners. Use of horizontal rectangles also makes it possible to avoid specifying the first interval which we will always suppose to be the one containing the left hand vertical part of the rectangle.

Remark 6. (Outer disks and shading). We occasionally omit the outer disk when describing a planar algebra element, especially in the case that there are no boundary points on the outer disk. Also, unless the shading is explicitly indicated in a picture, we follow the convention that the region adjacent to the boundary region marked with a $*$ is unshaded (white).

2.2. Graded algebra structures.

Definition 6. If P is a planar algebra we define a graded algebra GrP as follows. As a graded vector space $GrP = \bigoplus_{n=0}^{\infty} P_{n,+}$ and the g $\wedge : P_n \times P_m \to P_{m+n}$ is given by the tangle below which puts the element of P_n entirely to the left of the element of P_n:

$a \wedge b =$ 

(The shading in the picture above is determined by saying that the region adjacent to the marked interval on the outer box is unshaded (white); as before, $*$'s denote the marked intervals on the disks). Note that one could also define a dual structure changing $+$ to $-$ and changing the shading in the above figure.

As a graded algebra a subfactor planar algebra is just the free graded algebra on a certain graded vector space, as we shall see.

If $\mathcal{P}$ is a subfactor planar algebra let $\mathfrak{M}$ be the 2-sided ideal of $Gr\mathcal{P}$ spanned by all elements of degree 1 or more. $\mathfrak{M}$. Each graded piece of $\mathfrak{M}$ has an innner product as defined above. For each $n \geq 1$ let $\mathfrak{N}_n$ be the orthogonal complement of $(\mathfrak{M}^2)_n$ in $\mathfrak{M}_n$.

THEOREM 1. *With notation as above, $Gr\mathcal{P}$ is the free graded algebra generated freely by $\cup_{n=1}^{\infty}\mathfrak{N}_n$.*

PROOF. Let $\pi = (\pi_1, \pi_2, ..., \pi_k)$ be an ordered k-tuple of integers with $\pi_i \geq 1$ and $\sum_{j=1}^{k} \pi_k = n$. Then multiplication defines linear maps

$$mult_\pi : \mathfrak{N}_{\pi_1} \otimes \mathfrak{N}_{\pi_2} \otimes \cdots \otimes \mathfrak{N}_{\pi_k} \to \mathcal{P}_n.$$

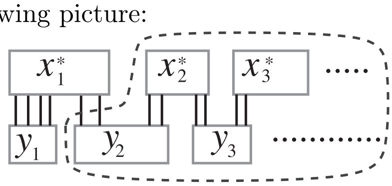
By induction the images of $mult_\pi$ span $\mathfrak{M}_n^2$ as π varies. So, together with $\mathfrak{N}_n$ they span $\mathcal{P}_n$. Thus the theorem follows from the two assertions:
i) Each $mult_\pi$ is injective.
ii) The images of the $mult_\pi$ are orthogonal for different π.

To see i), note that each $mult_\pi$ is an isometry if we give

$$\mathfrak{N}_{\pi_1} \otimes \mathfrak{N}_{\pi_2} \otimes \cdots \otimes \mathfrak{N}_{\pi_k}$$

the Hilbert space tensor product structure.

To see ii), let π and ρ be two distinct partitions of n as above. Suppose $\pi_1 > \rho_1$. Consider the following picture:



This is the inner product of an element $y_1 \otimes y_2 \otimes \cdots$ in $\mathfrak{N}_{\rho_1} \otimes \mathfrak{N}_{\rho_2} \otimes \cdots \otimes \mathfrak{N}_{\rho_k}$ with an element $x_1 \otimes x_2 \otimes \cdots$ in $\mathfrak{N}\pi_1 \otimes \mathfrak{N}_{\pi_2} \otimes \cdots \otimes \mathfrak{N}_{\pi_k}$. (Here $\pi_1 = 3, \pi_2 = 2$ and $\rho_1 = 2 = \rho_2 = \rho_3$). One may evaluate the tangle inside the dashed curve to obtain an element of $\mathfrak{M}_{\pi_1 - \rho_1}$. Thus the figure is actually the inner product of x_1 with an element of $\mathfrak{M}^2$, thus it is zero. So the images of $mult_\pi$ and $mult_\rho$ are orthogonal unless $\pi_1 = \rho_1$. Continuing in this way we see that the images of $mult_\pi$ and $mult_\rho$ are orthogonal unless $\pi = \rho$. $\square$

REMARK 7. Writing elements of $Gr\mathcal{P}$ as sums of products of elements orthogonal to $\mathfrak{M}^2$ times arbitrary elements gives, by an easy argument with generating functions,

$$\Psi_{\mathcal{P}}(z) = 1 - \frac{1}{\Phi_{\mathcal{P}}(z)}$$

where $\Psi_{\mathcal{P}}(z)$ is the generating function for $\dim(\mathfrak{M}/\mathfrak{M}^2)_n$, and $\Phi_{\mathcal{P}}(z)$ is the generating function for $\dim P_n$. In general, if Φ_n is the generating function for the dimensions of the graded vector space $\mathfrak{M}^n/\mathfrak{M}^{n+1}$ we have $\Phi_n = \Phi(\Phi_n - \Phi_{n+1})$, so that $\Phi_n = (1 - 1/\Phi)^n$.

Although the graded algebra structure is not commutative even up to a sign, the presence of the cyclic group action gives a kind of "cyclic commutativity" as follows, where ρ denotes the action of the counterclockwise rotation tangle on P_n:

PROPOSITION 1. *If $\mathcal{P}$ is a planar algebra then*

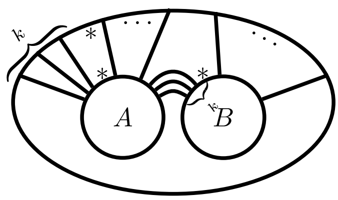
$$\rho^{\deg a}(a \wedge b) = b \wedge a$$

PROOF. Just draw the picture. □

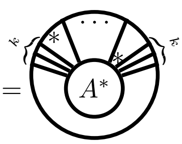
REMARK 8. The multiplication in an exterior algebra can be made to satisfy exactly the same commutativity formula by making the cyclic group act by the appropriate sign in each degree.

Besides $Gr\mathcal{P}$ we will need other "shifted" graded *-algebra structures on $\mathcal{P}$ in order to define a subfactor and analyze its tower.

DEFINITION 7. Given a planar algebra $P = (P_n)$ and an integer $k \geq 0$ we make $\oplus_{n=k}^{\infty} P_n$ into an associative (unital) *-algebra with multiplication $\wedge_k : P_m \times P_n \to P_{m+n-k}$ given by the following formula:

$A \wedge_k B =$ 

The involution (denoted by † to distinguish it from the usual involution $*$ on P_k) is given by

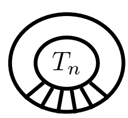
$A^\dagger =$ 

The shading in both figures above is determined by the condition that the marked boundary region $*$ is adjacent to an unshaded (white) region. Here A^* means $\phi(A)$ where ϕ is an orientation-reversing diffeomorphism (cf. Def. 4(2)).

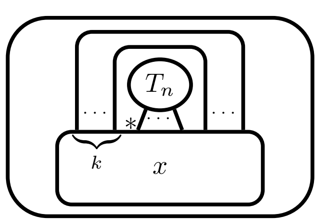
We denote this *-algebra by $Gr_k P$.

2.3. The traces Tr_k. Any planar algebra contains in a canonical way the Temperley-Lieb planar algebra TL. Indeed, TL is spanned by TL diagrams: a TL diagram is a diagram that has no inner disks, and all of whose strings connect points on the outer disk. Any such diagram is naturally an element of P.

DEFINITION 8. Let T_n be the sum of all TL diagrams having $2n$ points on the outer disk represented pictorially below (for $n = 3$):



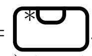
(The position of the $*$ is irrelevant, by since the set of TL diagrams is invariant under a rotation by $2\pi/n$). The trace $Tr_k(x)$ is defined for $x \in P_m$, $m \geq k$, and is valued in the zero box space of P:



where $n = m - k$ (in other words, there are k strings surrounding T_n).

LEMMA 1. *Tr_k is a trace on $Gr_k P$ if endowed with the multiplication $\wedge_k$.*

PROOF. This follows from the fact that the set of all TL diagrams on $2m$ points is invariant under rotations by $2\pi/m$. ☐

Before proceeding further, we consider an example. Let us assume that P is a subfactor planar algebra, so that in particular $P_{0,\pm}$ are one-dimensional and Tr_k is scalar-valued. Let $\cup$ be the following element of TL: $\cup =$ 

. Let us denote by Φ the moment generating function of $\cup$. Thus we let $\Phi(z)$ be the unique scalar defined by

$$\Phi(z) = \sum_{n=0}^{\infty} Tr_0(\underbrace{\cup \wedge_0 \cdots \wedge_0 \cup}_{n \text{ times}}) z^n.$$

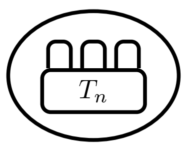
We shall presently compute $\Phi(z)$ by using planar algebra methods.

DEFINITION 9. Let T_n be the element of the planar algebra defined as the sum of all the Temperley-Lieb diagrams connecting the $2n$ boundary points,

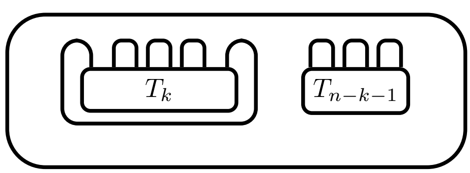
LEMMA 2.

$$\Phi(z) = \frac{1-(\delta-1)z}{2z}\left(1 - \sqrt{1 - \frac{4z}{(1-(\delta-1)z)^2}}\right).$$

PROOF. The trace of $\cup^n$ is given by the picture (corresponding to $n = 3$):



Group the TL diagrams in T_n according to where the first boundary point of $\cup^n$ is connected. Adding all those diagrams where it is connected to its nearest neighbor we get $\delta Tr_0(\cup^{n-1})$. Proceeding similarly we get, for $k = 1, 2, \ldots, n-1$, contributions of the form:



If the first term in the picture is rotated by one we may use the rotational invariance of T_k to see that it is just $Tr_0(\cup^k)$. Thus we have, for each $n > 0$,

$$Tr_0(\cup^n) = (\delta-1)Tr_0(\cup^{n-1}) + \sum_{k=0}^{n-1} Tr_0(\cup^k)Tr_0(\cup^{n-k-1})$$

Multiplying both sides by z^n and summing from $n = 1$ to ∞ we see that

$$\Phi - 1 = z(\delta-1)\Phi + z\Phi^2$$

Solving the quadratic equation and checking the first term to get the right solution we obtain our answer. ☐

The function Φ in Lemma 2 is that of a free Poisson random variable having R-transform $\delta(1-z)^{-1}$ (see [**Voi00**, p. 311]). We shall give an alternative computation using free convolution later in the paper (see Lemma 5).

The following lemma can be easily proved by drawing the appropriate pictures:

LEMMA 3. *The obvious linear embedding of P_k into $Gr_kP = P_k \oplus P_{k+1} \oplus \cdots$ is an algebra $*$-homomorphism from P_k endowed with its usual $*$-algebra structure of Remark 4, to Gr_kP taken with multiplication $\wedge_k$ and conjugation $\dagger$ as in Definition 7. Moreover, this embedding carries the trace Tr_k to the usual trace Tr on P_k.*

2.4. The planar algebra of a bipartite graph. Let $\Gamma = \Gamma_+ \cup \Gamma_-$ be a (locally finite) bipartite graph with adjacency matrix A_Γ possessing an eigenvector $\mu = \mu_v$ (v being a vertex of Γ) with $\mu_v > 0$ for all v and $A_\Gamma \mu = \delta\mu$. Note that although μ may be unbounded as a function of Γ, the ratios $\mu(v)/\mu(v')$, where v and v' are adjacent, are bounded by the eigenvector condition.

We shall denote by E the set of oriented edges of Γ, taken with all possible orientations. Thus $E = E_+ \cup E_-$, where E_+ consists of all edges of Γ oriented so as to start at a vertex in Γ_+ and end at a vertex in Γ_-, and E_- will consists of all edges of Γ oriented so as to start in Γ_- and end in Γ_+. For $e \in E$ we'll denote by e^o the edge with the opposite orientation.

In [**Jon00**] a planar algebra was associated with the above data with the property that closed strings may be removed multiplicatively as in remark 2. We quickly redo this planar algebra with a slightly different (but isomorphic) structure, emphasizing those elements that arise when Γ is infinite.

With Γ, μ as above we will define the planar algebra $P^\Gamma = P^\Gamma_{n,\pm}$, where $P^\Gamma_{n,\pm}$ is the vector space of bounded functions on loops on Γ of length $2n$ starting and ending in Γ_+ for the plus sign and Γ_- for the minus sign.

DEFINITION 10. (Spin State) Given a planar tangle T, and a bipartite graph Γ as above a *spin state* σ will be a function from the regions of T to the vertices of Γ, shaded regions being mapped to Γ_+ and unshaded ones to Γ_-, together with a function from the strings of T to the edges of Γ such that if a string S is part of the boundary of the regions R_1 and R_2 then $\sigma(S)$ is an edge connecting $\sigma(R_1)$ and $\sigma(R_2)$.

Note that a state σ determines a function $\ell_\sigma : \mathcal{D}_T \cup \{\text{boundary disc}\} \to \{\text{loops on } \Gamma\}$ in the obvious way — if we follow a disc of T around clockwise, the intervals, beginning at the distinguished one, touch regions of T to which σ has assigned vertices of Γ and the strings connected to the marked boundary points of a disc D have been assigned edges of Γ connecting the vertices on either side. We will call $\ell_\sigma(D)$ the loop *induced* on D by σ.

DEFINITION 11. (The curvature factor of a spin state.) Given a tangle and a spin state σ as above, define the curvature factor $c(\sigma)$ as follows. First isotope the tangle so that all discs are horizontal rectangles (with the first boundary interval on the left as in remark 5) and all marked points are on the top edges of the rectangles. Arrange also for all singularities of the y coordinate on the strings to be generic (maxima or minima). Near such a maximum (resp. minimum) we see regions above and below, one of which is convex, labeled by adjacent (on Γ) vertices v_{convex} and v_{concave} according to σ. Assign the number $\sqrt{\dfrac{\mu(v_{\text{convex}})}{\mu(v_{\text{concave}})}}$ to this singularity. Then the curvature factor is

$$c(\sigma) = \text{product over all maxima and minima of } \sqrt{\frac{\mu(v_{\text{convex}})}{\mu(v_{\text{concave}})}}.$$

DEFINITION 12. (The planar algebra of a bipartite graph.) We now define the action of a planar tangle T on $\mathcal{P}^\Gamma$. We are given a function $R : \mathcal{D}_T \to$ functions on $\{$loops on $\Gamma\}$ and we have to define a function on loops appropriate to the boundary of T in a multilinear way.

So, given a loop γ appropriate to the boundary, define

$$Z_T(R)(\gamma) = \sum_\sigma \left\{ \prod_{D \in \mathcal{D}_T} R(D)(\ell_\sigma(D)) \right\} c(\sigma)$$

where the sum runs over all σ which induce γ on the boundary of T.

The main thing to note in this definition is that the sum is finite since there are only a finite number of states inducing γ on the boundary, and it defines a bounded function since all the R are bounded and so is the factor $c(\sigma)$.

We leave it as an exercise to show that this definition of Z_T is compatible with the gluing of tangles and the *-structure, where the * of a loop is that loop read backwards; also that the eigenvector property of μ guarantees that contractible closed strings in tangles can be removed with a multiplicative factor of δ; also that this planar algebra structure is isomorphic to that of **[Jon99]**, the only change being in how tangles are isotoped in order to define the factor $c(\sigma)$. The reason for the change is that we are mostly dealing with the *graded* algebra, for which the isotopy we use is the most natural.

Each of the vector spaces $P^\Gamma_{n,\pm}$ is infinite dimensional if Γ is infinite. Moreover $P^\Gamma_{0,\pm}$ are the abelian von Neumann algebras $\ell^\infty(\Gamma_\pm)$ which act on the $P^\Gamma_{n,\pm}$. (Note that the graded product and the usual product are the same on these subalgebras). The trace tangle when applied to any element of the planar algebra $P^\Gamma_{0,\pm}$ produces an element of $\ell^\infty(\Gamma_\pm)$. We thus get a bilinear conditional expectation $\mathcal{E}$ from $P^\Gamma_{0,\pm}$ (taken with its usual product) onto $\ell^\infty(\Gamma_\pm)$.

The inner product tangles of definition 5 thus become $\ell^\infty(\Gamma_\pm)$-valued inner products, satisfying $\langle a, b\rangle = \mathcal{E}(a^*b)$. It will follow from a representation of the graph planar algebra on a Hilbert space that the conditional expectation $\mathcal{E}$ (and thus the inner product) is non-negative definite.

2.4.1. *Representing the planar algebra of a bipartite graph as loops.* In the next few sections, we shall work out several examples, which make explicit the operations of graded multiplication on P^Γ, and which will be useful in the rest of the paper. All of the facts mentioned below are straightforward consequences of the definition of the graph planar algebra.

We will sometimes use the word "loop" to also mean the planar algebra element given by the delta function on the set of all loops supported on the given loop.

As a matter of convenience, when inserting a loop into an internal disc of a tangle we will line up the edges of the loop with the boundary points of the disc, starting with the one first in clockwise order after $*$. This convention is useful, since given a string meeting the disc in question at a certain boundary point, any state σ which has a nonzero contribution to the sum Z_T of will have to assign the edge of this boundary point to that string.

For an edge e we'll write $s(e)$ for its starting vertex and $t(e)$ for its ending vertex. For a vertex v we'll write $\Gamma_+(v)$ for the set of all edges starting at v (i.e., $\Gamma_+(v) = \{e : s(e) = v\}$), and we'll denote by $\Gamma_-(v)$ the set of all edges that end at v. We'll also use the notation

$$\sigma(e) = \left[\frac{\mu(t(e))}{\mu(s(e))}\right]^{1/2}.$$

Let L_k^+ be the set of all loops of length $2k$ starting at an even vertex, and L_k^- be the set of all loops of length $2k$ starting at an odd vertex.

From now on, fix an integer t and consider the algebra $Gr_t P^\Gamma$ with its graded multiplication $\wedge_t$.

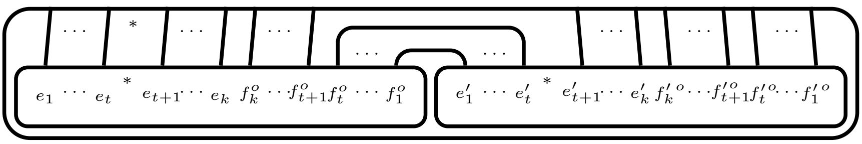
Let $a \in L_k^+$ be a loop,

$$a = e_{t+1} \cdots e_k f_k^o \cdots f_1^o e_1 \cdots e_t$$

where e_j and f_j are edges of Γ. Let

$$b = e'_{t+1} \cdots e'_k {f'_k}^o \cdots {f'_1}^o e'_1 \cdots e'_t$$

Then the graded product $a \wedge_t b$ is given by:



which translates into the following formula:

$$\begin{aligned} a \wedge_t b &= \delta_{s(f_1)=s(e'_1)} \prod_{j=1}^{t} \delta_{f_j = e'_j} \left[\frac{\mu(s(e'_j))}{\mu(t(e'_j))}\right]^{1/2} \cdot \\ &\quad \cdot e_{t+1} \cdots e_k f_k^o \cdots f_{t+1}^o \; e'_{t+1} \cdots e'_{k'} {f'}^o_{k'} \cdots {f'}^o_1 \cdots e_1 \cdots e_t. \end{aligned}$$

Apart from the Perron-Frobenious factors, $a \wedge_t b$ corresponds to a kind of amalgamated concatenation of paths, although the edges of the path should be cyclically permuted. If for a path $a \in L_k^\pm$ (parity according to t) we denote by $D_t(a)$ the path that starts at the $(t+1)$-st edge of a, then we have:

$$D_t(a) \wedge_t D_t(b) = \text{const} \cdot D_t(c)$$

where c is zero if the last t edges of a do not form the inverse of the path formed by the first t segments of b, and is the concatenation of a (with the last t segments removed) and b (with the first t segments removed) otherwise.

In particular, if $t = 0$, given two paths a, b in L_k^+ the graded multiplication $\wedge_0$ is just concatenation of paths (note that in this case D_t is the identity map).

The (usual) trace Tr is given by

$$Tr(e_1 \cdots e_k f_k^o \cdots f_1^o) = \prod_j \delta_{e_j = f_j} \sigma(e_j) s(e)$$

(where again the infinite sums are locally finite).

2.4.2. $TL \subset Gr_0 P^\Gamma$. Let us now set $t = 0$ and identify in terms of paths the element of $TL(k) \subset P^\Gamma_k \subset Gr_0 P^\Gamma$ corresponding to any TL picture. Suppose that we are given a box B with $2k$ boundary points (arranged so that all boundary points are at the top and $*$ is at position 0 from the top-left). Assume also that there are k non-crossing curves inside B which connect pairs of boundary points together. Let π be the associated non-crossing pairing. The associated element of the planar algebra is the function w_B on loops, defined on a loop a by

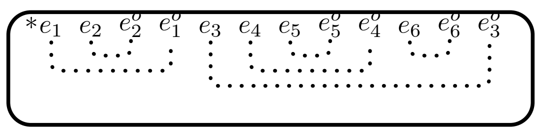
$$w_B(a) = \begin{cases} \sigma(e_1)\cdots\sigma(e_n) & \text{if } e_i = e_j^o \text{ whenever } i \overset{\pi}{\sim} j,\ i \neq j, \\ 0 & \text{otherwise.} \end{cases}$$

If $\pi = \{i_1, j_1\} \cup \cdots \cup \{i_k, j_k\}$ where $i_1 < i_2 < \cdots$ and $i_p < j_p$, then one can think of w_B as the following locally finite sum of delta functions:

$$w_B = \sum_{e_1\cdots e_{2k} \in L_k^+} \left\{ \prod \delta_{e_{i_p} = e^o_{j_p}} \sigma(e_{i_p}) \right\} (e_1 \cdots e_{2k}).$$

For example, the element $w_B \in Gr_0 P^\Gamma$ associated to the non-crossing pairing $\{1,4\},\{2,3\},\{5,12\},\{6,9\},\{7,8\},\{10,11\}$ (thus $k = 6$ and $t = 0$) is presented below:

$$w_B = \sum_{\substack{e_1,\ldots,e_6 : \\ e_1 e_2 e_2^o e_1^o e_3 e_4 e_5 e_5^o e_6 e_6^o e_3^o \in L_5^+}} \sigma(e_1)\cdots\sigma(e_6)$$



(The dotted lines are for illustration purposes only and are not part of the planar diagram). In this way, given a $TL(k)$ element B we get an associated element $w_B \in P^\Gamma_{k,+} \in Gr_k P^\Gamma$. This embedding is the canonical inclusion of the Temperley-Lieb planar algebra into P^Γ.

2.4.3. *The center-valued trace Tr_0 on $Gr_0 P^\Gamma$.* As before, we denote by T_k the element

$$T_k = \sum_{B \in TL(k)} w_B$$

obtained by summing over all $TL(k)$ diagrams.

Let $P_{0,\pm}$ be the zero-box space, i.e., as a linear space it is $\ell^\infty(\Gamma_\pm)$. The algebras $P^\Gamma_{0,\pm}$, when considered with the graded multiplication $\wedge_0$, are abelian, and are in the center of $Gr_0 P^\Gamma$. Recall that $\mathcal{E} : P^\Gamma_n \to P^\Gamma_0$ is a P^Γ_0-bilinear map determined by $\mathcal{E}(ab^*) = \langle a, b \rangle$; one can check that $\mathcal{E}(v) = \mu(v)v$, where v denotes the delta function at $\Gamma_\pm$.

The center-valued trace $Tr_0 : Gr_0 P^\Gamma \to P^\Gamma_0$ is given by the equation

$$Tr_0(x) = \langle x, T_k \rangle = \mathcal{E}(x \cdot T_k), \qquad v \in V^+, x \in P^\Gamma_k.$$

Here as before T_k is the sum of all TL diagrams.

LEMMA 4. *Let $v \in \Gamma$ and let $\phi_v : Gr_0 P^\Gamma \to \mathbb{C}$ be defined by $\phi_v(x)v = Tr_0(x)\wedge_0 v$ (i.e., the value of $Tr_0(x)$, viewed as a function on Γ). Let $x = e_1 \cdots e_{2k} \in L_k^+$ be a loop. Then if x starts at v,*

$$\phi_v(x) = \sum_{\pi \in NCP(2k)} \prod_{\{i,j\} \subset \pi} \sigma(e_i)\delta_{e_i = e_j^o},$$

where the sum is over all non-crossing pairings of $2k$ integers and the product is taken over all tuples $\{i,j\}$, $i<j$ which are paired by π. If x does not start at v, $\phi_v(x)=0$.

Furthermore, ϕ_v is uniquely determined by the recursive formula

$$\phi_v(x) = \sum_{x=ex_1e^ox_2} \sigma(e)\phi_{t(e)}(x_1)\phi_v(x_2)$$

and the formula $\phi_v(ef^o) = \delta_{e=f}\ \delta_{s(e)=v}\sigma(e)$.

We note that although the support of an element $a \in Gr_0P^\Gamma$, viewed as a function on paths, may not be finite, the support of $a \wedge_0 v$ is always finite, since this element is supported on paths of a fixed length starting and ending at v. Thus the value of ϕ_v is well-defined. Moreover, to know the value of ϕ_v, it is sufficient to know its value on elements of Gr_0P^Γ that have finite support.

Proof. Clearly, the recursive formula gives rise to a uniquely defined linear functional on all finitely-supported elements of Gr_0P^Γ (these elements are, of course, viewed as functions on paths). By the comments above, we shall therefore prove the lemma if we prove that both the functional ϕ_v and the functional

$$\phi'_v(x) = \delta_{s(e_1)=v} \sum_{\pi\in NCP(2k)} \prod_{\{i,j\}\subset\pi} \sigma(e_i)\delta_{e_i=e_j^o}$$

satisfy this recursive relation.

Let $\pi \in NCP(2k)$. Then 1 is paired with some integer q. Thus $NCP(2k) = \sqcup_{q>1} NC\{2,\dots,q-1\} \times NC\{q+1,\dots,2k\}$. Thus

$$\begin{aligned}\phi'_v(x) &= \delta_{s(e_1)=v} \sum_{\pi\in NCP(2k)} \prod_{\{i,j\}\subset\pi} \sigma(e_i)\delta_{e_i=e_j^o}\\ &= \sum_{q>1} \sum_{\substack{\pi_1\,\in\,NCP\{1,\dots,q-1\}\\ \pi_2\,\in\,NCP\{q+1,\dots,2l\}}} \delta_{e_1=e_q^o}\sigma(e_1) \prod_{\{i,j\}\subset\pi_1} \sigma(e_i)\delta_{e_i=e_j^o} \prod_{\{i,j\}\subset\pi_2} \sigma(e_i)\delta_{e_i=e_j^o}\\ &= \sum_{x=ex_1e^ox_2} \sigma(e)\phi'_{t(e)}(x_1)\phi'_v(x_2).\end{aligned}$$

Furthermore, $\phi'_v(ef^o)$ is given by the claimed formula. Thus ϕ'_v satisfies the recursive relation.

We now turn to showing that ϕ_v satisfies the same recursive relation. Note that $\phi_v(x)=0$ unless x starts at v.

Note that if $x = e_1\cdots e_{2k}$ and $y = f_1\cdots f_{2k}$ then $\langle x,y\rangle = 0$ unless $x=y^o$ (an opposite of a path is a path with the order of edges and also all edges reversed). Furthermore, if $x=y^o$, then

$$\langle x, x^o\rangle = s(e_1)\prod_{i=1}^{2k}\sigma(e_i)$$

The set TL of all Temperley-Lieb diagrams can be written as a union

$$TL(2k) = \sqcup_q TL\{2,\dots,q-1\}\times TL\{q+1,\dots,2k\}$$

in a manner similar to decomposing the partitions (q denotes the other endpoint of the string ending at 1). Let us assume that x starts at v. Let us denote by

$\overset{\frown}{1B_1q}B_2$ the diagram in which 1 is connected to q and $B_1 \in TL\{2,\ldots,q-1\}$, $B_2 \in TL\{q+1,\ldots,2k\}$. Then

$$\begin{aligned} Tr_0(x) &= \langle x, T_k\rangle = \sum_{B\in TL(2k)} \langle x, w_B\rangle \\ &= \sum_q \sum_{\substack{B_1 \in TL\{1,\ldots,q-1\} \\ B_2 \in TL\{q+1,\ldots,2k\}}} \langle x, w_{\overset{\frown}{1B_1q}B_2}\rangle. \end{aligned}$$

Now, recall that

$$w_B = \sum_{f_1\ldots f_{2k}\in L_k^+} \left\{\prod \delta_{e_{i_p}=e^o_{j_p}}\sigma(f_{i_p})\right\} f_1\cdots f_{2k},$$

so that

$$\begin{aligned} w_{\overset{\frown}{1B_1q}B_2} &= \\ \sum_{f_1\ldots f_{q-1}e\ldots f_{2k-1}e^o\in L_k^+} &\sigma(e) \prod_{\substack{\{i_p,j_p\}\subset B_1 \\ \text{or } \{i_p,j_p\}\subset B_2}} \delta_{f_{i_p}=f^o_{j_p}}\sigma(f_{i_p}) e f_1\cdots f_{q-1}e^o f_{q+1}\cdots f_{2k} \\ &= \sum_e \sigma(e) e w_{B_1} e^o w_{B_2}. \end{aligned}$$

Moreover,

$$\langle x, w_{\overset{\frown}{1B_1q}B_2}\rangle = 0$$

unless x has the form $x = ex_1e^ox_2$ with x_1 a loop having length $q-2$ and e an edge. In this case,

$$\langle x, w_{\overset{\frown}{1B_1q}B_2}\rangle = \langle x_1, w_{B_1}\rangle\langle x_2, w_{B_2}\rangle\sigma(e)\sigma(e^o)\sigma(e) = \langle x_1, w_{B_1}\rangle\langle x_2, w_{B_2}\rangle\sigma(e).$$

Lastly, if $v = s(e)$ then

$$Tr_0(ee^o) = \langle e, e\rangle = \sigma(e)v.$$

It follows that ϕ_v satisfies the same recursive formula as ϕ'_v and, in particular, $\phi_v = \phi'_v$. ☐

2.4.4. *Examples.* Let us denote by $\cup$ the element $\sum_{ee^o\in L^+}\sigma(e)ee^o$. Then

$$\begin{aligned} Tr_0(\cup) &= \textstyle\sum_{e\in E_+} \displaystyle\sum_f \mathcal{E}(ee^o\cdot ff^o)\sigma(e)\sigma(f) \\ &= \sum_e \mathcal{E}(ee^o\cdot ee^o)\left[\frac{\mu(t(e))}{\mu(s(e))}\right] = \sum_e \left[\frac{\mu(t(e))}{\mu(s(e))}\right] s(e) \\ &= \sum_{v\in\Gamma_+} v\frac{1}{\mu(v)}\sum_{s(e)=v}\mu(t(e)) = \sum_{v\in\Gamma_+}\delta v, \end{aligned}$$

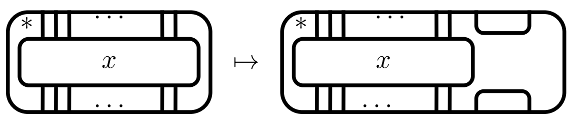
since $\sum_{s(e)=v}\mu(t(e)) = \sum_w \Gamma_{vv}\mu(w) = \delta\mu(v)$.

2.5. Planar subalgebras of P^Γ. It is not the planar algebras P^Γ that are of real interest, but some of their planar subalgebras, in particular those with finite dimensional P_n and 1-dimensional $P_{0,\pm}$ for which the inner product is thus scalar valued and inherits positive definiteness from P^Γ.

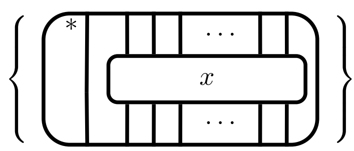
The following theorem, which follows from Popa's work on the theory of λ-lattices (see e.g. Theorem 2.9 (4) in [**PS03b**]), shows that any subfactor planar algebra is a sub-planar algebra of a planar algebra of a discrete bipartite graph.

THEOREM 2. *Let P be an (extremal) subfactor planar algebra, realized as the λ-lattice A_{ij} with principal graph Γ and associated Perron-Frobenius eigenvector μ. Let $\mathcal{A}_i^j$ be as in Theorem 2.9(4) in* [**PS03b**]. *Then:*
(a) The graph planar algebra P^Γ is the planar algebra of the inclusion $\mathcal{A}_{-1}^{-1} \subset \mathcal{A}_0^{-1}$; in other words, $(\mathcal{A}_{-1}^{-1})' \cap \mathcal{A}_k^{-1} = \mathcal{P}_k^\Gamma$;
(b) The isomorphism $P_{j,+} = A_{-1j} \cong (\mathcal{A}_{-1}^{0})' \cap \mathcal{A}_j^{-1} \subset (\mathcal{A}_{-1}^{-1})' \cap \mathcal{A}_j^{-1}$ gives rise to a planar algebra inclusion of P into P^Γ.

The algebras $\mathcal{A}_{-1}^{-1}$ and $\mathcal{A}_0^{-1}$ were constructed in [**PS03b**] as certain non-unital inductive limits of the algebra A_{ij}. Pictorially, this construction corresponds to e.g. taking $\mathcal{A}_0^{-1}$ to be the inductive limit of the algebras $\{P_k : k \text{ even}\}$ using the non-unital inclusion given by the following picture (the region containing $*$ is unshaded):

$P_k \ni$  $\in P_{k+2}$

The algebra $\mathcal{A}_{-1}^{-1}$ then consists of all diagrams having a vertical through-string on the left (again, region containing $*$ is unshaded):

$\mathcal{A}_{-1}^{-1} =$ 

3. A random matrix model for Tr_0

3.1. Random block matrices associated to the graph. Given Γ as above, let $\Gamma_\pm$ be the sets of its even and odd vertices. Let $A = A^+ \oplus A^-$ where $A^\pm = \ell^\infty(\Gamma_\pm) = P_{0,\pm}^\Gamma$. We shall denote by $E_\pm$ the set of edges of Γ which are positively or negatively oriented (according to the sign $\pm$). We shall make the convention that together with any edge $e \in E_\pm$ there is also its opposite edge $e^o \in E_\mp$.

We endow A with a (semi-finite) trace tr given on the minimal projections of A by the formula

$$tr(\delta_v) = \mu(v), \qquad v \in \Gamma.$$

Let N, M be integers. For each choice of M choose integers $\{M_v : v \in \Gamma f\}$ with the property that for each fixed vertex v, $M_v/M \to \mu(v)$ as $M \to \infty$.

In the foregoing, we will consider (infinite if the graph Γ is infinite) matrices whose entries are indexed by the set $\sqcup_{v\in\Gamma}\{1,\ldots,N\} \times \{1,\ldots,M_v\}$. Such an entry will be denoted $A_{ij\ mn\ vw}$, where $i, j \in \{1,\ldots,N\}$, $m \in \{1,\ldots,M_v\}$,

$n \in \{1, \ldots, M_w\}$ and $v, w \in \Gamma$. Given such a matrix $A = (A_{ij\ mn\ vw})$, we compute its trace as follows:

$$tr(A) = \sum_v \frac{1}{N} \sum_{1 \le i \le N} \frac{1}{M} \sum_{1 \le n \le M_v} A_{ii\ nn\ vv}.$$

Our matrices will be such that $A_{ij\ mn\ vw} = 0$ unless v, w belong to a finite set, so that the sum above is finite.

For $v \in \Gamma$ consider the diagonal matrix d_v given by

$$(d_v)_{ij\ mn\ uw} = \delta_{i=j}\delta_{m=n}\delta_{u=w=v}.$$

Note that the joint law of $(d_v : v \in \Gamma)$ converges as $M \to \infty$ to the joint law of $(\delta_v : v \in \Gamma)$.

Consider then for a positively oriented edge $e \in E_+$ from v to w the $NM_v \times NM_w$ matrix X_e defined as follows. The entry $X^e_{ij\ mn\ tu}$ is zero unless $t = v$ and $u = w$. Otherwise, $X^e_{ij\ mn\ vw}$ is (up to scaling) a random Gaussian matrix; in other words, the entries form a family of independent complex Gaussian random variables, each of variance $(\mu(s(e))\mu(t(e)))^{-1/2}(NM)^{-1}$. We shall moreover choose the matrices X_e in such a way that the entries of matrices corresponding to different positively oriented edges are independent. Thus the variables

$$\{X^e_{ij\ mn\ vw} : e \in E_+, v = s(e), w = t(e), 1 \le i, j \le N, 1 \le m \le M_v, 1 \le n \le M_w\}$$

are assumed to be independent.

For a negatively oriented edge f, set $X_f = X^*_{e^o}$. For a loop $w \in L^{\pm}_k, w = e_1 \cdots e_{2k}$, set $X_w = X_{e_1} \cdots X_{e_{2k}}$. Note that $w \mapsto X_w$ is a homomorphism from the algebra $(P^\Gamma, \wedge_0)$ to the algebra of random matrices.

3.2. Tr_0 via random matrices.

Proposition 2. *Let E denote the expected value of a random variable. Then the matrices X_e satisfy: (a) $d_v X_e d_w = \delta_{v=s(e)}\delta_{w=t(e)}X_e$; (b) $E(tr(X_e^* X_e)) = E(tr(X_e X_e^*))$ is independent of N and converges to $(\mu(s(e))\mu(t(e)))^{1/2}$ as $M \to \infty$; (c) For any $v \in V$, $w \in L^{\pm}_k$, $\lim_{M\to\infty} \lim_{N\to\infty} E(tr(d_v X_w)) = tr(\delta_v)Tr_0(w)(v)$ (here $Tr_0(w)(v)$ means the value of the function $Tr_0(w) \in \ell^\infty(\Gamma)$ at $v \in \Gamma$).*

Proof. (a) and (b) are both straightforward; note that

$$E(tr(X_e X_e^*)) = \frac{1}{(\mu(s(e))\mu(t(e)))^{1/2}} \frac{M_v M_w}{M^2} \to (\mu(s(e))\mu(t(e)))^{1/2}.$$

To see (c), we first note that if $w = ee^o$ then $E(tr(X_w)) \to (\mu(s(e))\mu(t(e))^{1/2}$ as $N \to \infty$ and then $M \to \infty$. On the other hand, $tr(Tr_0(w)) = \sigma(e)tr(w) = \sigma(e)\mu(s(e)) = (\mu(s(e))\mu(t(e)))^{1/2}$.

Denote by $\mathcal{E}$ the conditional expectation onto the algebra A. Then we have that if $v = s(e)$, $\mathcal{E}(X_e X_e^*)$ is a multiple of d_v. Since $\mathcal{E}$ is tr-preserving, we have

$$E(\mathcal{E}(X_e X_e^*)) = tr(v)^{-1} tr(\mathcal{E}(X_e X_e^*))\delta_v = \sigma(e) d_v.$$

In particular, we see that

$$E(\mathcal{E}(X_e X_f^*)) = \delta_{e=f}\ \delta_v\ \sigma(e).$$

It is known (see e.g. [**BG05, Shl96**]) that the variables $\{X_e : e \in \Gamma\}$ converge in distribution (jointly also with elements of A) to a family of A-valued semicircular variables with variance

$$\theta_e : \delta_w \mapsto \delta_{w=v}\delta_v\ \sigma(e).$$

Hence if $w = e_1 \cdots e_{2k}$, then for any $a \in A$,

$$\lim_{N\to\infty} \lim_{M\to\infty} tr(d_v(E(\mathcal{E}(X_w))) = tr\left(\delta_v \sum_{\pi\in NC(2k)} \prod_{\{i,j\}\subset\pi} \delta_{e_i=f_i}\sigma(e)\right).$$

By Lemma 4, we see that

$$tr(E(\mathcal{E}(X_w))d_v) \to tr(\delta_v)Tr_0(w)(v), \qquad \forall v \in V,$$

as claimed. ☐

Since the trace tr is positive and faithful on A we conclude that the center-valued trace Tr_0 is non-negative:

COROLLARY 1. *$Tr_0(x^* \wedge_0 x) \geq 0$ if $x \in P^\Gamma$.*

3.3. Another construction of random block matrices. Recall that a bipartite graph can be used as a Bratteli diagram to describe an inclusion of two algebras.

Let $B \subset C$ be an inclusion of multi-matrix algebras corresponding to the graph Γ. This means $B = \bigoplus_{v\in V_+} M_{k(v)\times k(v)}$ and $C = \bigoplus_{w\in V_-} M_{l(w)\times l(w)}$. In particular, each $v \in \Gamma_+$ corresponds to a central projection p_v in B (the unit of the v-th direct summand), and each $w \in \Gamma_-$ corresponds to a central projection $q_w \in C$. The inclusion $B \subset M$ is such that $p_v q_w = 0$ if there is no edge between v and w. If there are r edges between v and w, then $M_{k(v)\times k(v)} = p_v B p_v$ is included into $q_w C q_w = M_{l(w)\times l(w)}$ with index r. In particular, this means that $l(w) = rk(v)$ and also that we can choose r orthogonal projections $\{P^e\}_{s(e)=v,t(e)=w}$ in $q_w C q_w$ with the property that $P^e q_w C q_w P^e \overset{\phi_e}{\cong} M_{k(v)\times k(v)}$ and the inclusion of $p_v B p_v$ into $q_w C q_w$ is given by $x \mapsto \sum_{s(e)=v,t(e)=w} P^e \phi_j(x) P^e$. Choose also isometries $V_{e,f}$ so that $P^e V_{e,f} = V_{e,f} P^f$.

Let Tr be the semi-finite trace on $B \oplus C$ determined by the requirement that $Tr(p_v) = \mu(v)$, $Tr(q_w) = \mu(w)$.

Let Y be a semicircular element, free from $B \oplus C$ (this only makes sense in the case that $Tr(1) < \infty$; more precisely, we shall consider a large projection Q in the center of $B \oplus C$ and consider an element Y free from $Q(B \oplus C)Q$ with respect to $Tr(Q)^{-1}Tr(\cdot)$; our computations will not depend on Q once it is large enough).

To a positive edge e, we associate: (i) a central projection $p_{s(e)} \in B$; (ii) a projection $P^e \in q_{t(e)} C q_{t(e)} \subset C$.

Let $Y_e = (\mu(t(e))\mu(s(e)))^{-1/4} \sum_{s(f)=s(e),t(f)=t(e)} (p_{s(e)} Y P_e) V_{e,f}$ if $e \in E_+$ and $Y_e = Y_{e^o}^*$ if $e \in E_-$.

Note that $Y_e Y_f = 0$ unless $t(e) = s(f)$. We can think of Y_e as a limit of a "$\mu(s(e)) \times \mu(t(e))$" random block matrix, since its left and right support projections, p_v and q_w, have traces $\mu(s(e))$ and $\mu(t(e))$. In fact, one can model Y by a suitable GUE random matrix in the limit when its size goes to infinity, in which case the variables Y_e are indeed approximated in law by random blocks as their sizes go to infinity.

Furthermore, if $e \in E_+$,

$$\begin{aligned} Tr(p_v Y_e q_w Y_f^*) &= (\mu(t(e))\mu(s(e)))^{-1/2} Tr(p_v p_{s(e)}) Tr(\sum_{e',f'} V_{e,e'} q_w V_{f',f}) \\ &= (\mu(t(e))\mu(s(e)))^{-1/2} \delta_{v=s(e)} Tr(p_v) \delta_{w=t(e)} \delta_{v=f} Tr(q_w). \end{aligned}$$

Thus also $Tr(q_w Y_{f^o} p_v Y_{e^o}^*) = (\mu(t(e))\mu(s(e)))^{-1/2}\delta_{v=s(e)}Tr(p_v)\delta_{w=t(e)}\delta_{v=f}Tr(q_w)$. It follows that if we denote by $\mathcal{E}$ the conditional expectation onto the center of $B \oplus C$, then, keeping in mind that $Tr(\delta_v) = \mu(v)$,

$$\mathcal{E}(Y_e \delta_v Y_f) = (\mu(t(e))\mu(s(e)))^{-1/2}\delta_{e=f^o}\delta_{v=t(e)}\delta_{s(e)}\mu(t(v)) \;= \delta_{e=f^o}\delta_{v=t(e)}\delta_{s(e)}\sigma(e).$$

Thus the variables $\{Y_e : e \in E\}$ have the same joint $*$-distribution as the variables $\{X_e : e \in E\}$ that we constructed in the previous section.

4. The Fock Space Model

4.1. A Hilbert bimodule associated to a bipartite graph. Let Γ be a bipartite graph, as before. Consider the complex vector space H with basis given by the (oriented) edges E of the graph; we denote, as before, by E_+ the set of positively oriented edges. Then H is equipped with a natural conjugation which takes an edge to its opposite, $e \mapsto e^o$. The inner product on H is determined by requiring that $\langle e, f\rangle = 0$ unless $e = f$ and

$$\|e\|^2 = \left[\frac{\mu(s(e))}{\mu(t(e))}\right]^{1/2}.$$

(Note that $e \mapsto e^o$ is *not* isometric). As before, we shall use the notation

$$\sigma(e) = \left[\frac{\mu(t(e))}{\mu(s(e))}\right]^{1/2} = \|e\|^{-2}.$$

Let A denote the abelian algebra $A = \ell^\infty(\Gamma)$, where as before Γ denotes the set of vertices of Γ. Then H is naturally an A, A-bimodule: given e an edge in E, define

$$v \cdot e \cdot v' = \delta_{v=s(e)}\delta_{v'=t(e)}e.$$

Moreover, H has a natural A-valued inner product:

$$\langle e, f\rangle_A = \langle e, f\rangle s(e) = \langle e, f\rangle t(f^o).$$

4.2. The operators $c(e)$, the weight ϕ, and the A-valued conditional expectation E. We now consider the Fock space [**Pim97**]

$$\mathcal{F} = A \oplus \bigoplus_{k\geq 0} H^{\otimes_A k}$$

(here $\otimes_A$ denotes the relative bimodule tensor product). For $e \in E$ we consider the operator

$$\ell(e) : \mathcal{F} \to \mathcal{F}, \qquad \ell(e)\xi = e \otimes \xi.$$

Its adjoint is given by

$$\ell(e)^*(e_1 \otimes \cdots \otimes e_n) = \langle e, e_1\rangle_A e_2 \otimes \cdots \otimes e_n.$$

Note that the norm of this operator is given by

$$\|\ell(e)\| = \|\ell^*(e)\ell(e)\|^{1/2} = \|e\|.$$

Let also

$$c(e) = \ell(e) + \ell(e^o)^*.$$

Note that $c(e)^* = c(e^o)$.

Let $B(\mathcal{F})$ be the algebra of bounded adjointable operators on $\mathcal{F}$ and let $E : B(\mathcal{F}) \to A$ be the natural conditional expectation given by

$$E(X) = \langle 1_A, X1_A\rangle_A. \tag{1}$$

Each vertex $v \in \Gamma$ determines a state on $B(\mathcal{F})$ given by

$$\phi_v(X) = \delta_v \circ E(X),$$

where $\delta_v : A \to \mathbb{C}$ is the point evaluation at v. Then

$$\phi = \sum_v \phi_v$$

is a weight on $B(\mathcal{F})$. Note that ϕ is finite on all finite words in $c(e) : e \in E$, and therefore defines a semifinite weight on the von Neumann algebra $W^*(c(e) : e \in E)$ (in the GNS representation associated to ϕ).

LEMMA 5. *(i) The weight ϕ and the conditional expectation E are faithful on this algebra.*

(ii) The modular group of ϕ is determined by $\sigma_t^\phi(c(e)) = \left[\frac{\mu(t(e))}{\mu(s(e))}\right]^{it} c(e) = \sigma(e)^{2it}c(e)$.

(iii) Consider $\cup = \bigoplus_{v \text{ even}} \sum_{e\in\Gamma_+(v)} \sigma(e)c(e)c(e^o)$, where $\Gamma_+(v)$ denotes the set of all edges that start at v. Then for each v, the law of $\cup$ with respect to ϕ_v has no atoms and is the free Poisson law with R-transform $\delta(1-z)^{-1}$. In particular, $v\cup v$ is bounded for all v and thus the (possibly infinite) direct sum defining $\cup$ yields a bounded operator.

PROOF. The GNS vector space $\mathcal{F}_v$ associated to the state ϕ_v on the algebra $W^*(c(e) : e \in E)$ can be identified with the subspace of the Fock space $F(H) = \mathbb{C}v \oplus \bigoplus_{k\geq 1} H^{\otimes k}$ spanned by tensors of the form $e_1 \otimes \cdots \otimes e_n$, $e_j \in E$ for which $e_1 \cdots e_n$ form a path (i.e., are "composable": $s(e_j) = t(e_{j+1})$) and so that e_n starts at v. The identification takes the tensor $e_1 \otimes_A \cdots \otimes_A e_k$ to $e_1 \otimes \cdots \otimes e_k \in F(H)$. If we denote by $\hat{\ell}(e) : F(H) \to F(H)$ the operator $\hat{\ell}(e)\xi = e \otimes \xi$ and by $\hat{c}(e)$ the operator $\hat{c}(e) = \hat{\ell}(e) + \hat{\ell}(e^o)^*$, then we have

$$P\hat{c}(e)P = c(e), \qquad P : F(H) \to \mathcal{F}_v \text{ orthogonal projection.}$$

Let P be the set of all paths in Γ and $P(v)$ be the set of all paths starting at v. For a path $w = e_1 \cdots e_n \in P(v)$, let $c(w) = c(e_1) \cdots c(e_n)$ and similarly for $\hat{c}$. We then see that the joint laws associated to the vacuum expectation state of the variables

$$\{c(w) : w \in P(v)\} \text{ and } \{\hat{c}(w) : w \in P(v)\}$$

have the same law. Indeed, $\hat{c}(w)v = c(w)v$ if $w \in P_v$.

It follows that the von Neumann algebra generated by $(A, c(e) : e \in E)$ in the GNS representation π_v associated to ϕ_v can be embedded into the von Neumann algebra $W^*(\hat{c}(e) : e \in E)$ in such a way that the restriction of the state $\hat{\phi}_v = \langle v, \cdot v\rangle$ to the former algebra is exactly ϕ_v. But it is known **[Shl97]** that $\hat{\phi}_v$ is faithful, and so ϕ_v is faithful (on the image in the GNS construction π_v). Furthermore, the modular group of $\hat{\phi}_v$ is given by

$$\sigma_t^{\hat{\phi}_v}(\hat{c}(e)) = \left[\frac{\mu(t(e))}{\mu(s(e))}\right]^{it} \hat{c}(e).$$

It follows that

$$\sigma_t^{\phi_v}(\pi_v(c(e))) = \left[\frac{\mu(t(e))}{\mu(s(e))}\right]^{it} \pi_v(c(e)).$$

It is clear that the GNS vector space for the weight ϕ on $W^*(A, c(e) : e \in E)$ is just the direct sum of the GNS vector spaces for ϕ_v taken over all vertices v. Thus ϕ

is faithful and so E is faithful (on the possibly larger algebra $W^*(A, c(e) : e \in E)$). Thus (i) holds.

Let now

$$Y = \sum_{e \in \Gamma_+(v)} \sigma(e)\hat{c}(e)\hat{c}(e^o).$$

Then Y has the same law for $\hat{\phi}_v$ as does $\cup$ for ϕ_v. Note that $Y = \sum_{e \in \Gamma_+(v)} b(e)$, with $b(e) = \sigma(e)\hat{c}(e)\hat{c}(e^o)$. Thus $b(e)$ are free and so the law of Y satisfies

$$\mu_Y = \boxplus_{e \in \Gamma_+(v)} \mu_{b(e)}.$$

Now, for each e, $b(e)^{1/2}$ has the free Poisson distribution with R-transform $(\mu(t(e))/\mu(s(e))) \cdot (1-z)^{-1}$ (see [**Shl97**, Remark 4.4 on p. 347]). Thus the law of $b(e)$ has only one atom of mass $\alpha(e) = 1 - \mu(t(e))/\mu(s(e))$ at zero (this expression is to be interpreted as zero if it is negative). It follows from additivity of the R-transform [**VDN92**] that the law of Y is free Poisson with R-transform

$$(1-z)^{-1} \sum_{e \in \Gamma_+(v)} \frac{\mu(t(e))}{\mu(v)} = \frac{\delta}{1-z},$$

which will have an atom iff $\delta < 1$. Since $\delta \geq 1$, the law of Y has no atoms. Thus (iii) holds.

Finally, it is also clear that (ii) holds since a similar formula holds in the GNS representation of each ϕ_v and $\phi = \sum \phi_v$. □

LEMMA 6. *Let L be the set of all loops in Γ. Then the algebra $W^*(c(w) : w \in L)$ belongs to the fixed point of the modular group acting on the von Neumann algebra $W^*(c(w) : w \in P)$.*

PROOF. Since w is a loop, the factors $\mu(t(e))/\mu(s(e))$ associated to each factor in $c(w)$ cancel. □

4.3. The conditional expectation E realizes Tr_0. Let $Y_w = c(w)$, $w \in L_k^+$ (the set of all loops starting at a positive vertex and of length $2k$).

LEMMA 7. *Let $w \in L$ be a loop given by $w = e_1 \cdots e_n$. Then $\phi(Y_w) = \sum_{\pi \in NC(2k)} \prod_{\{i,j\} \subset \pi} \delta_{e_i = e_j^o} \sigma(e_i)$. In particular, $E(Y_w) = Tr_0(w)$.*

PROOF. $E(Y_w) = \sum_{\pi \in NC(2k)} \prod_{\{i,j\} \subset \pi} \delta_{e_i = e_j^o} \cdot E(Y_{e_i} Y_{e_j}) s(e_i)$. Moreover, $E(c(e_i)c(e_i)^*) = s(e_i)\sigma(e_i)$. The rest follows from Lemma 4. □

LEMMA 8. *Let L^+ be the set of loops starting at an even vertex. Consider $\mathcal{M}_0 = W^*(c(w) : w \in L)$ with its semi-finite weight ϕ. Then each even $v \in \Gamma$, defines a central projection in $\mathcal{M}_0$ and*

$$(\mathcal{M}_0, \phi) = \bigoplus_{v \ even} (v\mathcal{M}_0 v, \phi_v).$$

For each v, the algebra $v\mathcal{M}_0 v$ can be canonically embedded into a free group factor.

PROOF. If $w \in L$ is a loop starting at v, then $v'c(w) = c(w)v' = \delta_{v=v'} c(w)$. To see that $v \in \mathcal{M}_0$, note that v is the support projection of the element $v \cup v = \sum_{e \in \Gamma_+(v)} \sigma(e)c(e)c(e^o) \in \mathcal{M}_0$.

We have seen before that $\pi_v(W^*(c(w) : w \in P))$ with its state ϕ_v can be embedded into a free Araki-Woods factor associated to H and taken with its free quasi-free state, in a state-preserving way. The image of $\mathcal{M}_0$ under π_v is precisely

$v\mathcal{M}_0 v$, and this image clearly lies in the centralizer of the free quasi-free state. The free quasi-free state is almost-periodic (the modular group, restricted to $c(H)$ has as its eigenvectors the edges of Γ) and therefore the centralizer is a free group factor. □

Note that $\phi = \bigoplus \phi_v$ is faithful.

THEOREM 3. *Let $w \in L^+$ be a loop on Γ, starting at an even vertex. Then the map $w \mapsto c(w)$ extends to a trace-preserving embedding with dense range of $(Gr_0 P^\Gamma, \wedge_0, Tr_0)$ into $(\mathcal{M}_0, E)$. Thus Tr_0 is a faithful center-valued trace.*

PROOF. Clearly the theorem is true on elements of P_+^Γ that have finite support, i.e., are finite linear combinations of loops.

We have to check that this embedding makes sense for elements of P_+^Γ which, as functions on loops, have infinite support.

Let $w \in P_k^\Gamma$. Then for any $v \in \Gamma_+$, $\delta_v \wedge_0 w = w \wedge_0 \delta_v = \delta_v \wedge_0 w \wedge_0 \delta_v$ has finite support. Moreover, by assumption $\langle w, w^* \rangle \in P_0^\Gamma = \ell^\infty(\Gamma_+)$ has finite ℓ^∞ norm. But the value of $\langle w, w^* \rangle$ at v is exactly $\|c(\delta_v \wedge_0 w \wedge_0 \delta_v)\|^2_{L^2(\phi_v)}$ and is therefore uniformly bounded as a function of v. Moreover, note that each $c(\delta_v \wedge_0 w \wedge_0 \delta_v)$ belongs to the span of words of length $2k$ in operators $c(e) : e \in E_+$.

The eigenvector condition implies that the ratios $\mu(v)/\mu(w)$ for v, w adjacent are bounded, and also that the valence of the graph is bounded.

It follows that the linear dimension of the space of all loops of length k starting at a vertex v is uniformly bounded, by a constant independent of v. Moreover, the norms of the orthogonal basis for this space (consisting of the various loops) are bounded both above and below uniformly in v. Thus the restrictions of the operator norm and the $L^2(\phi)$-norm to the finite-dimensional linear span of loops of length k starting at v are equivalent, and the constants in the equivalence can be chosen to be uniform in v.

It follows that $vc(w)v$ is uniformly bounded in norm (independent of v).

Since the projections $v : v \in \Gamma_+$ are orthogonal, it follows that $c(w)$, defined as the ultraweakly-convergent sum $\sum vc(w)v$, is a bounded operator in $\mathcal{M}_0$.

Since the map $w \mapsto c(w)$ is bilinear over P_0^Γ, and is an algebra homomorphism when restricted to finite linear combinations of loops, it is easy to see that it is an algebra homomorphism on all of $Gr_0 P^\Gamma, \wedge_0$. □

4.4. The operator $\cup$. Let $\Gamma_+(v)$ denote the set of all edges starting at an even vertex v and let E_+ denote the set of all positively oriented edges (i.e., ones that start at an even vertex). Recall that

$$\cup = \sum_{e \in E_+} \sigma(e) c(e) c(e)^*.$$

If we let δ be the Perron-Frobenius eigenvalue, then

$$\begin{aligned}
(\cup) &= \sum_{e \in E_+} \sigma(e) \left(\ell(e)\ell(e^o) + (\ell(e)\ell(e^o))^* + \ell(e)\ell(e)^* + \sigma(e) \right) \\
&= 2 \sum_{e \in E_+} \sigma(e) \Re(\ell(e)\ell(e^o)) + \sum_v \sum_{e \in \Gamma_+(v)} \left[\frac{\mu(t(e))}{\mu(s(e))} \right] v + \sum_{e \in E_+} \sigma(e)\ell(e)\ell(e)^* \\
&= 2 \sum_{e \in \Gamma_+(v)} \sigma(e) \Re(\ell(e)\ell(e^o)) + \delta + \sum_{e \in E_+} \sigma(e)\ell(e)\ell(e)^*.
\end{aligned}$$

Here we used

$$\sum_{e\in\Gamma_+(v)} \mu(t(e)) = \sum_j \Gamma_{vj}\mu_j = \delta\mu(v)$$

so that

$$\sum_{e\in\Gamma_+(v)} \frac{\mu(t(e))}{\mu(s(e))} = \frac{1}{\mu(v)} \sum_{e\in\Gamma_+(v)} \mu(t(e)) = \frac{1}{\mu(v)}\delta\mu(v) = \delta.$$

Let $\mathcal{F}_+$ be the set of all vectors in $\mathcal{F}$ starting and ending in a positive vertex and let $A_+ = A\cap\mathcal{F}_+$. Since if $\zeta\in\mathcal{F}\ominus A_+$

$$\sum_{e\in\Gamma_+(v)} \sigma(e)\ell(e)\ell(e)^*\zeta = \zeta,$$

(because of the normalizations of the lengths of e, e^o we have that the sum $\sum_{e\in\Gamma_+(v)}\sigma(e)\ell(e)\ell(e)^*$ is the same as the sum $\sum_f \ell(f)\ell(f^*)$ where the summation is over an orthonormal basis). Thus

$$\cup|_{\mathcal{F}_+} = 2\sum_{e\in\Gamma_+(v)} \sigma(e)\Re(\ell(e)\ell(e^o)) + \delta + (1-P), \tag{2}$$

where $P:\mathcal{F}_+\to\mathcal{F}_+$ is the projection onto $A_+\subset\mathcal{F}_+$ and δ is the Perron-Frobenius eigenvalue.

As consequence, we note that we can now identify the position of $\mathcal{A}_v = \pi_v(W^*(Y)) = vW^*(Y)v$ inside of $v\mathcal{F}v = L^2(W^*(c(w) : w\in L), \phi_v)$ identified with a subspace of $\mathbb{C}v\oplus\bigoplus_{k\geq 1} H^{\otimes k}$:

Lemma 9. *Let v be an even vertex. Then $L^2(\mathcal{A}_v)$ is the closed linear span of the orthogonal system of vectors*

$$\xi^{\otimes k} = \left(\sum_{e\in\Gamma_+(v)} \sigma(e) e\otimes e^o\right)^{\otimes k}, \qquad k = 0, 1, \ldots.$$

Moreover,

$$\|\xi^{\otimes k}\|_2^2 = \delta^k,$$

where δ is the Perron-Frobenius eigenvalue.

Proof. We note that $Yv = \xi$. Moreover, the linear span of $\xi^{\otimes k}$ is clearly stable under the action of Y. Thus it is sufficient to prove that if $\xi^{\otimes r}$ for $r < k$ are in $L^2(\mathcal{A})$ then also $\xi^{\otimes k}\in L^2(\mathcal{A})$. But this follows from noting that $Y^k v = \xi^{\otimes k} + \zeta$, where ζ is a tensor of smaller degree in $L^2(\mathcal{A})$.

Furthermore,

$$\langle\xi,\xi\rangle = \sum_{e\in\Gamma_+(v)} \left[\frac{\mu(t(e))}{\mu(s(e))}\right] \|e\otimes e^o\|_2^2 = \sum_{e\in\Gamma_+(v)} \frac{\mu(t(e))}{\mu(v)} = \frac{1}{\mu(v)}\sum_j \Gamma_{vj}\mu_j = \delta.$$

∎

4.5. Relative commutant of $\cup$. Recall that the space $L^2(\mathcal{M}_0,\phi)$ admits the decomposition $L^2(\mathcal{M}_0,\phi)=\bigoplus_v L^2(v\mathcal{M}_0 v,\phi_v)=\bigoplus_v v\mathcal{F}v$. Let as before $Y=\cup$, $\mathcal{A}=W^*(Y)$ and $\mathcal{A}_v=v\mathcal{A}v$.

LEMMA 10. *(i) $\mathcal{A}_v$ is a singular MASA in $v\mathcal{M}_0v$.*
(ii) $W^*(\cup)'\cap\mathcal{M}_0=\bigoplus_{v\text{ even}}v\mathcal{A}v$.
(iii) Consider the algebra

$$\mathcal{N}_+=W^*(c(w):w\textit{ a path in }\Gamma\textit{ starting and ending at an even vertex}).$$

Then $\mathcal{A}'\cap\mathcal{N}_+=\bigoplus_{v\in\Gamma_+}v\mathcal{A}v$.

PROOF. We first note that any $v\in V$ commutes with $Y=\cup$. In particular, $v\in\mathcal{A}'\cap\mathcal{N}_+$. Hence $[Y,x]=0$ implies that $v[Y,x]w=[Y,vxw]=0$ for all $v,w\in V$. Hence $\mathcal{A}'\cap\mathcal{N}_+$ is the closure of $\sum_{v,w}(\mathcal{A}'\cap v\mathcal{N}_+w)$.

Consider the full Fock space $F(H)$ as in the proof of Lemma 5, where H is as before a Hilbert space having as basis edges of Γ. Thus $F(H)$ is spanned by all tensors of the form $e_{i_1}\otimes\cdots\otimes e_{i_m}$, where $e_{i_k}\in E$. Let $\tilde H=H\otimes H$, and consider $\tilde{\mathcal{F}}\subset F(H)$ given by $\tilde{\mathcal{F}}=\bigoplus_{k\geq 0}\tilde H^{\otimes k}$. Let $T=W^*(\hat\ell(e)+\hat\ell(e^o)^*:e\in E)$ acting on $F(H)$, and consider the subalgebra $Q\subset T$ given by

$$Q=W^*([\hat\ell(e)+\hat\ell(e^o)^*][\hat\ell(f)+\hat\ell(f^o)^*]:e,f\in E).$$

Then clearly $L^2(Q)\subset L^2(P)$ can be identified with $\tilde F\subset F(H)$. Furthermore, Q is invariant under the modular group associated to ϕ_v (the vector state associated to the vacuum vector in $\tilde{\mathcal{F}}\subset F(H)$). Thus the modular group of Q is the restriction to Q of the modular group of P.

Fix $v,w\in V$. Denote by λ_v the element

$$\sum_{e\in\Gamma_+(v)}\sigma(e)(\hat\ell(e)+\hat\ell(e^o)^*)(\hat\ell(e^o)+\hat\ell(e)^*)\in Q.$$

Denote by ρ_w the element

$$\sum_{e\in\Gamma_+(w)}\sigma(e)(\hat r(e)+\hat r(e^o)^*)(\hat r(e^o)+\hat r(e)^*)\in JQJ$$

(here $\hat r$ denotes the right creation operator and J is the modular conjugation). Note that $\rho_v=J\lambda_vJ$.

We now make the identification $U:L^2(v\mathcal{N}_+w)\hookrightarrow L^2(Q)$ obtained by sending a tensor $e_1\otimes_A\cdots\otimes_A e_{2n}$ associated to a *path* $e_1\cdots e_{2n}$ starting at v and ending at w to the tensor $e_1\otimes\cdots\otimes e_{2n}$. It is not hard to see that

$$\lambda_vU=UY,\qquad\rho_wU=UJYJ.$$

It follows that the laws of ρ_w and λ_v (with respect to the vacuum state on $\tilde{\mathcal{F}}(H)$) are the same as that of Y and have no atoms; thus $W^*(\lambda_v)$ and $W^*(\rho_w)$ are diffuse. In particular, if $\Xi\in L^2(v\mathcal{N}_+w)$ satisfies $Y\Xi=JYJ\Xi$, then $U\Xi$ satisfies $\lambda_vU\Xi=\rho_wU\Xi$.

Consequently, we would prove (iii) if we could show:
(a) if $v\neq w$, $\lambda_v\zeta=\rho_w\zeta$ for $\zeta\in UL^2(v\mathcal{N}_+w)$ only occurs if $\zeta=0$ and
(b) if $v=w$ and $\lambda_v\zeta=\rho_v\zeta$ for some $\zeta\in UL^2(v\mathcal{N}_+v)$, then $\zeta\in UL^2(v\mathcal{A}v)$.
Let $\xi_v=\sum_{e\in\Gamma_+(v)}\sigma(e)e\otimes e^o$.

Assume first that $u=v$. Let $K_v=\tilde H\ominus\mathbb{C}\xi_v$. Put $\mathcal{H}_v=\mathbb{C}\Omega\oplus\mathbb{C}\xi_v\oplus\mathbb{C}\xi_v^{\otimes 2}\oplus\cdots$. Then $\mathcal{H}_v=UL^2(v\mathcal{A}v)$ in such a way that the left and right multiplication by Y

on $L^2(v\mathcal{A}v)$ correspond to the actions of λ_v and ρ_v. In particular, $\mathcal{H}_v$ is invariant under both λ_v and ρ_v.

The image of $L^2(v\mathcal{N}_+v)$ under U lies in the closure of the direct sum

$$\mathcal{H}_v \oplus (\mathcal{H}_v \otimes K_v \otimes \mathcal{H}_v) \oplus (\mathcal{H}_v \otimes K_v \otimes \mathcal{H}_v \otimes K_v \otimes \mathcal{H}_v) \oplus \cdots .$$

(This direct sum is identified with a subspace $\tilde{\mathcal{F}}$ by identifying $\Omega \otimes \zeta$ and $\zeta \otimes \Omega$ with ζ if $\zeta \in F(H)$). Each direct summand in this sum is invariant under both ρ_v and λ_v since their actions respect the tensor product decompositions $\mathcal{H}_v \otimes K_v \otimes \cdots \otimes K_v \otimes \mathcal{H}_v$: ρ_v acts as $\mathrm{id} \otimes \rho_v|_{\mathcal{H}_v}$ and λ_v acts as $\lambda_v|_{\mathcal{H}_v} \otimes \mathrm{id}$.

Now, for any choice of orthonormal basis for $K_v \otimes \mathcal{H}_v \otimes \cdots \otimes K_v$, ζ_α, we have for all $h, g \in \mathcal{H}_v$:

$$\langle h \otimes \zeta_\alpha \otimes g, h' \otimes \zeta_{\alpha'} \otimes g' \rangle = \delta_{\alpha=\alpha'} \langle h, h' \rangle \langle g, g' \rangle$$

and consequently $\mathcal{H}_v \otimes K_v \otimes \cdots \otimes \mathcal{H}_v$ is isomorphic to an (infinite) multiple of $\mathcal{H}_v \otimes \mathcal{H}_v$ as a bimodule over $W^*(\lambda_v)$ acting on the left copy of $\mathcal{H}_v$ and $W^*(\rho_v) = JW^*(\lambda_v)J$ acting on the right copy of $\mathcal{H}_v$. Since the spectral measure of λ_v is non-atomic, it follows that there can be no vector Ξ contained in $\mathcal{H}_v \otimes K_v \otimes \cdots \otimes \mathcal{H}_v$ satisfying $\lambda_v \Xi = \rho_v \Xi$, since such a vector would give rise (via an isomorphism of $\mathcal{H}_v \otimes \mathcal{H}_v$ with Hilbert-Schmidt operators on this space) to a Hilbert-Schmidt operator on $\mathcal{H}_v$, commuting with λ_v.

Thus the only possible Ξ satisfying $\lambda_v \Xi = \rho_v \Xi$ and lying in the image of $UL^2(v\mathcal{N}_+v)$ must be contained in $\mathcal{H}_v = UL^2(v\mathcal{A}v)$. Thus we have proved (b).

To prove (a), we note that if $v \neq w$, and we let $K_{v,w} = \tilde{H} \ominus (\mathbb{C}\xi_v \oplus \mathbb{C}\xi_w)$, the image of $L^2(v\mathcal{N}_+w)$ lies in

$$\mathcal{H}_v \otimes \left(\bigoplus_{k \geq 0} \bigoplus_{u_1,\ldots,u_k \in \{v,w\}} K_{v,w} \otimes \mathcal{H}_{u_1} \otimes K_{v,w} \otimes \mathcal{H}_{u_2} \otimes \cdots \otimes K_{v_w} \right) \otimes \mathcal{H}_w$$

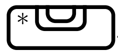
(once again identified with a subspace of $\tilde{\mathcal{F}}$ as before), which is isomorphic to an infinite multiple of $\mathcal{H}_v \otimes \mathcal{H}_w$ as a bimodule over $W^*(\lambda_v)$ acting on the left copy of $\mathcal{H}_v$ and $W^*(\lambda_w)$ acting on the right copy of $\mathcal{H}_w$. Once again, we see that there can be no vector Ξ satisfying $\lambda_v \Xi = \rho_w \Xi$ in this space. Thus (a) is also proved. Thus we have proved (iii).

Note that we have actually proved that $L^2(v\mathcal{M}_0 v, \phi_v)$ when viewed as a bimodule over $\mathcal{A}_v = W^*(vYv)$ is the direct sum of $L^2(\mathcal{A}_v)$ and an (infinite) multiple of the coarse $\mathcal{A}_v, \mathcal{A}_v$-bimodule $L^2(\mathcal{A}_v) \bar{\otimes} L^2(\mathcal{A}_v)$. Recall (see e.g. [**PS03a**] or [**FM77**]) that the normalizer of $\mathcal{A}_v$ is contained in its quasi-normalizer $\mathcal{NQ}(\mathcal{A}_v)$, which consists of those elements ζ in $v\mathcal{M}_0 v$ for which the associated bimodule $\overline{\mathcal{A}_v \zeta \mathcal{A}_v}^{L^2}$ is "discrete". This bimodule cannot be discrete if it contains a sub-bimodule isomorphic to a compression of the coarse $\mathcal{A}_v, \mathcal{A}_v$-bimodule. Thus the only $\zeta \in \mathcal{QN}(\mathcal{A}_v)$ must lie in $L^2(\mathcal{A}_v)$ and thus in $\mathcal{A}_v$. It follows that the normalizer of $\mathcal{A}_v$ in $v\mathcal{M}_0 v$ is contained in $\mathcal{A}_v$. Thus $\mathcal{A}_v$ is a singular MASA and so (i) follows. Now (ii) easily follows from (i). $\square$

4.6. The operator $\Cup$, relative commutant of $W^*(\cup, \Cup)$ and factoriality. We now consider the following sum

$$\Cup = \sum_{eff^o e^o \in L^+} \left[\frac{\mu(t(f))}{\mu(s(e))} \right]^{1/2} c(e)c(f)c(f)^* c(e)^*$$

taken over all loops that start at an even vertex. The pictorial representation of this planar algebra element is:

$\Cup =$  .

LEMMA 11. *Let v be a fixed even vertex. Assume that there is a path of length 2 from v to v not of the form $ee^o ff^o$. Then algebra $vW^*(\cup, \Cup)v$ has a trivial relative commutant inside of the algebra $v\mathcal{N}_+v$, where $\mathcal{N}_+ = W^*(c(w)$: w a path in Γ starting and ending at an even vertex).*

PROOF. Because of Lemma 10, we know that the relative commutant of $vW^*(\cup, \Cup)v$ inside of $v\mathcal{N}_+v$ is contained in $vW^*(\cup)v = \mathcal{A}_v$.

Let $\eta = e_1 \otimes f_1 \otimes f_1^o \otimes e_1^o$ where $f_1 \neq e_1^o$ and $e_1 f_1 f_1^o e_1^o$ is a path from v to v.

Set

$$Z = v \Cup v = \sum_{e, f^o} \sigma(e)\sigma(f)c(e)c(f)c(f^o)c(e^o),$$

where the sum is over all paths eff^oe^o from v to v. Then if $k, l > 0$, and $\xi \in L^2(vW^*(\cup)v) = L^2(\mathcal{A}_v)$ is as in Lemma 9, we have:

$$\begin{aligned}
\langle \eta \otimes \xi^{\otimes k}, [Z, \xi^{\otimes l}]\rangle &= \Big\langle \eta \otimes \xi^{\otimes k}, \sum_{e,f} \Big\{ \sigma(e)\sigma(f)c(e)c(f)c(f^o)c(e^o)\xi^{\otimes l} \\
&\qquad\qquad - \sigma(e)\sigma(f)\xi^{\otimes l}c(e)c(f)c(f^o)c(e^o)\Big\}\Big\rangle \\
&= \Big\langle \eta \otimes \xi^{\otimes k}, \sum_{e,f} \sigma(e)\sigma(f) e \otimes f \otimes f^o \otimes e^o \otimes \xi^{\otimes l} \\
&\qquad + \sum_{e,f} \sigma(e)\sigma(f) \frac{\mu(t(e))^{1/2}}{\mu(v)^{1/2}} e \otimes f \otimes f^o \otimes \xi^{\otimes(l-1)} \Big\rangle.
\end{aligned}$$

Thus

$$\langle \eta \otimes \xi^{\otimes k}, [Z, \xi^{\otimes l}]\rangle = \begin{cases} \frac{\mu(t(f_1))^{1/2}}{\mu(v)^{1/2}}, & l = k \\ \frac{\mu(t(f_1))^{1/2}}{\mu(v)^{1/2}} \cdot \frac{\mu(t(e_1))^{1/2}}{\mu(v)^{1/2}}, & l = k+1 \\ 0, & \text{otherwise.} \end{cases}$$

Thus if we consider

$$a = \sum \alpha_k \xi^{\otimes k} \in L^2(\mathcal{A}_v)$$

and assume that $[a, Z] = 0$ and $a \perp \mathbb{C}v$ (so that $\alpha_0 = 0$), we obtain

$$\begin{aligned}
0 &= \langle \eta \otimes \xi^{\otimes k}, [Z, a]\rangle \\
&= \frac{\mu(t(f_1))^{1/2}}{\mu(v)^{1/2}} \left(\alpha_k + \frac{\mu(t(e_1))^{1/2}}{\mu(v)^{1/2}} \alpha_{k+1} \right), \qquad k \geq 1.
\end{aligned}$$

Since the choice of e_1 was arbitrary, we find that

$$\alpha_k = -\frac{\mu(t(e))^{1/2}}{\mu(v)^{1/2}} \alpha_{k+1}, \qquad \forall e \in \Gamma_+(v).$$

If $a \neq 0$, not all α_k are zero; from this recursive relation we deduce that $\mu(t(e))$ are all equal to the same value, μ', independent of $e \in \Gamma_+(v)$ and that (after rescaling a by a non-zero constant) we may assume that $\alpha_{k+1} = (-1)^k \lambda^{-(k+1)}$ where $\lambda = (\mu(t(e))/\mu(v))^{1/2} = (\mu'/\mu(v))^{1/2}$.

On the other hand,

$$\sum \Gamma_{vj}\mu' = \delta\mu(v)$$

so that

$$(\sum \Gamma_{vj})\mu'/\mu(v) = (\sum \Gamma_{vj})\lambda^2 = \delta.$$

Thus if $N \geq 1$ is the valence of Γ at v, we find that $N\lambda^2 = \delta$, so that $\lambda^2 = \delta/N$.

Using the fact that $\|\xi\|_2^2 = \delta$, we compute:

$$\|a\|_2^2 = \sum_k |\alpha_k|^2 \|\xi^{\otimes k}\|_2^2 = \sum_k \lambda^{-2k}\delta^k = \sum_k (N/\delta)^k \delta^k = \sum_k N^k = \infty,$$

which is impossible. Thus $[Z, a] = 0$ forces $a \in \mathbb{C}v$. $\square$

LEMMA 12. *Let Γ be a connected bipartite graph with $N+1$ vertices, $v \in \Gamma$ even and assume that the hypothesis of Lemma 11 is not satisfied. Then the remaining vertices $e_1, \ldots, e_N$ of Γ are all connected to v by a single edge, and Γ has no other edges.*

We can now prove that the relative commutant of $W^*(\cup, \Cup)$ can be controlled, if the graph Γ is not too small. The cases we exclude are A_1 (a graph with a single vertex and no edges) and A_2 (a graph with exactly two vertices connected by a single edge). In these cases, $\cup$ and $\Cup$ commute (in fact, they are equal). In either of these cases, the Perron-Frobenius eigenvalue is 1, which is of little interest to us.

THEOREM 4. *Assume that $\Gamma \neq A_2$ and $\Gamma \neq A_1$ and let $v \in \Gamma$ even. Then (i) the relative commutant $(vW^*(\cup, \Cup)v)' \cap v\mathcal{M}_0 v$ is trivial. In particular, the center of $\mathcal{M}_0$ is the algebra $A_+ = \ell^\infty$(even vertices). (ii) $W^*(\cup, \Cup)' \cap \mathcal{N}_+ = P_0^\Gamma$ (where $\mathcal{N}_+ = W^*(c(w)$: w a path that starts and ends at an even vertex), and $P_0^\Gamma = \bigoplus_{v \text{ even}} v\mathbb{C}$).*

PROOF. Because of Lemma 11 and Lemma 12, it remains to consider the case in which Γ is a graph with $N+1 > 3$ vertices $v, e_1, \ldots, e_N$ with a single edge between v and each e_j and no other edges. Since $\Gamma = [1 \ \cdots \ 1]$, we find that $\|\Gamma\| = N$ and therefore $\delta = N$. Moreover, one can normalize the Perron-Frobenius eigenvector to be $\mu(e) = 1$ for all $e \in \{v, e_1, \ldots, e_n\}$.

Thus $\xi = \sum_j e_j \otimes e_j^o$, $Z = v \Cup v = \sum_i c(e_i)c(e_i^o)c(e_i)c(e_i^o)$.

Let $k > 1$. Then

$$\begin{aligned}
[Z, \xi^{\otimes k}] &= \sum_i c(e_i)c(e_i^o)c(e_i)c(e_i^o)\xi^{\otimes k} - \xi^{\otimes k}c(e_i)c(e_i^o)c(e_i)c(e_i^o) \\
&= \xi^{\otimes k-2} + 4\xi^{\otimes k-1} + 6\xi^{\otimes k} + 3\xi^{\otimes k+1} \\
&\quad + \sum_i e_i \otimes e_i^o \otimes e_i \otimes e_i^o \otimes \xi^{\otimes k-1} + \sum_i e_i \otimes e_i^o \otimes e_i \otimes e_i^o \otimes \xi^{\otimes k} \\
&\quad -\xi^{\otimes k-2} - 4\xi^{\otimes k-1} - 6\xi^{\otimes k} - 3\xi^{\otimes k+1} \\
&\quad - \sum_i \xi^{\otimes k-1} \otimes e_i \otimes e_i^o \otimes e_i \otimes e_i^o + \sum_i \xi^{\otimes k} e_i \otimes e_i^o \otimes e_i \otimes e_i^o \\
&= \sum_i e_i \otimes e_i^o \otimes e_i \otimes e_i^o \otimes (\xi^{\otimes k} + \xi^{\otimes k-1}) \\
&\quad -(\xi^{\otimes k} + \xi^{\otimes k-1}) \otimes \sum_i e_i \otimes e_i^o \otimes e_i \otimes e_i^o \\
&= \zeta \otimes (\xi^{\otimes k} + \xi^{\otimes k-1}) - (\xi^{\otimes k} + \xi^{\otimes k-1}) \otimes \zeta,
\end{aligned}$$

where we have set $\zeta = \sum_i e_i \otimes e_i^o \otimes e_i \otimes e_i^o$.

Let $\eta = e_1 \otimes e_1^o \otimes e_2 \otimes e_2^o$. Then $\eta \otimes \xi^{\otimes l} \perp \zeta \otimes \xi^{\otimes k}$ for all l, k. On the other hand, $\langle \eta \otimes \xi^{\otimes l}, \xi^{\otimes k} \otimes \zeta \rangle = \delta_{k=l} \|\xi^{\otimes k-1}\|$.

It follows that for any $k > 1$,

$$\begin{aligned}\langle \eta \otimes \xi^{\otimes l}, [Z, \xi^{\otimes k}] \rangle &= \langle \zeta \otimes (\xi^{\otimes k} + \xi^{\otimes k-1}) - (\xi^{\otimes k} + \xi^{\otimes k-1}) \otimes \zeta, \eta \otimes \xi^{\otimes l} \rangle \\ &= - \langle \xi^{\otimes k} \otimes \zeta, \eta \otimes \xi^{\otimes l} \rangle - \langle \xi^{\otimes k-1} \otimes \zeta, \eta \otimes \xi^{\otimes l} \rangle \\ &= - \delta_{l=k} \|\xi^{\otimes (l-1)}\| - \delta_{l=k-1} \|\xi^{\otimes (l-1)}\|.\end{aligned}$$

It follows that if $a = \sum \alpha_k \xi^{\otimes k} \in L^2(\mathcal{A}_v)$, and we assume that $[Z, a] = 0$, then we get for all $l \geq 2$:

$$\begin{aligned}0 = \langle [Z, a], \eta \otimes \xi^{\otimes l} \rangle &= \sum_k \alpha_k \langle [Z, \xi^{\otimes k}], \eta \otimes \xi^{\otimes l} \rangle \\ &= -\alpha_l \|\xi^{\otimes (l-1)}\| - \alpha_{l+1} \|\xi^{\otimes (l-1)}\|.\end{aligned}$$

It follows that α_k, $k \geq 2$, is a constant sequence. But the sequence $\{\alpha_k\}$ is in L^2 and thus must be zero.

It follows that $a \in L^2(\mathcal{A}_v)$ commutes with Z, then $a = \alpha_0 1 + \alpha_1 \cup$. But in this case, $[a, Z]\Omega = \alpha_1 [\cup, Z]\Omega$. The only tensors of degree 6 in $\cup Z \Omega$ are $\xi \otimes \zeta$, and the only terms of this degree in $Z \cup \Omega$ are $\zeta \otimes \xi$, which are not equal. Thus $[a, Z] = 0$ implies that also $\alpha_1 = 0$ and so a must be a scalar.

To see (ii), note first that any $v \in \Gamma_+$ is a projection in the relative commutant of $W^*(\cup, \Cup)' \cap \mathcal{N}_+$. Since the projections corresponding to different vertices are orthogonal, it follows that any element x in the relative commutant is a weakly-convergent infinite sum $\sum_{v \in \Gamma_+} vxv$, where $vxv \in vW^*(\cup, \Cup)'v \cap v\mathcal{N}_+ v = \mathbb{C}v$. $\square$

4.7. Factoriality of M_0. Let P be an (extremal) subfactor planar algebra embedded into the planar algebra of some graph Γ. Thus $(Gr_0 P, Tr_0)$ can be viewed as a subalgebra of $(Gr_0 P^\Gamma, Tr_0) \subset \mathcal{M}_0$. Moreover, $TL(1), TL(2) \subset Gr_0 P$, and so $\cup$ and $\Cup$ both belong to $Gr_0 P$. Therefore, the center of $W^*(Gr_0 P, Tr_0)$ is contained in the relative commutant of $W^*(\cup, \Cup)$ inside of $\mathcal{M}_0$. By Theorem 4, this relative commutant is the intersection of the algebra A_0 identified with the zero box space in P^Γ.

LEMMA 13. *Assume that the zero-box space of P is one-dimensional. Then $W^*(Gr_0 P, Tr_0) \cap A_+ = \mathbb{C}1_{A_+}$.*

PROOF. Note that tr_0 is the restriction to $W^*(Gr_0 P, \wedge_0, Tr_0)$ of the conditional expectation E from $\mathcal{M}_0$ onto A_+ (which is the center of $\mathcal{M}_0$). Since this conditional expectation is normal, if $z \in W^*(Gr_0 P, Tr_0) \cap A_+$, then $z = E(z)$. On the other hand, z is the limit (in the weak-operator topology) of some sequence $z_i \in Gr_0 P$. For each i, $E(z_i) = Tr_0(z_i)$ belongs to the zero-box space of P_+. Since $\cup \in P$ and the zero-box space is one-dimensional $E(z_i)$ must be a multiple of $\mathcal{E}(\cup) = \delta \mathbb{C} 1_{A_+}$ and hence $z = E(z) \in \mathbb{C}1_{A_+}$. $\square$

Thus if the zero box space of P is one-dimensional, and since $W^*(\cup)$ is diffuse, we automatically get:

THEOREM 5. *Let P be a planar algebra with one-dimensional zero box space and of index $\delta > 1$. Then $M_0 = W^*(Gr_0 P, Tr_0)$ is a type II_1 factor.*

Since $M_0 \subset \mathcal{M}_0$, we see by Lemma 8 that M_0 can be actually embedded into a direct sum of free group factors. In particular, M_0 has the Haagerup property and is R^ω-embeddable.

5. Higher relative commutants

5.1. The algebra $\mathfrak{M}_1$ and the trace ϕ_1. We now proceed to define the algebra $M_1 = W^*(Gr_1P, Tr_1)$, which will contain $M_0 = W^*(Gr_0P, Tr_0)$ as a subfactor.

Let us denote by $\mathfrak{M}_0$ the image of the algebra Gr_0P inside $M_0 \subset \mathcal{M}_0$ acting on the Fock space $\mathcal{F}$ as in the previous section.

We first recall from Section 2 that if we identify elements of P^Γ with paths, then the multiplication $\wedge_1$ on GrP_1^Γ can be expressed as follows. Let $w = e_1 \cdots e_n$ and $w' = e_1' \cdots e_m'$ be two paths. Denote by $D_1(w)$ the path obtained from w by following the path w, but starting at the first point of w (rather than its starting point). Then

$$D_1^{-1}(D_1(w) \wedge_1 D_1(w')) = \sigma(e_n)^{-1}\delta_{e_n=e_1'} e_1 \cdots e_{n-1}e_2' \cdots e_m'.$$

(note that the factor $\sigma(e_n)^{-1}$ is exactly the norm $\|e_n\|^2$).

To a path $w = e_1 \cdots e_n = e_1 w_0 e_n$, where $w_0 = e_2 \cdots e_{n-1}$ we associate the variable

$$c_1(w) = \ell(e_1)c(w_0)\ell(e_n^o)^* \in B(\mathcal{F}).$$

LEMMA 14. $Y^{(1)}_{D_1^{-1}(w)} Y^{(1)}_{D_1^{-1}(w')} = Y^{(1)}_{D_1^{-1}(w \wedge_1 w')}$.

PROOF. This follows from the relation $\ell(e)^*\ell(g) = \delta_{e=g}\|e\|^2$. □

Let us introduce the notation

$$\mathfrak{M}_1 = \operatorname{span}\{c_1(w) : w \in L_-\}$$

where L_- is the set of all loops starting at an odd vertex.

The vector space $\mathfrak{M}_1$ is an algebra with multiplication $\wedge_1$. Thus $w \mapsto c_1(D_1^{-1}(w))$ is a $*$-homomorphism from Gr_1P^Γ onto $\mathfrak{M}_1$. The unit of $\mathfrak{M}_1$ is the element

$$\sum_{e \in E_-} \sigma(e)\ell(e)\ell(e)^*$$

(here E_- is the set of all odd edges, i.e., ones *ending* at an even vertex).

LEMMA 15. *Let E_- be the set of all odd edges. Then the map*

$$i : Y \mapsto \sum_{e \in E_-} \sigma(e)\ell(e)Y\ell(e)^*$$

defines a unital $*$-homomorphism from the algebra $\mathfrak{M}_0$ to the algebra $\mathfrak{M}_1$.

PROOF. We note that

$$\begin{aligned} i(Y_w) \cdot i(Y_{w'}) &= \sum_{e \in E_-} \sigma(e)^2\|e\|^2\ell(e)Y_wY_{w'}\ell(e)^* \\ &= \sum_{e \in E_-} \sigma(e)\ell(e)Y_wY_{w'}\ell(e)^*. \end{aligned}$$

Thus i is a homomorphism. Moreover, i is clearly $*$-preserving. □

We now define a tracial weight ϕ_1 on $\mathfrak{M}_1$:

$$\phi_1(\ell(e)Y_w\ell(f)^*) = \delta^{-1}\delta_{e=f}\sigma(e)^{-1}\phi(Y_w)$$

(the first δ is the Perron-Frobenius eigenvalue; note that $e = f$ forces $Y_w \in \mathfrak{M}_0$). In other words,

$$\phi_1(X) = \delta^{-1}\sum_{f\in E_-}\sigma(f)^{-1}\phi(\langle f, Xf\rangle_A).$$

The last observation shows that $\phi_1(X)$ is a non-negative functional.

Moreover, for any $v \in \mathfrak{M}_0$,

$$\begin{aligned}\phi_1(i(v)) &= \delta^{-1}\sum_{f\in\Gamma_-(v)}\sigma(f)^{-1}\|f\|_2^2\\ &= \delta^{-1}\sum_{f\in\Gamma_-(v)}\left(\frac{\mu(s(f))}{\mu(t(f))}\right) = \delta^{-1}\sum_j \Gamma_{jv}\frac{\mu(j)}{\mu(v)} = \delta^{-1}\delta = 1.\end{aligned}$$

(here $\Gamma_-(v)$ denotes the set of edges ending at v).

Finally, ϕ_1 is a trace, since if w, w' are two loops of the form $e\hat{w}f^o$ and $e'\hat{w}'f'^o$ with $s(e) = t(f^o)$, $s(e') = t(f'^o)$ then

$$\begin{aligned}\phi_1(ww') &= \delta_{e=f'}\delta_{f=e'}\delta^{-1}\|e\|^4\|f\|^2\phi(\hat{w}\hat{w}')\\ &= \delta_{e'=f}\delta_{f'=e}\delta^{-1}\|e'\|^6\phi(\hat{w}'\hat{w}) = \phi_1(w'w)\end{aligned}$$

since $e = f'$ and $\|f'\| = \|e'\|$.

We finally note that

$$\begin{aligned}\phi_1(i(Y_w)) &= \sum_e \delta^{-1}\left(\frac{\mu(s(e))}{\mu(t(e))}\right)\phi(w)\\ &= \phi(w)\end{aligned}$$

because μ is an eigenvector for the graph matrix. We summarize these observations as the following

LEMMA 16. *The weight ϕ_1 is a semifinite faithful trace, and the inclusion $i : (\mathfrak{M}_0, \phi) \to (\mathfrak{M}_1, \phi_1)$ is trace-preserving.*

PROOF. Let us consider $Y = \sum_{g,h\in E_-}\ell(g)x_{g,h}\ell(h)^*$, $x_{g,h} \in W^*(c(e) : e \in \Gamma)$ with Y^*Y in the domain of ϕ_1. Then $Y^*Y = \sum_{g,h,g'}\ell(g)x_{g,h}x^*_{g',h}\ell(g')^*\|h\|^2$. Moreover,

$$\phi_1(YY^*) = \delta^{-1}\sum_{g,h}\|h\|^2\|g\|^4\phi(x_{g,h}x^*_{g,h})$$

and $x_{g,h}x^*_{g,h} \in \mathcal{M}_0$. Thus if $\phi_1(Y^*Y) = 0$, each of the positive terms in the sum above must be zero and so $\phi(x_{g,h}x^*_{g,h}) = 0$ for all g, h. It follows that $Y = 0$. $\square$

Define now the map $E_1 : \mathfrak{M}_1 \to \mathfrak{M}_0$ by

$$E_1(\ell(e)c(w)\ell(f)^*) = \delta_{e=f}\delta^{-1}\left(\frac{\mu(s(e))}{\mu(t(e))}\right)c(w).$$

Note that

$$E_1(i(Y)) = Y$$

and moreover

$$\begin{aligned}E_1(i(c(w))\ell(e)c(w)\ell(f)^*) &= Y_wE_1(\ell(e)c(w')\ell(f)^*)\\ E_1(\ell(e)c(w)\ell(f)^*i(c(w)')) &= E_1(\ell(e)c(w)\ell(f)^*)c(w')\end{aligned}$$

so that $i \circ E_1 : \mathfrak{M}_1 \to i(\mathfrak{M}_0) \subset \mathfrak{M}_1$ is an $\mathcal{A}_0$-linear projection. Moreover, we see that $\phi_1 = \phi \circ (i \circ E_1)$ so that $\mathcal{E}_1 = i \circ E_1$ is the trace-preserving conditional expectation from $\mathfrak{M}_1 \to \mathfrak{M}_0$. It follows that $\mathcal{E}_1$ extends also to the von Neumann algebra generated by $\mathfrak{M}_1$.

5.2. The algebras $\mathfrak{M}_n$ and traces ϕ_n. The algebras $\mathfrak{M}_n$ with semi-finite traces ϕ_k are defined in a similar way. The algebra $\mathfrak{M}_n$ is the linear span

$$\mathfrak{M}_n = \text{span}\{\ell(e_1)\cdots\ell(e_n)c(w)\ell(f_n)^*\cdots\ell(f_1)^* : e_1\cdots e_n w f_n^0 \cdots f_1^o \in L^{\pm}\}$$

where the parity of the loops is chosen to match the parity of n. For a loop $e_1 \cdots e_n w f_n^o \cdots f_1^o \in L_{\pm}$, we set

$$c_n(e_1\cdots e_n w f_n^o \cdots f_1^o) = \ell(e_1)\cdots\ell(e_n)c(w)\ell(f_n)^*\cdots\ell(f_1)^* \in \mathfrak{M}_n.$$

Then map $L^+ \ni w \mapsto c(D_k^{-1}(w))$ defines a $*$-homomorphism from $Gr_k P^\Gamma$ to $\mathfrak{M}_k$. (Recall that $D_k(w))$ denotes the loop obtained by replacing the starting point in w by the k-th point on the path w).

Let

$$\phi_n \quad = \quad \delta^{-n} \sum_{w=f_1\cdots f_n} \left(\frac{\mu(s(f_n))}{\mu(t(f_1))}\right)^{1/2} \phi \circ \langle \cdot f_n \otimes \cdots \otimes f_1, f_n \otimes \cdots \otimes f_1 \rangle_A.$$

The inclusion $i = i_n^{n-1} : \mathfrak{M}_{n-1} \to \mathfrak{M}_n$ is given by

$$c_{n-1}(w) \mapsto \sum_{ewe^o \in L} \sigma(e)^{-1}\ell(e)c_{n-1}(w)\ell(e)^*.$$

One can check that i is again a trace-preserving inclusion. The conditional expectation $E_n : \mathcal{M}_n \to \mathcal{M}_{n-1}$ is given by

$$E_n(\ell(e)c_{n-1}(w)\ell(f)^*) = \delta_{e=f}\delta^{-1}\left(\frac{\mu(s(e))}{\mu(t(e))}\right) c_{n-1}(w).$$

As before, set

$$\mathcal{E}_n = i \circ E_n : \mathfrak{M}_n \to i(\mathfrak{M}_{n-1}) \subset \mathfrak{M}_n.$$

It is not hard to check that this is the unique trace-preserving $\mathfrak{M}_{n-1}$ linear conditional expectation from $\mathfrak{M}_n$ to $i(\mathfrak{M}_{n-1})$ and that the trace ϕ_n is faithful (the argument is exactly the same as in the case $n = 1$). Moreover, one can easily check that $\mathcal{E}_n = Tr_n$ if we identify $\mathfrak{M}_n$ with $Gr_n P^\Gamma$.

Let us set

$$i_n^j = i_n^{n-1} \circ \cdots \circ i_{j+1}^j : \mathfrak{M}_j \to \mathfrak{M}_n, \qquad i_n = i_n^0.$$

Comparing these with the definitions of section 2 we get:

THEOREM 6. *The map $w \mapsto c_n(w)$ is a $*$-isomorphism from $Gr_k P^\Gamma$ onto $\mathfrak{M}_k$. The semifinite weight ϕ_k satisfies $\phi_k(v \wedge_0 Tr_k(x)) = \phi_k(v \wedge_0 x)$. In particular, the trace Tr_k is positive and faithful.*

5.3. Higher relative commutants. We now let

$$\mathcal{M}_k = W^*(Gr_k P^\Gamma, \phi_k) = W^*(\mathfrak{M}_k).$$

Given a planar subalgebra $P \subset P^\Gamma$, we'll denote by M the subalgebra of $\mathcal{M}_k$ generated by elements from P. In other words, $M_k = W^*(Gr_k P, Tr_k)$.

We'll denote by $\cup_n$ and $\Cup_n$ the images in $\mathfrak{M}_n$ of $\cup, \Cup \in \mathfrak{M}_0$. Note that $\cup_n, \Cup_n \in M_n$.

LEMMA 17. *Let $e_1 \cdots e_n f_1^o \cdots f_n^o$ be a loop in $L_\pm$ (parity according to n). Then the element $Z = \ell(e_1) \cdots \ell(e_n) \ell(f_1)^* \cdots \ell(f_n)^* \in W^*(\cup_n, \Cup_n)' \cap \mathfrak{M}_n \subset W^*(\cup_n, \Cup_n)' \cap \mathcal{M}_n$.*

The proof is a straightforward computation and is omitted.

LEMMA 18. *Let $P \subset P^\Gamma$ be a subfactor planar algebra with index, and let $M_n = W^*(Gr_n P, Tr_n)$ as above. Then $W^*(\cup_n, \Cup_n)' \cap M_n = P_{n,+}$.*

PROOF. Let Q_n be the set of all paths in Γ of length n ending at an even vertex (and starting at an even or odd vertex, according to the parity of n). For $w = e_1 \cdots e_n \in Q_n$, let $F_w = \ell(e_1) \cdots \ell(e_n)$. Then for any $Y \in \mathfrak{M}_n$,

$$\hat{Y}_{w,w'} = F_w^* Y F_{w'} \in \mathcal{N}_+. \tag{3}$$

Moreover,

$$Y = \sum_{w,w' \in Q_n} c_{w,w'} F_w \hat{Y}_{w,w'} F_{w'}^* \tag{4}$$

where $c_{w,w'}$ are some constants. Since the sum above is finite, it follows that equations (3) and (4) continue to hold whenever $Y \in \mathcal{M}_n$, i.e. after passing to weak limits.

Thus if $Y \in \mathcal{M}_n$, and we set $Z = F_w F_w^*$, $Z' = F_{w'} F_{w'}^*$, then $ZYZ' = F_w^* \hat{Y} F_w$, where $\hat{Y} \in \mathcal{N}_+$. Moreover, Y is equal to a fixed finite linear combination of terms $\{ZYZ' : w, w' \in Q_n\}$.

Let us assume now that $Y \in W^*(\cup_n, \Cup_n)' \cap \mathcal{M}_n$. Then by choosing Z, Z' as above, we see from Lemma 17 that $ZYZ' \in W^*(\cup_n, \Cup_n)' \cap \mathcal{M}_n$. Using equations (3) and (4), we conclude that Y is a finite linear combination of terms of the form

$$\ell(e_1) \cdots \ell(e_n) X \ell(f_n)^* \cdots \ell(f_1)^*, \qquad X \in \mathcal{N}_+,$$

and that each such term must belong to the relative commutant $W^*(\cup_n, \Cup_n)' \cap \mathcal{M}_k$.

We can thus assume that $Y = \ell(e_1) \cdots \ell(e_n) X \ell(f_n)^* \cdots \ell(f_1)^*$ with $X \in \mathcal{N}_+$. Then

$$[Y, i_n(\cup)] = \left(\frac{\mu(t(e_n))}{\mu(s(e_1))} \right)^{1/2} \ell(e_1) \cdots \ell(e_n) [X, \cup] \ell(f_n)^* \cdots \ell(f_1)^*$$

and similarly

$$[Y, i_n(\Cup)] = \left(\frac{\mu(t(e_n))}{\mu(s(e_1))} \right)^{1/2} \ell(e_1) \cdots \ell(e_n) [X, \Cup] \ell(f_n)^* \cdots \ell(f_1)^*.$$

Thus if Y is in the relative commutant of $i_n(\cup, \Cup) \cap \mathcal{M}_n$, then X must be in the relative commutant of $\{\cup, \Cup\}$ in $\mathcal{N}_+$, which we know to be A_+ (Theorem 4). It follows that

$$\begin{aligned} \{\cup, \Cup\}' \cap \mathcal{M}_n &\subset \operatorname{span}\{\ell(e_1) \cdots \ell(e_n) \ell(f_n)^* \cdots \ell(f_1)^* : e_1 \cdots e_n f_n^o \cdots f_1^o \in L_\pm\} \\ &= \{c_n(w) : w \in L_\pm \text{ a loop of length } 2n \text{ starting at an } \tfrac{\text{even}}{\text{odd}} \text{ vertex}\}. \end{aligned}$$

Since the reverse inclusion holds by Lemma 17, equality holds. In particular, $\{\cup_n, \Cup_n\}' \cap \mathcal{M}_n = \{\cup_n, \Cup_n\}' \cap \mathfrak{M}_n$. We now see from the definitions in section 2 that the latter algebra is exactly the planar algebra $P_{n,+}^\Gamma$ taken with its usual multiplication.

Thus it follows that

$$\{\cup, \Cup\}' \cap M_n = P_{n,+}^\Gamma \cap M_n.$$

We claim that the latter intersection is exactly $P_{n,+}$. To see this, write any $Y \in \mathcal{M}_n$ as

$$Y = \sum_{w,w' \in Q_n} c_{w,w'} F_w \hat{Y}_{w,w;} F_{w'}^*$$

with $\hat{Y}_{w,w'} = F_w^* Y F_{w'} \in \mathcal{N}_+$, as before. Let

$$E_n(Y) = \sum_{w,w' \in Q_n} c_{w,w'} F_w E(\hat{Y}_{w,w'}) F_w^*,$$

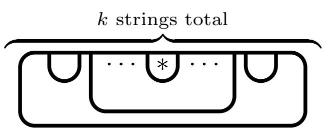
where $E : \mathcal{N}_+ \to A_+$ is the (normal) conditional expectation given by (1). Then E_n is a weakly-continuous map, and moreover $E_n(Y) = Y$ if $Y \in P^\Gamma_{n,+}$. Thus if $Y \in P^\Gamma_{n,+} \cap M_n$, then Y is the limit (in the weak operator topology) of a sequence $Y^{(j)} \in (P_{k_j,+}) \subset M_n$. But then $Y = E(Y) = \lim_k E(Y^{(k)})$. Since the zero-box space of P is one-dimensional, it follows that $E(Y^{(k)}) \in P_{n,+}$ (since then $E(\hat{Y}^{(k)}_{w,w'}) \in \mathbb{C}1_{A_+}$) and so $Y \in P_{n,+}$. Thus $P^\Gamma_{n,+} \cap M_n = P_{n,+}$ and the theorem is proved. $\square$

THEOREM 7. *Let $P \subset P^\Gamma$ be a subfactor planar subalgebra of index $\delta \neq 1$. Let $M_k = W^*(Gr_k P, tr_k)$. Then $M_0' \cap M_k = P_{k,+}$ as algebras (here $P_{k,+}$ is taken with ordinary multiplication) in a way that preserves Jones projections.*

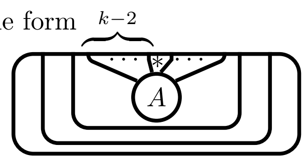
PROOF. Since $\cup_k, \Cup_k \in M_k$, Lemma 18 shows that $P_{k,+} \supset M_0' \cap M_k$. Thus it is enough to prove that $P_{k,+} \subset M_0' \cap M_k$. But this is immediate, since P_k commutes with $i_k(\mathfrak{M}_0)$ and thus with M_0. The correspondence takes Jones projections to Jones projections (as is immediate from the pictures). $\square$

LEMMA 19. *Let $\mathbf{e}_k \in P_{k,+}$, $k \geq 2$ be the Jones projection. Then $\mathbf{e}_k$ is the Jones projection for the inclusion $M_{k-2} \subset M_{k-1}$.*

PROOF. We first check that $\mathbf{e}_k \in M_{k-2}' \cap M_k$. Indeed,

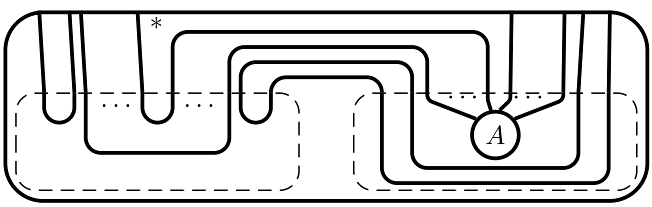
$\mathbf{e}_k = \delta^{-1}$ 

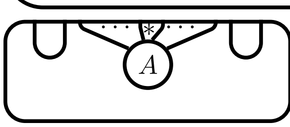
and since $x \in M_{k-2}$, it has the form

$x =$ 

.

We now compute:

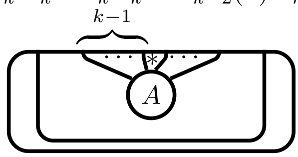
$\delta \mathbf{e}_k \wedge_k x \quad =$ 

$=$ 

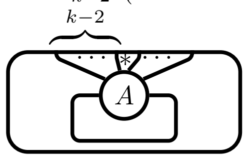
$= \delta x \wedge_k \mathbf{e}_k$ (by symmetry).

(here dashed lines indicate removed boxes).

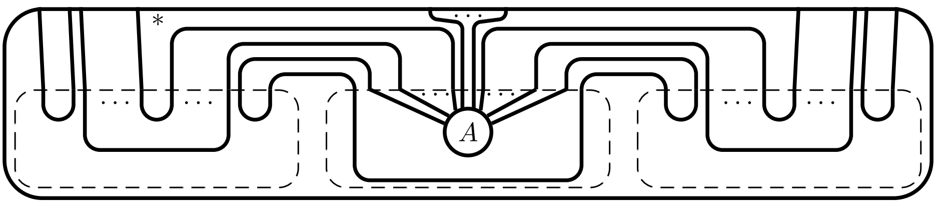
Next, we check that $\mathbf{e}_k \wedge_k x \wedge_k \mathbf{e}_k = \mathcal{E}_{k-2}(x) \wedge_k \mathbf{e}_k$ for

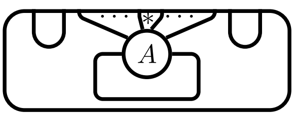
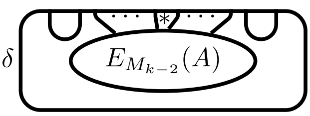
$x =$  $\in M_{k-1}.$

Note that it follows from the formula for $\mathcal{E}_{k-2}$ (or from an explicit computation using the trace) that

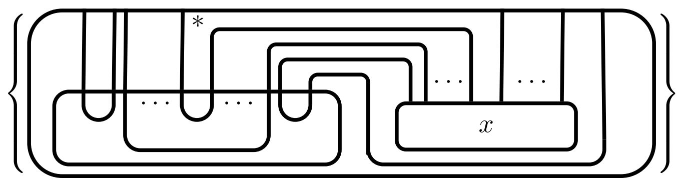
$E_{k-2}(A) = \delta^{-1}$ 

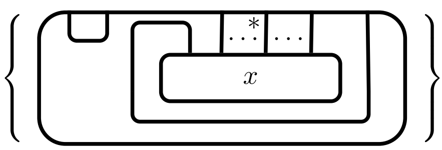
Now, we compute the product $\delta^2 \mathbf{e}_k \wedge_k x \wedge_k \mathbf{e}_k$:

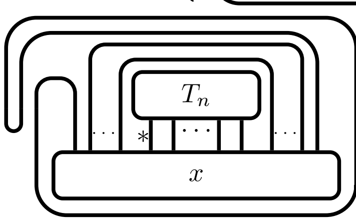
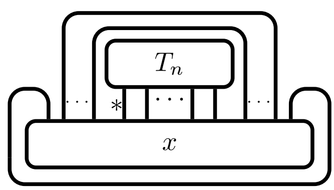


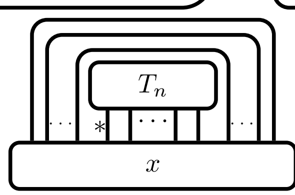
$=$  $= \delta$  $= \delta^2 \mathcal{E}_{k-2}(x) \wedge_k \mathbf{e}_k.$

Finally, we check that the trace Tr_k is λ-Markov. Let $x \in M_{k-1}$. Then:

$\delta Tr_k(x \wedge_k \mathbf{e}_k) = Tr_k$ 

$= Tr_k$ 

$=$  $=$ 

$=$  $= Tr_{k-1}(x).$

This completes the proof. □

LEMMA 20. *The algebras $M_0 \subset M_1 \subset M_2 \subset \cdots$ are exactly the tower obtained by iterating the basic construction for $M_0 \subset M_1$.*

PROOF. We first note that because of the Markov property and the Jones relations between the projections $\mathbf{e}_n$, the algebras $\hat{M}_n = \langle M, \mathbf{e}_1, \ldots, \mathbf{e}_{n-1} \rangle$, $n \geq 2$ are exactly the algebras appearing in the basic construction for $M_0 \subset M_1$. Hence clearly $\hat{M}_n \subset M_n$. Now suppose that for some n this inclusion were strict; choose smallest such n (necessarily > 1 since $M_0 = \hat{M}_0$ and $M_1 = \hat{M}_1$). Then the projection $\mathbf{e}_{n+1}$ is the Jones projection for $M_{n-1} \subset M_n$ and also for $M_{n-1} = \hat{M}_{n-1} \subset \hat{M}_n$. Thus the index of $M_{n-1} \subset M_n$ is the same as that of $M_{n-1} \subset \hat{M}_n$. But since $\hat{M}_n \subset M_n$, multiplicativity of index entails $[M_n : \hat{M}_n] = 1$ and thus $M = \hat{M}_n$, a contradiction. $\square$

5.4. The planar algebra structure on the higher relative commutants. At this stage we have constructed a (II_1) subfactor $M_0 \subset M_1$ and its tower M_k as the completions of $Gr_k P$. We have also shown that $M_0' \cap M_k$ is precisely subspace $P_k \subset Gr_k(P)$.

THEOREM 8. *The linear identification of P_k and $M_0' \cap M_j$ constructed in Theorem 7 is an isomorphism between P and the planar algebra of the subfactor $P(M_0 \subset M_1)$.*

In particular, any subfactor planar algebra can be naturally realized as the planar algebra of the II_1 subfactor $P(W^*(Gr_0 P, Tr_0) \subset W^*(Gr_1 P, Tr_1))$.

The second part of the theorem gives an alternative proof of a result of Popa [**Pop93, Pop95, PS03b**].

PROOF. We have seen in Theorem 7 that the restriction of the multiplication of $Gr_k(P)$ (hence M_k) to P_k is precisely the one given by the multiplication tangle. By [**Jon99**], to conclude that the planar algebra structure defined on P by this identification with the higher relative commutants for $M_0 \subset M_1$ is the same as the original planar algebra structure on P, we have to check the following.

1) That $M_0 \subset M_1$ is extremal (which means there is only one trace on the $M_0' \cap M_k$, that of M_k).

2) The Jones projections $\mathbf{e}_i$ of the tower are ($\frac{1}{\delta}$ times) the diagrammatic $\mathbf{e}_i$'s.

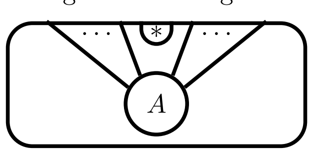
3) The inclusion of $M_0' \cap M_k$ in $M_0' \cap M_{k+1}$ is given by the appropriate tangle.

4) The trace on $M_0' \cap M_k$ given by restricting the trace on M_k is given by the appropriate tangle.

5) The projection from $M_0' \cap M_k$ onto $M_1' \cap M_k$ is given by the appropriate tangle.

For these, 1) follows from the definition of extremality in [**PP86, Pop94**] and a simple diagrammatic manipulation involving spherical invariance of the partition function. 2) was proved as part of Theorem 7. Properties 3) and 4) are just obvious pictures. The only one that requires any thought is 5), which we now prove.

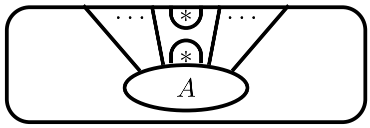
CLAIM 1. Any element in $M_1' \cap M_k$ is in the image of the map from $M_0' \cap M_{k-1}$ to $M_0' \cap M_k$ defined by the following annular tangle:

$A \mapsto$  .

(The shading is determined by the stars being in unshaded regions, the position of $*$ on the inside box being irrelevant.)

PROOF OF CLAIM. It is a simple diagrammatic calculation to show that the image of this tangle does indeed commute with M_1. On the other hand the tangle defines an injective map (the inverse tangle is obvious) and from general subfactor theory the dimensions of $M_0' \cap M_{k-1}$ and $M_1' \cap M_k$ are the same. □

CLAIM 2. If A is in $M_0' \cap M_k$, identified with P_k, then

$$E_{M_1'}(A) = \delta^{-1}$$  .

PROOF OF CLAIM. By extremality $E_{M_1'} = E_{M_1' \cap M_k}$ for elements of $M_0' \cap M_k$. Drawing the picture for $tr(AB)$ for $A \in M_0' \cap M_k$ and $B \in M_1' \cap M_k$, the result is visible. □

This concludes the proof of the Theorem. □

References

[BG05] F. Benaych-Georges. Rectangular random matrices, related free entropy and free Fisher's information. preprint arXiv.org:math/0512081, 2005.

[BIPZ78] E. Brézin, C. Itzykson, G. Parisi, and J. B. Zuber. Planar diagrams. *Comm. Math. Phys.*, 59(1):35–51, 1978.

[BJ97] D. Bisch and V. F. R. Jones. Algebras associated to intermediate subfactors. *Invent. Math.*, 128(1):89–157, 1997.

[CMŚS07] B. Collins, J. Mingo, P. Śniady, and R. Speicher. Second order freeness and fluctuations of random matrices. III. Higher order freeness and free cumulants. *Doc. Math.*, 12:1–70 (electronic), 2007.

[FM77] J. Feldman and C. C. Moore. Ergodic equivalence relations, cohomology, and von Neumann algebras I, II. *Trans. AMS*, 234:289–359, 1977.

[GMS06] A. Guionnet and E. Maurel-Segala. Combinatorial aspects of matrix models. *ALEA Lat. Am. J. Probab. Math. Stat.*, 1:241–279 (electronic), 2006.

[GS08] A. Guionnet and D. Shlyakhtenko. Free diffusion and matrix models with strictly convex potentials. *Geom. Func. Anal.*, 2008.

[Gui06] A. Guionnet. Random matrices and enumeration of maps. In *International Congress of Mathematicians. Vol. III*, pages 623–636. Eur. Math. Soc., Zürich, 2006.

[Jon83] V.F.R. Jones. Index for subfactors. *Invent. Math*, 72:1–25, 1983.

[Jon99] V.F.R. Jones. Planar algebras. Preprint, Berkeley, 1999.

[Jon00] V. F. R. Jones. The planar algebra of a bipartite graph. In *Knots in Hellas '98 (Delphi)*, volume 24 of *Ser. Knots Everything*, pages 94–117. World Sci. Publ., River Edge, NJ, 2000.

[Jon01] V.F.R. Jones. The annular structure of subfactors. In E. Ghys, P. de la Harpe, V.F.R. Jones, V. Sergiescu, and T. Tsuboi, editors, *Essays on Geometry and Related topics*, 401–463, volume 32 of *L'Enseignment Mathematique*, 2001.

[MS06] J. Mingo and R. Speicher. Second order freeness and fluctuations of random matrices. I. Gaussian and Wishart matrices and cyclic Fock spaces. *J. Funct. Anal.*, 235(1):226–270, 2006.

[MŚS07] J. Mingo, P. Śniady, and R. Speicher. Second order freeness and fluctuations of random matrices. II. Unitary random matrices. *Adv. Math.*, 209(1):212–240, 2007.

[Pim97] M. Pimsner. A class of C^*-algebras generalizing both Cuntz-Krieger algebras and crossed products by $\mathbb{Z}$. In D.-V. Voiculescu, editor, *Free Probability*, volume 12 of *Fields Institute Communications*, pages 189–212. American Mathematical Society, 1997.

[Pop93] S. Popa. Markov traces on universal Jones algebras and subfactors of finite index. *Invent. Math.*, 111:375–405, 1993.

[Pop94] S. Popa. Classification of amenable subfactors of type II. *Acta Math.*, 172(2):163–255, 1994.

[Pop95] S. Popa. An axiomatization of the lattice of higher relative commutants of a subfactor. *Invent. Math.*, 120(3):427–445, 1995.

[PP86] M. Pimsner and S. Popa. Entropy and index for subfactors. *Ann. Sci. École Norm. Sup. (4)*, 19(1):57–106, 1986.

[PS03a] S. Popa and D. Shlyakhtenko. Cartan subalgebras and bimodule decompositions of II_1 factors. *Math. Scand.*, 92(1):93–102, 2003.

[PS03b] S. Popa and D. Shlyakhtenko. Universal properties of $L(\mathbf{F}_\infty)$ in subfactor theory. *Acta Math.*, 191(2):225–257, 2003.

[Răd94] F. Rădulescu. Random matrices, amalgamated free products and subfactors of the von Neumann algebra of a free group, of noninteger index. *Invent. math.*, 115:347–389, 1994.

[Rae05] Iain Raeburn. *Graph algebras*, volume 103 of *CBMS Regional Conference Series in Mathematics*. Published for the Conference Board of the Mathematical Sciences, Washington, DC, 2005.

[Shl96] D. Shlyakhtenko. Random Gaussian band matrices and freeness with amalgamation. *Internat. Math. Res. Notices*, 20:1013–1026, 1996.

[Shl97] D. Shlyakhtenko. Free quasi-free states. *Pacific J. Math*, 177:329–368, 1997.

[Spe94] R. Speicher. Multiplicative functions on the lattice of non-crossing partitions and free convolution. *Math. Annalen*, 298:193–206, 1994.

[SU02] D. Shlyakhtenko and Y. Ueda. Irreducible subfactors of $L(F_\infty)$ of index $\lambda > 4$. *J. Reine Angew. Math*, 548:149–166, 2002.

[tH74] G. 't Hooft. A planar diagram theory for strong interactions. *Nucl. Phys. B*, 72:461–473, 1974.

[Voi85] D.-V. Voicuelscu. Symmetries of some reduced free product C^*-algebras, in "Operator Algebras and their Connections with Topology and Ergodic Theory", Lecture Notes in Math. 1132, Springer Verlag, 1985, 556–588.

[Voi91] D.-V. Voiculescu. Limit Laws for Random Matrices and Free Products. Inv. Math., 104:201–220, 1991.

[VDN92] D.-V. Voiculescu, K. Dykema, and A. Nica. *Free Random Variables*, volume 1 of *CRM monograph series*. American Mathematical Society, 1992.

[Voi00] Dan Voiculescu. Lectures on free probability theory. In *Lectures on probability theory and statistics (Saint-Flour, 1998)*, volume 1738 of *Lecture Notes in Math.*, pages 279–349. Springer, Berlin, 2000.

[Zvo97] A. Zvonkin. Matrix integrals and map enumeration: an accessible introduction. *Math. Comput. Modelling*, 26(8-10):281–304, 1997. Combinatorics and physics (Marseilles, 1995).

UMPA, ENS Lyon, 46 allée d'Italie, 69364 Lyon Cedex 07, FRANCE
E-mail address: aguionne@umpa.ens-lyon.fr

Department of Mathematics, UC Berkeley, Berkeley, CA 94720
E-mail address: vfr@math.berkeley.edu

Department of Mathematics, UCLA, Los Angeles, CA 90095
E-mail address: shlyakht@math.ucla.edu

Clay Mathematics Proceedings
Volume 11, 2010

The Tangent Groupoid and the Index Theorem

Nigel Higson

ABSTRACT. We present a proof of the index theorem for Dirac operators that is drawn from Connes' tangent groupoid approach, as described in his book *Noncommutative Geometry*.

1. Introduction

The algebra of linear partial differential operators on a smooth manifold is filtered by the usual concept of order. The *principal symbol* of an operator of order k is its class in the degree k part of the associated graded algebra, which might therefore be called the *principal symbol algebra*. The set of polynomials

$$p(x) = D_0 + xD_1 + \cdots + x^n D_n$$

with partial differential operator coefficients for which the order of the coefficient D_k is no more than k (for $0 \leq k \leq n$) is an algebra over $\mathbb{C}[x]$. The quotient by the ideal of polynomials that vanish at a nonzero $t \in \mathbb{R}$ is the algebra of partial differential operators. When t is zero the quotient is the principal symbol algebra. In this standard algebraic way, the algebra of linear partial differential operators is exhibited as a deformation of the principal symbol algebra.

The convolution C^*-algebra of Alain Connes' tangent groupoid is an analytic counterpart of this deformation. It exhibits the C^*-algebra generated by the smoothing operators on a smooth closed manifold as a deformation of the C^*-algebra generated by the smooth, compactly supported functions on the cotangent bundle (in comparison, the principal symbol algebra is the algebra of smooth functions on the cotangent bundle that are polynomial in each fiber).

There is a simple way to relate *elliptic* partial differential operators to Connes' C^*-algebra using spectral theory, and this makes the C^*-algebra of the tangent groupoid available for use as a tool in the index theory of elliptic operators. In a short section of his famous book [**Con94**, Section II.5], Connes sketches a proof of the Atiyah-Singer index theorem using the tangent groupoid and groupoid techniques. As he notes, his proof is closely related to the K-theory proof of Atiyah and Singer [**AS68a**], but it has the advantage of extending easily to more elaborate settings, for example to foliations.

2010 *Mathematics Subject Classification.* Primary 58J22; Secondary 58H05.
The author was partially supported by a US National Science Foundation grant.

Connes uses the tangent groupoid to build the analytic index map of Atiyah and Singer. Moreover Connes, like Atiyah and Singer, uses an embedding into Euclidean space to describe the index map topologically. But although these ingredients of the two proofs are the same, they are combined differently.

The aim of this paper is to explore this difference a bit further by presenting a proof of the index theorem that reduces Connes' proof to what is arguably its essence. The result is to my mind pleasingly simple, and although the argument does not actually mention groupoids at all, I hope it will help advertise the worth of Connes' C^*-algebraic and groupoid-theoretic points of view.

My friendship with Alain Connes goes back to our joint work on asymptotic morphisms and E-theory [**CH90**]. In some sense E-theory starts from the tangent groupoid, since the tangent groupoid is the basic example of an asymptotic morphism in the same way that the Dirac operator is the basic example of a Kasparov module. So it seems to me fitting to return to the subject here.

It goes without saying that Alain has taught me an immense amount over the years. Moreover, working with Alain gave my self-confidence a great boost at an early stage of my career, and it surely raised my standing in the eyes of others as well. So I owe him a great debt, and I am grateful that I am able to acknowledge that debt here.

2. Index for Families

In this section and the next I shall review some basic constructions in C^*-algebra K-theory and elliptic operator theory. Many of the likely readers of this paper will be familiar with these things, and they ought to proceed directly to Section 4, or even to the proof of the index theorem in Section 6. For those who are not so well-versed in these areas, I have tried to write enough to at least suggest that C^*-algebras are a very convenient setting in which to work out the analytic foundations of index theory.[1]

Let Y be a locally compact Hausdorff space and let $\mathcal{H}$ be a countably generated, continuous field of Hilbert spaces over Y. The reader is referred to [**Dix77**, Ch 10] or [**DD63**] for the definition, but the essence of it is to specify a family of sections deemed to be continuous, rather than derive the concept of continuous section from an overall topology on the bundle of Hilbert space fibers.

A *bounded (and adjointable) operator* on $\mathcal{H}$ is a uniformly bounded family $T = \{T_y \colon \mathcal{H}_y \to \mathcal{H}_y\}_{y\in Y}$ of operators on the fibers of $\mathcal{H}$ that, along with its adjoint family, maps continuous sections to continuous sections. The bounded operators form a C^*-algebra.

2.1. DEFINITION. The C^*-algebra of *compact operators* on $\mathcal{H}$ is generated by bounded operators of the form

$$T_y \colon v_y \mapsto \langle v'_y, v_y\rangle v''_y,$$

where v' and v'' are compactly supported continuous sections of $\mathcal{H}$.

The compact operators form a closed ideal within the bounded operators. Note that the compactness condition goes beyond compactness of the individual operators T_y. In the case of a constant field, bounded operators are bounded

[1] John Roe and I are preparing a book-length account of groupoids and index theory that has very much influenced what is written here.

$*$-strongly continuous families of operators, whereas compact operators are norm-continuous families of compact operators that vanish at infinity.

Suppose that the continuous field $\mathcal{H}$ is $\mathbb{Z}_2$-graded, and that on each $\mathcal{H}_y$ there is given an unbounded,[2] odd-graded, self-adjoint operator D_y such that:

(a) each resolvent operator $(D_y + iI)^{-1}$ is compact, and
(b) the family of resolvents is a compact operator on $\mathcal{H}$.

We aim to construct an index for the family $D = \{D_y\}_{y\in Y}$ in $K(Y)$ that reduces to the Fredholm index when Y is a point (in the graded context the Fredholm index is defined to be the dimension of the even-graded part of the kernel of D minus the dimension of the odd-graded part).

We aim to do so because throughout the proof of the index theorem we shall be working with families of Fredholm operators. But we shall begin by considering a single Hilbert space operator.

2.2. DEFINITION. Let D be an unbounded, odd-graded, self-adjoint operator on a $\mathbb{Z}_2$-graded Hilbert space $\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1$. Its *graph projection* is the orthogonal projection onto the graph of the part of D that maps $\mathcal{H}_0$ into $\mathcal{H}_1$.

The graph is a closed subspace of $\mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_1$, so the graph projection P_D is an operator on $\mathcal{H}$. If P_1 is the orthogonal projection onto $\mathcal{H}_1$, then it is easy to check that

$$P_D - P_1 = D(I + D^2)^{-1} + \gamma(I + D^2)^{-1},$$

where γ is the grading operator on $\mathcal{H}$. As a result, if the resolvent of D is compact, then $P_D - P_1$ is compact too.

The index of D can be recovered from P_D and P_1 using K-theory. In general, if the difference of two projections in a C^*-algebra lies in an ideal, then their formal difference determines an element in the K-theory of that ideal. In the case at hand, the K-theory of the compact operators is isomorphic to $\mathbb{Z}$, and the integer obtained from the formal difference $P_D - P_1$ is the index of D.

If D is an operator on a continuous field $\mathcal{H}$ as in (a) and (b) above, and if P_D and P_1 are obtained by applying the graph projection construction fiberwise, then the difference $P_D - P_1$ is compact, and so the formal difference determines an element in the K-theory of the compact operators on the continuous field $\mathcal{H}$.

If $\mathcal{H}$ is a *trivial* field, then this K-theory group is canonically isomorphic to $K(Y)$, and we therefore obtain an index class in $K(Y)$ as required. The case where $\mathcal{H}$ is not trivial can be handled in many ways. For example we can embed $\mathcal{H}$ as a continuous subfield of a trivial field $\mathcal{H}'$ (even as a summand; see [**DD63**, p. 259]) and then proceed as follows.

2.3. LEMMA. *Every compact operator T on $\mathcal{H}$ extends to a compact operator T' on $\mathcal{H}'$ by defining T'_y to be zero on the orthogonal complement of $\mathcal{H}_y$ in $\mathcal{H}'_y$.* □

The lemma gives a homomorphism from the compact operators on $\mathcal{H}$ to the compact operators on $\mathcal{H}'$ and hence a map from the K-theory of the compact operators on $\mathcal{H}$ into $K(Y)$.

2.4. DEFINITION. Let $D = \{D_y\}_{y\in Y}$ be an odd-graded operator on a $\mathbb{Z}_2$-graded continuous field of Hilbert spaces that satisfies conditions (a) and (b) above. Define the index class

$$\operatorname{Index}(D) \in K(Y)$$

[2]In accordance with standard usage, "unbounded" means "possibly unbounded."

by pushing forward the formal difference $P_D - P_1$ into $K(Y)$, as above.

2.5. LEMMA. *The index class is independent of the embedding into a trivial field that is used in its construction. It has the following properties:*

(a) $\operatorname{Index}(D)$ *is functorial: if* Z *is a closed subset of* Y*, then* $\operatorname{Index}(D)$ *maps to* $\operatorname{Index}(D\big|_Z)$ *under the restriction map from* $K(Y)$ *to* $K(Z)$.

(b) $\operatorname{Index}(D)$ *is additive: if* $D = D' \oplus D''$ *on* $\mathcal{H}' \oplus \mathcal{H}''$*, then*

$$\operatorname{Index}(D) = \operatorname{Index}(D') + \operatorname{Index}(D'').$$

(c) $\operatorname{Index}(D) = 0$ *if the operators* D_y *are invertible.*

(d) *If* D^{op} *denotes the operator on the field* $\mathcal{H}^{op}$ *obtained by reversing the grading on* $\mathcal{H}$*, then* $\operatorname{Index}(D^{op}) = -\operatorname{Index}(D)$. □

2.6. REMARK. If Y is compact, and if $\mathcal{H}$ is a continuous field of constant and finite fiber dimension, then

$$\operatorname{Index}(D) = [\mathcal{H}_0] - [\mathcal{H}_1] \in K(Y),$$

where $\mathcal{H}_0$ and $\mathcal{H}_1$ are the even and odd subfields of $\mathcal{H}$ (they are locally trivial, and hence are vector bundles). As a matter of fact, this additional property, together with the others, actually characterizes the index, although we shall not use this fact.

3. Functional Analysis for First Order Elliptic Operators

Let $\pi\colon X \to Y$ be a submersion between smooth manifolds. The manifolds may have boundaries, but if so, then we require that the boundary of X be the inverse image of the boundary of Y. The fibers $X_y = \pi^{-1}\{y\}$ are then smooth manifolds without boundary.

We shall assume that each fiber X_y is equipped with a smooth measure μ_y, and that if f is a smooth, compactly supported function on X, then the quantity $\int_{X_y} f(x)\, d\mu_y(x)$ is a smooth function of y.

Let S be a smooth Hermitian vector bundle over X and let S_y be its restriction to X_y. The Hilbert spaces $\mathcal{H}_y = L^2(X_y, S_y)$ form a continuous field of Hilbert spaces $\mathcal{H}$ over Y whose continuous sections are generated (in the sense of [**Dix77**, Proposition 10.2.3]) by the smooth compactly supported sections of S.

Let D_y be a first-order linear partial differential operator acting on the sections of S_y and suppose that the family $D = \{D_y\}$ is *smooth* in the sense that if u is a smooth section of S on X, then the section Du defined by

$$(Du)\big|_{X_y} = D_y(u\big|_{X_y})$$

is also smooth.

We shall assume that each D_y is formally self-adjoint. To apply the index construction of the last section we shall need to obtain from D_y an operator that is self-adjoint in the sense of Hilbert space theory (see for example [**Kat76**, Ch. 5]). For this the following concept is useful.

3.1. DEFINITION (See [**HR00**, Definition 10.2.8]). The manifold X is *complete with respect to* D if there is a smooth, proper function $g\colon X \to [0, \infty)$ such that the commutator $[D, g]$ is a uniformly bounded endomorphism of S.

3.2. PROPOSITION (See [**HR00**, Proposition 10.2.10]). *If* X *is complete with respect to* D, *then each formally self-adjoint operator* D_y *is essentially self-adjoint on the smooth, compactly supported sections of* S_y. □

Recall that an operator is essentially self-adjoint if its operator-theoretic closure is self-adjoint (see for example [**Kat76**, Ch. 5] again). From now on we shall assume that X is complete for D, and in a slight abuse of notation we shall write D_y when in operator-theoretic contexts we actually mean the closure of D_y. Form the resolvent family

$$r(D) := (D + iI)^{-1} = \{(D_y + iI)^{-1}\}_{y\in Y}.$$

It is certainly a bounded operator on the continuous field $\mathcal{H}$. To say more, we shall suppose from here on that each operator D_y is *elliptic*. We can then draw the following conclusion.

3.3. PROPOSITION. *If* f *is a smooth, compactly supported function on* X, *acting on the continuous field* $\mathcal{H}$ *as a family of multiplication operators, then* $f{\cdot}r(D)$ *is a compact endomorphism of* $\mathcal{H}$.

This is standard fare, but let us sketch a proof based on a C^*-algebra calculation.

3.4. LEMMA. *Let* A *be a* C^**-algebra that includes* $C_0(X)$ *as a* C^**-subalgebra and let* a *be an element of* A *that commutes with* $C_0(X)$. *Suppose that for every* $x \in X$ *and every* $\varepsilon > 0$ *there is some* $f \in C_0(X)$ *such that* $f(x) = 1$ *and* $\|f{\cdot}a\| < \varepsilon$. *Then* $f{\cdot}a = 0$ *for every* $f \in C_0(X)$. □

Let B be the C^*-algebra of bounded continuous functions from $(0, 1]$ into the bounded operators on $\mathcal{H}$, and let J be the ideal of functions whose distance to the compact operators converges to zero at 0. Let $A = B/J$. The operator-valued function $a\colon t \mapsto (tD + iI)^{-1}$ determines an element of A, and so does every constant operator-valued function $f\colon t \mapsto f$, for every $f \in C_0(X)$. It suffices to show that the product $f \cdot a \in A$ is zero since

$$\lim_{t\to 0} f{\cdot}(tD + iI)^{-1}(D + iI)^{-1} = -if{\cdot}(D + iI)^{-1}.$$

The elements a and f commute in A, so it suffices to verify the estimates in Lemma 3.4 for each given x and $\varepsilon > 0$. The product $f{\cdot}a \in A$ depends only on the restriction of D to a neighborhood of the support of f, so we may as well assume that D is compactly supported and elliptic near the support of f. Furthermore, by choosing f to have sufficiently small support, we may assume that $p\colon X \to Y$ is actually a trivial vector bundle (since any submersion is locally isomorphic to a trivial vector bundle).

Using the basic estimate for constant coefficient elliptic operators we can find f with sufficiently small support so that $f{\cdot}a$ is ε-close to $f{\cdot}a'$, where a' is defined in the same way as a, but using an operator D' that restricts to the same constant coefficient elliptic operator in each vector space fiber of p. A Fourier transform calculation then shows that for every $t \in (0, 1]$ the operator $f{\cdot}(tD' + iI)^{-1}$ is compact, and the proof of Proposition 3.3 is complete.

If X is compact, then we may choose $f \equiv 1$ in Proposition 3.3, and conclude from Section 2 that D has a well-defined index in the K-theory group $K(Y)$. Thus a smooth family of elliptic operators on the fibers of a submersion with compact fibers and compact base has a well-defined *families index* in $K(Y)$ (compare [**AS71**]). But in the proof of the index theorem that we shall present here the manifold X will not be compact.

As a substitute for compactness we shall work with operators of the form $D + E$, where E is a suitable smooth self-adjoint endomorphism of S. The operators in the family $D+E$ are still essentially self-adjoint because X is complete with respect to $D + E$. The compactness of the resolvent $r(D + E)$ is guaranteed (in the cases of concern to us) by the following calculation.

3.5. PROPOSITION. *Assume that S is $\mathbb{Z}_2$-graded and that D is odd-graded. Let E be a smooth, odd-graded self-adjoint endomorphism of the Hermitian bundle S over X. Assume that*

(a) *The square of E is a proper scalar function from X to $[0, \infty)$.*
(b) *The anticommutator $DE + ED$ is a uniformly bounded smooth endomorphism of S.*

Then $r(D + E)$ is a compact operator on the continuous field $\mathcal{H}$.

PROOF. We shall show that for every $\varepsilon > 0$ the family $r(D+E)$ lies within ε of a compact operator.

Choose a smooth, compactly supported real function f such that if $F = \gamma f$, where γ is the grading operator on S, then

$$\|(D + E + F)s\| \geq \varepsilon^{-1}\|s\|$$

for every compactly supported smooth section s. This is possible because first of all

$$(D + E + F)^2 = D^2 + (DE + ED) + [D, f]\gamma + E^2 + f^2.$$

It therefore suffices to choose f such that

$$E^2 + f^2 \geq \varepsilon^{-2} + \|DE + ED\| + \|[D, f]\|,$$

and this may be done because X is complete with respect to D.

The estimate implies that $\|r(D + E + F)\| < \varepsilon$. But then

$$r(D + E) - r(D + E + F) = r(D + E + F) \cdot \gamma f \cdot r(D + E),$$

and by Proposition 3.3 the right hand side is compact. □

3.6. REMARK. Obviously the hypotheses can be relaxed in various ways. But they are adequate for our purposes as they stand.

4. Deformation Spaces

This section describes the deformation construction that underlies both the tangent groupoid and the proof of the index theorem.

4.1. DEFINITION. Let M be a smooth, closed submanifold of a smooth manifold V without boundary. The *deformation space* $\mathbb{N}_V M$ associated to the inclusion of M into V is the set

$$\mathbb{N}_V M = N_V M \ \sqcup \ V \times (0, 1],$$

where $N_V M$ is the normal bundle of M in V.

4.2. REMARK. The deformation space concept is taken from algebraic geometry, where its counterpart is called the *deformation to the normal cone*. See [**BFM75**] or [**Ful98**, Chapter 5].

We equip $\mathbb{N}_V M$ with the weakest topology (i.e. the one with the fewest open sets) such that:

(a) The natural map $\mathbb{N}_V M \to V \times [0,1]$ that on $V{\times}(0,1]$ is the inclusion map, while on $N_V M$ is the projection to M, followed by inclusion into $V{\times}\{0\}$, is continuous.

(b) If $f\colon V \to \mathbb{R}$ is a smooth function that vanishes on M, then the function

$$\delta f\colon \mathbb{N}_V M \longrightarrow \mathbb{R}$$

defined by the formulas

$$\delta f(X) = X(f) \quad \text{and} \quad \delta f(\nu, t) = \frac{f(\nu)}{t}$$

is continuous.

This topology is Hausdorff and also locally Euclidean:

4.3. LEMMA. *If $x_1, \ldots, x_p, y_1, \ldots, y_q$ are smooth local coordinates on V such that the x_i restrict to local coordinates on M, whereas the y_j vanish on M, then the functions*

$$x_1, \ldots, x_p, \delta y_1, \ldots, \delta y_q, t$$

are local coordinates on the deformation space: they determine a homeomorphism from the open subset where they are defined to an open subset of $\mathbb{R}^{p+q} \times [0,1]$.

PROOF. Let U be the open subset of V on which the coordinates are defined. The would-be coordinate functions are then defined on $\mathbb{N}_U(M{\cap}U)$, which by the item (a) above is an open subset. (Incidentally, by x_j and t we mean the functions obtained by first applying the map in item (a), then composing with x_j and t.)

The associated map

$$\mathbb{N}_U(M{\cap}U) \to \mathbb{R}^p{\times}\mathbb{R}^q{\times}[0,1]$$

is continuous and one-to-one, and has open image. Viewing U as an open subset of $\mathbb{R}^p{\times}\mathbb{R}^q$ via the given coordinates, and identifying the normal bundle with $(M{\cap}U){\times}\mathbb{R}^q$, the inverse map is given by the formula

$$(u, \nu, t) \mapsto \begin{cases} (u, u + t\nu, t) & t \neq 0 \\ (u, \nu, 0) & t = 0. \end{cases}$$

To check its continuity one verifies that its compositions with the maps in (a) and (b) above are continuous. □

The coordinate charts given by the lemma in fact constitute an atlas for a smooth manifold structure on $\mathbb{N}_V M$. That smooth structure can be described a bit more invariantly as follows.

4.4. DEFINITION. Let X be a set and let $\mathcal{F} = \{f_\alpha\colon X \to V_\alpha\}$ be a family of functions from X into smooth manifolds. Let us say that a function $f\colon X \to \mathbb{R}$ is *smoothly composed from the family $\mathcal{F}$* if it has the form

$$X \xrightarrow{(f_{\alpha_1}, \ldots, f_{\alpha_k})} V_{\alpha_1} \times \cdots \times V_{\alpha_k} \xrightarrow{\quad h \quad} \mathbb{R},$$

where h is a smooth function on the product manifold.

4.5. PROPOSITION. *There is a unique smooth manifold structure on the deformation space $\mathbb{N}_V M$ for which the smooth real-valued functions on $\mathbb{N}_V M$ are precisely the functions that are locally smoothly composed from the functions in (a) and (b) above.*

PROOF. It suffices to show that if $x_1, \dots, x_p, y_1, \dots, y_q$ are local coordinates, as in the previous lemma, then every function on $\mathbb{N}_V M$ that is smoothly composed from the functions in (a) and (b) is locally smoothly composed from the functions

$$x_1, \dots, x_p, \delta y_1, \dots, \delta y_q, t.$$

In turn, it suffices to prove that every δf is so composed, and this is an immediate consequence of Taylor's theorem. □

4.6. EXAMPLE. Suppose that M is embedded in a smooth vector bundle W over M via a section $s\colon M \to W$. Each vector $w \in W$ determines a tangent vector at each point in the fiber containing W, and in this way the normal bundle $N_W M$ identifies with W itself. The product manifold $W \times [0,1]$ is mapped diffeomorphically onto the deformation space $\mathbb{N}_W M$ by the formulas

$$(w, 0) \mapsto w \in N_W M \quad \text{and} \quad (m, w, t) \mapsto (m, s(m) + tw, t) \in W \times (0, 1].$$

5. Spinor Bundles, the Thom Class and the Dirac Operator

Let V be a smooth, oriented euclidean vector bundle of rank $2k$ over a smooth manifold M. A *spinor bundle* for V is a smooth $\mathbb{Z}_2$-graded Hermitian vector bundle SV over M that is equipped with an $\mathbb{R}$-linear smooth bundle map

$$c\colon V \longrightarrow \operatorname{End}(SV)$$

such that:

(a) Each $c(v)$ is odd-graded, skew-adjoint, and satisfies

$$c(v)^2 = -\|v\|^2 1.$$

(b) The grading operator on SV is given by the local formula

$$\gamma = i^k v_1 \cdots v_{2k},$$

involving any oriented local orthonormal frame of V.

(c) The map c induces an isomorphism

$$c\colon \operatorname{Cliff}(V) \xrightarrow{\cong} \operatorname{End}(SV),$$

where $\operatorname{Cliff}(V)$ denotes the bundle of complex Clifford algebras associated to V.

See **[BD82]** or **[LM89]**, for example.

If M reduces to a single point, so that V is a single Euclidean vector space, then a spinor bundle for V, which is better named a *spinor vector space* S for V, exists and is unique up to isomorphism.

5.1. DEFINITION. Let V be an oriented, even-rank Euclidean vector bundle over a compact manifold M and let SV be a a spinor bundle for V. View the pullback of SV to the total space of V as a continuous field of (finite-dimensional) $\mathbb{Z}_2$-graded Hilbert spaces over the manifold V. The *Thom class* $\beta(SV) \in K(V)$ is the index class, in the sense of Section 2, of the family of operators

$$\left\{ E_v \colon S_{\pi(v)} V \to S_{\pi(v)} V \right\}_{v \in V}$$

where $E_v = c(v)\gamma$ is the *self-adjoint Clifford multiplication* operator associated to $v \in V$ (γ is the grading operator on SV).

5.2. REMARK. If V carries a Hermitian structure, then we can set $SV = \bigwedge_{\mathbb{C}}^{*} V$ and

$$c(v) = d(v) - d(v)^{*}\colon S_{\pi(v)}V \longrightarrow S_{\pi(v)}V,$$

where $d(v)(\omega) = v\wedge\omega$. The family $\{c(v)\gamma\}$ is homotopic to the family $\{d(v)+d(v)^{*}\}$, which agrees with the standard definition of Thom class.

There is a useful two-out-of-three principle for spinor bundles. We shall use it when we formulate the index theorem below.

5.3. LEMMA. *Let V' and V'' be oriented, even-rank euclidean vector bundles over M, and let $V = V' \oplus V''$.*

(a) *If SV' and SV'' are spinor bundles for V' and V'', then $SV' \hat{\otimes} SV''$, equipped with the action*

$$c(v', v'') = c(v')\hat{\otimes} I + I \hat{\otimes} c(v'')$$

is a spinor bundle for V.

(b) *If SV is a spinor bundle for V, and if SV' is a spinor bundle for V', then there exists a spinor bundle SV'' for V'' such that $SV' \hat{\otimes} SV''$, viewed as a spinor bundle for $V = V' \oplus V''$, is isomorphic to SV.*

PROOF. Part (a) is straightforward. As for (b), one can take

$$S'V = \mathrm{Hom}_{V''}(SV, SV''),$$

the $\mathbb{Z}_2$-graded Hermitian vector bundle of morphisms $SV \to SV''$ that commute with the actions of V'' on SV and SV''. $\square$

We shall also use the following concepts of *conjugate* spinor bundle and *opposite* spinor bundle (these terms are introduced for our current purposes only and are not standard).

5.4. DEFINITION. Let SV be a spinor bundle for a Euclidean vector bundle V of rank $2k$, and denote by $\overline{S}V$ the complex conjugate of the underlying Hermitian bundle. Since $\mathrm{End}(SV) = \mathrm{End}(\overline{S}V)$, we may equip $\overline{S}V$ with the same map

$$c\colon V \to \mathrm{End}(\overline{S}V)$$

with which SV is equipped. We shall also equip $\overline{S}V$ with the same grading as SV if k is even, and the opposite grading if k is odd. We obtain a spinor bundle for V, called the *conjugate* spinor bundle.

5.5. DEFINITION. Suppose that V' and V'' are oriented, even-rank euclidean vector bundles over M and that the direct sum

$$V = V' \oplus V''$$

is a trivial bundle. Let SV' be a spinor bundle for V'. A spinor bundle SV'' for V'' is *opposite* to the spinor bundle SV' if the tensor product

$$SV = \overline{S}V' \hat{\otimes} SV''$$

is isomorphic to the trivial spinor bundle for V.

5.6. REMARK. Lemma 5.3 guarantees that opposite spinor bundles exist.

Now let M be an oriented Riemannian manifold without boundary and let SM be a spinor bundle for the tangent bundle. A *Dirac operator* for SM is a first-order, formally self-adjoint, odd-graded linear partial differental operator

$$D\colon C_c^\infty(M, SM) \longrightarrow C_c^\infty(M, SM)$$

such that

$$[D, f]\, u = c\big(\operatorname{grad} f\big)\, u.$$

for every smooth function f on M (viewed on the left as a multiplication operator on sections of SM). Every Dirac operator D is elliptic and so if M is closed, then (the self-adjoint extension of) D is Fredholm.

5.7. EXAMPLE. If M is a Hermitian complex manifold, then the Dolbeault operator $D = \bar\partial + \bar\partial^*$ (see [**AS68b**, §4]) is a Dirac operator in the above sense. The manifold need not be Kähler.

The K-theory formulation of the index theorem uses an embedding of M into a Euclidean space. Let us assume for simplicity that M is embedded into a Euclidean vector space V *isometrically*. We can always adjust the metric on M and spinor bundle so that this is so, without altering the index of the Dirac operator, so this is no real restriction.[3]

5.8. THEOREM (Atiyah and Singer). *Let M be a closed submanifold of dimension $2k$ in an oriented, even-dimensional Euclidean space V and let S be a spinor vector space for V. If D is a Dirac operator on M, acting on sections of a spinor bundle SM, and if the normal bundle N_VM is equipped with a spinor bundle SN opposite to SM, then*

$$\operatorname{Index}(D) \cdot \beta(S) = (-1)^k \iota_*\big(\beta(SN)\big) \in K(V),$$

where $\iota_*\colon K(N_VM) \to K(V)$ is the map induced by a tubular neighborhood inclusion of N_VM into V.

5.9. REMARK. As it stands, the theorem is an identity in $K(V)$. An application of the Chern character yields the formula

$$\operatorname{Index}(D) = (-1)^{l-k} \int_{N_VM} \operatorname{ch}\big(\beta(SN)\big),$$

where $\dim(V) = 2l$, since $\int_V \operatorname{ch}(\beta(S)) = (-1)^l$. Characteristic class computations then give the more familiar formula

$$\operatorname{Index}(D) = \int_M e^{\frac{1}{2}c_1(L(SM))}\, \hat{A}(TM),$$

where $L(SM) = \operatorname{Hom}_V(\overline{S}M, SM)$ is the *canonical line bundle* of maps from $\overline{S}M$ to SM commuting with the action of TM. Compare [**AS68b**] or [**LM89**, Appendix D].

6. Proof of the Index Theorem

Let M be a smooth closed manifold that is embedded as a submanifold of a finite-dimensional real vector space V. Define a map

$$p\colon M \times V \times [0,1] \longrightarrow \mathbb{N}_VM$$

where $\mathbb{N}_VM$ is the deformation space of Section 4, by the formulas

$$p(m, v, 0) = p_m(v),$$

[3]Even without appealing the existence of isometric embeddings.

where p_m is the projection from V onto V/T_mM (which is the fiber of the normal bundle N_VM over the point $m \in M$), and, if $t \neq 0$,

$$p(m, v, t) = (m + tv, t).$$

6.1. LEMMA. *The map* p *is a submersion.* □

6.2. REMARK. If we think of $M\times V$ as a trivial vector bundle over M, and the diagonal embedding of M into $M\times V$ as a section, then the diffeomorphism from $M\times V\times[0,1]$ to $\mathbb{N}_{M\times V}M$, given in Example 4.6, identifies the submersion p with the map from $\mathbb{N}_{M\times V}M$ onto $\mathbb{N}_VM$ that is induced from the projection of $M\times V$ onto V.

Suppose now that SM is a Hermitian bundle on M (soon to be a spinor bundle) and that D is a first-order linear partial differential operator acting on the sections of SM (soon to be a Dirac operator, but for the moment not necessarily even elliptic).

Pull back the bundle SM to $M\times V\times[0,1]$. Define a smooth family of operators on the fibers of p, acting on the sections of this pullback bundle, as follows.

(a) If $(v,t) \in V\times(0,1] \subseteq \mathbb{N}_VM$, then the fiber of p over (v,t) is the manifold

$$\left\{(m, t^{-1}(v-m), t) : m \in M\right\} \subseteq M\times V\times[0,1]$$

and so the coordinate projection onto M identifies the fiber with the manifold M. We define $D_{(v,t)} = tD$.

(b) If $(m,v) \in N_VM \subseteq \mathbb{N}_VM$, then the fiber of the map p over X is the manifold

$$\{m\} \times T_mM \times \{0\} \subseteq M\times V\times[0,1].$$

Let D_m be the *model operator* on T_mM, obtained from D by freezing the coefficients at m and dropping order zero terms. We define $D_{(m,v)} = -D_m$.

6.3. LEMMA. *The operators above form a smooth family of elliptic operators. Moreover the manifold* $M\times V\times[0,1]$ *is complete with respect to this family.*

PROOF. If we embed the bundle S over M as a summand of a trivial bundle, then we can reduce the lemma to the case where S is trivial, in which case the original operator D is a system of operators on scalar functions. This allows us to further reduce to the cases where D is either a vector field or multiplication by a function f on M. In the latter case the family is multiplication by the smooth function $(m,v,t) \mapsto tf(m)$ on $M\times V\times[0,1]$. In the former case, if X is a vector field on M, then the associated family of operators is given by the smooth vector field

$$X_{(m,v,t)} = (tX_m, -X_m, 0)$$

on $M\times V\times[0,1]$, where we identify the tangent space of the product manifold at (m,v,t) with $T_mM \times V \times \mathbb{R}$ and we consider T_mM as a subspace of V via the given embedding of M.

As for completeness, if $g\colon V \to [0,\infty)$ is any smooth proper function, then its composition with the second coordinate projection on $M\times V\times[0,1]$ is a smooth proper function on the product manifold whose commutator with D is uniformly bounded. □

Assume now that V is even-dimensional, oriented and Euclidean. Assume that M has dimension $2k$, that it is oriented, and that it is equipped with the Riemannian metric it inherits as a submanifold of V. Let SM be a spinor bundle for TM and D be a Dirac operator for SM.

Fix a spinor space S for V and let $E\colon V \to \operatorname{End}(S)$ be self-adjoint Clifford multiplication, as in Definition 5.1. Consider E as a function

$$E\colon M\times V\times[0,1] \longrightarrow \operatorname{End}(S)$$

via the coordinate projection onto V.

Form the tensor product $SM\hat{\otimes}S$ with fibers $S_mM\hat{\otimes}S$. Form the operator $D\hat{\otimes}I$ on sections of $SM\hat{\otimes}S$ over $M\times V\times[0,1]$, and also the self-adjoint endomorphism $I\hat{\otimes}E$.

6.4. LEMMA. *The anticommutator of* $D\hat{\otimes}I$ *and* $I\hat{\otimes}E$ *is a uniformly bounded endomorphism of* $SM\hat{\otimes}S$*, while* $(I\hat{\otimes}E)^2$ *is a proper function on* $M\times V\times[0,1]$.

PROOF. The square of $I\hat{\otimes}E$ is the scalar function $\|v\|^2$, which is certainly a proper function on $M\times V\times[0,1]$. The anticommutator of $D\hat{\otimes}I$ and $I\hat{\otimes}E$ is the same as the commutator of $D\otimes I$ and $I\otimes E$ on the sections of $SM\otimes S$. ☐

According to Proposition 3.5 there is therefore an index class

$$\operatorname{Index}(D\hat{\otimes}I + I\hat{\otimes}E) \in K(\mathbb{N}_VM).$$

We shall prove the index theorem by computing the restriction of this index class to the closed subsets N_VM and V of $\mathbb{N}_VM$, where the latter is embedded as $V\times\{1\}$. We shall calculate that

$$\operatorname{Index}(D\hat{\otimes}I + I\hat{\otimes}E)\big|_{N_VM} = (-1)^k\beta(N_VM) \in K(N_VM)$$

and

$$\operatorname{Index}(D\hat{\otimes}I + I\hat{\otimes}E)\big|_{V} = \operatorname{Index}(D)\cdot\beta(V) \in K(V),$$

and then the index formula will follow from the following calculation.

6.5. LEMMA. *Let* $\iota\colon N_VM \to V$ *be a tubular neighborhood embedding associated to the embedding of* M *into* V *as a closed submanifold. The diagram*

$$\begin{array}{ccc} K(\mathbb{N}_VM) & = & K(\mathbb{N}_VM) \\ \downarrow & & \downarrow \\ K(N_VM) & \xrightarrow[\iota_*]{} & K(V), \end{array}$$

in which the vertical maps are given by the inclusions of N_VM *and* V *into* $\mathbb{N}_VM$*, is commutative.*

PROOF. Since $\mathbb{N}_VM = N_VM \sqcup V\times(0,1]$, and since $V\times(0,1]$ is contractible in the sense of locally compact spaces, the inclusion of N_VM into $\mathbb{N}_VM$ as a closed subset induces an isomorphism

$$K(\mathbb{N}_VM) \xrightarrow{\cong} K(N_VM)$$

by restriction of K-theory classes. It follows that the open inclusion of the tubular neighborhood $W = \iota(N_VM)$ into V induces an isomorphism

$$K(\mathbb{N}_WM) \xrightarrow{\cong} K(\mathbb{N}_VM).$$

The lemma therefore reduces to the case where $V = W$, and now the calculation in Example 4.6, plus the homotopy invariance of K-theory, completes the proof. □

We shall now calculate $\mathrm{Index}(D\hat{\otimes}I + I\hat{\otimes}E)|_{N_V M}$. Recall from Definition 5.5 that the opposite spinor bundle SN for the normal bundle $N_V M$ is defined so that there is an isomorphism of spinor bundles

$$M\times S \cong \overline{S}M \hat{\otimes} SN$$

for the trivial bundle $M\times V$. As a result there is an isomorphism

$$SM\hat{\otimes}S \cong \big(SM\hat{\otimes}\overline{S}M\,\big) \hat{\otimes} SN.$$

The continuous field of Hilbert spaces on which $(D\hat{\otimes}I + I\hat{\otimes}E)|_{N_V M}$ acts can therefore be written as the field with fiber

$$L^2(T_m M, S_m M)\hat{\otimes}S \cong L^2(T_m M, S_m M\hat{\otimes}\overline{S}_m M) \hat{\otimes} S_m N$$

over $(m,\nu) \in N_V M$. If we define an operator B_m on the first tensor factor on the right hand side by

$$B_m = -D_m\hat{\otimes}I + I\hat{\otimes}E_m,$$

where the function E_m on $T_m M$ is self-adjoint Clifford multiplication, and if we denote by E_ν self-adjoint Clifford multiplication by a normal vector ν on $S_m N$, then

$$(D\hat{\otimes}I + I\hat{\otimes}E)_{(m,\nu)} \cong B_m\hat{\otimes}I + I\hat{\otimes}E_\nu$$

(one should be aware that the descriptions on the left and right use different tensor product decompositions).

We shall now compute the index of the family on the right hand side. The first step is the following lemma, in which we shall use the canonical isomorphisms

$$S_m M\hat{\otimes}\overline{S}_m M \cong \mathrm{End}(S_m M) \cong \mathrm{Cliff}(T_m M),$$

so as to view B_m as an operator on $L^2(T_m M, \mathrm{Cliff}(T_m M))$. Note that according to our conventions, the first isomorphism in the display is grading-preserving if k is even and grading-reversing if k is odd.

6.6. LEMMA. *The kernel of* (*the closure of*) B_m *is spanned by the function*

$$\nu \mapsto \exp(-\tfrac{1}{2}\|\nu\|^2)I \in \mathrm{Cliff}(T_m M).$$

On the orthogonal complement of the kernel, B_m^2 *is bounded below by* 2.

PROOF. We compute that

$$B_m^2 = \Delta + \|\nu\|^2 + (N - 2k)$$

where Δ is the Laplace operator and N is the *number operator* that acts as pI on all monomials $e_{i_1}\cdots e_{i_p}$. The lemma therefore follows from the well-known eigenvalue theory of the quantum harmonic oscillator $\Delta + \|\nu\|^2$. See for example **[GJ87**, p. 12]. □

Now form the one-dimensional continuous field of Hilbert spaces

$$\mathcal{K}_m = \ker(B_m) \subseteq L^2(T_m M, S_m M\hat{\otimes}\overline{S}_m M).$$

It is purely even-graded if k is even, and purely odd-graded if k is odd. The section given in the lemma trivializes $\mathcal{K}$, and as a result

$$\begin{aligned} L^2(T_mM, S_mM\hat{\otimes}\overline{S}_mM)\hat{\otimes} S_mN &\cong \mathcal{K}_m\hat{\otimes}S_mN \oplus \mathcal{K}_m^{\perp}\hat{\otimes}S_mN \\ &\cong S_mN \oplus \mathcal{K}_m^{\perp}\hat{\otimes}S_mN, \end{aligned}$$

where the isomorphism between the first summands is grading-preserving or grading-reversing, according as k is even or odd. In the final direct sum decomposition the operator $B_m\hat{\otimes}I + I\hat{\otimes}E_v$ acts as the self-adjoint Clifford multiplication operator E_v on S_mN and as an invertible operator on $\mathcal{K}_m^{\perp}\hat{\otimes}S_mN$, since

$$(B_m\hat{\otimes}I + I\hat{\otimes}E_v)^2 = B_m^2\hat{\otimes}I + I\hat{\otimes}E_v^2,$$

while $B_m^2 \geq 2$ on $\mathcal{K}_m^{\perp}$. Using the additivity of the index, together with the triviality of the index of the second summand, we find that

$$\begin{aligned} \mathrm{Index}(D\hat{\otimes}I + I\hat{\otimes}E)\big|_{N_VM} &= \mathrm{Index}(B\hat{\otimes}I + I\hat{\otimes}E) \\ &= (-1)^k\beta(SN) \in K(N_VM), \end{aligned}$$

as required.

It remains to compute $\mathrm{Index}(D\hat{\otimes}I + I\hat{\otimes}E)|_V$. This is quite simple. The map

$$\begin{gathered} q\colon M\times V\times[0,1] \longrightarrow V\times[0,1] \\ q\colon (m,v,t) \mapsto (tm+v,t) \end{gathered}$$

is a submersion, and every fiber

$$q^{-1}\{(v,t)\} = \big\{(m, v-tm, t) : m\in M\big\} \subseteq M\times V\times[0,1]$$

is isomorphic to M via the projection to M. Construct the smooth family D that is the Dirac operator on each fiber, and then form the family

$$D\hat{\otimes}I + I\hat{\otimes}E,$$

acting on sections of $SM\hat{\otimes}S$ by using the same self-adjoint Clifford multiplication endomorphism as E as before. We are re-using notation, but this is not especially reckless because the restriction to $V \cong V\times\{1\}$ of the new family is identical to the same restriction of the old one. However the restriction to $V \cong V\times\{0\}$ of the new family, which has the same index as the restriction to $V\times\{1\}$ by homotopy invariance of K-theory, is the family of operators

$$D\hat{\otimes}I + I\hat{\otimes}E_v\colon L^2(M,SM)\hat{\otimes}S \longrightarrow L^2(M,SM)\hat{\otimes}S$$

Decompose $L^2(M,SM)$ into the kernel of D, direct sum its orthogonal complement, and decompose $L^2(M,SM)\hat{\otimes}S$ accordingly. On the second summand the above operators are uniformly bounded below by the first positive eigenvalue in the spectrum of D. On the first summand the operators are

$$I\hat{\otimes}E_v\colon\ \ker(D)\hat{\otimes}S \longrightarrow \ker(D)\hat{\otimes}S$$

Taking into account the grading on $\ker(D)$ we find that

$$\mathrm{Index}(D\hat{\otimes}I + I\hat{\otimes}E)\big|_V = \mathrm{Index}(D)\cdot\beta(S),$$

as required.

7. Some Remarks on the Tangent Groupoid

The proof of the index theorem given in Section 6 generalizes in a number of simple ways. For instance we can introduce a coefficient vector bundle F on M, and if D_F is a Dirac-type operator acting on $F \otimes SM$, then we find that

$$\mathrm{Index}(D_F) \cdot \beta(S) = (-1)^k \iota_* \big([F] \cdot \beta(SN) \big) \in K(V).$$

In addition, since the proof deals with families anyway, it extends directly to a proof of the Atiyah-Singer index theorem for families of Dirac-type operators.

Other cases can be handled too, but after a certain point it becomes more conceptual and otherwise more appropriate to invest in groupoid theory.

The tangent groupoid $\mathbb{T}M$ is the deformation space associated to the diagonal embedding of M into $M \times M$. This embedding is the inclusion of the units into the pair groupoid of M, and same construction, when applied to the inclusion of the unit space into other groupoids, produces other tangent groupoids. For example when applied to the foliation groupoid of a foliated manifold it produces the leafwise tangent groupoid of the foliation.

But let us focus on the classical index theorem and indicate how the proof of the index theorem presented in Section 6 is drawn from Connes' argument in [**Con94**, Section II.5]. Assuming M is embedded in a vector space V, Connes constructs a homomorphism from the groupoid $\mathbb{T}M$ into V, and hence an action of $\mathbb{T}M$ on the manifold V by translations [**Con94**, p. 105]. Associated to this there is a crossed product groupoid $\mathbb{T}M \ltimes V$, and in fact a family of crossed product groupoids, since the translations can be scaled by $s \in [0, 1]$.

When $s = 0$ the translation action is trivial and the groupoid $\mathbb{T}M \ltimes V$ is the tangent groupoid $\mathbb{T}M$ times the parameter space V. It exhibits an elliptic operator as a deformation of its symbol, more or less as indicated in the introduction.

Connes' key observation is that when $s = 1$ the crossed product $\mathbb{T}M \ltimes V$ is Morita equivalent to a *space*, that is, to a groupoid comprised entirely of units. The space in question is $\mathbb{N}_V M$ and in fact if we set $X = M \times V \times [0, 1]$, then

$$\mathbb{T}M \ltimes V \cong \big\{ (x_1, x_2) \in X \times X \, : \, p(x_1) = p(x_2) \big\},$$

where p is the submersion from X onto $\mathbb{N}_V M$ defined in Section 6. The proof we presented is based on the fact that if D is a Dirac operator, then the symbol-to-operator deformation at $s = 0$ that is encoded by the tangent groupoid deforms as s varies from 0 to 1, to the family that we studied in Section 6.

For general operators the analysis of the deformation as s varies is more difficult, and a Bott periodicity argument is required, as in the final part of Connes' proof [**Con94**, p. 106].

References

[AS68a] M. F. Atiyah and I. M. Singer. The index of elliptic operators. I. *Ann. of Math. (2)*, 87:484–530, 1968.

[AS68b] M. F. Atiyah and I. M. Singer. The index of elliptic operators. III. *Ann. of Math. (2)*, 87:546–604, 1968.

[AS71] M. F. Atiyah and I. M. Singer. The index of elliptic operators. IV. *Ann. of Math. (2)*, 93:119–138, 1971.

[BD82] P. Baum and R. G. Douglas. K homology and index theory. In *Operator algebras and applications, Part I (Kingston, Ont., 1980)*, volume 38 of *Proc. Sympos. Pure Math.*, pages 117–173. Amer. Math. Soc., Providence, R.I., 1982.

[BFM75] P. Baum, W. Fulton, and R. MacPherson. Riemann-Roch for singular varieties. *Inst. Hautes Études Sci. Publ. Math.*, (45):101–145, 1975.

[CH90] A. Connes and N. Higson. Déformations, morphismes asymptotiques et K-théorie bivariante. *C. R. Acad. Sci. Paris Sér. I Math.*, 311(2):101–106, 1990.

[Con94] A. Connes. *Noncommutative geometry*. Academic Press Inc., San Diego, CA, 1994.

[DD63] J. Dixmier and A. Douady. Champs continus d'espaces hilbertiens et de C^*-algèbres. *Bull. Soc. Math. France*, 91:227–284, 1963.

[Dix77] J. Dixmier. *C^*-algebras*. North-Holland Publishing Co., Amsterdam, 1977. Translated from the French by Francis Jellett, North-Holland Mathematical Library, Vol. 15.

[Ful98] W. Fulton. *Intersection theory*, volume 2 of *Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics]*. Springer-Verlag, Berlin, second edition, 1998.

[GJ87] J. Glimm and A. Jaffe. *Quantum physics*. Springer-Verlag, New York, second edition, 1987. A functional integral point of view.

[HR00] N. Higson and J. Roe. *Analytic K-homology*. Oxford Mathematical Monographs. Oxford University Press, Oxford, 2000.

[Kat76] T. Kato. *Perturbation theory for linear operators*. Springer-Verlag, Berlin, second edition, 1976. Grundlehren der Mathematischen Wissenschaften, Band 132.

[LM89] H. B. Lawson, Jr. and M.-L. Michelsohn. *Spin geometry*, volume 38 of *Princeton Mathematical Series*. Princeton University Press, Princeton, NJ, 1989.

PENNSYLVANIA STATE UNIVERSITY, UNIVERSITY PARK, PA 16802, USA
E-mail address: higson@math.psu.edu

Clay Mathematics Proceedings
Volume 11, 2010

Le Théorème de Périodicité en K-Théorie Hermitienne

Max Karoubi

Cet article est dédié à Alain Connes, pour sa contribution admirable à la beauté des mathématiques et sa générosité à nous la faire découvrir.

La périodicité de Bott joue un rôle primordial en K-théorie topologique. Elle est d'ailleurs liée intimement au théorème d'Atiyah-Singer et plus généralement à la géométrie non commutative. Dans deux articles précédents [**K1**] et [**K2**], nous avons démontré l'analogue de ce théorème en K-théorie hermitienne pour des anneaux <u>discrets</u> avec (anti)involution $a \mapsto \overline{a}$, sous l'hypothèse qu'il existe un élément λ du centre de A tel que $\lambda + \overline{\lambda} = 1$ (on dit alors que 1 est scindé dans A). Si l'anneau est commutatif et muni de l'involution triviale, ceci introduit l'hypothèse que 2 est inversible dans A.
Si cette dernière hypothèse est anodine pour les algèbres de Banach, il n'en est pas de même pour des anneaux importants comme l'anneau de groupe $\mathbb{Z}\pi$, où π est un groupe discret. Une difficulté rencontrée pour l'étude de ce type d'anneau est la divergence entre les notions de forme quadratique et de forme hermitienne. Dans cet article, nous développons une théorie qui dépasse cette dichotomie et qui est déjà présente dans le travail fondamental de Ranicki [**R**]. Grâce à cette théorie le théorème de périodicité peut être démontré pour tout anneau. Nous montrons par exemple que les groupes de Witt supérieurs d'un corps fini de caractéristique 2 sont tous isomorphes à $\mathbb{Z}/2$ (exemple 5.14).
Les méthodes de cet article sont beaucoup inspirées de celles de [**K1**] et [**K2**] que nous adaptons à notre propos, ce qui nous permet d'être relativement bref pour certaines démonstrations. Un autre ingrédient essentiel est un cup-produit entre formes quadratiques défini par Clauwens [**C**]. Celui-ci permet de définir le morphisme de périodicité dans le cas général. L'article de Clauwens ayant été écrit dans un contexte différent, nous reprenons dans un appendice les lemmes essentiels dont nous avons besoin pour nos démonstrations.
Résumons brièvement les différentes parties de cet article en commençant par le théorème principal qui sera démontré dans le § 5.

Théorème. *Soit A un anneau quelconque. On a alors une équivalence d'homotopie naturelle*

$$ {}_{\varepsilon}V^{\text{él}}(A) \approx \Omega_{-\varepsilon} U^{\text{él}}(A) $$

2010 *Mathematics Subject Classification.* Primary 19G38.

où les espaces $V^{\text{él}}$ et $U^{\text{él}}$ représentent les fibres homotopiques des foncteurs oubli et hyperbolique respectivement pour des catégories de modules quadratiques appropriées.

(1) **Description de différents types de formes hermitiennes.** Après des rappels sur les définitions classiques utilisées, nous introduisons un nouveau type de groupe orthogonal, dit “élargi” : cf. 1.6/7. Si 1 est scindé dans A, celui-ci coïncide avec le groupe orthogonal sur l'anneau des nombres duaux associé à A, soit $A[e]/e^2$, noté simplement $A(e)$ dans la suite de l'article.

(2) **Les groupes de Grothendieck et Bass en $K-$théorie hermitienne.** Nous montrons comment les théorèmes principaux en $K-$théorie hermitienne restent valables dans le cas “élargi”. Nous précisons aussi les notations utilisées, en suivant partiellement la terminologie du livre de Bak [**B1**]. Par exemple, la notation “L”, utilisée en [**K1**] et [**K2**], est abandonnée et remplacée par la notation “KQ”, pour éviter toute ambigüité avec les groupes de chirurgie.

(3) **Les groupes $_\varepsilon\mathbf{K}\mathcal{Q}_n(\mathbf{A})$ pour $n > 0$ et $n < 0$.** Les définitions essentielles sont contenues dans ce paragraphe, en utilisant des idées bien connues en $K-$théorie algébrique. Le théorème 3.2 permet de comparer les théories “max” et “min”, suivant la terminologie de Bak. Nous montrons aussi comment les techniques de Quillen se transcrivent dans notre situation en une description plus géométrique des éléments de $_\varepsilon K\mathcal{Q}_n(A)$.

(4) **Cup-produits en $K-$théorie hermitienne. Le cup-produit de Clauwens.** Le cup-produit en $K-$théorie hermitienne est défini à l'aide de sa description en termes de fibrés plats. Un cup-produit plus subtil, dû essentiellement à Clauwens, est défini en 4.3 (cf. aussi l'appendice). Nous montrons comment tous ces produits sont reliés entre eux dans le théorème 4.8.

(5) **Le théorème fondamental de la $K-$théorie hermitienne pour des anneaux arbitraires.** Dans ce paragraphe, nous généralisons les résultats principaux de [**K1**] et [**K2**] (cf. le théorème 5.2 et la remarque 5.11). La relation avec les groupes de Witt est faite dans le théorème 5.10.

(6) **Les groupes de Witt stabilisés.** En utilisant les résultats précédents, nous introduisons une théorie nouvelle de groupes de Witt “stabilisés” généralisant ceux définis en [**K4**]. Ses propriétés fondamentales sont décrites en 6.1. Une généralisation dans le cadre des schémas a été proposée par M. Schlichting [**S**] en supposant 2 inversible.

(7) **Appendice.** Les lemmes de Clauwens.

Remerciements. Ce travail a été essentiellement accompli pendant le programme thématique sur la théorie de l'homotopie en 2007, organisé au Fields Institute à Toronto. Je remercie également A. Ranicki pour avoir attiré mon attention sur l'article de Clauwens [**C**], J. Berrick pour la démonstration du lemme 4 en appendice, plus simple que le lemme original de Clauwens, ainsi que le referee et M. Schlichting pour des commentaires pertinents après une première version de cet article.

1. Description des différents types de formes hermitiennes et quadratiques.

1.1. Soit A un anneau muni d'une (anti)involution $a \mapsto \overline{a}$ (on dit alors que A est un anneau hermitien) et soit[1] $\varepsilon = \pm 1$. Nous désignons par $\mathcal{P}(A)$ la catégorie des $A-$modules (à droite) qui sont projectifs de type fini (les morphismes étant restreints aux isomorphismes). Si E est un objet de $\mathcal{P}(A)$, son dual E^* est le groupe formé des applications additives $f : E \to A$ telles que $f(x\lambda) = \overline{\lambda} f(x)$, où $\lambda \in A$ et $x \in E$. C'est en fait un objet de $\mathcal{P}(A)$, la structure de A-module à droite étant définie par la formule $(f \cdot \lambda)(x) = f(x)\lambda$. Le module E et son bidual E^{**} sont isomorphes canoniquement grâce à la correspondance $x \mapsto (f \mapsto \overline{f(x)})$. Nous identifierons E à E^{**} par cet isomorphisme. Par ailleurs, si $f : E \to F$ est un morphisme dans $\mathcal{P}(A)$, son transposé ${}^t f : F^* \to E^*$ est défini par la formule classique ${}^t(f)(g) = g \cdot f$ et on a ${}^t({}^t f) = f$, compte tenu des isomorphismes canoniques entre les modules E, F et leurs biduaux respectifs.

1.2. Nous définissons une forme ε-hermitienne sur E comme un morphisme $\phi : E \to E^*$ tel que ${}^t\phi = \varepsilon\phi$, où ${}^t\phi : (E^*)^* \cong E \to E^*$. La forme ϕ est dite "non dégénérée" si c'est un isomorphisme. Il convient de remarquer que la donnée de ϕ équivaut à celle d'une application $\mathbb{Z}-$bilinéaire

$$\chi : E \times E \to A$$

telle que $\chi(x\lambda, y\mu) = \overline{\lambda}\chi(x, y)\mu$ si λ et $\mu \in A$, x et $y \in E$. La correspondance est donnée par la formule classique suivante :

$$\chi(x, y) = \phi(y)(x)$$

La condition de ε-symetrie (${}^t\phi = \varepsilon\phi$) se traduit par l'identité

$$\chi(y, x) = \varepsilon\overline{\chi(x, y)}$$

Dans cet article, les formes hermitiennes ϕ qui nous intéressent sont paires : elles s'écrivent sous la forme

$$\phi = \phi_0 + \varepsilon^t\phi_0$$

Il convient de noter que ϕ_0 n'est pas déterminé par cette formule. Si ϕ_1 est un autre choix et si pose $\gamma = \phi_0 - \phi_1$, on a ${}^t\gamma = -\varepsilon\gamma$.

1.3. Les formes hermitiennes paires sont les objets d'une catégorie notée[2] ${}_\varepsilon\mathcal{Q}^{\max}(A)$, définie de la manière suivante : un morphisme

$$(E, \phi) \to (F, \psi)$$

est un isomorphisme f entre les A-modules sous-jacents tel que le diagramme suivant commute :

$$\begin{tikzcd} E \arrow[r, "f"] \arrow[d, "\phi"'] & F \arrow[d, "\psi"] \\ E^* & F^* \arrow[l, "{}^t f"] \end{tikzcd}$$

1. On pourrait choisir plus généralement un élément ε du centre de A tel que $\varepsilon\overline{\varepsilon} = 1$. Cependant, on se ramène à ce cas en remplaçant A par $M_2(A)$, l'algèbre des matrices 2×2 à coefficients dans A, munie d'une involution adéquate (cf. 1.10).

2. En suivant la terminologie de Bak [**B1**].

1.4. De manière parallèle, en suivant Tits [**T**] et Wall [**W**], on définit une forme ε-quadratique non dégénérée sur E comme une classe de morphismes

$$\phi_0 : E \to E^*$$

tels que $\phi_0 + \varepsilon^t\phi_0 = \phi$ soit une forme hermitienne non dégénérée. Plus précisément, la classe de ϕ_0 est définie modulo l'addition par un morphisme du type ${}^t\gamma - \varepsilon\gamma$. Les formes ε-quadratiques sont aussi les objets d'une catégorie notée[3] ${}_\varepsilon\mathcal{Q}^{\min}(A)$. Un morphisme

$$(E, \phi_0) \to (F, \psi_0)$$

est un isomorphisme f entre les A-modules sous-jacents tel qu'il existe γ, morphisme de E dans E^*, vérifiant l'identité

$${}^tf \cdot \psi_0 \cdot f = \phi_0 + \gamma - \varepsilon^t\gamma \qquad (S)$$

1.5. Remarques. Si A est un corps muni de l'involution triviale, il est facile de voir que la catégorie des 1-formes quadratiques est équivalente à la catégorie usuelle : il suffit de poser

$$q(x) = \phi_0(x)(x)$$

Cette remarque justifie la définition abstraite introduite dans 1.3.

Par ailleurs, si 1 est scindé dans A (cf. l'introduction), la catégorie des modules ε-hermitiens est équivalente à celle des modules ε-quadratiques : avec les définitions ci-dessus il suffit de poser $\gamma = \lambda({}^tf \cdot \psi_0 \cdot f - \phi_0)$. Ce cas se présente notamment si 2 est inversible dans A.

1.6. Nous allons maintenant introduire une troisième catégorie qui jouera un rôle important dans notre travail et qui sera notée $\mathcal{Q}^{\text{él}}(A)$ ("él" pour "élargi" ; cf. la fin de 1.7). Les objets sont quasiment les mêmes que ceux de la catégorie ${}_\varepsilon\mathcal{Q}^{\min}(A)$ précédente, sauf que l'on considère les ϕ_0 comme donnés dans la structure (on ne considère pas seulement les <u>classes</u> de tels ϕ_0). Un morphisme de (E, ϕ_0) vers (F, ψ_0) est défini par un couple (f, γ), tel que l'identité (S) ci-dessus soit satisfaite. La loi de composition des morphismes s'explicite ainsi

$$(f, \gamma).(g, \zeta) = (f \cdot g, \zeta + {}^tg \cdot \gamma \cdot g) \qquad (C)$$

ce qui est cohérent avec l'identité (S).

1.7. Il est instructif de décrire plus précisément le groupe des automorphismes d'un objet dans chacune des trois catégories. Si (E, ϕ) est un objet de ${}_\varepsilon\mathcal{Q}^{\max}(A)$, **le groupe unitaire** ${}_\varepsilon O^{\max}(E, \phi)$ est défini par des isomorphismes $f : E \to E$ tels que

$${}^tf \cdot \phi \cdot f = \phi$$

Si on note $f^* = \phi^{-1} \cdot {}^tf \cdot \phi$ l'opérateur adjoint de f, il revient au même d'écrire

$$f^* \cdot f = \mathrm{Id}_E \text{ (ou } f \cdot f^* = \mathrm{Id}_E)$$

Le **groupe orthogonal** ${}_\varepsilon O^{\min}(E, \phi_0)$ est défini par des isomorphismes $f : E \to E$ tels qu'il existe γ, morphisme de E dans E^*, vérifiant l'identité

$${}^tf \cdot \phi_0 \cdot f = \phi_0 + \gamma - \varepsilon^t\gamma \qquad (E)$$

Il est clair que ${}_\varepsilon O^{\min}(E, \phi_0)$ est un sous-groupe de ${}_\varepsilon O^{\max}(E, \phi)$ pour $\phi = \phi_0 + \varepsilon^t\phi_0$, la forme hermitienne associée à ϕ_0. Il est facile de voir que les groupes ${}_\varepsilon O^{\max}(E, \phi)$

3. cf. la note précédente.

et ${}_\varepsilon O^{\min}(E, \phi_0)$ coïncident si 1 est scindé dans A.

Finalement, le **groupe orthogonal élargi** ${}_\varepsilon O^{\text{él}}(E, \phi_0)$ est défini par des couples (f, γ) vérifiant l'identité (E) ci-dessus. La loi de composition est donnée par l'identité (C) écrite aussi plus haut. On a un épimorphisme

$$ {}_\varepsilon O^{\text{él}}(E, \phi_0) \to {}_\varepsilon O^{\min}(E, \phi_0) $$

dont le noyau est égal au groupe abélien ${}_\varepsilon S(E)$ formé des morphismes $\gamma : E \to E^*$ tels que ${}^t\gamma = \varepsilon\gamma$. Nous obtenons ainsi une extension de groupes non triviale en général

$$ 1 \longrightarrow {}_\varepsilon S(E) \longrightarrow {}_\varepsilon O^{\text{él}}(E, \phi_0) \longrightarrow {}_\varepsilon O^{\min}(E, \phi_0) \longrightarrow 1 $$

Celle-ci justifie la terminologie adoptée de “groupe orthogonal élargi”.

Pour les calculs, il est commode d'identifier E à son dual par la forme hermitienne ϕ associée à ϕ_0. Le morphisme γ est alors remplacé par un endomorphisme $u = \phi^{-1} \cdot \gamma$ de E. On peut de même remplacer ϕ_0 par $\psi = \phi^{-1} \cdot \phi_0$. On a alors $\psi^* = \phi^{-1} \cdot {}^t\psi \cdot \phi = \phi^{-1} \cdot ({}^t\phi_0 \cdot \overline{\varepsilon}\phi^{-1} \cdot \phi) = \phi^{-1} \cdot (\phi - \phi_0) = 1 - \psi$. La relation (E) ci-dessus s'écrit alors $f^* \cdot \psi \cdot f = \psi + u - u^*$ ou encore $f^{-1}.\psi.f = \psi + u - u^*$ puisque f est unitaire.

Grâce à cette traduction, la loi de composition dans ${}_\varepsilon O^{\text{él}}(E, \phi_0)$ s'écrit simplement

$$ (f, u) \cdot (g, v) = (f \cdot g, v + g^* \cdot u \cdot g) = (f \cdot g, v + g^{-1} \cdot u \cdot g) $$

Le noyau ${}_\varepsilon S(E)$ de l'homomorphisme surjectif ${}_\varepsilon O^{\text{él}}(E, \phi_0) \to {}_\varepsilon O^{\min}(E, \phi_0)$ s'identifie à l'ensemble des morphismes auto-adjoints de E, noté simplement $S(E)$. L'extension précédente s'écrit alors de manière équivalente

$$ 1 \to S(E) \longrightarrow {}_\varepsilon O^{\text{él}}(E) \longrightarrow {}_\varepsilon O^{\min}(E) \longrightarrow 1 $$

Dans cette extension, le groupe ${}_\varepsilon O^{\min}(E)$ opère à droite sur $S(E)$ par la formule suivante :

$$ (u, g) \mapsto g^{-1} \cdot u \cdot g $$

1.8. Si 1 est scindé dans A, on peut définir une section s de cette extension en posant

$$ s(g) = (g, \lambda(g^* \cdot \psi \cdot g - \psi)) = (g, \lambda(g^{-1} \cdot \psi \cdot g - \psi)) $$

Il en résulte que le groupe orthogonal élargi s'identifie au produit semi-direct du groupe orthogonal ${}_\varepsilon O(E)$ par le groupe additif $S(E)$, grâce à l'action définie cidessus. Une autre façon de voir les choses est d'introduire l'anneau des nombres duaux $A(e)$ avec $e^2 = 0$ et $\overline{e} = -e$ puis d'étendre les scalaires à $A(e)$. Nous savons déjà que le groupe orthogonal ${}_\varepsilon O^{\min}(E)$ s'identifie au groupe unitaire ${}_\varepsilon O^{\max}(E)$. Par ailleurs, l'épimorphisme

$$ O^{\max}(E(e)) \to O^{\max}(E) $$

a comme noyau l'ensemble des matrices unitaires du type $1+ue$, c'est-à-dire vérifiant l'identité $(1 + ue)(1 - u^*e) = 1 + (u - u^*)e = 1$, soit $u = u^*$. Le groupe unitaire opère sur ce noyau par l'action à droite définie par la même action : $(u, g) \mapsto g^{-1}ug$. Il en résulte que le groupe orthogonal élargi $O^{\text{él}}(E)$ s'identifie à $O^{\max}(E(e))$ en tant que produit semi-direct.

1.9. Considérons le cas particulier où $A = B \times B^{op}$, B^{op} étant l'anneau opposé à B, l'involution permutant les facteurs du produit. Si nous posons $\lambda = (1,0)$, on a $\lambda + \overline{\lambda} = 1$, ce qui montre que 1 est scindé dans A. Il est facile de voir que la donnée d'un A-module hermitien équivaut à celle d'un B-module. Les catégories ${}_\varepsilon\mathcal{Q}^{\max}(A)$ et ${}_\varepsilon\mathcal{Q}^{\min}(A)$ sont donc toutes les deux équivalentes à la catégorie $\mathcal{P}(B)$ (avec les isomorphismes comme morphismes). D'après 1.7, nous en déduisons que les catégories ${}_\varepsilon\mathcal{Q}^{\text{él}}(A)$ et $\mathcal{P}(B(e))$ sont équivalentes. En effet, nous avons montré en 1.7 que le groupe des automorphismes d'un objet de ${}_\varepsilon\mathcal{Q}^{\text{él}}(A)$ est le même que celui des automorphismes du B-module correspondant, vu comme un objet de $B(e)$ par extension des scalaires. Puisque les classes d'isomorphie d'objets de $\mathcal{P}(B(e))$ coïncident avec les classes d'isomorphie d'objets de $\mathcal{P}(B)$, l'assertion résulte de considérations générales sur les équivalences de catégories.

1.10. Rappelons maintenant la définition du foncteur hyperbolique classique

$$H : \mathcal{P}(A) \to {}_\varepsilon\mathcal{Q}^{\min}(A)$$

Si E est un objet de $\mathcal{P}(A)$, $H(E)$ est le A-module $E \oplus E^*$ muni de la forme quadratique

$$\varphi_0 : E \oplus E^* \to (E \oplus E^*)^* \approx E^* \oplus E$$

définie par la matrice

$$\varphi_0 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

Si u est un isomorphisme dans la catégorie $\mathcal{P}(A)$, on définit $H(u) = g = u \oplus {}^t u^{-1}$. On vérifie que ${}^t g.\varphi_0.g = \varphi_0$ et que $H(u)$ est donc bien un isomorphisme dans la catégorie ${}_\varepsilon\mathcal{Q}^{\min}(A)$. On peut décrire ce foncteur de manière plus conceptuelle en considérant l'anneau $\Lambda = M_2(A)$ des matrices 2×2 à coefficients dans A et où l'involution est définie par

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} \overline{d} & \overline{b} \\ \overline{c} & \overline{a} \end{pmatrix}$$

L'équivalence de Morita démontrée dans [**B1**] §9 montre que les catégories ${}_\varepsilon\mathcal{Q}^{\min}(\Lambda)$ et ${}_\varepsilon\mathcal{Q}^{\min}(A)$ sont équivalentes. On démontre par la même méthode que les catégories ${}_\varepsilon\mathcal{Q}^{\max}(\Lambda)$ et ${}_\varepsilon\mathcal{Q}^{\max}(A)$ d'une part et les catégories ${}_\varepsilon\mathcal{Q}^{\text{él}}(\Lambda)$ et ${}_\varepsilon\mathcal{Q}^{\text{él}}(A)$ d'autre part sont équivalentes. Le foncteur hyperbolique $\mathcal{P}(A) \to {}_\varepsilon\mathcal{Q}^{\min}(A)$ est alors induit par l'homomorphisme d'anneaux $A \times A^{op} \to M_2(A)$ défini par

$$(a,b) \mapsto \begin{pmatrix} a & 0 \\ 0 & \overline{b} \end{pmatrix}$$

D'après 1.8, cette méthode a l'avantage de définir un nouveau foncteur hyperbolique de $\mathcal{P}(A)$ dans la catégorie plus fine ${}_\varepsilon\mathcal{Q}^{\text{él}}(A)$ par la composition des foncteurs évidents suivants induits par des morphismes d'anneaux ou des équivalences de Morita :

$$\mathcal{P}(A) \to \mathcal{P}(A(e)) \sim {}_\varepsilon\mathcal{Q}^{\text{él}}(A \times A^{op}) \to {}_\varepsilon\mathcal{Q}^{\text{él}}(M_2(A)) \sim {}_\varepsilon\mathcal{Q}^{\text{él}}(A)$$

On procède de même pour le foncteur "oubli" ${}_\varepsilon\mathcal{Q}^{\text{él}}(A) \to \mathcal{P}(A)$ qui est la composition

$${}_\varepsilon\mathcal{Q}^{\text{él}}(A) \to {}_\varepsilon\mathcal{Q}^{\text{él}}(A \times A^{op}) \sim \mathcal{P}(A(e)) \to \mathcal{P}(A)$$

2. Les groupes de Grothendieck et Bass en $K-$théorie hermitienne.

2.1. Aux catégories précédentes ${}_\varepsilon\mathcal{Q}^{\max}(A)$, ${}_\varepsilon\mathcal{Q}^{\min}(A)$ et ${}_\varepsilon\mathcal{Q}^{\text{él}}(A)$, nous pouvons associer trois groupes de $K-$théorie hermitienne notés ${}_\varepsilon K\mathcal{Q}^{\max}(A)$, ${}_\varepsilon K\mathcal{Q}^{\min}(A)$ et ${}_\varepsilon K\mathcal{Q}^{\text{él}}(A)$ respectivement, reliés par des homomorphismes canoniques

$$ {}_\varepsilon K\mathcal{Q}^{\text{él}}(A) \xrightarrow{u} {}_\varepsilon K\mathcal{Q}^{\min}(A) \xrightarrow{v} {}_\varepsilon K\mathcal{Q}^{\max}(A) \ , $$

Il est clair que u est un isomorphisme et que v est surjectif. Par ailleurs, nous pouvons définir l'analogue du groupes de Bass $K_1(A)$ en $K-$théorie hermitienne. Dans ce but, le lemme suivant, dont la démonstration est détaillée dans [**KV**] p. 61 par exemple, est essentiel.

2.2. Lemme. *Tout module ε-quadratique est facteur direct d'un module hyperbolique.*

2.3. Puisque tout module projectif de type fini est facteur direct d'un module libre du type A^n, on voit que les groupes classiques qui jouent le rôle de $GL_n(A)$ sont les groupes d'automorphismes de modules hyperboliques du type $H(A^n)$ dans chacune des trois catégories concernées.
Plus précisément, écrivons $E = M \oplus M^*$ (on considèrera le cas où $M = A^n$ un peu plus tard). La forme quadratique associée est définie par la matrice ϕ_0 précédente avec ψ comme forme hermitienne associée, soit

$$ \phi_0 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \ \phi = \begin{pmatrix} 0 & 1 \\ \varepsilon & 0 \end{pmatrix} $$

Si $f : E \to E$ est un homomorphisme défini par une matrice $f = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, son adjoint est la matrice $f^* = \begin{pmatrix} {}^t d & \varepsilon {}^t b \\ \varepsilon {}^t c & {}^t a \end{pmatrix}$.
Dans le cas où $M = A^n$, il convient de remplacer la notation ${}^t u$ par ${}^t\overline{u}$, si on écrit u comme une matrice $n \times n$. En effet, la conjugaison résulte de l'identification de A^n avec son dual $(A^n)^*$.

2.4. Notations. On désigne par ${}_\varepsilon O^{\max}_{n,n}(A)$ (resp. ${}_\varepsilon O^{\min}_{n,n}(A)$, ${}_\varepsilon O^{\text{él}}_{n,n}(A)$) le groupe unitaire (resp. orthogonal, orthogonal élargi) associé au module hyperbolique $(A)^n \oplus (A^n)^*$.

2.5. Exemple. Supposons que A soit un corps muni de l'involution triviale et que $\varepsilon = 1$. Le fait que f soit unitaire ($f \in {}_1O^{\max}_{n,n}(A)$) se traduit par les identités suivantes (où a, b, c et d sont des matrices $n \times n$) :

$$ \begin{aligned} a \cdot {}^t d + b \cdot {}^t c &= 1 \\ a \cdot {}^t b + b \cdot {}^t a &= 0 \\ c \cdot {}^t d + d \cdot {}^t c &= 0 \\ c \cdot {}^t b + d \cdot {}^t a &= 1 \end{aligned} $$

L'automorphisme f est orthogonal ($f \in {}_1O^{\min}_{n,n}(A)$) s'il existe en outre des matrices h et k telles que $a \cdot {}^t b = h - {}^t h$ et $c \cdot {}^t d = k - {}^t k$.

Pour décrire un élément du groupe orthogonal élargi, il faut se donner en outre un endomorphisme défini par une matrice $2n \times 2n$,

$$u = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$$

liée à f et à la forme ϕ_0 (cf. 1.6/7). Plus précisément, le couple (f, u) doit vérifier l'identité suivante

$$\begin{pmatrix} {}^t d \cdot a & {}^t b \cdot d \\ {}^t c \cdot a & {}^t c \cdot b \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} - \begin{pmatrix} {}^t\delta & {}^t\beta \\ {}^t\gamma & {}^t\alpha \end{pmatrix}$$

Elle résulte de l'équation (E) en 1.7, à condition d'identifier $E \oplus E^*$ à $E^* \oplus E$ (avec $E = A^n$).

2.6. Revenons au cas général d'un anneau A quelconque. Pour simplifier, nous écrirons ${}_\varepsilon O_{n,n}(A)$ au lieu de ${}_\varepsilon O^{\max}_{n,n}(A)$, ${}_\varepsilon O^{\min}_{n,n}(A)$, ${}_\varepsilon O^{\text{él}}_{n,n}(A)$ en revenant à ces notations spécifiques lorsqu'il sera nécessaire de distinguer les trois groupes. De même, nous utiliserons la terminologie uniforme "groupe orthogonal" au lieu de "groupe unitaire", "groupe orthogonal" ou "groupe orthogonal élargi", lorsque nos considérations s'appliquent aux trois variantes. Avec ces conventions, le groupe orthogonal infini ${}_\varepsilon O(A)$ est défini comme la limite inductive des groupes ${}_\varepsilon O_{n,n}(A)$ avec les inclusions évidentes. En suivant l'exemple du groupe linéaire, nous définissons le "groupe de Bass" ${}_\varepsilon K\mathcal{Q}_1(A)$ comme le quotient de ${}_\varepsilon O(A)$ par le sous-groupe des commutateurs $[{}_\varepsilon O(A), {}_\varepsilon O(A)]$. Le fait que ce sous-groupe soit parfait résulte de considérations bien connues sur la stabilisation des matrices qu'on peut résumer par des identités générales. La première est la suivante :

$$\begin{pmatrix} \alpha\beta\alpha^{-1}\beta & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha^{-1} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \beta^{-1} \end{pmatrix} \begin{pmatrix} \alpha^{-1} & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \beta^{-1} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \beta \end{pmatrix}$$

Par ailleurs, modulo le sous-groupe des commutateurs, une matrice du type

$$\begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha^{-1} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

peut aussi s'écrire

$$\begin{pmatrix} \alpha & 0 & 0 \\ 0 & \alpha^{-1} & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & \alpha & 0 \\ 0 & 0 & \alpha^{-1} \\ 1 & 0 & 0 \end{pmatrix}$$

qui est le commutateur suivant

$$\begin{pmatrix} \alpha & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha^{-1} & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Toutes ces identités (qui sont vraies dans le cadre plus général de catégorie monoïdales symétriques) démontrent bien que $[{}_\varepsilon O(A), {}_\varepsilon O(A)]$ est parfait. Pour chacune des trois théories considérées, on utilisera les notations ${}_\varepsilon K\mathcal{Q}^{\max}_1(A)$, ${}_\varepsilon K\mathcal{Q}^{\min}_1(A)$, ${}_\varepsilon K\mathcal{Q}^{\text{él}}_1(A)$ ou simplement ${}_\varepsilon K\mathcal{Q}_1(A)$.

2.7. Théorème . *Considérons un carré cartésien d'anneaux hermitiens (avec φ_1 surjectif)*

$$\begin{tikzcd} A \arrow[r, "\psi_1"] \arrow[d, "\varphi_2"'] & A_1 \arrow[d, "\varphi_1"] \\ A_2 \arrow[r, "\psi_2"'] & A' \end{tikzcd}$$

On a alors une suite exacte (dite de Mayer-Vietoris) entre les groupes de K-théorie hermitienne

$$\begin{aligned} {}_{\varepsilon}K\mathcal{Q}_1(A) \longrightarrow {}_{\varepsilon}K\mathcal{Q}_1(A_1) \oplus {}_{\varepsilon}K\mathcal{Q}_1(A_2) \longrightarrow {}_{\varepsilon}K\mathcal{Q}_1(A') \longrightarrow {}_{\varepsilon}K\mathcal{Q}(A) \longrightarrow \cdots \\ \cdots \longrightarrow {}_{\varepsilon}K\mathcal{Q}(A_1) \oplus {}_{\varepsilon}K\mathcal{Q}(A_2) \longrightarrow {}_{\varepsilon}K\mathcal{Q}(A') \end{aligned}$$

Démonstration. Ce théorème classique peut être démontré de diverses manières. L'une d'entre eux est esquissée dans le livre de Milnor [**M**] et détaillée dans celui de Bak [**B1**]. Une autre démonstration est indiquée dans [**KV**] p. 68-70 (elle s'applique dans les trois situations). Le point important est de remarquer qu'un élément du sous-groupe des commutateurs $[{}_{\varepsilon}O(A'), {}_{\varepsilon}O(A')]$ se relève en un élément de ${}_{\varepsilon}O(A_1)$. Ceci est démontré grâce au lemme de Whitehead classique adapté au cas hermitien (cf. [**KV**] théorème 2.6 par exemple).

2.8. Dans [**B1**] p. 191, Bak démontre une suite exacte intéressante reliant les groupes ${}_{\varepsilon}K\mathcal{Q}^{\max}$ et ${}_{\varepsilon}K\mathcal{Q}^{\min}$. Elle s'écrit

$${}_{\varepsilon}K\mathcal{Q}_1^{\min}(A) \longrightarrow {}_{\varepsilon}K\mathcal{Q}_1^{\max}(A) \longrightarrow {}_{\varepsilon}\Xi(A) \longrightarrow {}_{\varepsilon}K\mathcal{Q}^{\min}(A) \longrightarrow {}_{\varepsilon}K\mathcal{Q}^{\max}(A)$$

Le groupe de 2-torsion ${}_{\varepsilon}\Xi(A)$ est explicité ainsi. Nous définissons d'abord $\Gamma = \Gamma(A)$ comme l'ensemble des éléments a de A tels que $\overline{a} = \varepsilon a$ et Λ comme le sous-groupe de Γ formé des $b - \varepsilon\overline{b}$. Alors ${}_{\varepsilon}\Xi(A)$ est le quotient de $\Gamma/\Lambda \otimes_A \Gamma/\Lambda$ par le sous-groupe engendré par tous les éléments de la forme

$$\{a\otimes b - b\otimes a\} \text{ et } \{a\otimes b - a\otimes ba\overline{b}\}$$

Dans la définition du produit tensoriel $\Gamma/\Lambda\otimes_A\Gamma/\Lambda$, l'action à droite de A sur Γ/Λ est $(\gamma, a)\mapsto\overline{a}\gamma a$. L'action à gauche est définie de manière similaire par $(a,\gamma)\mapsto a.\gamma.\overline{a}$ Un théorème plus général est en fait énoncé dans [**B1**] en utilisant des "formes paramètres" arbitraires Γ et Λ.

2.9. Remarque. La suite exacte précédente permet de définir un invariant des formes quadratiques proche de l'invariant de Arf en considérant des corps de caractéristique 2 (cf. [**B2**]). Dans ce cas, le groupe $K\mathcal{Q}_1^{\max}(A)$ est réduit à 0, $K\mathcal{Q}^{\max}(A) \cong \mathbb{Z}$ et le noyau de la flèche

$$K\mathcal{Q}^{\min}(A) \to K\mathcal{Q}^{\max}(A) = \mathbb{Z}$$

s'identifie ainsi au groupe $\Xi(A)$ précédent : c'est le quotient de $A\otimes_{\mathbb{Z}}A$ par le sous-groupe engendré par les relations $\{a\otimes b - b\otimes a\}$, $\{a\otimes b - a\otimes b^2a\}$ et $\{c^2a\otimes b - a\otimes c^2b\}$. L'invariant de Arf classique est obtenu par l'application $a\otimes b\mapsto a\cdot b$: elle est à valeurs dans le quotient G de F par le sous-groupe additif engendré par les relations $\{a^2 - a\}$. Cette application de $\Xi(A)$ dans G admet une rétraction induite par l'application $a\mapsto 1\otimes a$.

3. Les groupes de $K-$théorie hermitienne ${}_\varepsilon KQ_n(A)$ pour $n < 0$ et $n > 0$.

3.1. Pour définir les groupes ${}_\varepsilon KQ_n$ pour $n < 0$, nous suivons le même schéma qu'en K-théorie algébrique [**KV**]. De manière précise, si on pose $n = -m$, on pose $KQ_n(A) = KQ(S^m A)$, où $S^m A$ est la $m^{\text{ième}}$ suspension de l'anneau A. Notons que l'isomorphisme

$$ {}_\varepsilon KQ^{\text{él}}(A) \cong {}_\varepsilon KQ^{\min}(A) $$

implique par suspensions itérées l'isomorphisme

$$ {}_\varepsilon KQ^{\text{él}}_{-m}(A) \cong {}_\varepsilon KQ^{\min}_{-m}(A) $$

Le théorème suivant est moins évident.

3.2. Théorème . *L'homomorphisme*

$$ {}_\varepsilon KQ^{\min}_n(A) \to {}_\varepsilon KQ^{\max}_n(A) $$

est surjectif pour $n = 0$, bijectif pour $n < 0$.

Démonstration. La surjectivité pour tout n est une conséquence immédiate des définitions (car nous considérons des formes hermitiennes paires). Par induction sur n, il suffit de démontrer l'injectivité pour $n = -1$. Pour cela, écrivons la suite exacte 2.8, en remplaçant A par sa suspension SA et son cône CA. On obtient alors un diagramme commutatif

$$\begin{tikzcd}
{}_\varepsilon KQ_1^{\min}(CA) \arrow[r] \arrow[d] & {}_\varepsilon KQ_1^{\max}(CA) \arrow[r] \arrow[d] & {}_\varepsilon \Xi(CA) \arrow[r] \arrow[d] & {}_\varepsilon KQ^{\min}(CA) \arrow[r] \arrow[d] & {}_\varepsilon KQ^{\max}(CA) \arrow[d] \\
{}_\varepsilon KQ_1^{\min}(SA) \arrow[r] & {}_\varepsilon KQ_1^{\max}(SA) \arrow[r] & {}_\varepsilon \Xi(SA) \arrow[r] & {}_\varepsilon KQ^{\min}(SA) \arrow[r] & {}_\varepsilon KQ^{\max}(SA)
\end{tikzcd}$$

Puisque le cône d'un anneau est "flasque" (il existe un foncteur τ de la catégorie $\mathcal{P}(CA)$ dans elle-même tel que $\tau \oplus \mathrm{Id}$ soit isomorphe à τ), ses groupes de $K-$théorie hermitienne sont réduits à 0, ce qui implique que ${}_\varepsilon \Xi(CA)$ est aussi égal à 0. Pour démontrer l'injectivité de la flèche ${}_\varepsilon KQ^{\min}_{-1}(A) \to {}_\varepsilon KQ^{\max}_{-1}(A)$, il suffit donc de montrer que l'homomorphisme ${}_\varepsilon \Xi(CA) \to {}_\varepsilon \Xi(SA)$ est surjectif, ce qui est une conséquence du lemme suivant.

3.3. Lemme. *Notons $\Gamma(R)$ le groupe Γ défini en 2.8 pour tout anneau R. Alors l'homomorphisme canonique*

$$ \Gamma(CA) \to \Gamma(SA) $$

est surjectif.

Démonstration. Un élément de $\Gamma(SA)$ est défini par une matrice infinie M telle que sur chaque ligne et chaque colonne il n'existe qu'un nombre fini d'éléments non nuls et telle que ${}^t\overline{M} = \varepsilon M$ modulo une matrice finie. Soient a_{ij} les éléments (en nombre fini) de la matrice M tels que $\overline{a_{ij}} \neq \varepsilon a_{ji}$. Si on remplace ces éléments par 0, on trouve une matrice N dans CA qui est ε-hermitienne et dont la classe dans SA est égale à celle de M.

3.4. Définissons maintenant les groupes ${}_\varepsilon K\mathcal{Q}_n$ pour $n > 0$, ce qui est plus délicat. En principe, il suffit de copier la construction + de Quillen à l'espace $B_\varepsilon O(A)$, ce qui est possible car le sous-groupe des commutateurs $[{}_\varepsilon O(A), {}_\varepsilon O(A)]$ est parfait. On définit alors ${}_\varepsilon K\mathcal{Q}_n(A)$ comme le $n^{\text{ième}}$ groupe d'homotopie de $B_\varepsilon O(A)^+$ (pour $n > 0$). En fait, nous disposons de trois groupes de K-théorie hermitienne

$$ {}_\varepsilon K\mathcal{Q}_n^{\max}(A),\ {}_\varepsilon K\mathcal{Q}_n^{\min}(A) \text{ et } {}_\varepsilon K\mathcal{Q}_n^{\text{él}}(A) $$

associés respectivement aux groupes ${}_\varepsilon O^{\max}(A)$, ${}_\varepsilon O^{\min}(A)$ et ${}_\varepsilon O^{\text{él}}(A)$. Conformément à la philosophie de cet article, nous adopterons la notation uniforme ${}_\varepsilon K\mathcal{Q}_n(A)$ pour ne pas compliquer l'exposition, lorsqu'il n'y a pas de risque de confusion. Ces groupes sont difficiles à calculer en général, comme d'ailleurs les groupes $K_n(A)$ de Quillen dont ils sont la généralisation. Nous verrons cependant que, dans une certaine mesure, les "groupes de Witt supérieurs" ${}_\varepsilon W_n(A) = \operatorname{Coker}(K_n(A) \to {}_\varepsilon K\mathcal{Q}_n(A))$ sont plus accessibles.

3.5. Comme il est bien connu, il existe d'autres définitions des foncteurs K_n et ${}_\varepsilon K\mathcal{Q}_n$ équivalentes à la construction + de Quillen. La construction dite "$S^{-1}S$" (due aussi à Quillen) est détaillée dans le cadre hermitien dans [**K1**] §1 et nous l'utiliserons pour la preuve de 4.6. Il existe aussi une définition en termes de A-fibrés plats qui est détaillée dans [**K2**] p. 42 et c'est celle que nous utiliserons essentiellement ici. Rappelons-là brièvement dans le cadre que nous intéresse.
On définit un A-fibré hermitien "virtuel" sur un CW-complexe X comme la donnée d'une fibration acyclique $Y \to X$ et d'un A-fibré plat E sur Y, la fibre étant un A-module projectif de type fini muni d'une forme hermitienne dans l'un des trois sens que nous avons donnés à ce terme (ceci veut dire que les fonctions de transition du fibré sur Y sont des fonctions localement constantes dans chacune des trois catégories "max", "min" ou "él" concernées).

Deux tels fibrés virtuels

$$ E \to Y \to X \text{ et } E' \to Y' \to X $$

sont dits équivalents s'il existe un fibré virtuel $E_1 \to Y_1 \to X$ et un diagramme commutatif

$$\begin{tikzcd} Y \arrow[r, "f"] \arrow[d, "\sigma"'] & X \arrow[d, "f'"] \\ Y_1 \arrow[ur, "f_1"] & Y' \arrow[l, "\sigma'"] \end{tikzcd}$$

tel que $\sigma^*(E_1) \cong E$ et $\sigma'^*(E_1) \cong E'$.

En suivant le même schéma qu'en [**K2**] p. 42-50, on montre que le groupe de Grothendieck construit avec ces fibrés virtuels est isomorphe au groupe défini par les classes d'homotopie de X dans ${}_\varepsilon K\mathcal{Q}_0(A) \times B_\varepsilon O(A)^+$, noté ${}_\varepsilon K\mathcal{Q}_A(X)$, et qui est une "théorie cohomologique" en X. Si X est une sphère de dimension $n \geq 0$, on retrouve ainsi ${}_\varepsilon K\mathcal{Q}_n(A)$ comme le conoyau de la flèche évidente ${}_\varepsilon K\mathcal{Q}_0(A) \to {}_\varepsilon K\mathcal{Q}_A(X)$.

On peut définir le spectre de la $K-$théorie hermitienne par la même méthode qu'en $K-$théorie algébrique. Ainsi, dans [**K1**], on démontre l'analogue du théorème de Gersten-Wagoner [**W**] en $K-$théorie hermitienne : on a une équivalence d'homotopie (non naturelle) entre $\Omega(B_\varepsilon O(SA)^+)$ et ${}_\varepsilon K\mathcal{Q}_0(A) \times B_\varepsilon O(A)^+$ (la même démonstration s'applique dans les trois cas considérés ici). Plus précisément, on définit le Ω-spectre de la K-théorie hermitienne ${}_\varepsilon\mathbf{KQ(A)}_*$ par les formules suivantes :

$$\begin{array}{ll} {}_\varepsilon\mathbf{KQ(A)_n} = \mathbf{\Omega(B_\varepsilon O(S^{n+1}A)^+)} & \text{pour } n \geq 0 \\ {}_\varepsilon\mathbf{KQ(A)_n} = \mathbf{\Omega^{-n}(B_\varepsilon O(A)^+)} & \text{pour } n < 0 \end{array}$$

En fait, ce spectre n'est qu'un langage commode. Pour pouvoir définir des cup-produits en K-théorie hermitienne, nous nous servirons plutôt de la théorie cohomologique associée en termes de fibrés virtuellement plats comme nous l'avons explicité plus haut. D'ailleurs, une situation analogue se présente en $K-$théorie topologique, où les opérations sont plus aisément définies sur les fibrés vectoriels plutôt que sur la grassmannienne infinie.

4. Cup-produits en $K-$théorie hermitienne. Le cup-produit de Clauwens.

4.1. L'avantage du point de vue des fibrés plats est une définition très simple du cup-produit. Celui-ci est explicité dans [**K2**] à partir d'un morphisme $\mathbb{Z}$-bilinéaire

$$\varphi : A \times B \to C$$

vérifiant la propriété de multiplicativité suivante :

$$\varphi(aa', bb') = \varphi(a,b)\varphi(a',b')$$

Le cup-produit s'écrit alors sous la forme d'un accouplement bilinéaire

$$K_A(X) \times K_B(Y) \to K_{A\otimes B}(X \times Y)$$

où la flèche est simplement induite par le produit tensoriel des fibrés virtuellement plats. Si X est un espace muni d'un point base P, il est commode d'introduire la "K-théorie réduite" $\widetilde{K}_A(X) = Ker[K_A(X) \to K_A(P) = K_0(A)]$. Le produit précédent induit alors un "cup-produit réduit"

$$\widetilde{K}_A(X) \times \widetilde{K}_B(Y) \to \widetilde{K}_{A\otimes B}(X \wedge Y)$$

En particulier, si X (resp. Y) est une sphère S^n (resp. S^p) avec n et $p \geq 0$, on en déduit le cup-produit usuel en K-théorie algébrique (cf. aussi [**L**]).

4.2. Le même schéma s'applique en $K-$théorie hermitienne[4]. Par exemple, compte tenu des signes de symétrie, les cup-produits classiques sont schématisés par des accouplements

$${}_\varepsilon K\mathcal{Q}^{\max} \times {}_\eta K\mathcal{Q}^{\min} \to {}_{\varepsilon\eta}K\mathcal{Q}^{\min}$$

et

$${}_\varepsilon K\mathcal{Q}^{\max} \times {}_\eta K\mathcal{Q}^{\text{él}} \to {}_{\varepsilon\eta}K\mathcal{Q}^{\text{él}}$$

De manière précise, si nous considérons une ε-forme hermitienne paire $\phi = \phi_0 + \varepsilon\,{}^t\phi_0$ sur un A-module E et une forme η-quadratique définie par une classe de de morphismes ψ_0 sur un B-module F, alors $\phi \otimes \psi_0$ est une classe de forme $\varepsilon\eta$-quadratique sur $E \otimes F$. En outre, si α (resp. β) est un morphisme unitaire (resp. orthogonal) de E (resp. F), il est facile de voir que $\alpha \otimes \beta$ est un morphisme orthogonal de $E \otimes F$. De manière analogue, si (β, γ) est un morphisme dans la catégorie ${}_\eta\mathcal{Q}^{\text{él}}$, le

4. Á condition de supposer en outre que $\overline{\varphi(a,b)} = \varphi(\overline{a}, \overline{b})$

couple $(\alpha\otimes\beta,\alpha\otimes\gamma)$ définit un morphisme dans la catégorie ${}_{\varepsilon\eta}\mathcal{Q}^{\text{él}}$, ce qui définit le deuxième accouplement.
Ces deux cup-produits, définis en termes de modules, s'étendent naturellement aux fibrés plats ou virtuellement plats dans les catégories concernées (il convient de noter cependant que ψ_0 n'est pas donné dans la structure pour le premier accouplement mais seulement sa classe fibre par fibre). En considérant des fibrés plats sur des sphères homologiques, on définit ainsi des accouplements

$$ {}_{\varepsilon}K\mathcal{Q}_n^{\max}(A)\times{}_{\eta}K\mathcal{Q}_p^{\min}(B)\to{}_{\varepsilon\eta}K\mathcal{Q}_{n+p}^{\min}(C)$$

et

$$ {}_{\varepsilon}K\mathcal{Q}_n^{\max}(A)\times{}_{\eta}K\mathcal{Q}_p^{\text{él}}(B)\to{}_{\varepsilon\eta}K\mathcal{Q}_{n+p}^{\text{él}}(C)$$

4.3. Nous allons maintenant introduire un autre cup-produit plus subtil, dû essentiellement à Clauwens [**C**]. Celui-ci a été écrit par Clauwens pour les catégories de modules mais il s'étend aisément aux "bonnes" catégories de fibrés virtuellement plats munis de formes quadratiques. De manière précise, considérons d'abord la catégorie ${}_{\eta}Q^{\text{él}}(B)$ et la catégorie $Q^{\text{él}_0}(A\,[s])$ formée des $A\,[s]$-modules provenant de A par extension des scalaires, l'involution de A et la transformation $s\mapsto 1-s$.
Un objet de ${}_{\varepsilon}\mathcal{Q}^{\text{él}_0}(A[s])$ peut être décrit comme un couple (E,θ), où E est un objet de $\mathcal{P}(A)$ et θ une forme ε-quadratique sur $E\otimes_{\mathbb{Z}}\mathbb{Z}[s]$ s'écrivant sous la forme $\sum\theta_n s^n$, où θ_n est un morphisme de E vers E^*.
Considérons maintenant un objet (F,δ) de ${}_{\eta}\mathcal{Q}^{\text{él}}(B)$, où δ est une forme η -quadratique non dégénérée sur F avec $\Delta=\delta+\eta^t\delta$ comme forme hermitienne associée. Sur $E\otimes F$ on peut alors considérer la forme $\varepsilon\eta$-quadratique définie par la formule suivante

$$\kappa=\sum\theta_n\otimes\Delta(\Delta^{-1}\delta)^n$$

Cette formule se simplifie si on identifie F et son dual par l'isomorphisme Δ, ce qui revient à remplacer $\Delta^{-1}\delta$ par δ. On peut de même identifier E à E^* par l'isomorphisme $\theta_0+\sum_{n=0}^{\infty}{}^t\theta_n$. Le foncteur de dualité $f\mapsto{}^tf$ est alors remplacée par le foncteur d'adjonction $f\mapsto f^*$. Un avantage de cette formulation est aussi de se débarrasser des signes de symétrie. La formule précédente s'écrit alors sous une forme plus simple

$$\kappa=\sum\theta_n\otimes\delta^n$$

avec $\delta^*=1-\delta$. En quelques lemmes fondamentaux (cf. [**C**] p. 43 et 44 et aussi l'appendice, où on écrit ϕ au lieu de $\Delta^{-1}\delta$ pour éviter toute confusion), Clauwens montre que l'accouplement précédent

$$Obj({}_{\varepsilon}\mathcal{Q}^{\text{él}_0}(A[s]))\times Obj({}_{\eta}\mathcal{Q}^{\text{él}}(B))\to Obj({}_{\varepsilon\eta}\mathcal{Q}^{\text{él}}(A\otimes B))$$

est bien défini sur les classes d'isomorphie de modules quadratiques élargis. En fait, Clauwens considère dans son article des modules libres mais sa méthode est plus générale, comme nous l'explicitons dans l'appendice. Nous pouvons même aller un peu plus loin en interprétant cette correspondance comme un produit tensoriel. De manière précise, si E' est un A-module "élargi" et si F est un B-module élargi, F peut être vu come un $\mathbb{Z}\,[s]$-module par l'action de l'endomorphisme $(\Delta^{-1}\delta)$ ci-dessus. La correspondance $(E',F)\mapsto E'\otimes_{\mathbb{Z}[s]}F$ définit alors un accouplement plus général

$$Ob({}_{\varepsilon}Q^{\text{él}}(A\,[s])\times Ob({}_{\eta}Q^{\text{él}}(B))\longrightarrow Ob({}_{\epsilon\eta}Q^{\text{él}}(A\otimes B))$$

En particulier, nous pouvons définir un cup-produit remarquable

$$ {}_{\varepsilon}KQ^{\text{él}}(A\,[s]) \times {}_{\eta}KQ^{\text{él}}(B) \longrightarrow {}_{\varepsilon\eta}KQ^{\text{él}}(A \otimes B) $$

Dans les considérations précédentes, nous aurions pu remplacer la catégorie $\mathcal{Q}^{\text{él}}$ par la catégorie plus simple $\mathcal{Q}^{\min}$. La raison pour travailler dans la catégorie $\mathcal{Q}^{\text{él}}$ est notre souhait de généraliser l'acccouplement défini sur les groupes $K\mathcal{Q}_0$ aux groupes $K\mathcal{Q}_n$ définis dans le §3 pour $n > 0$. Si nous choisissons la définition de la $K-$théorie hermitienne en termes de fibrés plats, il nous faut montrer par exemple que la classe d'isomorphie de la forme quadratique κ définie plus haut ne dépend que des classes de θ et de δ. Les lemmes de Clauwens (redémontrés en appendice) montrent la nécessité de se donner le morphisme γ dans la formule (S) en 1.4. Grâce à ce nouveau point du vue, on peut étendre le cup-produit précédent aux groupes de K-théorie supérieurs (dans la catégorie "él") soit

$$ {}_{\varepsilon}KQ_n^{\text{él}}(A\,[s]) \times {}_{\eta}KQ_p^{\text{él}}(B) \longrightarrow {}_{\varepsilon\eta}KQ_{n+p}^{él}(A \otimes B) $$

4.4. Ce cup-produit sur les groupes $KQ^{\text{él}}$ supérieurs nécessite plus d'explications. En effet, il convient de remarquer que les accouplements précédents entre fibrés virtuellement plats sont simplement définis au niveau des classes d'isomorphie d'objets ; ils ne sont pas fonctoriels par rapport à la 2^{e} variable. Un point qui mérite d'être vérifié avec soin est donc la locale trivialité du produit tensoriel de fibrés plats (dans la catégorie "él") sur un espace X. Soit donc (U_i) un recouvrement trivialisant de X pour des fibrés plats E et F (dans la catégorie "él"). On a donc des trivialisations $\varphi_i : E_{U_i} \longrightarrow T_i$ et $\psi_i : F_{U_i} \longrightarrow T_i'$, où T et T' sont des fibrés triviaux. Grâce aux lemmes de Clauwens adaptés à la catégorie des fibrés plats (voir l'appendice, notamment le lemme 7.4), nous déduisons des trivialisations φ_i et ψ_i une trivialisation de $E_{U_i} \otimes F_{U_i} = (E \otimes F)_{U_i}$, c'est à dire un isomorphisme explicite

$$ \gamma_i : (E \otimes F)_{U_i} \longrightarrow T_i \otimes T_i' $$

De ces différents isomorphismes (lorsque i varie), on déduit des fonctions de transitions $\gamma_j \cdot \gamma_i^{-1}$ pour le fibré $E \otimes F$, assez compliquées cependant. Ainsi les lemmes démontrés dans l'appendice pour des classes d'isomorphie de modules se transcrivent sans problème aux classes d'isomorphie de fibrés plats.

4.5. Au début de son article (théorème 1, p. 42), Clauwens montre que modulo l'addition de A-modules hyperboliques (voir l'appendice pour un énoncé précis), on peut se ramener au cas où θ est "linéaire", i.e. du type $\theta = gs$. En d'autres termes, $\theta_n = 0$, à l'exception de θ_1 qui est égal à g. Puisque la forme hermitienne associée $gs + \varepsilon^t g(1-s)$ est un isomorphisme, ceci implique que ${}^t g = \varepsilon g(1+N)$, où N est un endomorphisme nilpotent de E (un tel g est dit "presque hermitien"). Dans ce cas, la formule pour la forme quadratique κ ci-dessus est très simple : on trouve

$$ \kappa = g \otimes \delta $$

(si on identifie F à son dual par Δ) En d'autres termes, l'accouplement précédent sur les groupes $K\mathcal{Q}^{\text{él}}$ généralise (pour $N = 0$) l'accouplement classique entre les formes hermitiennes (non nécessairement paires) et les formes quadratiques. Un cas particulier important est le cup-produit

$$ {}_{1}K\mathcal{Q}_1^{\text{él}}(S\mathbb{Z}\,[s]) \times {}_{\eta}K\mathcal{Q}_p^{\text{él}}(B) \longrightarrow {}_{\eta}K\mathcal{Q}_{1+p}^{él}(S\mathbb{Z} \otimes B) = {}_{\eta}\mathcal{K}Q_{1+p}^{\text{él}}(SB) $$

4.6. Théorème . *Soit u_1 l'élement de ${}_1K\mathcal{Q}_1^{\text{él}_0}(S\mathbb{Z}[s]) = {}_1K\mathcal{Q}_1^{\text{él}}(S\mathbb{Z}[s])$ correspondant à l'élément unité dans $K\mathcal{Q}_0(\mathbb{Z}[s]) = \mathbb{Z}\times\mathbb{Z}$ (cf. [**C**], p. 47). Alors le cup-produit par u_1 induit un isomorphisme entre ${}_\eta K\mathcal{Q}_p^{\text{él}}(B)$ et ${}_\eta K\mathcal{Q}_{1+p}^{\text{él}}(SB)$*

Démonstration. Elle est analogue à celle en $K-$théorie algébrique ou hermitienne classique (cf. [**K1**] p. 224).

4.7. Rappelons par ailleurs qu'un autre cup-produit plus simple a été défini en 4.2 :

$$ {}_\varepsilon K\mathcal{Q}_n^{\max}(A) \times {}_\eta K\mathcal{Q}_p^{\text{él}}(B) \to {}_{\varepsilon\eta}K\mathcal{Q}_{n+p}^{\text{él}}(A\otimes B)$$

Ces deux produits sont reliés ainsi :

4.8. Théorème . *Le cup-produit de Clauwens est partiellement associatif dans le sens suivant. Pour trois anneaux B, C et D, on a le diagramme commutatif (avec $n = n_1 + n_2$, $\varepsilon = \varepsilon_1\varepsilon_2$)*

$$\begin{tikzcd}
{}_{\varepsilon_1}K\mathcal{Q}_{n_1}^{\max}(C) \times {}_{\varepsilon_2}K\mathcal{Q}_{n_2}^{\text{él}}(D[s]) \times {}_\eta K\mathcal{Q}_p^{\text{él}}(B) \arrow[r] \arrow[d] & {}_{\varepsilon_1}K\mathcal{Q}_{n_1}^{\max}(C) \times {}_{\varepsilon_2\eta}K\mathcal{Q}_{n_2+p}^{\text{él}}(D\otimes B) \arrow[d] \\
{}_{\varepsilon_1\varepsilon_2}K\mathcal{Q}_{n_1+n_2}^{\text{él}}((C\otimes D)[s]) \times {}_\eta K\mathcal{Q}_p^{\text{él}}(B) \arrow[r] & {}_{\varepsilon\eta}K\mathcal{Q}_{n+p}^{\text{él}}(C\otimes D\otimes B)
\end{tikzcd}$$

Démonstration. C'est une conséquence directe de la formule donnée en 4.3. Nous devons multiplier les deux membres de la formule par la même forme hermitienne paire avant et après avoir fait le produit tensoriel par $\Delta(\Delta^{-1}\delta)^n$.

4.9. Remarque. Pour les degrés négatifs, nous avons seulement à considérer des modules sur des suspensions itérées des anneaux considérés. La notion de forme quadratique élargie est alors inutile dans les démonstrations. On peut même se limiter aux formes hermitiennes paires pour les degrés < 0 d'après 3.2.

4.10. Remarque. Si 1 est scindé dans A (par exemple si 2 est inversible), on a des isomorphismes $K\mathcal{Q}_n^{\text{él}}(A) \cong K\mathcal{Q}_n^{\max}(A(e)) \cong K\mathcal{Q}_n^{\min}(A(e))$ avec $\overline{e} = -e$.

5. Le théorème fondamental de la $K-$théorie hermitienne pour des anneaux arbitraires

5.1. Dans ce paragraphe, nous allons désigner le spectre de la K-théorie hermitienne ainsi que celui de la K-théorie algébrique par des caractères gras. De manière précise, $\mathbf{K}(\mathbf{A})$ représentera le spectre de la K-théorie algébrique usuelle ; celui de la K-théorie hermitienne sera représenté par l'un des trois spectres ${}_\varepsilon\mathbf{K}\mathcal{Q}^{\max}(\mathbf{A})$, ${}_\varepsilon\mathbf{K}\mathcal{Q}^{\min}(\mathbf{A})$ ou ${}_\varepsilon\mathbf{K}\mathcal{Q}^{\text{él}}(\mathbf{A})$, suivant la théorie considérée. En particulier, les foncteurs "oubli" et hyperbolique induisent des morphismes

$$ {}_\varepsilon\mathbf{K}\mathcal{Q}^{\text{él}}(\mathbf{A}) \to \mathbf{K}(\mathbf{A}) \text{ et } \mathbf{K}(\mathbf{A}) \to {}_{-\varepsilon}\mathbf{K}\mathcal{Q}^{\text{él}}(\mathbf{A})$$

dont les fibres homotopiques respectives seront notées ${}_\varepsilon\mathbf{V}^{\text{él}}(\mathbf{A})$ et ${}_{-\varepsilon}\mathbf{U}^{\text{él}}(\mathbf{A})$. L'énoncé suivant généralise le théorème de [**K2**] (p. 260).

5.2. Théorème . *Nous avons une équivalence d'homotopie naturelle*

$$ {}_\varepsilon\mathbf{V}^{\text{él}}(\mathbf{A}) \approx \mathbf{\Omega}_{-\varepsilon}\mathbf{U}^{\text{él}}(\mathbf{A})$$

5.3. Remarques. Le théorème est évident lorsque $A = B \times B^{op}$, une situation déjà considérée dans les paragraphes précédents. Dans ce cas, les spectres ${}_\varepsilon\mathbf{V}^{\text{él}}(\mathbf{A})$ et $\mathbf{\Omega}_{-\varepsilon}\mathbf{U}^{\text{él}}(\mathbf{A})$ coïncident tous les deux avec la fibre homotopique du morphisme évident $\mathbf{K}(\mathbf{B}(\mathbf{e})) \to \mathbf{K}(\mathbf{B})\mathbf{x}\mathbf{K}(\mathbf{B})$.

Par ailleurs, si 1 est scindé dans A, nous retrouvons le théorème fondamental de la K-théorie hermitienne énoncé dans [**K2**] p. 260 (cf. la remarque 5.11 un peu plus loin). La démonstration du théorème 5.2 va être en fait calquée sur celle de [**K2**]. Nous mentionnerons simplement ici les modifications à y apporter.

5.4. Rappelons d'abord le principe général de la démonstration dans [**K2**] que nous appliquerons à plusieurs reprises : un morphisme d'anneaux hermitiens $f : A \to B$ induit une application entre spectres

$$ {}_\varepsilon\mathbf{K}\mathcal{Q}^{\text{él}}(\mathbf{A}) \to {}_\varepsilon\mathbf{K}\mathcal{Q}^{\text{él}}(\mathbf{B}) $$

dont nous pouvons interpréter la fibre homotopique d'après un argument adapté de Wagoner [**W**]. Pour cela, on considère le produit fibré d'anneaux

$$\begin{tikzcd} R \arrow[r] \arrow[d] & CB \arrow[d] \\ SA \arrow[r] & SB \end{tikzcd}$$

d'où on déduit la fibration homotopique

$$ {}_\varepsilon\mathbf{K}\mathcal{Q}^{\text{él}}(\mathbf{R}) \longrightarrow {}_\varepsilon\mathbf{K}\mathcal{Q}^{\text{él}}(\mathbf{SA}) \longrightarrow {}_\varepsilon\mathbf{K}\mathcal{Q}^{\text{él}}(\mathbf{SB}) $$

car ${}_\varepsilon\mathbf{K}\mathcal{Q}^{\text{él}}(\mathbf{CB})$ est contractile. L'espace des lacets de ${}_\varepsilon\mathbf{K}\mathcal{Q}^{\text{él}}(\mathbf{R})$ est donc la fibre homotopique recherchée du morphisme

$$ {}_\varepsilon\mathbf{K}\mathcal{Q}^{\text{él}}(\mathbf{A}) \to {}_\varepsilon\mathbf{K}\mathcal{Q}^{\text{él}}(\mathbf{B}) $$

Deux cas importants peuvent être considérés. Dans le premier, le morphisme est $A \times A^{op} \to M_2(A)$ et dans le second $A \to A \times A^{op}$, tous les deux définis en 1.8. Si nous désignons [5] par U_A (resp. V_A) l'anneau R obtenu dans ces deux cas, nous voyons que ${}_\varepsilon\mathbf{U}^{\text{él}}(\mathbf{A})$ est homotopiquement équivalent à $\mathbf{\Omega}_\varepsilon\mathbf{K}\mathcal{Q}^{\text{él}}(\mathbf{U_A})$ et que ${}_\varepsilon\mathbf{V}^{\text{él}}(\mathbf{A})$ est homotopiquement équivalent à $\mathbf{\Omega}\mathbf{K}_\varepsilon\mathcal{Q}^{\text{él}}(\mathbf{V_A})$.

5.5. Nous souhaitons définir une application

$$ {}_\varepsilon\mathbf{V}^{\text{él}}(\mathbf{SA}) \to {}_{-\varepsilon}\mathbf{U}^{\text{él}}(\mathbf{A}) $$

L'idée, déjà présente dans [**K2**], est d'inclure cette application dans le diagramme suivant

$$\begin{tikzcd} {}_\varepsilon\mathbf{K}\mathcal{Q}^{\text{él}}(\mathbf{A}) \arrow[r] \arrow[d] & \mathbf{K}(\mathbf{A}) \arrow[r] \arrow[d] & {}_\varepsilon\mathbf{V}^{\text{él}}(\mathbf{SA}) \arrow[r] \arrow[d] & {}_\varepsilon\mathbf{K}\mathcal{Q}^{\text{él}}(\mathbf{SA}) \arrow[r] \arrow[d, "\sigma"] & \mathbf{K}(\mathbf{SA}) \arrow[r] \arrow[d] & {}_\varepsilon\mathbf{V}^{\text{él}}(\mathbf{S^2A}) \arrow[d] \\ {}_{-\varepsilon}\mathbf{D}^{\text{él}}(\mathbf{A}) \arrow[r] & \mathbf{K}(\mathbf{A}) \arrow[r] & {}_{-\varepsilon}\mathbf{U}^{\text{él}}(\mathbf{A}) \arrow[r] & {}_{-\varepsilon}\mathbf{D}^{\text{él}}(\mathbf{SA}) \arrow[r] & \mathbf{K}(\mathbf{SA}) \arrow[r] & {}_{-\varepsilon}\mathbf{U}^{\text{él}}(\mathbf{SA}) \end{tikzcd}$$

La théorie ${}_{-\varepsilon}\mathbf{D}^{\text{él}}(\mathbf{A})$ est ici la fibre homotopique de l'application $\mathbf{K}(\mathbf{A}) \to {}_{-\varepsilon}\mathbf{U}^{\text{él}}(\mathbf{A})$ qui est induite par le morphisme d'anneaux $A \times A^{op} \to M_2(A)$

5. En fait, pour la $K-$théorie, c'est à dire la $K-$théorie hermitienne de $A \times A^{op}$, nous devons remplacer l'anneau des nombres duaux $A(e)$ par A, comme il a été précisé en 1.8.

décrit précédemment. Pour compléter ce diagramme, nous utilisons un élément remarquable de ${}_{-1}D_0^{\max}(\mathbb{Z})$ et effectuons le "cup-produit" par cet élément pour définir une application naturelle $\sigma : {}_\varepsilon\mathbf{K}\mathcal{Q}^{\text{él}}(\mathbf{A}) \to {}_{-\varepsilon}\mathbf{D}^{\text{él}}(\mathbf{A})$. Les détails sont explicités en [**K2**] §2.3-8 (le fait que 1 soit éventuellement scindé dans A n'est pas nécessaire pour cet argument, comme il a été déjà souligné dans [**K2**]).

5.6. Nous procédons de manière symétrique pour construire une application en sens inverse ${}_{-\varepsilon}\mathbf{U}^{\text{él}}(\mathbf{A}) \to {}_\varepsilon\mathbf{V}^{\text{él}}(\mathbf{SA})$. Elle s'insère dans le diagramme commutatif suivant

$$\begin{tikzcd}
{} \arrow[r] & \mathbf{\Omega}_{-\varepsilon}\mathbf{K}\mathcal{Q}^{\text{él}}(\mathbf{A}) \arrow[r] \arrow[d] & {}_{-\varepsilon}\mathbf{U}^{\text{él}}(\mathbf{A}) \arrow[r] \arrow[d] & \mathbf{K}(\mathbf{A}) \arrow[r] \arrow[d] & {}_{-\varepsilon}\mathbf{K}\mathcal{Q}^{\text{él}}(\mathbf{A}) \arrow[r] \arrow[d, "\theta"] & {}_{-\varepsilon}\mathbf{U}^{\text{él}}(\mathbf{A}) \arrow[d] \\
{} \arrow[r] & {}_\varepsilon\mathbf{E}^{\text{él}}(\mathbf{SA}) \arrow[r] & {}_\varepsilon\mathbf{V}^{\text{él}}(\mathbf{SA}) \arrow[r] & \mathbf{K}(\mathbf{A}) \arrow[r] & {}_\varepsilon\mathbf{E}^{\text{él}}(\mathbf{S^2A}) \arrow[r] & {}_\varepsilon\mathbf{V}^{\text{él}}(\mathbf{S^2A})
\end{tikzcd}$$

La théorie ${}_\varepsilon\mathbf{E}^{\text{él}}(\mathbf{A})$ est ici la fibre homotopique de l'application composée

$${}_\varepsilon\mathbf{V}^{\text{él}}(\mathbf{A}) \to {}_\varepsilon\mathbf{V}^{\text{él}}(\mathbf{SA} \times \mathbf{SA^{op}}) = \mathbf{K}(\mathbf{A}(\mathbf{e})) \to \mathbf{K}(\mathbf{A})$$

Pour compléter le diagramme, nous devons définir une application

$$\theta : {}_{-\varepsilon}\mathbf{K}\mathcal{Q}^{\text{él}}(\mathbf{A}) \to {}_\varepsilon\mathbf{E}^{\text{él}}(\mathbf{S^2A})$$

L'idée nouvelle par rapport à [**K2**] est d'utiliser maintenant le cup-produit de Clauwens (écrit de manière relative pour la theorie E), soit

$${}_{-1}E^{\text{él}}_{-2}(\mathbb{Z}[s]) \times {}_{-\varepsilon}K\mathcal{Q}^{\text{él}}_n(A) \to {}_\varepsilon E^{\text{él}}_{n-2}(A)$$

(avec $\overline{s} = -s$).

Ceci se traduit au niveau des spectres par l'application θ. L'élément de ${}_{-1}E^{\text{él}}_{-2}(\mathbb{Z}[s]) = {}_{-1}K\mathcal{Q}^{\text{él}}_{-2}(\mathbb{Z}[s]) = {}_{-1}K\mathcal{Q}^{\min}_{-2}(\mathbb{Z}[s])$ avec lequel est effectué le cup-produit est écrit de manière explicite dans [**K1**] p. 243 par une matrice à 30 termes avec un léger changement de notations (remplacer la lettre λ par s). Nous devons ensuite plonger l'algèbre des polynômes laurentiens en les deux variables z et t dans la double suspension de $\mathbb{Z}[s]$.

Pour terminer la démonstration du théorème 5.2, nous devons montrer que les deux compositions

$${}_\varepsilon\mathbf{V}^{\text{él}}(\mathbf{SA}) \longrightarrow {}_{-\varepsilon}\mathbf{U}^{\text{él}}(\mathbf{A}) \longrightarrow {}_\varepsilon\mathbf{V}^{\text{él}}(\mathbf{SA}) \quad \text{et} \quad {}_{-\varepsilon}\mathbf{U}^{\text{él}}(\mathbf{A}) \longrightarrow {}_\varepsilon\mathbf{V}^{\text{él}}(\mathbf{SA}) \longrightarrow {}_{-\varepsilon}\mathbf{U}^{\text{él}}(\mathbf{A})$$

sont des équivalences d'homotopie. Nous nous référons de nouveau à [**K2**] p 273-277 pour le détail des arguments. Le point essentiel est l'associativité partielle du cup-produit établi en 4.7 qui remplace l'associativité usuelle utilisée en [**K2**]. En effet, de cette associativité partielle, on déduit des diagrammes commutatifs

$$\begin{tikzcd}
{}_1K\mathcal{Q}^{\text{él}}_0(\mathbb{Z}[s]) \times {}_\varepsilon K\mathcal{Q}^{\text{él}}_n(A) \arrow[r] \arrow[d] & {}_\varepsilon K\mathcal{Q}^{\text{él}}_n(A) \arrow[d] \\
{}_{-1}D^{\text{él}}_0(\mathbb{Z}[s]) \times {}_\varepsilon K\mathcal{Q}^{\text{él}}_n(A) \arrow[r] \arrow[d] & {}_{-\varepsilon}D^{\text{él}}_n(A) \arrow[d] \\
{}_1K\mathcal{Q}^{\text{él}}_0(\mathbb{Z}[s]) \times {}_\varepsilon K\mathcal{Q}^{\text{él}}_n(A) \arrow[r] & {}_\varepsilon K\mathcal{Q}^{\text{él}}_n(A)
\end{tikzcd}$$

Ce raisonnement montre que la composition

$${}_\varepsilon\mathbf{V}^{\text{él}}(\mathbf{SA}) \longrightarrow {}_{-\varepsilon}\mathbf{U}^{\text{él}}(\mathbf{A}) \longrightarrow {}_\varepsilon\mathbf{V}^{\text{él}}(\mathbf{SA})$$

est une équivalence d'homotopie. On démontre de même la commutativité du diagramme

$$\begin{array}{ccc}
{}_{-1}D_0^{\text{él}}(\mathbb{Z}[s]) \times {}_{\varepsilon}K\mathcal{Q}_n^{\text{él}}(A) & \longrightarrow & {}_{-\varepsilon}D_n^{\text{él}}(A) \\
\downarrow & & \downarrow \\
{}_{1}K\mathcal{Q}_0^{\text{él}}(\mathbb{Z}[s]) \times {}_{\varepsilon}K\mathcal{Q}_n^{\text{él}}(A) & \longrightarrow & {}_{\varepsilon}K\mathcal{Q}_n(A) \\
\downarrow & & \downarrow \\
{}_{-1}D_0^{\text{él}}(\mathbb{Z}[s]) \times {}_{\varepsilon}K\mathcal{Q}_n^{\text{él}}(A) & \longrightarrow & {}_{-\varepsilon}D_n^{\text{él}}(A)
\end{array}$$

ce qui montre que la composition en sens inverse

$$ {}_{-\varepsilon}\mathbf{U}^{\text{él}}(\mathbf{A}) \longrightarrow {}_{\varepsilon}\mathbf{V}^{\text{él}}(\mathbf{SA}) \longrightarrow {}_{-\varepsilon}\mathbf{U}^{\text{él}}(\mathbf{A})$$

est aussi une équivalence d'homotopie.

5.7. Remarque. Si nous nous intéressons uniquement aux "groupes de Witt élargis"

$$ {}_{\varepsilon}W_n^{\text{él}}(A) = \operatorname{Coker}(K_n(A) \to {}_{\varepsilon}K\mathcal{Q}_n^{\text{él}}(A))$$

les arguments précédents se simplifient considérablement (avec un résultat moins fort cependant ; à comparer avec 5.9 et 6.6). Le cup-produit par les éléments $u_2 \in {}_{-1}W_2^{\max}(\mathbb{Z})$ et $u_{-2} \in {}_{-1}W_{-2}^{\text{él}}(\mathbb{Z}[s])$, associés aux éléments construits en 5.5 et 5.6, définissent des homomorphismes

$$ {}_{\varepsilon}W_n^{\text{él}}(A) \to {}_{-\varepsilon}W_{n+2}^{\text{él}}(A) \text{ et } {}_{-\varepsilon}W_{n+2}^{\text{él}}(A) \to {}_{\varepsilon}W_n^{\text{él}}(A)$$

dont la composition (à isomorphisme près) est la multiplication par 4 (en utilisant des arguments de $K-$théorie topologique : cf. [**K1**], p. 251). Notons que ${}_{\varepsilon}W_n^{\text{él}}(A)$ est isomorphe à ${}_{\varepsilon}W_n^{\min}(A)$ si $n \leq 0$ et à ${}_{\varepsilon}W_n^{\max}(A)$ si $n < 0$. Le groupe de Witt "stabilisé" que nous définirons dans le §6 utilisera de manière essentielle le deuxième cup-produit.

5.8. Comme il a été explicité en [**K2**] p. 278, le théorème 5.2 implique une suite exacte à 12 termes dont les termes sont définis ainsi. Le "cogroupe de Witt" ${}_{\varepsilon}\overline{W_n}^{\text{él}}(A)$ est le noyau de la flèche oubli

$$ {}_{\varepsilon}K\mathcal{Q}_n^{\text{él}}(A)) \to K_n(A)$$

Nous définissons le groupe $k_n(A)$ (resp. $\overline{k}_n(A)$) comme le groupe de cohomologie de Tate pair (resp. impair) de $\mathbb{Z}/2$ opérant sur $K_n(A)$.

5.9. Théorème . *Avec les définitions précédentes, nous avons une suite exacte à 12 termes où, pour simplifier, nous écrivons F pour $F(A)$ en général, F étant l'un des foncteurs $W^{\text{él}}$, $\overline{W}^{\text{él}}$, $k^{\text{él}}$ ou $\overline{k}^{\text{él}}$*

$$\cdots \longrightarrow k_{n+1} \longrightarrow {}_{-\varepsilon}W_{n+2}^{él} \longrightarrow {}_{\varepsilon}\overline{W}_n^{él} \longrightarrow \overline{k}_{n+1}^{él} \longrightarrow {}_{-\varepsilon}\overline{W}_{n+1}^{él} \longrightarrow {}_{-\varepsilon}W_{n+1}^{él}$$

$$\longrightarrow k_{n+1} \longrightarrow {}_{\varepsilon}W_{n+2}^{él} \longrightarrow {}_{-\varepsilon}\overline{W}_n^{él} \longrightarrow \overline{k}_{n+1} \longrightarrow {}_{\varepsilon}\overline{W}_{n+1}^{él} \longrightarrow {}_{\varepsilon}W_{n+1}^{él} \longrightarrow k_{n+1} \ldots$$

5.10. Théorème . *Supposons que 1 soit scindé dans A (par exemple que 2 soit inversible). Les homomorphismes naturels*

$$ {}_{\varepsilon}W_n^{\text{él}}(A) \to {}_{\varepsilon}W_n(A) \text{ et } {}_{\varepsilon}\overline{W}_n^{\text{él}}(A) \to {}_{\varepsilon}\overline{W}_n(A) $$

sont alors des isomorphismes.

Démonstration. En raisonnant par récurrence sur n, c'est une conséquence immédiate de 5.9 et du théorème 4.3 de [**K2**] (voir aussi la remarque suivante).

5.11. Remarque. Si 1 est scindé dans A, nous avons un diagramme commutatif de spectres

$$\begin{array}{ccc} {}_{\varepsilon}\mathbf{V}^{\text{él}}(\mathbf{A}) & \approx & \mathbf{\Omega}_{-\varepsilon}\mathbf{U}^{\text{él}}(\mathbf{A}) \\ \uparrow & & \uparrow \\ {}_{\varepsilon}\mathbf{V}(\mathbf{A}) & \approx & \mathbf{\Omega}_{-\varepsilon}\mathbf{U}(\mathbf{A}) \end{array}$$

où les flèches verticales sont des monomorphismes scindés. On voit ainsi que le théorème 5.2 implique le théorème fondamental de [**K2**] p. 260. Nous profitons de cette occasion pour combler une lacune dans sa démonstration : elle supposait implicitement que ${}_{1}W(\mathbb{Z}[s]) \approx \mathbb{Z}$, un résultat dû aussi à Clauwens ([**C**] p. 47).

5.12. Nous allons conclure ce paragraphe par un calcul explicite de groupes de Witt dans des situations qui ne sont pas envisagées en [**K2**]. Nous remarquons d'abord que par la même méthode, nous pouvons définir en bas degrés des morphismes de périodicité

$$ {}_{\varepsilon}U^{\min}(A) \longrightarrow {}_{-\varepsilon}V^{\min}(SA) \quad \text{et} \quad {}_{-\varepsilon}V^{\min}(SA) \longrightarrow {}_{\varepsilon}U^{\min}(A) $$

inverses l'un de l'autre à isomorphisme près, en sorte que le diagramme suivant commute

$$\begin{array}{ccccc} {}_{\varepsilon}U^{\text{él}}(A) & \longrightarrow & {}_{-\varepsilon}V^{\text{él}}(SA) & \longrightarrow & {}_{\varepsilon}U^{\text{él}}(A) \\ \downarrow & & \downarrow & & \downarrow \\ {}_{\varepsilon}U^{\min}(A) & \longrightarrow & {}_{-\varepsilon}V^{\min}(SA) & \longrightarrow & {}_{\varepsilon}U^{\min}(A) \end{array}$$

En effet, la sophistication des fibrés plats n'est pas nécessaire dans cette situation. Par ailleurs, puisque ${}_{\varepsilon}K\mathcal{Q}^{\text{él}}(B)$ est isomorphe à ${}_{\varepsilon}K\mathcal{Q}^{\min}(B)$ pour tout anneau B, on déduit du diagramme précédent un isomorphisme ${}_{\varepsilon}U^{\text{él}}(A) \overset{\cong}{\to} {}_{\varepsilon}U^{\min}(A)$. Nous avons enfin le diagramme commutatif suivant de suites exactes

$$\begin{array}{ccccccccc} 0 & \longrightarrow & {}_{\varepsilon}W_1^{\text{él}}(A) & \longrightarrow & {}_{\varepsilon}U^{\text{él}}(A) & \longrightarrow & K(A) & \longrightarrow & {}_{\varepsilon}K\mathcal{Q}^{\text{él}}(A) \\ & & \downarrow & & \downarrow & & \downarrow & & \downarrow \\ 0 & \longrightarrow & {}_{\varepsilon}W_1^{\min}(A) & \longrightarrow & {}_{\varepsilon}U^{\min}(A) & \longrightarrow & K(A) & \longrightarrow & {}_{\varepsilon}K\mathcal{Q}^{\min}(A) \end{array}$$

Puisque les trois flèches de droite verticales sont des isomorphismes, nous en déduisons le théorème suivant

5.13. Théorème . *Pour tout anneau A, l'homomorphisme naturel*

$$ {}_{\varepsilon}W_1^{\text{él}}(A) \to {}_{\varepsilon}W_1^{\min}(A) $$

est un isomorphisme.

5.14. Exemple. Soit $A = \mathbb{F}_q$ un corps fini de caractéristique 2. D'après Quillen, les groupes $K_n(\mathbb{F}_q)$ sont des groupes finis d'ordre impair à l'exception de $K_0(\mathbb{F}_q) = \mathbb{Z}$. On a $W_0(\mathbb{F}_q) = \mathbb{Z}/2$, isomorphisme défini par l'invariant de Arf et $W_1(\mathbb{F}_q) = \mathbb{Z}/2$, isomorphisme défini par l'invariant de Dickson. Ici les groupes de Witt sont ceux calculés avec la forme paramètre min (c'est-à-dire ceux associés à des formes quadratiques).
Par ailleurs, la suite exacte des 12 (théorème 5.11) se réduit en fait à une suite à 6 termes, car $\varepsilon = 1 = -1$. Si on utilise le théorème précédent, on en déduit que les groupes de Witt élargis $W_n^{\text{él}}(\mathbb{F}_q)$ sont égaux à $\mathbb{Z}/2$ pour tout $n \in \mathbb{Z}$.

6. Les groupes de Witt stabilisés

6.1. Remarque. Ce paragraphe est une extension aux anneaux quelconques des idées développées dans une Note aux Comptes Rendus [**K4**]. Une autre extension aux schémas est décrite dans [**S**].

6.2. Nous nous plaçons dans la catégorie des anneaux discrets A avec involution $a \mapsto \overline{a}$ (nous ne supposons pas la commutativité ni l'existence d'un élément unité). Les groupes de Witt stabilisés ${}_\varepsilon\mathcal{W}_n(A)$, avec $\varepsilon = \pm 1$ et $n \in \mathbb{Z}$, que nous définirons plus loin, vérifient les propriétés suivantes

1) Exactitude. Pour toute suite exacte d'anneaux discrets avec involution

$$0 \longrightarrow A' \longrightarrow A \longrightarrow A'' \longrightarrow 0$$

nous avons une suite exacte naturelle des groupes $\mathcal{W}$

$$\longrightarrow {}_\varepsilon\mathcal{W}_{n+1}(A) \longrightarrow {}_\varepsilon\mathcal{W}_{n+1}(A") \longrightarrow {}_\varepsilon\mathcal{W}_n(A') \longrightarrow {}_\varepsilon\mathcal{W}_n(A) \longrightarrow {}_\varepsilon\mathcal{W}_n(A'') \longrightarrow$$

2) Periodicité. Nous avons un isomorphisme naturel

$${}_\varepsilon\mathcal{W}_n(A) \cong {}_{-\varepsilon}\mathcal{W}_{n+2}(A)$$

et par conséquent une périodicité 4 par rapport à l'indice n.

3) Invariance par extension nilpotente. Si I est un idéal nilpotent dans A, la projection $A \to A/I$ induit un isomorphisme

$${}_\varepsilon\mathcal{W}_n(A) \cong {}_\varepsilon\mathcal{W}_n(A/I)$$

En d'autres termes ${}_\varepsilon\mathcal{W}_n(I) = 0$ pour un anneau nilpotent.

4) Invariance homotopique. Si 1 est scindé dans A (en particulier si 2 est inversible) l'extension polynomiale $A \to A[t]$ (où $\overline{t} = t$) induit un isomorphisme

$${}_\varepsilon\mathcal{W}_n(A) \cong {}_\varepsilon\mathcal{W}_n(A[t])$$

5) Normalisation. Si A est unitaire, il existe un homomorphisme naturel

$$\Theta : {}_\varepsilon W_n(A) \to {}_\varepsilon\mathcal{W}_n(A)$$

où ${}_\varepsilon W_n(A)$ est le groupe de Witt classique [**K1**] construit avec les formes quadratiques. Celui-ci induit un isomorphisme

$${}_\varepsilon W_n(A) \otimes_\mathbb{Z} \mathbb{Z}' \cong {}_\varepsilon\mathcal{W}_n(A) \otimes_\mathbb{Z} \mathbb{Z}'$$

où $\mathbb{Z}' = \mathbb{Z}[1/2]$.
Si A est noethérien régulier, l'homomorphisme Θ est un isomorphisme lorsque $n \leq 0$. Si on suppose en outre que 2 est inversible dans A, les ${}_1\mathcal{W}_n(A)$, n mod 4, sont les groupes de Witt triangulés de Balmer [Ba].

6.3. Pour démontrer l'existence d'une telle théorie, nous allons essentiellement utiliser les résultats du paragraphe précédent sur la périodicité en K-théorie hermitienne. Rappelons que dans [**K2**] p. 243 nous avons défini un élément remarquable u_{-2} dans

$$_{-1}K\mathcal{Q}_{-2}(\mathbb{Z}[s]) = {}_{-1}K\mathcal{Q}^{\text{él}}_{-2}(\mathbb{Z}[s])$$

défini par une matrice antisymétrique ayant 30 éléments et à coefficients dans l'anneau des polynômes laurentiens à deux variables $\mathbb{Z}[s][t, u, t^{-1}, u^{-1}]$. Cet élément nous a déjà servi dans le §5 pour définir la flèche $_{-\varepsilon}U^{\text{él}}_{n+1}(A) \to {}_{\varepsilon}V^{\text{él}}_{n}(A)$.

6.4. Dans le paragraphe 4, nous avons défini pour tout anneau unitaire A un cup-produit

$$_{-1}K\mathcal{Q}^{\text{él}}_{-2}(\mathbb{Z}[s]) \times {}_{\varepsilon}K\mathcal{Q}^{\text{él}}_{n}(A) \to {}_{-\varepsilon}K\mathcal{Q}^{\text{él}}_{n-2}(A)$$

Puisque nous sommes seulement intéressés aux valeurs de n qui sont ≤ 0, nous pouvons remplacer les groupes $K\mathcal{Q}^{\text{él}}_{n}$ par $K\mathcal{Q}^{\min}_{n}$(et même $K\mathcal{Q}^{\max}_{n}$ pour $n < 0$), que nous noterons simplement $K\mathcal{Q}_n$. En outre, l'homomorphisme de périodicité (défini par le cup-produit avec u_{-2})

$$\beta : {}_{\varepsilon}K\mathcal{Q}_n(A) \to {}_{-\varepsilon}K\mathcal{Q}_{n-2}(A)$$

composé à gauche par la flèche oubli $_{-\varepsilon}K\mathcal{Q}_{n-2}(A) \to K_{n-2}(A)$ ou composé à droite par la flèche hyperbolique $K_n(A) \to {}_{\varepsilon}K\mathcal{Q}_n(A)$ est réduit à 0 (car la K-théorie de la suspension d'un anneau noethérien régulier est triviale). Par conséquent, la limite inductive du système de groupes de $K-$théorie hermitienne

$$_{\varepsilon}K\mathcal{Q}_n(A) \longrightarrow {}_{-\varepsilon}K\mathcal{Q}_{n-2}(A) \longrightarrow {}_{\varepsilon}K\mathcal{Q}_{n-4}(A) \longrightarrow {}_{-\varepsilon}K\mathcal{Q}_{n-6}(A) \longrightarrow \cdots$$

est aussi la limite inductive du système de groupes de Witt associés

$$_{\varepsilon}W_n(A) \longrightarrow {}_{-\varepsilon}W_{n-2}(A) \longrightarrow {}_{\varepsilon}W_{n-4}(A) \longrightarrow {}_{-\varepsilon}W_{n-6}(A) \longrightarrow \cdots$$

Cette limite est par définition le groupe de Witt stabilisé $_{\varepsilon}\mathcal{W}_n(A)$ que nous souhaitions définir. Notons que grâce à l'excision en K-théorie et en K-théorie hermitienne en degrés ≤ 0, nous pouvons étendre cette définition aux anneaux non nécessairement unitaires en définissant $_{\varepsilon}K\mathcal{Q}_n(A)$ comme le noyau de $_{\varepsilon}K\mathcal{Q}_n(A^+) \to {}_{\varepsilon}K\mathcal{Q}_n(\mathbb{Z})$, où A^+ est l'anneau A (considéré comme une $\mathbb{Z}$-algèbre) après addition d'un élément unité. La définition de $_{\varepsilon}\mathcal{W}_n(A)$ pour A non unitaire est tout à fait analogue. De ces considérations et de l'excision pour les groupes $K\mathcal{Q}_n$ si $n \leq 0$, nous déduisons la première propriété des groupes de Witt stabilisés :

6.5. Théorème . *A toute suite exacte d'anneaux discrets avec involution*

$$0 \longrightarrow A' \longrightarrow A \longrightarrow A" \longrightarrow 0$$

nous pouvons associer naturellement une suite exacte des groupes de Witt stabilisés

$$\cdots \longrightarrow {}_{\varepsilon}\mathcal{W}_n(A') \longrightarrow {}_{\varepsilon}\mathcal{W}_n(A) \longrightarrow {}_{\varepsilon}\mathcal{W}_n(A'') \longrightarrow {}_{\varepsilon}\mathcal{W}_{n-1}(A') \longrightarrow {}_{\varepsilon}\mathcal{W}_{n-1}(A) \longrightarrow \cdots$$

6.6. L'isomorphisme $_{\varepsilon}\mathcal{W}_n(A) \cong {}_{-\varepsilon}\mathcal{W}_{n+2}(A)$ et la périodicité 4 se déduisent immédiatement des définitions.

6.7. Théorème (normalisation). *Soit A un anneau noethérien régulier unitaire. Alors le groupe de Witt stabilisé ${}_1\mathcal{W}_0(A)$ (resp.${}_{-1}W_0(A)$) coïncide avec le groupe de Witt classique des formes quadratiques (resp. (-1)-quadratiques). En outre, pour tout anneau A unitaire, les homomorphismes canoniques*

$${}_\varepsilon W_n^{\text{él}}(A) \longrightarrow {}_\varepsilon W_n^{\min}(A) \longrightarrow {}_\varepsilon W_n^{\max}(A) \longrightarrow {}_\varepsilon \mathcal{W}_n(A)$$

induisent des isomorphismes en tensorisant par $\mathbb{Z}' = \mathbb{Z}[1/2]$.

Démonstration. Puisque les groupes de K-théorie négative de A sont triviaux si A est noethérien régulier, la suite exacte à 12 termes décrite en 5.9 montre que les flèches de la suite

$${}_\varepsilon W_0(A) \longrightarrow {}_{-\varepsilon} W_{-2}(A) \longrightarrow {}_\varepsilon W_{-4}(A) \longrightarrow {}_{-\varepsilon} W_{-6}(A) \longrightarrow \cdots$$

sont des isomorphismes. Par exemple, si $\varepsilon = 1$ et si A est le corps à 2 éléments, nous trouvons le groupe $\mathbb{Z}/2$ (qui est détecté par l'invariant de Arf).
Par ailleurs si A est un anneau quelconque, en utilisant la localisation en K-théorie hermitienne, nous avons construit en [**K1**] deux éléments dans ${}_{-1}W_2^{\max}(\mathbb{Z})$ et ${}_{-1}W_{-2}^{\max}(\mathbb{Z})$ dont le cup-produit dans ${}_1W^{\max}(\mathbb{Z})$ est une puissance de 2. Les premiers isomorphismes se démontrent en se ramenant par périodicité aux degrés négatifs. Le dernier isomorphisme résulte de la suite exacte à 12 termes démontrée en 5.9.

6.8. Théorème (invariance par extension nilpotente). *Si I est un idéal nilpotent dans A, la projection $A \to A/I$ induit un isomorphisme*

$${}_\varepsilon \mathcal{W}_n(A) \cong {}_\varepsilon \mathcal{W}_n(A/I)$$

Par conséquent, ${}_\varepsilon \mathcal{W}_n(I) = 0$ pour tout idéal nilpotent I.

Démonstration. Sans restreindre la généralité, nous pouvons supposer que A est unitaire. Dans ce cas, il est bien connu que tout module projectif de type fini sur A/I provient d'un module projectif E sur A par extension des scalaires et qu'il est donc du type E/I. Par conséquent, la forme ε-hermitienne sur A/I est donnée par un isomorphisme

$$\varphi : E/I \to (E/I)^*$$

Puisque φ est paire, nous pouvons l'écrire sour la forme $\varphi_0 + \varepsilon^t\varphi_0$. Soit $\widetilde{\varphi}_0$ un homomorphisme $E \to E^*$ tel que $\widetilde{\varphi}_0 = \varphi_0 \mod I$. Alors $\varphi = \widetilde{\varphi_0} + \varepsilon^t\widetilde{\varphi_0}$ est une forme ε- hermitienne non dégénérée $E \to E^*$ qui est un relevé de φ. Ceci montre que le morphisme ${}_\varepsilon K\mathcal{Q}_0(A) \to {}_\varepsilon K\mathcal{Q}_0(A/I)$ est surjectif pour tout idéal nilpotent I (aussi bien pour $K\mathcal{Q}^{\max}$ que pour $K\mathcal{Q}^{\min}$). Il en est donc de même de

$${}_\varepsilon K\mathcal{Q}_n(A) \to {}_\varepsilon K\mathcal{Q}_n(A/I)$$

pour $n \leq 0$ en considérant des suspensions itérées (él, max et min coïncident en degrés $n < 0$; cf. 3.2). La surjectivité de l'homomorphisme ${}_\varepsilon \mathcal{W}_n(A) \to {}_\varepsilon \mathcal{W}_n(A/I)$ en résulte.
L'injectivité du morphisme ${}_\varepsilon \mathcal{W}_n(A) \to {}_\varepsilon \mathcal{W}_n(A/I)$ est plus délicate à montrer. En raisonnant par récurrence sur le degré de nilpotence de I, nous pouvons d'abord supposer que $I^2 = 0$. Par ailleurs, nous savons que tout module muni d'une forme hermitienne paire est facteur direct d'un module hyperbolique. C'est donc l'image d'un projecteur auto-adjoint p, soit $p^2 = p$ et $p^* = p$ dans un $H(A^n)$.

Enfin, sans restreindre la généralité (puisque nous stabilisons), nous pouvons supposer que A est la suspension SR d'un anneau R et que $I = SJ$ où J est un idéal de R tel que $J^2 = 0$. La démonstration de l'injectivité se résume alors à la solution du problème suivant : nous considérons deux projecteurs auto-adjoints p_0 et p_1 dans un module hyperbolique sur $A = SR$ tels que leurs images $\mod I$, soient $\overline{p}_0$ et $\overline{p}_1$ sont conjuguées. Puisque ${}_\varepsilon K\mathcal{Q}_1(SR) \cong {}_\varepsilon K\mathcal{Q}_0(R)$ en général et que le morphisme ${}_\varepsilon K\mathcal{Q}_0(R) \to {}_\varepsilon K\mathcal{Q}_0(R/J)$ est surjectif comme nous l'avons vu précédemment, nous pouvons supposer sans restreindre la généralité [6] que $\overline{p}_0 = \overline{p}_1$ ou encore $p_1 = p_0 + \sigma$, où σ appartient à I. De l'identité $(p_1)^2 = p_1$ et de l'égalité $I^2 = 0$, nous déduisons les relations suivantes :

$$\begin{aligned} \sigma &= p_0\sigma + \sigma p_0 \\ \sigma p_0 \sigma &= 0 \\ \sigma^2 = \sigma^2 p_0 &= p_0\sigma^2 \end{aligned}$$

Considérons maintenant l'endomorphisme $\alpha = 1 - p_0 - p_1 + 2p_0p_1$. Puisque $\alpha \equiv 1 \mod I$, c'est un isomorphisme. Par ailleurs, il vérifie la relation $\alpha p_1 = p_0\alpha$. Nous allons maintenant montrer que $\alpha\alpha^* = 1$. Pour cela, on remarque que α s'écrit aussi

$$\alpha = 1 - \sigma + 2p_0\sigma$$

et, grâce aux identités précédentes, un calcul direct montre bien que

$$\alpha\alpha^* = (1 - \sigma + 2p_0\sigma)(1 - \sigma + 2\sigma p_0) = 1$$

Les projecteurs p_0 et p_1 sont ainsi conjugués par un automorphisme unitaire et déterminent par conséquent la même classe de forme hermitienne paire [7].

6.9. Théorème (invariance homotopique). *Soit A un anneau unitaire tel que 1 soit scindé dans A. Il existe donc un élément λ dans le centre de A tel que $\lambda + \overline{\lambda} = 1$. L'extension polynomiale $A \to A[t]$ (avec $\overline{t} = t$) induit alors un isomorphisme*

$$ {}_\varepsilon\mathcal{W}_n(A) \cong {}_\varepsilon\mathcal{W}_n(A[t])$$

Démonstration. Il suffit de démontrer le théorème pour $n = 0$. Celui-ci est déjà connu pour 2 inversible dans A (voir [**O**] pour une preuve simple). Cependant, il existe des anneaux où 2 n'est pas inversible et où 1 est scindé, par exemple le corps fini $\mathbb{F}_4$ muni de l'involution non triviale. Pour traiter ce cas plus général, nous devons réexaminer la preuve classique. En fait, le seul point qui mérite une précision dans cette preuve est le lemme suivant.

6.10. Lemme. *Soit A un anneau avec λ dans le centre de A tel que $1 = \lambda + \overline{\lambda}$. Soit E un A-module muni d'une forme ε-hermitienne et soit $\alpha = 1 + \nu t$ un élément de $GL(E \otimes \mathbb{Z}[t])$ avec ν nilpotent et auto-adjoint. Alors α peut être écrit sous la forme $\gamma(t)^*\gamma(t)$, où $\gamma(t)$ est un polynôme en t dans l'anneau engendré par λ et ν.*

Démonstration. Nous allons construire par récurrence sur n un polynôme de degré

6. La surjectivité de l'homomorphisme ${}_\varepsilon K\mathcal{Q}_1^{\min}(\Lambda) \to {}_\varepsilon K\mathcal{Q}_1^{\min}(\Lambda/I)$ implique la surjectivité de l'homomorphisme ${}_\varepsilon O^{\min}(\Lambda) \to {}_\varepsilon O^{\min}(\Lambda/I)$.

7. D'après 2.10, il revient au même de considérer des formes hermitienne paires ou des formes quadratiques dans les groupes stabilisés.

au plus n dans l'anneau engendré par ν et λ, soit $\gamma_n(t) = 1+a_1t+a_2t^2+\cdots+a_nt^n$, tel que $\gamma_n(t)^*\gamma_n(t) \equiv 1+\nu t \mod (\nu t)^{n+1}$. Pour $n=1$, nous posons $\gamma_1(t) = 1+\lambda\nu t$. Si γ_n est construit, nous avons $\gamma_n(t)^*\gamma_n(t) = 1+\nu t+b_{n+1}(\nu t)^{n+1} \mod (\nu t)^{n+2}$ avec $b_{n+1} = \overline{B1}_{n+1}$. Nous posons alors $\gamma_{n+1}(t) = (1-\lambda b_{n+1}(\nu t)^{n+1})\gamma_n(t)$ pour obtenir l'identité requise

$$\gamma_{n+1}(t)^*\gamma_{n+1}(t) \equiv 1+\nu t \mod (\nu t)^{n+2}$$

6.11. Exemple. Si A est un corps fini de caractéristique 2, il est facile de montrer que les groupes de Witt stabilisés $\mathcal{W}_n(A)$ sont tous isomorphes à $\mathbb{Z}/2$. Ils coïncident en fait avec les groupes $W_n^{\text{él}}(A)$ en tout degré.

6.12. Remarque. Ces groupes de Witt stabilisés ont été généralisés aux schémas par M. Schlichting [**S**]. Dans cette généralité, on doit cependant supposer 2 inversible.

7. Les lemmes de Clauwens

7.1. Lemme. *La forme hermitienne associée à la forme quadratique κ définie en 4.3 est non dégénérée.*

Démonstration. Nous suivons les simplications de notation indiquées en 4.3 en remplaçant notamment δ par ϕ tel que $\phi+\phi^* = 1$. Nous pouvons donc écrire

$$\kappa = \sum \theta_n \otimes \phi^n$$

qu'il est plus suggestif de noter $\theta(\phi)$. Nous avons alors

$$\kappa+\kappa^* = \sum \theta_n\otimes\phi^n + \sum(\theta_n)^*\otimes(\phi)^{*n} = \sum \theta_n\otimes\phi^n + \sum(\theta_n)^*\otimes(1-\phi)^n$$

Par ailleurs, on sait que le polynôme en s défini par $\sum \theta_n\otimes s^n + \sum(\theta_n)^*\otimes(1-s)^n$ est inversible (c'est la forme hermitienne H associée à θ). Il en résulte évidemment que $\kappa+\kappa^*$ est inversible. On peut aussi l'écrire $H(\phi)$ avec un abus d'écriture évident.

7.2. Lemme. *Si on change $\theta = \sum \theta_n s^n$ en $\theta+Z-Z^*$, les formes quadratiques associés κ et κ' sont équivalentes.*

Démonstration. La forme quadratique $\theta = \sum \theta_n s^n$ est modifiée en

$$\sum \theta_n s^n + \sum \sigma_n s^n - \sum(\sigma_n)^*(1-s)^n$$

Par conséquent κ est modifiée en $\kappa+\sigma(\phi)-(\sigma(\phi))^*$ (remplacer s par ϕ).

7.3. Lemme. *Modulo l'image de $K\mathcal{Q}(A)$ dans $K\mathcal{Q}^{\text{él}_0}(A[s])$ (et même d'une forme hyperbolique sur A), tout élément de ce dernier groupe peut être représenté par une forme linéaire en s.*

Démonstration. Soit $\theta = \sum_0^N \theta_n s^n$ une forme quadratique de degré N. L'identité suivante et un raisonnement par récurrence sur N montre qu'on peut réduire le degré de θ à 0 ou 1

$$\begin{pmatrix} 1 & -s & (\theta_N)^*(1-s)^{N-1} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \theta & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1+s & 1 & 0 \\ \theta_N s^{N-1} & 0 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \theta - \theta_N s^N & 0 & -s \\ \theta_N s^{N-1} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Si θ s'écrit $\theta_0 + \theta_1 s$, on peut aussi éliminer le terme constant en écrivant que θ est équivalente à

$$\theta_0 + \theta_1 s - \theta_0(1-s) + (\theta_0)^* s = (\theta_1 + \theta_0 + (\theta_0)^*)s$$

ce qui démontre le lemme.

Le lemme précédent nous montre qu'il suffit de vérifier la validité du produit de Clauwens défini en 4.3 (mêmes notations), dans le cas où θ est une forme linéaire en s, soit σs avec σ presque symétrique, i.e. $\sigma^* = \sigma(1+N)$, avec N nilpotent. Il nous faut montrer ensuite que le cup-produit de Clauwens ne dépend que de la forme quadratique associée à δ (ou l'endomorphisme ϕ grâce à l'identification de F à son dual). Rappelons qu'on a aussi identifié E à son dual par l'isomorphisme $\theta_0 + \sum_{n=0}^{\infty} {}^t\theta_n$.

Si on pose $G = E \otimes F$, la transposée ${}^t f$ d'une application f de G dans son dual s'identifie également à son application adjointe f^* (cf. les remarques faites en 4.3).

7.4. Lemme. *Soit ϕ et ζ deux endomorphismes de F tels que $\phi + \phi^* = 1$. Pour tout entier $p \geq 0$, il existe alors un isomorphisme f_p de G sur son dual tel que*

$$(f_p)^*(\sigma\otimes\phi)f_p = \sigma\otimes(\phi + \zeta - \zeta^*) + Z_p - (Z_p)^* \mod (\sigma N^{p+1}\otimes 1)$$

où $N = \sigma^{-1}\sigma^* - 1$ est nilpotent et où l'expression mod $(\sigma N^{p+1}\otimes 1)$ signifie une somme de morphismes du type $\sigma N^{p+1}\otimes\kappa_{p+1} + \sigma N^{p+2}\otimes\kappa_{p+2} + \cdots$ (qui est finie puisque N est nilpotent).

Démonstration. Puisque $\sigma^* = \sigma + \sigma N$, on a $\sigma^* N^k \otimes 1 = \sigma N^k \otimes 1 \mod (\sigma N^{k+1}\otimes 1)$. On a de même $N^{*k}\sigma N^r = \sigma N^{r+k} \mod \sigma N^{r+k+1}\otimes 1$. Nous allons maintenant construire f_p et Z_p par récurrence sur p. Pour $p=0$, on pose $f_0 = 1$ et $Z_0 = -\sigma\otimes\zeta$. Pour définir f_{p+1} à partir de f_p, on écrit

$$(f_p)^*(\sigma\otimes\phi)f_p - [\sigma\otimes(\phi+\zeta-\zeta^*) + Z_p - (Z_p)^*] = -\sigma N^{p+1}\otimes\kappa_{p+1} \mod (\sigma N^{p+1}\otimes 1)$$

On pose alors $U = N_{p+1}\otimes\kappa_{p+1}$ et $f_{p+1} = f_p + U$ et $Z_{p+1} = Z_p + U^*(\sigma\otimes\phi)$

En travaillant mod $(\sigma N^{p+2}\otimes 1)$, on obtient les identités suivantes

$$(f_{p+1})^*(\sigma\otimes\phi)f_{p+1} - [\sigma\otimes(\phi+\zeta-\zeta^*) + Z_{p+1} - (Z_{p+1})^*]$$
$$= (f_{p+1})^*(\sigma\otimes\phi)f_{p+1} - (f_p)^*(\sigma\otimes\phi)f_p - (Z_{p+1} - Z_p) + ((Z_{p+1})^* - (Z_p)^*) - \sigma N^{p+1}\otimes\kappa$$
$$= U^*(\sigma\otimes\phi) + (\sigma\otimes\phi)U - U^*(\sigma\otimes\phi)) + (\sigma^*\otimes\phi^*)U - \sigma N^{p+1}\otimes\kappa = \sigma N^{p+1}(\phi+\phi^*-1)\kappa$$
$$= 0 \mod \sigma N^{p+2}\otimes 1$$

Ceci achève la démonstration du lemme.

Références

[B1] BAK A. : K-theory of forms. Annals of Math Studies 98, Princeton N.J. Princeton University Press (1981).

[B2] BAK A. : Arf 's theorem for trace noetherian and other rings. Journal of Pure and Applied Algebra 14 (1979), 1-20

[BA] BALMER P. : An introduction to triangular Witt groups and a survey of applications. Algebraic and arithmetic theory of quadratic forms, 3158, Contemp. Math., 344, Amer. Math. Soc. (2004).

[C] CLAUWENS F.J.-B.J : The K-theory of almost symmetric forms. Topological structures II. Mathematical Centre Tracts 115, (1975) 41-49.

[K1] KAROUBI M. : Théorie de Quillen et homologie du groupe orthogonal. Annals of Maths 112, (1980) 207-282.

[K2] KAROUBI M. : Le théorème fondamental de la K-théorie hermitienne. Annals of Maths 112, (1980) 259-282.

[K3] KAROUBI M. : Homologie cyclique et K-théorie. Astérisque N° 149 (1987).

[K4] KAROUBI M. : Stabilization of the Witt group. Comptes Rendus Acad. Sci. Paris 362, (2006) 165-168.

[KV] KAROUBI M. et VILLAMAYOR O. : K-théorie algébrique et K-théorie topologique II. Math. Scand. 32, (1973) 57-86.

[L] LODAY J.-L. : K-théorie algébrique et représentations de groupes. Ann. Sci. Ec. Norm. Sup., (1976) 309-377.

[M] MILNOR J. : Introduction to Algebraic K-theory. Ann. of Math Studies 197. Princeton N.J. Princeton University Press (1974).

[O] OJANGUREN M. : On Karoubi's theorem : $W(A) = W(A[t])$. Arch. Math. Basel 43, (1984) 328-331.

[R] RANICKI A. : Algebraic L-theory I. Foundations. Proc. London Math. Soc. (3) 27 (1973) 101-125.

[S] SCHLICHTING M. : (en préparation)

[T] TITS J. : Formes quadratiques, groupes orthogonaux et algèbres de Clifford. Invent. Math. 5, (1968)19–41.

[Wa] WAGONER J.B. : Delooping classifying spaces in Algebraic K-theory. Topology 11, (1972) 349-370.

[W] WALL C.T.C. : On the axiomatic foundations of the theory of Hermitian forms. Proc. Cambridge Philod. Soc. 67 (1970) 243-250.

Université Paris 7, Mathématiques, Case 7012, 175/179, rue du Chevaleret 75250 Paris, France

E-mail address: max.karoubi@gmail.com

Clay Mathematics Proceedings
Volume 11, 2010

A Short Survey of Cyclic Cohomology

Masoud Khalkhali

Dedicated with admiration and affection to Alain Connes

Abstract. This is a short survey of some aspects of Alain Connes' contributions to cyclic cohomology theory in the course of his work on noncommutative geometry over the past 30 years.

Contents

1. Introduction

Cyclic cohomology was discovered by Alain Connes no later than 1981 and in fact it was announced in that year in a conference in Oberwolfach [**5**]. I have reproduced the text of his abstract below. As it appears in his report, one of Connes' main motivations to introduce cyclic cohomology theory came from index theory on foliated spaces. Let $(V, \mathcal{F})$ be a compact foliated manifold and let $V/\mathcal{F}$ denote the space of leaves of $(V, \mathcal{F})$. This space, with its natural quotient topology, is, in general, a highly singular space and in noncommutative geometry one usually replaces the quotient space $V/\mathcal{F}$ with a noncommutative algebra $A = C^*(V, \mathcal{F})$ called the foliation algebra of $(V, \mathcal{F})$. It is the convolution algebra of the holonomy groupoid of the foliation and is a C^*-algebra. It has a dense subalgebra $\mathcal{A} = C^\infty(V, \mathcal{F})$ which plays the role of the algebra of smooth functions on $V/\mathcal{F}$. Let D be a transversally elliptic operator on $(V, \mathcal{F})$. The analytic index of D, index$(D) \in$

2010 *Mathematics Subject Classification.* Primary 58B34; Secondary 19D55, 16T05, 18G30, 58J42.

$K_0(A)$, is an element of the K-theory of A. This should be compared with the family index theorem [**1**] where the analytic index of a family of fiberwise elliptic operators is an element of the K-theory of the base. Connes realized that to identify this class by a cohomological expression it would be necessary to have a noncommutative analogue of the Chern character, i.e., a map from $K_0(\mathcal{A})$ to a, then unknown, cohomology theory for the noncommutative algebra $\mathcal{A}$. This theory, now known as cyclic cohomology, would then play the role of the noncommutative analogue of de Rham homology of currents for smooth manifolds. Its dual version, *cyclic homology*, corresponds, in the commutative case, to de Rham cohomology.

Connes arrived at his definition of cyclic cohomology by a careful analysis of the algebraic structures deeply hidden in the (super)traces of products of commutators of operators. These expressions are directly defined in terms of an elliptic operator and its parametrix and give the index of the operator when paired with a K-theory class. In his own words [**5**]:

"The transverse elliptic theory for foliations requires as a preliminary step a purely algebraic work, of computing for a noncommutative algebra $\mathcal{A}$ the cohomology of the following complex: n-cochains are multilinear functions $\varphi(f^0, \dots, f^n)$ of $f^0, \dots, f^n \in \mathcal{A}$ where

$$\varphi(f^1, \dots, f^0) = (-1)^n \varphi(f^0, \dots, f^n)$$

and the boundary is

$$\begin{aligned} b\varphi(f^0, \dots, f^{n+1}) &= \varphi(f^0 f^1, \dots, f^{n+1}) - \varphi(f^0, f^1 f^2, \dots, f^{n+1}) + \cdots \\ &\quad + (-1)^{n+1} \varphi(f^{n+1} f^0, \dots, f^n). \end{aligned}$$

The basic class associated to a transversally elliptic operator, for $\mathcal{A} =$ the algebra of the foliation, is given by:

$$\varphi(f^0, \dots, f^n) = Trace\,(\varepsilon F[F, f^0][F, f^1] \cdots [F, f^n]), \quad f^i \in \mathcal{A}$$

where

$$F = \begin{pmatrix} 0 & Q \\ P & 0 \end{pmatrix}, \quad \varepsilon = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

and Q is a parametrix of P. An operation

$$S : H^n(\mathcal{A}) \to H^{n+2}(\mathcal{A})$$

is constructed as well as a pairing

$$K(\mathcal{A}) \times H(\mathcal{A}) \to \mathbb{C}$$

where $K(\mathcal{A})$ is the algebraic K-theory of A. It gives the index of the operator from its associated class φ. Moreover $\langle e, \varphi \rangle = \langle e, S\varphi \rangle$, so that the important group to determine is the inductive limit $H_p = \varinjlim H^n(\mathcal{A})$ for the map S. Using the tools of homological algebra the groups $H^n(\mathcal{A}, \mathcal{A}^*)$ of Hochschild cohomology with coefficients in the bimodule $\mathcal{A}^*$ are easier to determine and the solution of the problem is obtained in two steps:
1) the construction of a map

$$B : H^n(\mathcal{A}, \mathcal{A}^*) \to H^{n-1}(\mathcal{A})$$

and the proof of a long exact sequence

$$\cdots \to H^n(\mathcal{A}, \mathcal{A}^*) \xrightarrow{B} H^{n-1}(\mathcal{A}) \xrightarrow{S} H^{n+1}(\mathcal{A}) \xrightarrow{I} H^{n+1}(\mathcal{A}, \mathcal{A}^*) \to \cdots$$

where I is the obvious map from the cohomology of the above complex to the Hochschild cohomology;
2) the construction of a spectral sequence with E_2 term given by the cohomology of the degree -1 differential $I \circ B$ on the Hochschild groups $H^n(\mathcal{A}, \mathcal{A}^*)$ and which converges strongly to a graded group associated to the inductive limit.

This purely algebraic theory is then used. For $\mathcal{A} = C^\infty(V)$ one gets the de Rham homology of currents, and for the pseudo-torus, i.e. the algebra of the Kronecker foliation, one finds that the Hochschild cohomology depends on the Diophantine nature of the rotation number while the above theory gives H_p^0 of dimension 2 and H_p^1 of dimension 2, as expected, but from some remarkable cancellations."

A full exposition of these results later appeared in two IHES preprints [**6**], and were eventually published as [**9**]. With the appearance of [**9**] one could say that the first stage of the development of noncommutative geometry and specially cyclic cohomology reached a stage of maturity. In the next few sections I shall try to give a quick and concise survey of some aspects of cyclic cohomology theory as they were developed in [**9**]. The last two sections are devoted to developments in the subject after [**9**] arising from the work of Connes.

It is a distinct honor and a great pleasure to dedicate this short survey of cyclic cohomology theory as a small token of our friendship to Alain Connes, the originator of the subject, on the occasion of his 60th birthday. It inevitably only covers part of what has been done by Alain in this very important corner of noncommutative geometry. It is impossible to cover everything, and in particular I have left out many important topics developed by him including, among others, the Godbillon-Vey invariant and type III factors [**8**], the transverse fundamental class for foliations [**8**], the Novikov conjecture for hyperbolic groups [**18**], entire cyclic cohomology [**10**], and multiplicative characteristic classes [**12**]. Finally I would like to thank Farzad Fathi zadeh for carefully reading the text and for several useful comments, and Arthur Greenspoon who kindly edited the whole text.

2. Cyclic cohomology

Cyclic cohomology can be defined in several ways, each shedding light on a different aspect of it. Its original definition [**5, 9**] was through a remarkable subcomplex of the Hochschild complex that we recall first. By algebra in this paper we mean an associative algebra over $\mathbb{C}$. For an algebra $\mathcal{A}$ let

$$C^n(\mathcal{A}) = \mathrm{Hom}(\mathcal{A}^{\otimes(n+1)}, \mathbb{C}), \quad n = 0, 1, \ldots,$$

denote the space of $(n+1)$-linear functionals on $\mathcal{A}$. These are our *n-cochains.* The Hochschild differential $b : C^n(\mathcal{A}) \to C^{n+1}(\mathcal{A})$ is defined as

$$\begin{aligned}(b\varphi)(a_0, \ldots, a_{n+1}) \quad &= \quad \sum_{i=0}^{n} (-1)^i \varphi(a_0, \ldots, a_i a_{i+1}, \ldots, a_{n+1}) \\ &\quad + (-1)^{n+1} \varphi(a_{n+1} a_0, \ldots, a_n).\end{aligned}$$

The cohomology of the complex $(C^*(\mathcal{A}), b)$ is the Hochschild cohomology of $\mathcal{A}$ with coefficients in the bimodule $\mathcal{A}^*$.

The following definition is fundamental and marks the departure from Hochschild cohomology in [**5, 9**]:

Definition 2.1. *An n-cochain $\varphi \in C^n(\mathcal{A})$ is called cyclic if*

$$\varphi(a_n, a_0, \ldots, a_{n-1}) = (-1)^n \varphi(a_0, a_1, \ldots, a_n)$$

for all $a_0, \ldots, a_n$ in $\mathcal{A}$. The space of cyclic n-cochains will be denoted by $C^n_\lambda(\mathcal{A})$.

Just why, of all possible symmetry conditions on cochains, the cyclic property is a reasonable choice is at first glance not at all clear.

Lemma 2.1. *The space of cyclic cochains is invariant under the action of b, i.e., $b\,C^n_\lambda(\mathcal{A}) \subset C^{n+1}_\lambda(\mathcal{A})$ for all $n \geq 0$.*

To see this one introduces the operators $\lambda : C^n(\mathcal{A}) \to C^n(\mathcal{A})$ and $b' : C^n(\mathcal{A}) \to C^{n+1}(\mathcal{A})$ by

$$\begin{aligned}
(\lambda\varphi)(a_0, \ldots, a_n) &= (-1)^n \varphi(a_n, a_0, \ldots, a_{n-1}), \\
(b'\varphi)(a_0, \ldots, a_{n+1}) &= \sum_{i=0}^{n} (-1)^i \varphi(a_0, \ldots, a_i a_{i+1}, \ldots, a_{n+1}),
\end{aligned}$$

and checks that $(1-\lambda)b = b'(1-\lambda)$. Since $C^*_\lambda(\mathcal{A}) = \mathrm{Ker}\,(1-\lambda)$, the lemma is proved.

We therefore have a subcomplex of the Hochschild complex, called the *cyclic complex* of $\mathcal{A}$:

$$C^0_\lambda(\mathcal{A}) \xrightarrow{b} C^1_\lambda(\mathcal{A}) \xrightarrow{b} C^2_\lambda(\mathcal{A}) \xrightarrow{b} \cdots . \tag{1}$$

Definition 2.2. *The cohomology of the cyclic complex* (1) *is the cyclic cohomology of $\mathcal{A}$ and will be denoted by $HC^n(\mathcal{A})$, $n = 0, 1, 2, \ldots$.*

And that is Connes' first definition of cyclic cohomology. A cocycle for the cyclic cohomology group $HC^n(\mathcal{A})$ is called a *cyclic n-cocycle* on $\mathcal{A}$. It is an $(n+1)$-linear functional φ on $\mathcal{A}$ which satisfies the two conditions:

$$(1-\lambda)\varphi = 0, \quad \text{and} \quad b\varphi = 0.$$

The inclusion of complexes

$$(C^*_\lambda(\mathcal{A}), b) \hookrightarrow (C^*(\mathcal{A}), b) \tag{2}$$

induces a map I from cyclic cohomology to Hochschild cohomology:

$$I : HC^n(\mathcal{A}) \longrightarrow HH^n(\mathcal{A}), \quad n = 0, 1, 2, \ldots .$$

A closer inspection of the long exact sequence associated to (2), yields *Connes' long exact sequence* relating Hochschild cohomology to cyclic cohomology. This is however easier said than done. The reason is that to identify the cohomology of the quotient one must use another long exact sequence, and combine the two long exact sequences to obtain the result. To simplify the notation, let us denote the Hochschild and cyclic complexes by C and C_λ, respectively. Then (2) gives us an exact sequence of complexes

$$0 \to C_\lambda \to C \xrightarrow{\pi} C/C_\lambda \to 0. \tag{3}$$

Its associated long exact sequence is

$$\cdots \longrightarrow HC^n(\mathcal{A}) \longrightarrow HH^n(\mathcal{A}) \longrightarrow H^n(C/C_\lambda) \longrightarrow HC^{n+1}(\mathcal{A}) \longrightarrow \cdots \tag{4}$$

We need to identify the cohomology groups $H^n(C/C_\lambda)$. To this end, consider the short exact sequence of complexes

$$0 \longrightarrow C/C_\lambda \xrightarrow{1-\lambda} (C, b') \xrightarrow{N} C_\lambda \longrightarrow 0, \tag{5}$$

where the operator N is defined by

$$N = 1 + \lambda + \lambda^2 + \cdots + \lambda^n : C^n \longrightarrow C^n.$$

The relations $(1-\lambda)b = b'(1-\lambda)$, $N(1-\lambda) = (1-\lambda)N = 0$, and $bN = Nb'$ show that $1-\lambda$ and N are morphisms of complexes in (5). As for the exactness of (5), the only nontrivial part is to show that $\ker(N) \subset \operatorname{im}(1-\lambda)$, which can be verified. Assuming $\mathcal{A}$ is unital, the middle complex (C, b') in (5) can be shown to be exact with a contracting homotopy $s : C^n \to C^{n-1}$ defined by $(s\varphi)(a_0, \ldots, a_{n-1}) = (-1)^{n-1}\varphi(a_0, \ldots, a_{n-1}, 1)$, which satisfies $b's + sb' = \mathrm{id}$. The long exact sequence associated to (5) looks like

(6)

$$\cdots \longrightarrow H^n(C/C_\lambda) \longrightarrow H^n_{b'}(C) \longrightarrow HC^n(\mathcal{A}) \longrightarrow H^{n+1}(C/C_\lambda) \longrightarrow H^{n+1}_{b'}(C) \longrightarrow \cdots$$

Since $H^n_{b'}(C) = 0$ for all n, it follows that the connecting homomorphism

$$\delta : HC^{n-1}(\mathcal{A}) \to H^n(C/C_\lambda) \tag{7}$$

is an *isomorphism* for all $n \geq 0$. Using this in (4), we obtain *Connes' long exact sequence* relating Hochschild and cyclic cohomology:

$$\cdots \longrightarrow HC^n(\mathcal{A}) \xrightarrow{I} HH^n(\mathcal{A}) \xrightarrow{B} HC^{n-1}(\mathcal{A}) \xrightarrow{S} HC^{n+1}(\mathcal{A}) \longrightarrow \cdots . \tag{8}$$

The operators B and S play a prominent role in noncommutative geometry. As we shall see, the operator B is the analogue of de Rham's differential in the noncommutative world, while the *periodicity operator* S is closely related to Bott periodicity in topological K-theory. Remarkably, there is a formula for B on the level of cochains given by $B = NB_0$, where $B_0 : C^n \to C^{n-1}$ is defined by

$$B_0\varphi(a_0, \ldots, a_{n-1}) = \varphi(1, a_0, \ldots, a_{n-1}) - (-1)^n\varphi(a_0, \ldots, a_{n-1}, 1).$$

Using the relations $(1-\lambda)b = b'(1-\lambda)$, $(1-\lambda)N = N(1-\lambda) = 0$, $bN = Nb'$, and $sb' + b's = 1$, one shows that

$$bB + Bb = 0, \quad \text{and} \quad B^2 = 0. \tag{9}$$

Using the periodicity operator S, the *periodic cyclic cohomology* of $\mathcal{A}$ is then defined as the direct limit of cyclic cohomology groups under the operator S:

$$HP^i(\mathcal{A}) := \varinjlim HC^{2n+i}(\mathcal{A}), \qquad i = 0,\, 1.$$

Notice that since S has degree 2, there are only two periodic groups. These periodic groups have better stability properties compared to cyclic cohomology groups. For example, they are homotopy invariant, and they pair with K-theory.

A much deeper relationship between Hochschild and cyclic cohomology groups is encoded in Connes' (b, B)-*bicomplex* and the associated Connes spectral sequence that we shall briefly recall now. Consider the relations (9). The (b, B)-bicomplex of a unital algebra $\mathcal{A}$, denoted by $\mathcal{B}(\mathcal{A})$, is the bicomplex

$$
\begin{array}{ccccc}
\vdots & & \vdots & & \vdots \\
C^2(\mathcal{A}) & \xrightarrow{B} & C^1(\mathcal{A}) & \xrightarrow{B} & C^0(\mathcal{A}) \\
{\scriptstyle b}\uparrow & & {\scriptstyle b}\uparrow & & \\
C^1(\mathcal{A}) & \xrightarrow{B} & C^0(\mathcal{A}) & & \\
{\scriptstyle b}\uparrow & & & & \\
C^0(\mathcal{A}) & & & &
\end{array}
$$

As usual, there are two spectral sequences attached to this bicomplex. The following fundamental result of Connes [**9**] shows that the spectral sequence obtained from filtration by rows converges to cyclic cohomology. Notice that the E^1 term of this spectral sequence is the Hochschild cohomology of $\mathcal{A}$.

THEOREM 2.1. *The map $\varphi \mapsto (0, \ldots, 0, \varphi)$ is a quasi-isomorphism of complexes*

$$(C_\lambda^*(\mathcal{A}),\, b) \to (\operatorname{Tot}\mathcal{B}(\mathcal{A}),\, b + B).$$

This is a consequence of the vanishing of the E^2 term of the second spectral sequence (filtration by columns) of $\mathcal{B}(\mathcal{A})$. To prove this, Connes considers the short exact sequence of b-complexes

$$0 \longrightarrow \operatorname{Im} B \longrightarrow \operatorname{Ker} B \longrightarrow \operatorname{Ker} B/\operatorname{Im} B \longrightarrow 0,$$

and proves that ([**9**], Lemma 41), the induced map

$$H_b(\operatorname{Im} B) \longrightarrow H_b(\operatorname{Ker} B)$$

is an isomorphism. This is a very technical result. It follows that $H_b(\operatorname{Ker} B/\operatorname{Im} B)$ vanishes. To take care of the first column one appeals to the fact that $\operatorname{Im} B \simeq \operatorname{Ker}(1-\lambda)$ is the space of cyclic cochains.

We give an alternative proof of Theorem (2.1) above. To this end, consider the *cyclic bicomplex* $\mathcal{C}(\mathcal{A})$ defined by

$$
\begin{array}{ccccccc}
\vdots & & \vdots & & \vdots & & \\
C^2(\mathcal{A}) & \xrightarrow{1-\lambda} & C^2(\mathcal{A}) & \xrightarrow{N} & C^2(\mathcal{A}) & \xrightarrow{1-\lambda} & \cdots \\
\uparrow{\scriptstyle b} & & \uparrow{\scriptstyle -b'} & & \uparrow{\scriptstyle b} & & \\
C^1(\mathcal{A}) & \xrightarrow{1-\lambda} & C^1(\mathcal{A}) & \xrightarrow{N} & C^1(\mathcal{A}) & \xrightarrow{1-\lambda} & \cdots \\
\uparrow{\scriptstyle b} & & \uparrow{\scriptstyle -b'} & & \uparrow{\scriptstyle b} & & \\
C^0(\mathcal{A}) & \xrightarrow{1-\lambda} & C^0(\mathcal{A}) & \xrightarrow{N} & C^0(\mathcal{A}) & \xrightarrow{1-\lambda} & \cdots
\end{array}
$$

The total cohomology of $\mathcal{C}(\mathcal{A})$ is isomorphic to cyclic cohomology:

$$H^n(\operatorname{Tot}\mathcal{C}(\mathcal{A})) \simeq HC^n(\mathcal{A}), \quad n \geq 0.$$

This is a consequence of the simple fact that the rows of $\mathcal{C}(\mathcal{A})$ are exact except in degree zero, where their cohomology coincides with the cyclic complex $(C_\lambda^*(\mathcal{A}), b)$.

So it suffices to show that $\mathrm{Tot}\,\mathcal{B}(\mathcal{A})$ and $\mathrm{Tot}\,\mathcal{C}(\mathcal{A})$ are quasi-isomorphic. This can be done by explicit formulas. Consider the chain maps

$$\begin{aligned} I : \mathrm{Tot}\,\mathcal{B}(\mathcal{A}) &\to \mathrm{Tot}\,\mathcal{C}(\mathcal{A}), \quad I = \mathrm{id} + Ns, \\ J : \mathrm{Tot}\,\mathcal{C}(\mathcal{A}) &\to \mathrm{Tot}\,\mathcal{B}(\mathcal{A}), \quad J = \mathrm{id} + sN. \end{aligned}$$

It can be directly verified that the following operators define chain homotopy equivalences:

$$\begin{aligned} g : \mathrm{Tot}\,\mathcal{B}(A) &\to \mathrm{Tot}\,\mathcal{B}(A), \quad g = Ns^2B_0, \\ h : \mathrm{Tot}\,\mathcal{C}(A) &\to \mathrm{Tot}\,\mathcal{C}(\mathcal{A}), \quad h = s. \end{aligned}$$

To give an example of an application of the spectral sequence of Theorem (2.1), let me recall Connes' computation of the *continuous cyclic cohomology* of the topological algebra $\mathcal{A} = C^\infty(M)$, i.e., the algebra of smooth complex valued functions on a closed smooth n-dimensional manifold M. This example is important since, apart from its applications, it clearly shows that cyclic cohomology is a noncommutative analogue of de Rham homology.

The continuous analogues of Hochschild and cyclic cohomology for topological algebras are defined as follows [**9**]. Let $\mathcal{A}$ be a topological algebra. A *continuous cochain* on $\mathcal{A}$ is a jointly continuous multilinear map $\varphi : \mathcal{A} \times \mathcal{A} \times \cdots \times \mathcal{A} \to \mathbb{C}$. By working with just continuous cochains, as opposed to all cochains, one obtains the continuous analogues of Hochschild and cyclic cohomology groups. In working with algebras of smooth functions (both in the commutative and noncommutative case), it is essential to use this continuous analogue.

The topology of $C^\infty(M)$ is defined by the sequence of seminorms

$$\|f\|_n = \sup |\partial^\alpha f|, \quad |\alpha| \leq n,$$

where the supremum is over a fixed, finite, coordinate cover for M. Under this topology, $C^\infty(M)$ is a locally convex, in fact nuclear, topological algebra. Similarly one topologizes the space of p-forms on M for all $p \geq 0$. Let

$$\Omega_p M := \mathrm{Hom}_{\mathrm{cont}}(\Omega^p M, \mathbb{C})$$

denote the continuous dual of the space of $p-$forms on M. Elements of $\Omega_p M$ are called *de Rham p-currents*. By dualizing the de Rham differential d, we obtain a differential $d^* : \Omega_* M \to \Omega_{*-1} M$, and a complex, called the *de Rham complex of currents* on M:

$$\Omega_0 M \xleftarrow{d^*} \Omega_1 M \xleftarrow{d^*} \Omega_2 M \xleftarrow{d^*} \cdots .$$

The homology of this complex is the *de Rham homology* of M and we denote it by $H_*^{dR}(M)$.

It is easy to check that for any m-current C, closed or not, the cochain

$$\varphi_C(f_0, f_1, \ldots, f_m) := \langle C,\, f_0 df_1 \cdots df_m \rangle, \tag{10}$$

is a continuous Hochschild cocycle on $C^\infty(M)$. Now if C is closed, then one checks that φ_C is a cyclic m-cocycle on $C^\infty(M)$. Thus we obtain natural maps

$$\Omega_m M \to HH^m_{cont}(C^\infty(M)) \tag{11}$$

and

$$Z_m M \to HC^m_{cont}(C^\infty(M)), \tag{12}$$

where $Z_m(M) \subset \Omega_m M$ is the space of closed m-currents on M. For example, if M is oriented and C represents its orientation class, then

$$\varphi_C(f_0, f_1, \ldots, f_n) = \int_M f_0 df_1 \cdots df_n, \tag{13}$$

which is easily checked to be a cyclic n-cocycle on $\mathcal{A}$.

In [**9**], using an explicit resolution, Connes shows that (11) is a quasi-isomorphism. Thus we have a natural isomorphism between space of de Rham currents on M and the continuous Hochschild cohomology of $C^\infty(M)$:

$$HH^i_{\text{cont}}(C^\infty(M)) \simeq \Omega_i M \qquad i = 0, 1, \ldots \tag{14}$$

To compute the continuous cyclic homology of $\mathcal{A}$, one first observes that under the isomorphism (14) the operator B corresponds to the de Rham differential d^*. More precisely, for each integer $n \geq 0$ there is a commutative diagram:

$$\begin{array}{ccc} \Omega_{n+1} M & \xrightarrow{\mu} & C^{n+1}(\mathcal{A}) \\ \Big\downarrow{\scriptstyle d^*} & & \Big\downarrow{\scriptstyle B} \\ \Omega_n M & \xrightarrow{\mu} & C^n(\mathcal{A}) \end{array}$$

where $\mu(C) = \varphi_C$ and φ_C is defined by (10). Then, using the spectral sequence of Theorem (2.1) and the isomorphism (14), Connes obtains [**9**]:

$$HC^n_{\text{cont}}(C^\infty(M)) \simeq Z_n(M) \oplus H^{\text{dR}}_{n-2}(M) \oplus \cdots \oplus H^{\text{dR}}_k(M), \tag{15}$$

where $k = 0$ if n is even and $k = 1$ is n is odd. For the continuous periodic cyclic cohomology he obtains

$$HP^k_{\text{cont}}(C^\infty(M)) \simeq \bigoplus_i H^{\text{dR}}_{2i+k}(M), \quad k = 0, 1. \tag{16}$$

We shall also briefly recall Connes' computation of the Hochschild and cyclic cohomology of smooth noncommutative tori [**9**]. This result already appeared in Connes' Oberwolfach report [**5**]. When θ is rational, the smooth noncommutative torus $\mathcal{A}_\theta$ can be shown to be Morita equivalent to $C^\infty(T^2)$, the algebra of smooth functions on the 2-torus. One can then use Morita invariance of Hochschild and cyclic cohomology to reduce the computation of these groups to those for the algebra $C^\infty(T^2)$. This takes care of rational θ. So we can assume θ is irrational and we denote the generators of $\mathcal{A}_\theta$ by U and V with the relation $VU = \lambda UV$, where $\lambda = e^{2\pi i\theta}$.

Recall that an irrational number θ is said to satisfy a Diophantine condition if $|1 - \lambda^n|^{-1} = O(n^k)$ for some positive integer k.

Proposition 2.1. *([**9**]) Let $\theta \notin \mathbb{Q}$. Then*
a) One has $HH^0(\mathcal{A}_\theta) = \mathbb{C}$,
b) If θ satisfies a Diophantine condition then $HH^i(\mathcal{A}_\theta)$ is 2-dimensional for i=1 and is 1-dimensional for $i = 2$,
c) If θ does not satisfy a Diophantine condition, then $HH^i(\mathcal{A}_\theta)$ are infinite dimensional non-Hausdorff spaces for $i = 1, 2$.

Remarkably, for all values of θ, the periodic cyclic cohomology is finite dimensional and is given by

$$HP^0(\mathcal{A}_\theta) = \mathbb{C}^2, \qquad HP^1(\mathcal{A}_\theta) = \mathbb{C}^2.$$

An explicit basis for these groups are given by cyclic 1-cocycles

$$\varphi_1(a_0, a_1) = \tau(a_0\delta_1(a_1)), \quad \text{and} \quad \varphi_1(a_0, a_1) = \tau(a_0\delta_2(a_1))$$

and cyclic 2-cocycles

$$\varphi(a_0, a_1, a_2) = \tau(a_0(\delta_1(a_1)\delta_2(a_2) - \delta_2(a_1)\delta_1(a_2))), \quad \text{and} \quad S\tau,$$

where $\delta_1, \delta_2 : \mathcal{A}_\theta \to \mathcal{A}_\theta$ are the canonical derivations defined by

$$\delta_1(\sum a_{mn}U^mV^n) = \sum ma_{mn}U^mV^n, \qquad \delta_2(U^mV^n) = \sum na_{mn}U^mV^n,$$

and $\tau : \mathcal{A}_\theta \to \mathbb{C}$ is the canonical trace.

A noncommutative generalization of formulas like (13) was introduced in [**9**] and played an important role in the development of cyclic cohomology theory in general. It gives a geometric meaning to the notion of a cyclic cocycle over an algebra and goes as follows. Let (Ω, d) be a differential graded algebra. A *closed graded trace* of dimension n on (Ω, d) is a linear map

$$\int : \Omega^n \longrightarrow \mathbb{C}$$

such that

$$\int d\omega = 0, \quad \text{and} \int [\omega_1, \omega_2] = 0,$$

for all ω in Ω^{n-1}, ω_1 in Ω^i, ω_2 in Ω^j and $i + j = n$. An n dimensional cycle over an algebra $\mathcal{A}$ is a triple $(\Omega, \int, \rho)$, where $\int$ is an n-dimensional closed graded trace on (Ω, d) and $\rho : A \to \Omega_0$ is an algebra homomorphism. Given a cycle $(\Omega, \int, \rho)$ over $\mathcal{A}$, its *character* is the cyclic n-cocycle on $\mathcal{A}$ defined by

$$\varphi(a_0, a_1, \ldots, a_n) = \int \rho(a_0)d\rho(a_1)\cdots d\rho(a_n). \tag{17}$$

Conversely one shows that all cyclic cocycles are obtained in this way.

Once one has the definition of cyclic cohomology, it is not difficult to formulate a dual notion of *cyclic homology* and a pairing between the two. Let $C_n(\mathcal{A}) = \mathcal{A}^{\otimes(n+1)}$. The analogues of the operators b, b' and λ are easily defined on $C_*(\mathcal{A})$ and are usually denoted by the same letters, as we do here. For example $b : C_n(\mathcal{A}) \to C_{n-1}(\mathcal{A})$ is defined by

$$\begin{aligned} b(a_0 \otimes \cdots \otimes a_n) &= \sum_{i=0}^{n-1}(-1)^i(a_0 \otimes \cdots \otimes a_ia_{i+1} \otimes \cdots \otimes a_n) && (18)\\ &+ (-1)^n(a_na_0 \otimes \cdots \otimes a_{n-1}). && (19) \end{aligned}$$

Let

$$C_n^\lambda(\mathcal{A}) := C_n(\mathcal{A})/\mathrm{Im}(1 - \lambda).$$

The relation $(1 - \lambda)b' = b(1 - \lambda)$ shows that the operator b is well defined on $C_*^\lambda(\mathcal{A})$. The complex $(C_*^\lambda(\mathcal{A}), b)$ is called the *homological cyclic complex* of $\mathcal{A}$ and its homology, denoted by $HC_n(\mathcal{A}), n = 0, 1, \ldots,$ is the *cyclic homology* of $\mathcal{A}$. The evaluation map $\langle\varphi, (a_0 \otimes \cdots \otimes a_n)\rangle \mapsto \varphi(a_0, \ldots, a_n)$ clearly defines a degree zero pairing $HC^*(\mathcal{A}) \otimes HC_*(\mathcal{A}) \to \mathbb{C}$. Many results of cyclic cohomology theory, such as Connes' long exact sequence and spectral sequence, and Morita invariance, continue to hold for cyclic homology theory with basically the same proofs.

Another important idea of Connes in the 1980's was the introduction of *entire cyclic cohomology* of Banach algebras [**10**]. This allows one to deal with algebras of functions on infinite dimensional (noncommutative) spaces such as those appearing in constructive quantum field theory. These algebras typically don't carry finitely summable Fredholm modules, but in some cases have so-called θ-summable Fredholm modules. In [**10**] Connes extends the definition of Chern character to such Fredholm modules with values in entire cyclic cohomology.

After the appearance of [**9**], cyclic (co)homology theory took on many lives and was further developed along distinct lines, including a purely algebraic one, with a big impact on algebraic K-theory. The cyclic cohomology of many algebras was later computed including the very important case of group algebras [**2**] and groupoid algebras. For many of these more algebraic aspects of the theory we refer to [**33, 32**] and references therein.

3. From K-homology to cyclic cohomology

As I said in the introduction, Connes' original motivation for the development of cyclic cohomology was to give a receptacle for a noncommutative Chern character map on the *K-homology* of noncommutative algebras. The cycles of K-homology can be represented by, even or odd, *Fredholm modules*. Here we just focus on the odd case, and we refer to [**9, 13**] for the even case. Given a Hilbert space $\mathcal{H}$, let $\mathcal{L}(\mathcal{H})$ denote the algebra of bounded linear operators on $\mathcal{H}$, and $\mathcal{K}(\mathcal{H})$ denote the algebra of compact operators. Also, for $1 \leq p < \infty$, let $\mathcal{L}^p(\mathcal{H})$ denote the Schatten ideal of p-summable operators. By definition, $T \in \mathcal{L}^p(\mathcal{H})$ if $|T|^p$ is a trace class operator.

Definition 3.1. *An odd Fredholm module over a unital algebra $\mathcal{A}$ is a pair $(\mathcal{H}, F)$ where*
1. $\mathcal{H}$ is a Hilbert space endowed with a representation

$$\pi : \mathcal{A} \longrightarrow \mathcal{L}(\mathcal{H}),$$

2. $F \in \mathcal{L}(\mathcal{H})$ is a bounded selfadjoint operator with $F^2 = I$,
3. For all $a \in \mathcal{A}$ we have

$$[F, \pi(a)] = F\pi(a) - \pi(a)F \in \mathcal{K}(\mathcal{H}). \tag{20}$$

A Fredholm module $(\mathcal{H}, F)$ is called *p-summable* if, instead of (20), we have the stronger condition:

$$[F, \ \pi(a)] \in \mathcal{L}^p(\mathcal{H}) \tag{21}$$

for all $a \in \mathcal{A}$.

To give a simple example, let $A = C(S^1)$ be the algebra of continuous functions on the circle and let A act on $\mathcal{H} = L^2(S^1)$ as multiplication operators. Let $F(e_n) = e_n$ if $n \geq 0$ and $F(e_n) = -e_n$ for $n < 0$, where $e_n(x) = e^{2\pi i n x}, n \in \mathbb{Z}$, denotes the standard orthonormal basis of $\mathcal{H}$. Clearly F is selfadjoint and $F^2 = I$. To show that $[F, \pi(f)]$ is a compact operator for all $f \in C(S^1)$, notice that if $f = \sum_{|n| \leq N} a_n e_n$ is a finite trigonometric sum then $[F, \pi(f)]$ is a finite rank operator and hence is compact. In general we can uniformly approximate a continuous function by a trigonometric sum and show that the commutator is compact for any continuous f. This shows that $(\mathcal{H}, F)$ is an odd Fredholm module over $C(S^1)$. This Fredholm module is not p-summable for any $1 \leq p < \infty$. If we restrict it to the subalgebra $C^\infty(S^1)$ of smooth functions, then it can be checked that $(\mathcal{H}, F)$ is in fact p-summable for all $p > 1$, but is not 1-summable even in this case.

Now let me describe Connes' noncommutative Chern character from K-homology to cyclic cohomology. Let $(\mathcal{H},\, F)$ be an odd p-summable Fredholm module over an algebra $\mathcal{A}$. For any odd integer $2n-1$ such that $2n \geq p$, Connes defines a cyclic $(2n-1)$-cocycle φ_{2n-1} on $\mathcal{A}$ by [**9**]

$$\varphi_{2n-1}(a_0, a_1, \ldots, a_{2n-1}) = \operatorname{Tr}\left(F[F,\, a_0][F,\, a_1]\cdots[F,\, a_{2n-1}]\right), \tag{22}$$

where Tr denotes the operator trace and instead of $\pi(a)$ we simply write a. Notice that by our p-summability assumption, each commutator is in $\mathcal{L}^p(\mathcal{H})$ and hence, by Hölder inequality for Schatten class operators, their product is in fact a trace class operator as soon as $2n \geq p$. One checks by a direct computation that φ_{2n-1} is a cyclic cocycle.

The next proposition shows that these cyclic cocycles are related to each other via the periodicity S-operator of cyclic cohomology. This is probably how Connes came across the periodicity operator S in the first place.

Proposition 3.1. *For all n with $2n \geq p$ we have*

$$S\varphi_{2n-1} = -(n + \tfrac{1}{2})\, \varphi_{2n+1}.$$

By rescaling φ_{2n-1}'s, one obtains a well defined element in the periodic cyclic cohomology. The (unstable) odd *Connes-Chern character* $\mathrm{Ch}^{2n-1} = \mathrm{Ch}^{2n-1}(\mathcal{H},\, F)$ of an odd finitely summable Fredholm module $(\mathcal{H},\, F)$ over $\mathcal{A}$ is defined by rescaling the cocycles φ_{2n-1} appropriately. Let

$$\mathrm{Ch}^{2n-1}\,(a_0, \ldots, a_{2n-1}) := (-1)^n 2(n - \tfrac{1}{2})\cdots\tfrac{1}{2}\operatorname{Tr}\left(F[F, a_0][F, a_1]\cdots[F, a_{2n-1}]\right).$$

Definition 3.2. *The Connes-Chern character of an odd p-summable Fredholm module $(\mathcal{H}, F)$ over an algebra $\mathcal{A}$ is the class of the cyclic cocycle Ch^{2n-1} in the odd periodic cyclic cohomology group $HP^{odd}(\mathcal{A})$.*

By the above Proposition, the class of Ch^{2n-1} in $HP^{odd}(\mathcal{A})$ is independent of the choice of n.

Let us compute the character of the Fredholm module of the above Example with $\mathcal{A} = C^\infty(S^1)$. By the above definition, $\mathrm{Ch}^1\,(\mathcal{H},\, F) = [\varphi_1]$ is the class of the following cyclic 1-cocycle in $HP^{odd}(\mathcal{A})$:

$$\varphi_1\,(f_0, f_1) = \operatorname{Tr}\left(F[F, f_0][F, f_1]\right).$$

One can identify this cyclic cocycle with a local formula. We claim that

$$\varphi_1\,(f_0, f_1) = \frac{4}{2\pi i}\int f_0 df_1, \quad \text{for all } f_0, f_1 \in \mathcal{A}.$$

By linearity, It suffices to check this relation for basis elements $f_0 = e_m, f_1 = e_n$ for all $m, n \in \mathbb{Z}$, which is easy to do.

The duality, that is, the bilinear pairing, between K-theory and K-homology is defined through the Fredholm index. More precisely there is an *index pairing* between odd (resp. even) Fredholm modules over $\mathcal{A}$ and the algebraic K-theory group $K_1^{\mathrm{alg}}(\mathcal{A})$ (resp. $K_0(\mathcal{A})$). We shall describe it only in the odd case at hand. Let $(\mathcal{H},\, F)$ be an odd Fredholm module over $\mathcal{A}$ and let $U \in \mathcal{A}^\times$ be an invertible element in $\mathcal{A}$. Let $P = \frac{F+1}{2} : \mathcal{H} \to \mathcal{H}$ be the projection operator defined by F. One checks that the operator

$$PUP : P\mathcal{H} \to P\mathcal{H}$$

is a Fredholm operator. In fact, using the compactness of commutators $[F, a]$, one checks that $PU^{-1}P$ is an inverse for PUP modulo compact operators, which of course implies that PUP is a Fredholm operator. The index pairing is then defined as

$$\langle(\mathcal{H},\, F),\, [U]\rangle := \operatorname{index}(PUP),$$

where the index on the right hand side is the Fredholm index. If the invertible U happens to be in $M_n(\mathcal{A})$ we can apply this definition to the algebra $M_n(\mathcal{A})$ by noticing that $(\mathcal{H}\otimes\mathbb{C}^n, F\otimes 1)$ is a Fredholm module over $M_n(\mathcal{A})$ in a natural way. The resulting map can be shown to induce a well defined additive map

$$\langle(\mathcal{H},\, F),\, -\rangle : K_1^{\text{alg}}(\mathcal{A}) \to \mathbb{C}.$$

Notice that this map is purely topological in the sense that to define it we did not have to impose any finite summability, i.e., smoothness, condition on the Fredholm module.

Going back to our example and choosing $f : S^1 \to GL_1(\mathbb{C})$ a continuous function on S^1 representing an element of $K_1^{\text{alg}}(C(S^1))$, the index pairing $\langle[(\mathcal{H},\, F)],\, [f]\rangle = \operatorname{index}(PfP)$ can be explicitly calculated. In fact in this case a simple homotopy argument gives the index of the *Toeplitz operator* $PfP : P\mathcal{H} \to P\mathcal{H}$ in terms of the winding number of f around the origin:

$$\langle[(H,\, F)],\, [f]\rangle = -W(f, 0).$$

Of course, when f is smooth the winding number can be computed by integrating the 1-form $\frac{1}{2\pi i}\frac{dz}{z}$ over the curve defined by f:

$$W(f, 0) = \frac{1}{2\pi i}\int f^{-1}df = \frac{1}{2\pi i}\varphi(f^{-1}, f)$$

where φ is the cyclic 1-cocycle on $C^\infty(S^1)$ defined by $\varphi(f, g) = \int f dg$. This is a special case of a very general index formula proved by Connes [**9**] in a fully noncommutative situation:

Proposition 3.2. *Let $(\mathcal{H},\, F)$ be an odd p-summable Fredholm module over an algebra $\mathcal{A}$ and let $2n-1$ be an odd integer such that $2n \geq p$. If u is an invertible element in A then*

$$\operatorname{index}(PuP) = \frac{(-1)^n}{2^{2n}}\varphi_{2n-1}(u^{-1}, u, \ldots, u^{-1}, u),$$

where the cyclic cocycle φ_{2n-1} is defined by

$$\varphi_{2n-1}(a_0, a_1, \ldots, a_{2n-1}) = \operatorname{Tr}(F[F, a_0][F, a_1]\cdots[F, a_{2n-1}]).$$

The above index formula can be expressed in a more conceptual manner once Connes' Chern character in K-theory is introduced. In [**4, 9**], Connes shows that the Chern-Weil definition of Chern character on topological K-theory admits a vast generalization to a noncommutative setting. For a noncommutative algebra $\mathcal{A}$ and each integer $n \geq 0$, he defined pairings between cyclic cohomology and K-theory:

$$HC^{2n}(\mathcal{A}) \otimes K_0(\mathcal{A}) \longrightarrow \mathbb{C}, \qquad HC^{2n+1}(\mathcal{A}) \otimes K_1^{\text{alg}}(\mathcal{A}) \longrightarrow \mathbb{C}. \tag{23}$$

These pairings are compatible with the periodicity operator S in cyclic cohomology in the sense that

$$\langle[\varphi],\, [e]\rangle = \langle S[\varphi], [e]\rangle,$$

for all cyclic cocycles φ and K-theory classes $[e]$, and thus induce a pairing

$$HP^i(\mathcal{A}) \otimes K_i^{\text{alg}}(\mathcal{A}) \longrightarrow \mathbb{C}, \quad i = 0, 1$$

between periodic cyclic cohomology and K-theory.

We briefly recall its definition. Let φ be a cyclic $2n$-cocycle on $\mathcal{A}$ and let $e \in M_k(\mathcal{A})$ be an idempotent representing a class in $K_0(\mathcal{A})$. The pairing $HC^{2n}(\mathcal{A}) \otimes K_0(\mathcal{A}) \longrightarrow \mathbb{C}$ is defined by

$$\langle [\varphi],\, [e] \rangle = (n!)^{-1}\, \tilde{\varphi}(e, \ldots, e), \tag{24}$$

where $\tilde{\varphi}$ is the 'extension' of φ to $M_k(\mathcal{A})$ defined by the formula

$$\tilde{\varphi}(m_0 \otimes a_0, \ldots, m_{2n} \otimes a_{2n}) = \operatorname{tr}(m_0 \cdots m_{2n})\varphi(a_0, \ldots, a_{2n}). \tag{25}$$

It can be shown that $\tilde{\varphi}$ is a cyclic n-cocycle as well.

The formulas in the *odd case* are as follows. Given an invertible matrix $u \in M_k(\mathcal{A})$, representing a class in $K_1^{\text{alg}}(\mathcal{A})$, and an odd cyclic $(2n-1)$-cocycle φ on $\mathcal{A}$, the pairing is given by

$$\langle [\varphi],\, [u] \rangle := \frac{2^{-(2n+1)}}{(n-\frac{1}{2}) \cdots \frac{1}{2}}\, \tilde{\varphi}(u^{-1} - 1, u - 1, \ldots, u^{-1} - 1, u - 1). \tag{26}$$

Any cyclic cocycle can be represented by a *normalized* cocycle for which $\varphi(a_0, \ldots, a_n) = 0$ if $a_i = 1$ for some i. When φ is normalized, formula (26) reduces to a particularly simple form:

$$\langle [\varphi],\, [u] \rangle = \frac{2^{-(2n+1)}}{(n-\frac{1}{2}) \cdots \frac{1}{2}}\, \tilde{\varphi}(u^{-1}, u, \ldots, u^{-1}, u). \tag{27}$$

Using the pairing $HC^{2n-1}(\mathcal{A}) \otimes K_1^{\text{alg}}(\mathcal{A}) \to \mathbb{C}$ and the definition of $\mathrm{Ch}^{2n-1}(H,\, F)$, the above index formula in Proposition (3.2) can be written as

$$\operatorname{index}(PuP) = \langle \mathrm{Ch}^{2n-1}(H,\, F),\, [u] \rangle, \tag{28}$$

or in its stable form

$$\operatorname{index}(PuP) = \langle \mathrm{Ch}^{odd}(H,\, F),\, [u] \rangle.$$

This equality amounts to the equality *Topological Index = Analytic Index* in a fully noncommutative setting.

An immediate consequence of the index formula (28) is an integrality theorem for numbers defined by the right hand side of (28). This should be compared with classical integrality results for topological invariants of manifolds that are established through the Atiyah-Singer index theorem. An early nice application was Connes' proof of the idempotent conjecture for group C^*-algebras of *free groups* in [**9**]. Among other applications I should mention a mathematical treatment of integral quantum Hall effect, and most recently to quantum computing in the work of Mike Freedman and collaborators [**29**].

4. Cyclic modules

With the introduction of the *cyclic category* Λ in [**7**], Connes took another major step in conceptualizing and generalizing cyclic cohomology far beyond its original inception. We already saw in the last section three different definitions of the cyclic cohomology of an algebra through explicit complexes. The original motivation of [**7**] was to define the cyclic cohomology of algebras as a derived functor. Since the category of algebras and algebra homomorphisms is not an additive category, the standard abelian homological algebra is not applicable here. Let k be a unital commutative ring. In [**7**], an abelian category Λ_k of cyclic k-modules is defined that can be thought of as the 'abelianization' of the category of k-algebras. Cyclic cohomology is then shown to be the derived functor of the *functor of traces*, as we shall explain in this section. More generally Connes defined the notion of a *cyclic object* in an abelian category and its cyclic cohomology [**7**].

Later developments proved that this extension of cyclic cohomology was of great significance. Apart from earlier applications, we should mention the recent work [**16**] where the abelian category of cyclic modules plays a role similar to that of the category of motives for noncommutative geometry. Another recent example is the cyclic cohomology of Hopf algebras [**20, 21, 30, 31**], which cannot be defined as the cyclic cohomology of an algebra or a coalgebra but only as the cyclic cohomology of a cyclic module naturally attached to the given Hopf algebra and a coefficient system (see the last section for more on Hopf cyclic cohomology). Let us briefly sketch the definition of the cyclic category Λ.

Recall that the simplicial category Δ is a small category whose objects are the totally ordered sets

$$[n] = \{0 < 1 < \cdots < n\}, \qquad n = 0, 1, 2, \ldots,$$

and whose morphisms $f : [n] \to [m]$ are order preserving, i.e. monotone non-decreasing, maps $f : \{0, 1, \ldots, n\} \to \{0, 1, \ldots, m\}$. Of particular interest among the morphisms of Δ are *faces* δ_i and *degeneracies* σ_j,

$$\delta_i : [n-1] \to [n], \quad i = 0, 1, \ldots, n,$$
$$\sigma_j : [n+1] \to [n], \qquad j = 0, 1, \ldots, n.$$

By definition δ_i is the unique injective morphism missing i and σ_j is the unique surjective morphism identifying j with $j+1$.

The *cyclic category* Λ has the same set of objects as Δ and in fact contains Δ as a subcategory. Morphisms of Λ are generated by simplicial morphisms and new morphisms $\tau_n : [n] \to [n], n \geq 0$, defined by $\tau_n(i) = i+1$ for $0 \leq i < n$ and $\tau_n(n) = 0$. We have the following extra relations:

$$\begin{aligned} \tau_n \delta_i &= \delta_{i-1} \tau_{n-1}, \quad \tau_n \delta_0 = \delta_n, && 1 \leq i \leq n, \\ \tau_n \sigma_i &= \sigma_{i-1} \tau_{n+1}, \quad \tau_n \sigma_0 = \sigma_n \tau_{n+1}^2 && 1 \leq i \leq n, \\ \tau_n^{n+1} &= \mathrm{id}. \end{aligned}$$

It can be shown that the classifying space $B\Lambda$ of the small category Λ is homotopy equivalent to the classifying space of the circle S^1 [**7**].

A *cyclic object* in a category $\mathcal{C}$ is a functor $\Lambda^{\mathrm{op}} \to \mathcal{C}$. A *cocyclic object* in $\mathcal{C}$ is a functor $\Lambda \to \mathcal{C}$. For any commutative unital ring k, we denote the category of cyclic k-modules by Λ_k. A morphism of cyclic k-modules is a natural transformation between the corresponding functors. It is clear that Λ_k is an abelian category. More generally, if $\mathcal{A}$ is an abelian category then the category $\Lambda\mathcal{A}$ of cyclic objects in $\mathcal{A}$ is itself an abelian category.

Let Alg_k denote the category of unital k-algebras and unital algebra homomorphisms. There is a functor

$$\natural : \mathrm{Alg}_k \longrightarrow \Lambda_k, \quad A \mapsto A^\natural,$$

defined by

$$A_n^\natural = A^{\otimes(n+1)}, \qquad n \geq 0,$$

with face, degeneracy and cyclic operators given by

$$\begin{aligned}
\delta_i(a_0 \otimes a_1 \otimes \cdots \otimes a_n) &= a_0 \otimes \cdots \otimes a_i a_{i+1} \otimes \cdots \otimes a_n,\\
\delta_n(a_0 \otimes a_1 \otimes \cdots \otimes a_n) &= a_n a_0 \otimes a_1 \otimes \cdots \otimes a_{n-1},\\
\sigma_i(a_0 \otimes a_1 \otimes \cdots \otimes a_n) &= a_0 \otimes \cdots \otimes a_i \otimes 1 \otimes \cdots \otimes a_n,\\
\tau_n(a_0 \otimes a_1 \otimes \cdots \otimes a_n) &= a_n \otimes a_0 \otimes \cdots \otimes a_{n-1}.
\end{aligned}$$

A unital algebra map $f : A \to B$ induces a morphism of cyclic modules $f^\natural : A^\natural \to B^\natural$ by $f^\natural(a_0 \otimes \cdots \otimes a_n) = (f(a_0) \otimes \cdots \otimes f(a_n))$.

This functor $\natural$ embeds the non-additive category of k-algebras into the abelian category of cyclic k-modules. A first main observation of [**7**] is that

$$\mathrm{Hom}_{\Lambda_k}(A^\natural,\ k^\natural) \simeq T(A),$$

where $T(A)$ is the space of traces from $A \to k$. To a trace τ one associate the cyclic map $(f_n)_{n\geq 0}$, where

$$f_n(a_0 \otimes a_1 \otimes \cdots \otimes a_n) = \tau(a_0 a_1 \cdots a_n), \quad n \geq 0.$$

It can be easily shown that this defines a one to one correspondence.

Now we can state the following fundamental theorem of Connes [**7**] which greatly extends the above observation and shows that cyclic cohomology is a derived functor, in fact an Ext functor, provided that we work in the category of cyclic modules:

THEOREM 4.1. *Let k be a field of characteristic zero. For any unital k-algebra A, there is a canonical isomorphism*

$$HC^n(A) \simeq \mathrm{Ext}^n_{\Lambda_k}(A^\natural, k^\natural), \qquad \text{for all } n \geq 0.$$

Apart from their applications in the study of cyclic cohomology of algebras and Hopf algebras (about the latter see the next section), cyclic modules have also come to play an important role in applications of noncommutative geometry to number theory. They play a role similar to that of motives in algebraic geometry. Let me briefly explain this point.

The program outlined by Connes, Consani and Marcolli in [**16**] aims at creating an environment where something like Weil's proof of the Riemann hypothesis for function fields can be repeated in the characteristic zero case. Among other things, they produce an analogue of the Frobenius automorphism in characteristic zero in this paper. Since Connes' trace formula is over the noncommutative *adèles class space* [**14**], the geometric setting is that of noncommutative geometry and they must go far beyond what is done so far in noncommutative geometry and import many ideas from modern algebraic geometry to noncommutative geometry. To achieve this, as a first step, good analogues of *étale cohomology*, the *category of motives*, and *correspondences* in noncommutative geometry must be introduced. Happily it turns out that Connes' category of cyclic modules and the closely related bivariant cyclic homology, as well as KK-theory, are quite useful in this regard.

The construction of the Frobenius in characteristic zero follows a very general process that combines cyclic homology with quantum statistical mechanics in a novel way. Starting from a pair (A, φ) of an algebra and a state φ (a noncommutative space endowed with a 'probability measure'), they proceed by invoking the canonical one-parameter group of automorphisms $\sigma = \sigma_\varphi$ and consider the extremal equilibrium states Σ_β at inverse temperatures $\beta > 1$. Under suitable conditions, there is an algebra map

$$\rho : A \rtimes_\sigma \mathbb{R} \to \mathcal{S}(\Sigma_\beta \times \mathbb{R}_+^*) \otimes \mathcal{L},$$

where $\mathcal{L}$ denotes the algebra of trace class operators. The cyclic module $D(A, \varphi)$ is defined as the cokernel of the induced map by $\mathrm{Tr} \circ \rho$ on the cyclic modules of these two algebras. The dual multiplicative group $\mathbb{R}_+^*$ acts on $D(A, \varphi)$ and, in examples coming from number theory, replaces Frobenius in characteristic zero. The three steps involved in the construction of $D(A, \varphi)$ are called *cooling*, *distillation*, and *dual action* in the paper.

A remarkable property of the cyclic category Λ, not shared by the simplicial category, is its *self-duality* in the sense that there is a natural isomorphism of categories $\Lambda \simeq \Lambda^{\mathrm{op}}$ [**7**]. Roughly speaking, the duality functor $\Lambda^{\mathrm{op}} \longrightarrow \Lambda$ acts as the identity on objects of Λ and exchanges face and degeneracy operators while sending the cyclic operator to its inverse. Thus to a cyclic (resp. cocyclic) module one can associate a cocyclic (resp. cyclic) module by applying the duality isomorphism. This duality plays an important role in Hopf cyclic cohomology.

5. The local index formula and beyond

In practice, computing Connes-Chern characters defined by formulas like (22) is rather difficult since they involve the ordinary operator trace and are non-local. Thus one needs to compute the class of this cyclic cocycle by a *local formula*. This is rather similar to passing from the McKean-Singer formula for the index of an elliptic operator to a local cohomological formula involving integrating a locally defined differential form, i.e., the Atiyah-Singer index formula. The solution of this problem was arrived at in two stages. First, in [**13**], Connes gave a partial answer by giving a local formula for the *Hochschild class* of the Chern character, and then Connes and Moscovici gave a formula that captures the full cyclic cohomology class of the character by a local formula [**19**]. Broadly speaking, the ideas involved amount to going from *noncommutative differential topology* to *noncommutative spectral geometry*, and need the introduction of two new concepts.

In the first place, a noncommutative analogue of integration was found by Connes by replacing the operator trace by the Dixmier trace [**11**], and, secondly, one refines the topological notion of Fredholm module by the metric notion of a *spectral triple*, or K-cycles as they were originally named in [**13**]. Developing the necessary tools to handle this local index formula, shaped, more or less, the second stage of the development of noncommutative geometry after the appearance of the landmark papers [**9**]. One can say that while in its first stage noncommutative geometry was influenced by differential and algebraic topology, especially index theory, the Novikov conjecture and the Baum-Connes conjecture, in this second stage it was chiefly informed by spectral geometry.

We start with a quick review of the *Dixmier trace* and the *noncommutative integral*, following [**13**] closely. For a compact operator T, let $\mu_n(T), n = 1, 2, \ldots,$ denote the sequence of eigenvalues of $|T| = (T^*T)^{\frac{1}{2}}$ written in decreasing order. Thus, by the minimax principle, $\mu_1(T) = ||T||$, and in general

$$\mu_n(T) = \inf ||T|_V||, \quad n \geq 1,$$

where the infimum is over the set of subspaces of codimension $n - 1$, and $T|_V$ denotes the restriction of T to the subspace V. The natural domain of the Dixmier trace is the set of operators

$$\mathcal{L}^{1,\infty}(\mathcal{H}) := \{T \in \mathcal{K}(\mathcal{H});\ \sum_{1}^{N} \mu_n(T) = O(\log N)\}.$$

Notice that trace class operators are automatically in $\mathcal{L}^{1,\infty}(\mathcal{H})$. The Dixmier trace of an operator $T \in \mathcal{L}^{1,\infty}(\mathcal{H})$ measures the *logarithmic divergence* of its ordinary trace. More precisely, we are interested in taking some kind of limit of the bounded sequence

$$\sigma_N(T) = \frac{\sum_1^N \mu_n(T)}{\log N}$$

as $N \to \infty$. The problem of course is that, while by our assumption the sequence is bounded, the usual limit may not exists and must be replaced by a carefully chosen 'generalized limit'.

To this end, let $\mathrm{Trace}_\Lambda(T), \Lambda \in [1, \infty)$, be the piecewise affine interpolation of the partial trace function $\mathrm{Trace}_N(T) = \sum_1^N \mu_n(T)$. Recall that a state on a C^*-algebra is a non-zero positive linear functional on the algebra. Let $\omega : C_b[e, \infty) \to \mathbb{C}$ be a normalized state on the algebra of bounded continuous functions on $[e, \infty)$ such that $\omega(f) = 0$ for all f vanishing at ∞. Now, using ω, the Dixmier trace of a positive operator $T \in \mathcal{L}^{1,\infty}(\mathcal{H})$ is defined as

$$\mathrm{Tr}_\omega(T) := \omega(\tau_\Lambda(T)),$$

where

$$\tau_\Lambda(T) = \frac{1}{\log \Lambda} \int_e^\Lambda \frac{\mathrm{Trace}_r(T)}{\log r} \frac{dr}{r}$$

is the Cesàro mean of the function $\frac{\mathrm{Trace}_r(T)}{\log r}$ over the multiplicative group $\mathbb{R}_+^*$. One then extends Tr_ω to all of $\mathcal{L}^{1,\infty}(\mathcal{H})$ by linearity.

The resulting linear functional Tr_ω is a positive trace on $\mathcal{L}^{(1,\infty)}(\mathcal{H})$. It is easy to see from its definition that if T actually happens to be a trace class operator then $\mathrm{Tr}_\omega(T) = 0$ for all ω, i.e., the Dixmier trace is invariant under perturbations by trace class operators. This is a very useful property and makes Tr_ω a flexible tool in computations. The Dixmier trace, Tr_ω, in general depends on the limiting procedure ω; however, for the class of operators T for which $\mathrm{Lim}_{\Lambda \to \infty} \tau_\Lambda(T)$ exit, it is independent of the choice of ω and is equal to $\mathrm{Lim}_{\Lambda\to\infty}\tau_\Lambda(T)$. One of the main results proved in [**11**] is that if M is a closed n-dimensional manifold, E is a smooth vector bundle on M, P is a pseudodifferential operator of order $-n$ acting between L^2-sections of E, and $H = L^2(M, E)$, then $P \in \mathcal{L}^{(1,\infty)}(\mathcal{H})$ and, for any choice of ω, $\mathrm{Tr}_\omega(P) = n^{-1}\mathrm{Res}(P)$. Here Res denotes Wodzicki's noncommutative residue. For example, if D is an elliptic first order differential operator, $|D|^{-n}$ is a pseudodifferential operator of order $-n$ and, for any bounded operator a, the Dixmier trace $\mathrm{Tr}_\omega(a|D|^{-n})$ is independent of the choice of ω.

The second ingredient of the local index formula is the notion of *spectral triple* [**13**]. Spectral triples provide a refinement of Fredholm modules. Going from Fredholm modules to spectral triples is similar to going from the conformal class of a Riemannian metric to the metric itself. Spectral triples simultaneously provide a notion of *Dirac operator* in noncommutative geometry, as well as a Riemannian type *distance function* for noncommutative spaces.

To motivate the definition of a spectral triple, we recall that the Dirac operator $D\!\!\!/$ on a compact Riemannian Spin^c manifold acts as an unbounded selfadjoint operator on the Hilbert space $L^2(M,S)$ of L^2-spinors on the manifold M. If we let $C^\infty(M)$ act on $L^2(M,S)$ by multiplication operators, then one can check that for any smooth function f, the commutator $[D,f]=Df-fD$ extends to a bounded operator on $L^2(M,S)$. Now the geodesic distance d on M can be recovered from the following beautiful *distance formula* of Connes [**13**]:

$$d(p,q)=\mathrm{Sup}\{|f(p)-f(q)|;\, \|\,[D,f]\,\|\leq 1\},\qquad \forall p,q\in M.$$

The triple $(C^\infty(M),L^2(M,S),\not{D})$ is a commutative example of a spectral triple. Its general definition, in the odd case, is as follows. This definition should be compared with Definition (3.1).

Definition 5.1. *Let $\mathcal{A}$ be a unital algebra. An odd spectral triple on $\mathcal{A}$ is a triple $(\mathcal{A},\mathcal{H},D)$ consisting of a Hilbert space $\mathcal{H}$, a selfadjoint unbounded operator $D:\mathrm{Dom}(D)\subset\mathcal{H}\to\mathcal{H}$ with compact resolvent, i.e., $(D+\lambda)^{-1}\in\mathcal{K}(\mathcal{H})$, for all $\lambda\notin\mathbb{R}$, and a representation $\pi:\mathcal{A}\to\mathcal{L}(\mathcal{H})$ of $\mathcal{A}$ such that for all $a\in\mathcal{A}$, the commutator $[D,\pi(a)]$ is defined on $\mathrm{Dom}(D)$ and extends to a bounded operator on $\mathcal{H}$.*

The finite summability assumption (21) for Fredholm modules has a finer analogue for spectral triples. For simplicity we shall assume that D is invertible (in general, since $\mathrm{Ker}\,D$ is finite dimensional, by restricting to its orthogonal complement we can always reduce to this case). A spectral triple is called *finitely summable* if for some $n\geq 1$

$$|D|^{-n}\in\mathcal{L}^{1,\infty}(\mathcal{H}). \tag{29}$$

A simple example of an odd spectral triple is $(C^\infty(S^1),L^2(S^1),D)$, where D is the unique selfadjoint extension of the operator $-i\frac{d}{dx}$. Eigenvalues of $|D|$ are $|n|, n\in\mathbb{Z}$, which shows that, if we restrict D to the orthogonal complement of constant functions, then $|D|^{-1}\in\mathcal{L}^{1,\infty}(L^2(S^1))$.

Given a spectral triple $(\mathcal{A},\mathcal{H},D)$, one obtains a Fredholm module $(\mathcal{A},\mathcal{H},F)$ by choosing $F=\mathrm{Sign}\,(D)=D|D|^{-1}$. Connes' Hochschild character formula gives a local expression for the Hochschild class of the Connes-Chern character of $(\mathcal{A},\mathcal{H},F)$ in terms of D itself. For this one has to assume that the spectral triple $(\mathcal{A},\mathcal{H},D)$ is *regular* in the sense that for all $a\in\mathcal{A}$,

$$a \text{ and } [D,a]\in\cap\,\mathrm{Dom}(\delta^k)$$

where the derivation δ is given by $\delta(x)=[|D|,x]$.

Now, assuming (29) holds, Connes defines an $(n+1)-$linear functional φ on $\mathcal{A}$ by

$$\varphi(a^0,a^1,\ldots,a^n)=\mathrm{Tr}_\omega(a^0[D,a^1]\cdots[D,a^n]|D|^{-n}).$$

It can be shown that φ is a Hochschild n-cocycle on $\mathcal{A}$. We recall that a Hochschild n-cycle $c\in Z_n(A,A)$ is an element $c=\sum a^0\otimes a^1\otimes\cdots\otimes a^n\in A^{\otimes(n+1)}$ such that its Hochschild boundary $b(c)=0$, where b is defined by (18). The following result, known as *Connes' Hochschild character formula*, computes the Hochschild class of the Chern charcater by a local formula, i.e., in terms of φ:

Theorem 5.1. *Let $(\mathcal{A},\mathcal{H},D)$ be a regular spectral triple. Let $F=\mathrm{Sign}\,(D)$ denote the sign of D and $\tau_n\in HC^n(\mathcal{A})$ denote the Connes-Chern charcater of $(\mathcal{H},F)$. For every n-dimensional Hochschild cycle $c=\sum a^0\otimes a^1\otimes\cdots\otimes a^n\in Z_n(\mathcal{A},\mathcal{A})$, one has*

$$\langle\tau_n,c\rangle=\sum\varphi(a^0,a^1,\ldots,a^n).$$

Identifying the full cyclic cohomology class of the Connes-Chern character of $(\mathcal{A}, \mathcal{H}, D)$ by a local formula is the content of Connes-Moscovici's local index formula. For this we have to assume the spectral triple satisfies another technical condition. Let $\mathcal{B}$ denote the subalgebra of $\mathcal{L}(\mathcal{H})$ generated by operators $\delta^k(a)$ and $\delta^k([D, a])$, $k \geq 1$. A spectral triple is said to have a discrete *dimension spectrum* Σ if $\Sigma \subset \mathbb{C}$ is discrete and for any $b \in \mathcal{B}$ the function

$$\zeta_b(z) = \mathrm{Trace}(b|D|^{-z}), \qquad \mathrm{Re}\, z > n,$$

extends to a holomorphic function on $\mathbb{C} \setminus \Sigma$. It is further assumed that Σ is simple in the sense that $\zeta_b(z)$ has only simple poles in Σ.

The local index formula of Connes and Moscovici [**19**] is given by the following Theorem (we have used the formulation in [**15**]):

THEOREM 5.2. *1. The equality*

$$⨍ P = \mathrm{Res}_{z=0}\, \mathrm{Trace}(P|D|^{-z})$$

defines a trace on the algebra generated by $\mathcal{A}$, $[D, \mathcal{A}]$, and $|D|^z$, $z \in \mathbb{C}$.
2. There are only a finite number of non-zero terms in the following formula which defines the odd components $(\varphi_n)_{n=1,3,\ldots}$ of an odd cyclic cocycle in the (b, B) bicomplex of $\mathcal{A}$: For each odd integer n let

$$\varphi_n(a^0, \ldots, a^n) := \sum_k c_{n,k} ⨍ a^0[D, a^1]^{(k_1)} \cdots [D, a^n]^{(k_n)}|D|^{-n-2|k|}$$

where $T^{(k)} := \nabla^k$ and $\nabla(T) = D^2T - TD^2$, k is a multi-index, $|k| = k_1 + \cdots + k_n$ and

$$c_{n,k} := (-1)^{|k|}\sqrt{2i}(k_1! \cdots k_n!)^{-1}((k_1 + 1) \cdots (k_1 + k_2 + \cdots k_n))^{-1}\Gamma(|k| + \frac{n}{2}).$$

3. The pairing of the cyclic cohomology class $(\varphi_n) \in HC^*(\mathcal{A})$ with $K_1(\mathcal{A})$ gives the Fredholm index of D with coefficients in $K_1(\mathcal{A})$.

As is indicated in part 1) of the above Theorem, a regular spectral triple necessarily defines a trace on its underlying algebra by the formula $a \in \mathcal{A} \mapsto ⨍ a = \mathrm{Res}_{z=0}\, \mathrm{Trace}(a|D|^{-z})$. Thus, to deal with 'type III algebras' which carry no non-trivial traces, the notion of spectral triple must be modified. In [**25**] Connes and Moscovici define a notion of *twisted spectral triple*, where the twist is afforded by an algebra automorphism (related to the modular automorphism group). More precisely, one postulates that there exists an automorphism σ of $\mathcal{A}$ such that the twisted commutators

$$[D, a]_\sigma := Da - \sigma(a)D$$

are bounded operators for all $a \in \mathcal{A}$. They show that, in the twisted case, the Dixmier trace induces a twisted trace on the algebra $\mathcal{A}$, but surprisingly, under some regularity conditions, the Connes-Chern character of the phase space lands in ordinary cyclic cohomology. Thus its pairing with ordinary K-theory makes sense, and it can be recovered as the index of Fredholm operators. This suggests the significance of developing a local index formula for twisted spectral triples, *i.e.* finding a formula for a cocycle, cohomologous to the Connes-Chern character in the (b, B)-bicomplex, which is given in terms of twisted commutators and residue functionals. I beleive that this new theme of twisted spectral triples, and type III noncommutative geometry in general, will dominate the subject in near future.

For example, very recently a local index formula has been proved for a class of twisted spectral triples by Henri Moscovici [**34**] that can be found in the present volume. This class is obtained by twisting an ordinary spectral triple $(\mathcal{A}, \mathcal{H}, D)$ by a subgroup G of conformal similarities of the triple, *i.e.* the set of all unitary operators $U \in \mathcal{U}(\mathcal{H})$ such that $U\mathcal{A}U^* = \mathcal{A}$, and $UDU^* = \mu(U)D$, with $\mu(U) > 0$. It is shown that the crossed product algebra $\mathcal{A} \rtimes G$ admits an automorphism σ, given by the formula $\sigma(aU) = \mu(U)^{-1}aU$, for all $a \in \mathcal{A}, U \in G$, and $(\mathcal{A} \rtimes G, \mathcal{H}, D)$ is a twisted spectral triple. The analogue of the noncommutative residue on the circle, for algebras of formal twisted pseudodifferential symbols, is constructed in [**27**].

A very recent development related to (twisted) spectral triples is the noncommutative Gauss-Bonnet theorem of Connes and Tretkoff for the noncommutative two-torus A_θ [**26**]. In classical geometry a *spectral zeta function* is associated to the Laplacian $\Delta_g = d^*d$ of a Riemann surface with metric g:

$$\zeta(s) = \sum_j \lambda_j^{-s}, \ \ \mathrm{Re}(s) > 1,$$

where the λ_j' s are the nonzero eigenvalues of Δ_g. This zeta function has a meromorphic continuation with no pole at 0, and the Gauss-Bonnet theorem for surfaces can be expressed as

$$\zeta(0) + \mathrm{Card}\{j|\lambda_j = 0\} = \frac{1}{12\pi}\int_\Sigma R = \frac{1}{6}\chi(\Sigma),$$

where R is the curvature and $\chi(\Sigma)$ is the Euler-Poincaré characteristic.

It is this formulation of the Gauss-Bonnet theorem in spectral terms that admits a generalization to noncommutative geometry. Let A_θ denote the C^*-algebra of the noncommutative torus with parameter $\theta \in \mathbb{R} \setminus \mathbb{Q}$ and let $\tau : A_\theta \to \mathbb{C}$ denote its faithful normalized trace. One can define an inner product

$$\langle a, b\rangle = \tau(b^*a), \ \ a, b \in A_\theta,$$

and complete A_θ with respect to this inner product to obtain a Hilbert space $\mathcal{H}_0$. More generally, for any smooth selfadjoint element $h = h^* \in A_\theta$ one defines an inner product $\langle a, b\rangle_\varphi = \tau(b^*ae^{-h})$, where the positive linear functional $\varphi = \varphi_h$ is defined by

$$\varphi(a) = \tau(ae^{-h}), \quad a \in A_\theta.$$

Let $\mathcal{H}_\varphi$ denote the completion of A_θ with respect to this conformally equivalent metric.

Using the canonical derivations δ_1 and δ_2 of A_θ, one can introduce a complex structure on A_θ by defining

$$\partial = \delta_1 + i\delta_2, \ \ \partial^* = \delta_1 - i\delta_2.$$

These operators can be considered as unbounded operators on $\mathcal{H}_0$ and ∂^* is the adjoint of ∂. Then the *unperturbed Laplacian* on A_θ is given by

$$\Delta = \partial^*\partial = \delta_1^2 + \delta_2^2.$$

In general we can consider the unbounded operator $\partial = \delta_1 + i\delta_2 : \mathcal{H}_\varphi \to \mathcal{H}^{(1,0)}$, where $\mathcal{H}^{(1,0)}$ is the completion of the linear span of elements of the form $a\partial b$ with $a, b \in A_\theta^\infty$. Let ∂_φ^* denote its adjoint. Then the Laplacian for the conformally equivalent metric $\langle a, b\rangle_\varphi$ is given by $\Delta' = \partial_\varphi^*\partial$.

In [**26**], Connes and Tretkoff show that the value at 0 of the zeta function associated to this Laplacian Δ' is an invariant of the conformal class of the metric on A_θ, i.e. of h. A natural problem here is to extend this result by considering the most general complex structure on A_θ of the form $\partial = \delta_1 + \tau\delta_2$, where τ is a complex number with $\mathrm{Im}(\tau) > 0$. This problem is now solved in full generality in [**28**].

6. Hopf cyclic cohomology

A major development in cyclic cohomology theory in the last ten years was the introduction of *Hopf cyclic cohomology* for Hopf algebras by Connes and Moscovici [**20**]. As we saw in Section 5, the local index formula gives the Connes-Chern character of a regular spectral triple $(\mathcal{A}, \mathcal{H}, D)$ as a cyclic cocycle in the (b, B)-bicomplex of the algebra $\mathcal{A}$. For spectral triples of interest in transverse geometry [**20**], this cocycle is *differentiable* in the sense that it is in the image of the Connes-Moscovici characteristic map χ_τ defined below (31), with $H = \mathcal{H}_1$ a Hopf algebra and $\mathcal{A} = \mathcal{A}_\Gamma$, a noncommutative algebra, whose definitions we shall recall in this section. To identify this cyclic cocycle in terms of characteristic classes of foliations, they realized that it would be extremely helpful to show that it is the image of a polynomial in some universal cocycles for a cohomology theory for a universal Hopf algebra, and this gave birth to Hopf cyclic cohomology and to the universal Hopf algebra $H = \mathcal{H}_1$. This is similar to the situation for classical characteristic classes of manifolds, which are pullbacks of group cohomology classes.

The Connes-Moscovici characteristic map can be formulated in general terms as follows. Let H be a Hopf algebra acting as *quantum symmetries* of an algebra $\mathcal{A}$, i.e., $\mathcal{A}$ is a left H-module, and the algebra structure of $\mathcal{A}$ is compatible with the coalgebra structure of H in the sense that the multiplication $\mathcal{A} \times \mathcal{A} \to \mathcal{A}$ and the unit map $\mathbb{C} \to \mathcal{A}$ of $\mathcal{A}$ are morphisms of H-modules. A common terminology to describe this situation is to say that $\mathcal{A}$ is a left H-module algebra. Using Sweedler's notation for the coproduct of H, $\Delta(h) = h^{(1)} \otimes h^{(2)}$ (summation is understood), this latter compatibility condition can be expressed as

$$h(ab) = h^{(1)}(a)h^{(2)}(b), \quad \text{and} \quad h(1) = \varepsilon(h)1,$$

for all $h \in H$ and $a, b \in \mathcal{A}$. In general one should think of such actions of Hopf algebras as the noncommutative geometry analogue of the action of differential operators on a manifold.

It is also important to extend the notion of trace to allow *twisted traces*, such as KMS states in quantum statistical mechanics, as well as the idea of *invariance* of a (twisted) trace. The general setting introduced in [**20**] is the following. Let $\delta : H \to \mathbb{C}$ be a character of H, i.e. a unital algebra map, and $\sigma \in H$ be a grouplike element, i.e. it satisfies $\Delta\sigma = \sigma \otimes \sigma$. A linear map $\tau : A \to \mathbb{C}$ is called δ-*invariant* if for all $h \in H$ and $a \in \mathcal{A}$,

$$\tau(h(a)) = \delta(h)\tau(a),$$

and is called a σ-*trace* if for all a, b in $\mathcal{A}$,

$$\tau(ab) = \tau(b\sigma(a)).$$

Now for $a, b \in \mathcal{A}$, let

$$\langle a,\, b\rangle := \tau(ab).$$

Let τ be a σ-trace on $\mathcal{A}$. Then τ is δ-invariant if and only if the *integration by parts formula* holds. That is, for all $h \in H$ and $a, b \in A$,

$$\langle h(a), b\rangle = \langle a, \widetilde{S}_\delta(h)(b)\rangle. \tag{30}$$

Here S denotes the antipode of H and the δ-*twisted antipode* $\widetilde{S}_\delta : H \to H$ is defined by $\widetilde{S}_\delta = \delta * S$, i.e.

$$\widetilde{S}_\delta(h) = \delta(h^{(1)})S(h^{(2)}),$$

for all $h \in H$. Loosely speaking, (30) says that the formal adjoint of the differential operator h is $\widetilde{S}_\delta(h)$. Following [**20, 21**], we say that (δ, σ) is a *modular pair* if $\delta(\sigma) = 1$, and a *modular pair in involution* if in addition we have

$$\widetilde{S}_\delta^2(h) = \sigma h \sigma^{-1},$$

for all h in H. The importance of this notion will become clear in the next paragraph.

Now, for each $n \geq 0$, the Connes-Moscovici *characteristic map*

$$\chi_\tau : H^{\otimes n} \longrightarrow C^n(\mathcal{A}), \tag{31}$$

is defined by

$$\chi_\tau(h_1 \otimes \cdots \otimes h_n)(a_0 \otimes \cdots \otimes a_n) = \tau(a_0 h_1(a_1) \cdots h_n(a_n)).$$

Notice that the right hand side of (31) is the cocyclic module that (its cohomology) defines the cyclic cohomology of the algebra $\mathcal{A}$. The main question about (31) is whether one can promote the collection of linear spaces $\{H^{\otimes n}\}_{n\geq 0}$ to a cocyclic module such that the characteristic map χ_τ turns into a morphism of cocyclic modules. We recall that the face, degeneracy, and cyclic operators for $\{C^n(\mathcal{A})\}_{n\geq 0}$ are defined by:

$$\begin{aligned}
\delta_i \varphi(a_0, \ldots, a_{n+1}) &= \varphi(a_0, \ldots, a_i a_{i+1}, \ldots, a_{n+1}), \quad i = 0, \ldots, n,\\
\delta_{n+1} \varphi(a_0, \ldots, a_{n+1}) &= \varphi(a_{n+1} a_0, a_1, \ldots, a_n),\\
\sigma_i \varphi(a_0, \ldots, a_n) &= \varphi(a_0, \ldots, a_i, 1, \ldots, a_n), \quad i = 0, \ldots, n,\\
\tau_n \varphi(a_0, \ldots, a_n) &= \varphi(a_n, a_0, \ldots, a_{n-1}).
\end{aligned}$$

The relation $h(ab) = h^{(1)}(a)h^{(2)}(b)$ shows that, in order for χ_τ to be compatible with face operators, the face operators δ_i on $H^{\otimes n}$, at least for $0 \leq i < n$, must involve the coproduct of H. In fact if we define, for $0 \leq i \leq n$, $\delta_i^n : H^{\otimes n} \to H^{\otimes (n+1)}$, by

$$\begin{aligned}
\delta_0(h_1 \otimes \cdots \otimes h_n) &= 1 \otimes h_1 \otimes \cdots \otimes h_n,\\
\delta_i(h_1 \otimes \cdots \otimes h_n) &= h_1 \otimes \cdots \otimes h_i^{(1)} \otimes h_i^{(2)} \otimes \cdots \otimes h_n,\\
\delta_{n+1}(h_1 \otimes \cdots \otimes h_n) &= h_1 \otimes \cdots \otimes h_n \otimes \sigma,
\end{aligned}$$

then we have, for all $i = 0, 1, \ldots, n$

$$\chi_\tau \delta_i = \delta_i \chi_\tau.$$

Notice that the last relation is a consequence of the σ-trace property of τ. Similarly, the relation $h(1_A) = \varepsilon(h)1_A$ shows that the degeneracy operators on $H^{\otimes n}$ should involve the counit of H. We thus define

$$\sigma_i(h_1 \otimes \cdots \otimes h_n) = h_1 \otimes \cdots \otimes \varepsilon(h_i) \otimes \cdots \otimes h_n.$$

It is very hard, on the other hand, to come up with a correct formula for the *cyclic operator* $\tau_n : H^{\otimes n} \to H^{\otimes n}$. Compatibility with χ_τ demands that

$$\tau(a_0\tau_n(h_1 \otimes \cdots \otimes h_n)(a_1 \otimes \cdots \otimes a_n)) = \tau(a_n h_1(a_0)h_2(a_1)\cdots h_n(a_{n-1})),$$

for all a_i's and h_i's. For $n = 1$, the integration by parts formula (30), combined with the σ-trace property of τ, shows that

$$\tau(a_1 h(a_0)) = \tau(h(a_0)\sigma(a_1)) = \tau(a_0 \tilde{S}_\delta(h)\sigma(a_1)).$$

This suggests that we should define $\tau_1 : H \to H$ by

$$\tau_1(h) = \tilde{S}_\delta(h)\sigma.$$

Note that the required cyclicity condition for τ_1, $\tau_1^2 = 1$, is equivalent to the involution condition $\widetilde{S}_\delta^2(h) = \sigma h \sigma^{-1}$ for the pair (δ, σ). This line of reasoning can be extended to all $n \geq 0$ and gives us:

$$\begin{aligned}
\tau(a_n h_1(a_0)\cdots h_n(a_{n-1})) &= \tau(h_1(a_0)\cdots h_n(a_{n-1})\sigma(a_n)) \\
&= \tau(a_0 \tilde{S}_\delta(h_1)(h_2(a_1)\cdots h_n(a_{n-1})\sigma(a_n))) \\
&= \tau(a_0 \tilde{S}_\delta(h_1)\cdot(h_2 \otimes \cdots \otimes h_n \otimes \sigma)(a_1 \otimes \cdots \otimes a_n)).
\end{aligned}$$

This suggests that the *Hopf-cyclic operator* $\tau_n : H^{\otimes n} \to H^{\otimes n}$ should be defined as

$$\tau_n(h_1 \otimes \cdots \otimes h_n) = \tilde{S}_\delta(h_1)\cdot(h_2 \otimes \cdots \otimes h_n \otimes \sigma),$$

where $\cdot$ denotes the diagonal action defined by

$$h \cdot (h_1 \otimes \cdots \otimes h_n) := h^{(1)}h_1 \otimes h^{(2)}h_2 \otimes \cdots \otimes h^{(n)}h_n.$$

The remarkable fact, proved by Connes and Moscovici [**20, 21**], is that endowed with the above face, degeneracy, and cyclic operators, $\{H^{\otimes n}\}_{n\geq 0}$ is a cocyclic module. The proof is a very clever and complicated tour de force of Hopf algebra identities.

The resulting cyclic cohomology groups, which depend on the choice of a modular pair in involution (δ, σ), are denoted by $HC^n_{(\delta,\sigma)}(H)$, $n = 0, 1, \ldots$. The characteristic map (31) clearly induces a map between corresponding cyclic cohomology groups

$$\chi_\tau : HC^n_{(\delta,\sigma)}(H) \to HC^n(\mathcal{A}).$$

Under this map Hopf cyclic cocycles are mapped to cyclic cocycles on $\mathcal{A}$. Very many of the interesting cyclic cocycles in noncommutative geometry are obtained in this fashion. Using the above discussed cocyclic module structure of $\{H^{\otimes n}\}_{n\geq 0}$, we see that a Hopf cyclic n-cocycle is an element $x \in H^{\otimes n}$ which satisfies the relations

$$bx = 0, \quad (1 - \lambda)x = 0,$$

where $b : H^{\otimes n} \to H^{\otimes(n+1)}$ and $\lambda : H^{\otimes n} \to H^{\otimes n}$ are defined by

$$\begin{aligned}
b(h^1 \otimes \cdots \otimes h^n) &= 1 \otimes h_1 \otimes \cdots \otimes h_n \\
&+ \sum_{i=1}^{n}(-1)^i h_1 \otimes \cdots \otimes h_i^{(1)} \otimes h_i^{(2)} \otimes \cdots \otimes h_n \\
&+ (-1)^{n+1}h_1 \otimes \cdots \otimes h_n \otimes \sigma, \\
\lambda(h_1 \otimes \cdots \otimes h_n) &= (-1)^n \tilde{S}_\delta(h_1)\cdot(h_2 \otimes \cdots \otimes h_n \otimes \sigma).
\end{aligned}$$

The characteristic map (31) has its origins in Connes' earlier work on noncommutative differential geometry [**4**], and on his work on the transverse fundamental class of foliations [**8**]. In fact in these papers some interesting cyclic cocycles were defined in the context of actions of Lie algebras and (Lie) groups. Both examples can be shown to be special cases of the characteristic map. For example let $\mathcal{A} = \mathcal{A}_\theta$ denote the smooth algebra of coordinates for the noncommutative torus with parameter $\theta \in \mathbb{R}$. The abelian Lie algebra $\mathbb{R}^2$ acts on $\mathcal{A}_\theta$ via canonical derivations δ_1 and δ_2. The standard trace τ on $\mathcal{A}_\theta$ is invariant under the action of $\mathbb{R}^2$, i.e., we have $\tau(\delta_1(a)) = \tau(\delta_2(a)) = 0$ for all $a \in \mathcal{A}_\theta$. Then one can directly check that under the characteristic map (31) the two dimensional generator of the Lie algebra homology of $\mathbb{R}^2$ is mapped to the following cyclic 2-cocycle on $\mathcal{A}_\theta$ first defined in [**4**]:

$$\varphi(a_0, a_1, a_2) = \frac{1}{2\pi i}\tau(a_0(\delta_1(a_1)\delta_2(a_2) - \delta_2(a_1)\delta_1(a_2))).$$

For a second example let G be a discrete group and c be a normalized group n-cocycle on G with trivial coefficients. Here by normalized we mean $c(g_1, \ldots, g_n) = 0$ if $g_i = e$ for some i. Then one checks that the following is a cyclic n-cocycle on the group algebra $\mathbb{C}G$ [**8**]:

$$\varphi(g_0, g_1 \ldots, g_n) = \begin{cases} c(g_1, g_2 \ldots, g_n) & \text{if } g_0 g_1 \ldots g_n = 1 \\ \quad 0 & \text{otherwise} \end{cases}$$

After an appropriate dual version of Hopf cyclic cohomology is defined, one can show that this cyclic cocycle can also be defined via (31).

The most sophisticated example of the characteristic map (31), so far, involves the Connes-Moscovici Hopf algebra $\mathcal{H}_1$ and its action on algebras of interest in transverse geometry. In fact, as we shall see, $\mathcal{H}_1$ acts as quantum symmetries of various objects of interest in noncommutative geometry, like the frame bundle of the 'space' of leaves of codimension one foliations or the 'space' of modular forms modulo the action of Hecke correspondences.

To describe $\mathcal{H}_1$, let $\mathfrak{g}_{aff}$ denote the Lie algebra of the group of affine transformations of the line with linear basis X and Y and the relation $[Y, X] = X$. Let $\mathfrak{g}$ be an abelian Lie algebra with basis $\{\delta_n;\ n = 1, 2, \ldots\}$. Its universal enveloping algebra $U(\mathfrak{g})$ should be regarded as the continuous dual of the pro-unipotent group of orientation preserving diffeomorphisms φ of $\mathbb{R}$ with $\varphi(0) = 0$ and $\varphi'(0) = 1$. It is easily seen that $\mathfrak{g}_{aff}$ acts on $\mathfrak{g}$ via

$$[Y, \delta_n] = n\delta_n, \quad [X, \delta_n] = \delta_{n+1},$$

for all n. Let $\mathfrak{g}_{CM} := \mathfrak{g}_{aff} \rtimes \mathfrak{g}$ be the corresponding semidirect product Lie algebra. As an algebra, $\mathcal{H}_1$ coincides with the universal enveloping algebra of the Lie algebra $\mathfrak{g}_{CM}$. Thus $\mathcal{H}_1$ is the universal algebra generated by $\{X, Y, \delta_n; n = 1, 2, \ldots\}$ subject to the relations

$$[Y, X] = X, \;\; [Y, \delta_n] = n\delta_n, \;\; [X, \delta_n] = \delta_{n+1}, \;\; [\delta_k, \delta_l] = 0,$$

for $n, k, l = 1, 2, \ldots$. We let the counit of $\mathcal{H}_1$ coincide with the counit of $U(\mathfrak{g}_{CM})$. Its coproduct and antipode, however, will be certain deformations of the coproduct and antipode of $U(\mathfrak{g}_{CM})$ as follows. Using the universal property of $U(\mathfrak{g}_{CM})$, one checks that the relations

$$\Delta Y = Y \otimes 1 + 1 \otimes Y, \quad \Delta\delta_1 = \delta_1 \otimes 1 + 1 \otimes \delta_1,$$
$$\Delta X = X \otimes 1 + 1 \otimes X + \delta_1 \otimes Y,$$

determine a unique algebra map $\Delta : \mathcal{H}_1 \to \mathcal{H}_1 \otimes \mathcal{H}_1$. Note that Δ is not cocommutative and it differs from the coproduct of the enveloping algebra $U(\mathfrak{g}_{CM})$. Similarly, one checks that there is a unique antialgebra map, the antipode, $S : \mathcal{H}_1 \to \mathcal{H}_1$ determined by the relations

$$S(Y) = -Y, \;\; S(X) = -X + \delta_1 Y, \;\; S(\delta_1) = -\delta_1.$$

The first realization of $\mathcal{H}_1$ was through its action as quantum symmetries of the 'frame bundle' of the noncommutative space of leaves of codimension one foliations. More precisely, given a codimension one foliation $(V, \mathcal{F})$, let M be a smooth transversal for $(V, \mathcal{F})$. Let $A = C_0^\infty(F^+M)$ denote the algebra of smooth functions with compact support on the bundle of positively oriented frames on M and let $\Gamma \subset Diff^+(M)$ denote the holonomy group of $(V, \mathcal{F})$. One has a natural prolongation of the action of Γ to $F^+(M)$ by

$$\varphi(y, y_1) = (\varphi(y), \varphi'(y)(y_1)).$$

Let $A_\Gamma = C_0^\infty(F^+M) \rtimes \Gamma$ denote the corresponding crossed product algebra. Thus the elements of A_Γ consist of finite linear combinations (over $\mathbb{C}$) of terms fU_φ^* with $f \in C_0^\infty(F^+M)$ and $\varphi \in \Gamma$. Its product is defined by

$$fU_\varphi^* \cdot gU_\psi^* = (f \cdot \varphi(g))U_{\psi\varphi}^*.$$

There is an action of $\mathcal{H}_1$ on A_Γ, given by [**20, 23**]:

$$Y(fU_\varphi^*) = y_1 \frac{\partial f}{\partial y_1} U_\varphi^*, \quad X(fU_\varphi^*) = y_1 \frac{\partial f}{\partial y} U_\varphi^*,$$

$$\delta_n(fU_\varphi^*) = y_1^n \frac{d^n}{dy^n}(log \frac{d\varphi}{dy}) fU_\varphi^*.$$

Once these formulas are given, it can be checked, by a long computation, that A_Γ is indeed an $\mathcal{H}_1$-module algebra. To define the corresponding characteristic map, Connes and Moscovici defined a modular pair in involution $(\delta, 1)$ on $\mathcal{H}_1$ and a δ-invariant trace on A_Γ as we shall describe next.

Let δ denote the unique extension of the modular character

$$\delta : \mathfrak{g}_{aff} \to \mathbb{R}, \quad \delta(X) = 1, \; \delta(Y) = 0,$$

to a character $\delta : U(\mathfrak{g}_{aff}) \to \mathbb{C}$. There is a unique extension of δ to a character, denoted by the same symbol $\delta : \mathcal{H}_1 \to \mathbb{C}$. Indeed, the relations $[Y, \delta_n] = n\delta_n$ show that we must have $\delta(\delta_n) = 0$, for $n = 1, 2, \ldots$. One can then check that these relations are compatible with the algebra structure of $\mathcal{H}_1$. The algebra $A_\Gamma = C_0^\infty(F^+(M) \rtimes \Gamma$ admits a δ-invariant trace $\tau : A_\Gamma \to \mathbb{C}$ given by [**20**]:

$$\tau(fU_\varphi^*) = \int_{F^+(M)} f(y, y_1) \frac{dy dy_1}{y_1^2}, \quad \text{if } \varphi = 1,$$

and $\tau(fU_\varphi^*) = 0$, otherwise. Now, using the δ-invariant trace τ and the above defined action $\mathcal{H}_1 \otimes A_\Gamma \to A_\Gamma$, the characteristic map (31) takes the form

$$\chi_\tau : HC^*_{(\delta,1)}(\mathcal{H}_1) \longrightarrow HC^*(A_\Gamma).$$

This map plays a fundamental role in transverse index theory in [**20**].

The Hopf algebra $\mathcal{H}_1$ shows its beautiful head in number theory as well. To give an indication of this, I shall briefly discuss the *modular Hecke algebras* and actions of $\mathcal{H}_1$ on them as they were introduced by Connes and Moscovici in [**23, 24**]. For each $N \geq 1$, let

$$\Gamma(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{Z}); \;\; \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \bmod N \right\}$$

denote the level N congruence subgroup of $\Gamma(1) = SL(2,\mathbb{Z})$. Let $\mathcal{M}_k(\Gamma(N))$ denote the space of modular forms of level N and weight k and

$$\mathcal{M}(\Gamma(N)) := \bigoplus_k \mathcal{M}_k(\Gamma(N))$$

be the graded algebra of modular forms of level N. Finally, let

$$\mathcal{M} := \lim_{\overrightarrow{N}} \mathcal{M}(\Gamma(N))$$

denote the algebra of modular forms of all levels, where the inductive system is defined by divisibility. The group

$$G^+(\mathbb{Q}) := GL^+(2,\mathbb{Q}),$$

acts on $\mathcal{M}$ through its usual action on functions on the upper half-plane (with corresponding weight):

$$(f,\alpha) \mapsto f|_k\alpha(z) = \det(\alpha)^{k/2}(cz+d)^{-k} f(\alpha\cdot z),$$
$$\alpha = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \alpha\cdot z = \frac{az+b}{cz+d}.$$

The simplest modular Hecke algebra is the crossed-product algebra

$$\mathcal{A} = \mathcal{A}_{G^+(\mathbb{Q})} := \mathcal{M} \rtimes G^+(\mathbb{Q}).$$

Elements of this (noncommutative) algebra will be denoted by finite sums $\sum fU_\gamma^*$, $f \in \mathcal{M}$, $\gamma \in G^+(\mathbb{Q})$. $\mathcal{A}$ can be thought of as the algebra of noncommutative coordinates on the noncommutative quotient space of modular forms modulo Hecke correspondences.

Now consider the operator X of degree two on the space of modular forms defined by

$$X := \frac{1}{2\pi i}\frac{d}{dz} - \frac{1}{12\pi i}\frac{d}{dz}(\log\Delta)\cdot Y,$$

where

$$\Delta(z) = (2\pi)^{12}\eta^{24}(z) = (2\pi)^{12} q \prod_{n=1}^{\infty}(1-q^n)^{24}, \quad q = e^{2\pi i z},$$

η is the Dedekind eta function, and Y is the grading operator

$$Y(f) = \frac{k}{2}\cdot f, \quad \text{for all } f \in \mathcal{M}_k.$$

It is shown in [**23**] that there is a unique action of $\mathcal{H}_1$ on $\mathcal{A}_{G^+(\mathbb{Q})}$ determined by

$$X(fU_\gamma^*) = X(f)U_\gamma^*, \quad Y(fU_\gamma^*) = Y(f)U_\gamma^*,$$
$$\delta_1(fU_\gamma^*) = \mu_\gamma\cdot f(U_\gamma^*),$$

where

$$\mu_\gamma(z) = \frac{1}{2\pi i}\frac{d}{dz}\log\frac{\Delta|\gamma}{\Delta}.$$

This action is compatible with the algebra structure, i.e., $\mathcal{A}_{G^+(\mathbb{Q})}$ is an $\mathcal{H}_1$-module algebra. Thus one can think of $\mathcal{H}_1$ as quantum symmetries of the noncommutative space represented by $\mathcal{A}_{G^+(\mathbb{Q})}$.

More generally, for any congruence subgroup Γ, an algebra $A(\Gamma)$ is constructed in [**23**] that contains as subalgebras both the algebra of Γ-modular forms and the Hecke ring at level Γ. There is also a corresponding action of $\mathcal{H}_1$ on $A(\Gamma)$.

The Hopf cyclic cohomology groups $HC^n_{(\delta,\sigma)}(H)$ are computed in several cases in [**20**]. Of particular interest for applications to transverse index theory and number theory is the (periodic) cyclic cohomology of $\mathcal{H}_1$. It is shown in [**20**] that the periodic groups $HP^n_{(\delta,1)}(\mathcal{H}_1)$ are canonically isomorphic to the Gelfand-Fuchs cohomology, with trivial coefficients, of the Lie algebra $\mathfrak{a}_1$ of formal vector fields on the line:

$$HP^*_{(\delta,1)}(\mathcal{H}_1) = \bigoplus_{i\geq 0} H^{*+2i}_{GF}(\mathfrak{a}_1, \mathbb{C}).$$

This result is very significant in that it relates the Gelfand-Fuchs construction of characteristic classes of smooth manifolds to a noncommutative geometric construction of the same via $\mathcal{H}_1$. Connes and Moscovici also identified certain interesting elements in the Hopf cyclic cohomology of $\mathcal{H}_1$. For example, it can be directly checked that the elements $\delta_2' := \delta_2 - \frac{1}{2}\delta_1^2$ and δ_1 are Hopf cyclic 1-cocycles for $\mathcal{H}_1$, and

$$F := X \otimes Y - Y \otimes X - \delta_1 Y \otimes Y$$

is a Hopf cyclic 2-cocycle. Under the characteristic map (31) and for $A = A_\Gamma$ these Hopf cyclic cocycles are mapped to the Schwarzian derivative, the Godbillon-Vey cocycle, and the transverse fundamental class of Connes [**8**], respectively. See [**24**] for detailed calculations as well as relations with modular forms and modular Hecke algebras. Very recently the unstable cyclic cohomology groups of $\mathcal{H}_1$, and a series of other Hopf algebras attached to pseudogroups of geometric structures, were fully computed in [**35, 36**]. In particular it is shown that the groups $HC^n_{(\delta,\sigma)}(\mathcal{H}_1)$ are finite dimensional for all n.

The notion of modular pair in involution (δ, σ) for a Hopf algebra might seem rather ad hoc at a first glance. This is in fact not the case and the concept is very natural and fundamental. For example, it is shown in [**21**] that *coribbon Hopf algebras* and compact quantum groups are endowed with canonical modular pairs of the form $(\delta, 1)$ and, dually, *ribbon Hopf algebras* have canonical modular pairs of the type $(1, \sigma)$. The fundamental importance of modular pairs in involution was further elucidated when *Hopf cyclic cohomology with coefficients* was introduced in [**30, 31**]. It turns out that some very stringent conditions have to be imposed on an H-module M in order for M to serve as a coefficient (local system) for Hopf cyclic cohomology theory. Such modules are called *stable anti-Yetter-Drinfeld modules*. More precisely, a (left-left) anti-Yetter-Drinfeld H-module is a left H-module M which is simultaneously a left H-comodule such that

$$\rho(hm) = h^{(1)}m^{(-1)}S(h^{(3)}) \otimes h^{(2)}m^{(0)},$$

for all $h \in H$ and $m \in M$. Here $\rho : M \to H \otimes M$, $\rho(m) = m^{(-1)} \otimes m^{(0)}$ is the comodule structure map of M. M is called stable if in addition we have

$$m^{(-1)}m^{(0)} = m,$$

for all $m \in M$. Given a stable anti-Yetter-Drinfeld (SAYD) module M over H, one can then define the Hopf cyclic cohomology of H with coefficients in M. One-dimensional SAYD modules correspond to Connes-Moscovici's modular pairs in involution. More precisely, there is a one-to-one correspondence between modular pairs in involution (δ, σ) on H and SAYD module structures on $M = \mathbb{C}$, the ground field, defined by

$$h.r = \delta(h)r, \quad r \mapsto \sigma \otimes r,$$

for all $h \in H$ and $r \in \mathbb{C}$. Thus a modular pair in involution can be regarded as an 'equivariant line bundle' over the noncommutative space represented by the Hopf algebra H.

References

[1] M. F. Atiyah and I. M. Singer, The index of elliptic operators. IV. Ann. of Math. (2), 93 1971, 119-138.

[2] D. Burghelea, The cyclic homology of the group rings. Comment. Math. Helv. **60** (1985), no. 3, 354–365.

[3] A. Connes, A survey of foliations and operator algebras. Operator algebras and applications, Part I (Kingston, Ont., 1980), pp. 521–628, Proc. Sympos. Pure Math., 38, Amer. Math. Soc., Providence, R.I., 1982.

[4] A. Connes, C^* algébres et géométrie différentielle. (French) C. R. Acad. Sci. Paris Sér. A-B **290** (1980), no. 13, A599–A604.

[5] A. Connes, Spectral sequence and homology of currents for operator algebras. Oberwolfach Tagungsber., 41/81, Funktionalanalysis und C^*-Algebren, 27-9/3-10, 1981.

[6] A. Connes, Noncommutative differential geometry. Chapter I: The Chern character in K homology, Preprint IHES octobre 1982 ; Chapter II: de Rham homology and noncommutative algebra, Preprint IHES février 1983.

[7] A. Connes, Cohomologie cyclique et foncteurs Ext^n . (French) (Cyclic cohomology and functors Ext^n) C. R. Acad. Sci. Paris Sr. I Math. **296** (1983), no. 23, 953–958.

[8] A. Connes, Cyclic cohomology and the transverse fundamental class of a foliation. Geometric methods in operator algebras (Kyoto, 1983), 52–144, Pitman Res. Notes Math. Ser., 123, Longman Sci. Tech., Harlow, 1986.

[9] A. Connes, Noncommutative differential geometry. Chapter I: The Chern character in K homology; Chapter II: de Rham homology and noncommutative algebra, Publ. Math. IHES no. 62 (1985), 41-144.

[10] A. Connes, Entire cyclic cohomology of Banach algebras and characters of θ-summable Fredholm modules. K-Theory 1 (1988), no. 6, 519–548.

[11] A. Connes, The action functional in noncommutative geometry. Comm. Math. Phys. 117 (1988), no. 4, 673–683.

[12] A. Connes and M. Karoubi, Caractère multiplicatif d'un module de Fredholm. (French) [Multiplicative character of a Fredholm module] K-Theory 2 (1988), no. 3, 431–463.

[13] A. Connes, Noncommutative geometry. Academic Press, Inc., San Diego, CA, 1994.

[14] A. Connes, Trace formula in noncommutative geometry and the zeros of the Riemann zeta function. Selecta Math. (N.S.) 5 (1999), no. 1, 29–106.

[15] A. Connes, Noncommutative geometry year 2000. Highlights of mathematical physics (London, 2000), 49–110, Amer. Math. Soc., Providence, RI, 2002.

[16] A. Connes, C. Consani and M. Marcolli, Noncommutative geometry and motives: the thermodynamics of endomotives, Advances in Math. 214 (2) (2007).

[17] A. Connes and M. Marcolli, Noncommutative geometry, quantum fields and motives. American Mathematical Society Colloquium Publications, 55. American Mathematical Society, Providence, RI; Hindustan Book Agency, New Delhi, 2008. xxii+785 pp.

[18] A. Connes and H. Moscovici, Cyclic cohomology, the Novikov conjecture and hyperbolic groups. Topology 29 (1990), no. 3, 345–388.

[19] A. Connes and H. Moscovici, The local index formula in noncommutative geometry. Geom. Funct. Anal. **5** (1995), no. 2, 174–243.

[20] A. Connes and H. Moscovici, Hopf algebras, cyclic cohomology and the transverse index theorem, Comm. Math. Phys. **198** (1998), no. 1, 199–246.

[21] A. Connes and H. Moscovici, Cyclic cohomology and Hopf algebra symmetry. Conference Moshé Flato 1999 (Dijon). Lett. Math. Phys. **52** (2000), no. 1, 1–28.

[22] A. Connes and H. Moscovici, Rankin-Cohen brackets and the Hopf algebra of transverse geometry. Mosc. Math. J. 4 (2004), no. 1, 111–130.

[23] A. Connes and H. Moscovici, Modular Hecke algebras and their Hopf symmetry. Mosc. Math. J. 4 (2004), no. 1, 67–109, 310.

[24] A. Connes and H. Moscovici, Background independent geometry and Hopf cyclic cohomology, math.QA/0505475.

[25] A. Connes and H. Moscovici, Type III and spectral triples. Traces in number theory, geometry and quantum fields, 57–71, Aspects Math., E38, Friedr. Vieweg, Wiesbaden, 2008.
[26] A. Connes and P. Tretkoff, The Gauss-Bonnet theorem for the noncommutative two torus, arXiv:0910.0188v1.
[27] F. Fathi Zadeh and M. Khalkhali, The algebra of formal twisted pseudodifferential symbols and a noncommutative residue, arXiv:0810.0484.
[28] F. Fathi Zadeh and M. Khalkhali, The Gauss-Bonnet theorem for noncommutative two tori with a general conformal structure, arXiv:1005.4947. To appear in Journal of Noncommutative Geometry.
[29] M. Freedman, Towards quantum characteristic classes, Lecture at MSRI, August 11, 2008, available at http://www.msri.org/communications/vmath/VMathVideos/VideoInfo/3895/-showvideo.
[30] P. Hajac, M. Khalkhali, B. Rangipour, and Y. Sommerhäuser, Stable anti-Yetter-Drinfeld modules. C. R. Math. Acad. Sci. Paris 338 (2004), no. 8, 587–590.
[31] P. Hajac, M. Khalkhali, B. Rangipour, and Y. Sommerhäuser, Hopf-cyclic homology and cohomology with coefficients. C. R. Math. Acad. Sci. Paris 338 (2004), no. 9, 667–672.
[32] M. Khalkhali, Basic Noncommutative Geometry, European Mathematical Society Publishing House, 2009.
[33] J. L. Loday, Cyclic Homology. Springer-Verlag, second edition (1998).
[34] H. Moscovici, Local index formula and twisted spectral triples, arXiv:0902.0835. This volume.
[35] H. Moscovici, and B. Rangipour, Cyclic cohomology of Hopf algebras of transverse symmetries in codimension 1. Adv. Math. 210 (2007), no. 1, 323–374.
[36] H. Moscovici, and B. Rangipour Hopf algebras of primitive Lie pseudogroups and Hopf cyclic cohomology. Adv. Math. 220 (2009), no. 3, 706–790.

Mathematics Department, The University of Western Ontario, London, Ontario, N6A 5B7, Canada
E-mail address: masoud@uwo.ca

Clay Mathematics Proceedings
Volume 11, 2010

The Core Hopf Algebra

Dirk Kreimer

Abstract. We study the core Hopf algebra underlying the renormalization Hopf algebra.

1. Introduction

In a recent study of the role of limiting mixed Hodge structure, Spencer Bloch and the author introduced the core Hopf algebra on one-particle irreducible graphs (1PI graphs, also dubbed core graphs in [**1**]). It is a Hopf algebra which contains the renormalization Hopf algebra as a quotient algebra. One can also view it as the renormalization Hopf algebra of a field theory formulated in infinite dimensions, as then any graph which has closed loops is superficially divergent, and any sum over all superficially divergent 1PI graphs reduces to a sum over 1PI graphs.

In this short contribution, we introduce the core Hopf algebra in examples and discuss its larger role in quantum field theory. Formal proofs are to be found in future work. Our main task is to outline some intriguing aspects of the Hopf algebra structure underlying perturbation theory, going far beyond the problem of renormalization. I feel these ideas are a fitting tribute to my earlier papers with Alain on the subject [**4, 5, 6**], and I report on these ideas here for the first time in public in deep respect for Alain's contributions to science, and in deep gratitude for his friendship.

2. The core Hopf algebra

The basic formula for the Hopf algebra of a renormalizable field theory is

$$\Delta(\Gamma) = \Gamma \otimes \mathbb{I} + \mathbb{I} \otimes \Gamma + \sum_{\gamma=\cup\gamma_i, \omega(\gamma_i)\leq 0} \gamma \otimes \Gamma/\gamma. \tag{1}$$

Here, the sum runs over disjoint unions of superficially divergent one-particle irreducible graphs, and Γ/γ is obtained by shrinking in Γ each component γ_i of $\gamma = \cup_i \gamma_i$ to a point. A component γ_i is superficially divergent if $\omega(\gamma_i) := \mathrm{b}(\gamma_i)D - w(\gamma_i) \leq 0$. Here, b gives the first Betti number, the number of independent cycles, D is the dimension of spacetime and $w(\gamma_i)$ the sum of the scaling weights of internal edges and vertices of γ_i. A scaling weight of an edge is the dimensionality (in units of

2010 *Mathematics Subject Classification.* Primary 81Q30; Secondary 81T15, 81T18, 16T05.
Work supported in parts by grant NSF-DMS/0603781. Author supported by CNRS.

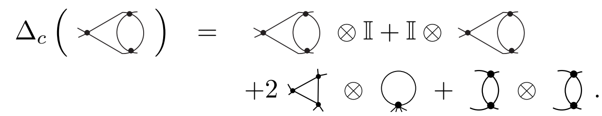
mass) of the coefficient of the quadratic field monomial generating this covariance, and minus that dimensionality for a vertex. So for example, in non-Abelian gauge theory in four dimensions of spacetime, a propagating gauge-field has a scaling weight of two (as any boson in four dimensions) and the three-gluon vertex has scaling weight -1, as it involves three gauge fields and one spacetime derivative. See [**1**] for notation and details. This renormalization Hopf algebra can be easily augmented to take care of the quantum numbers which label external legs, incorporating formfactors and kinematics of Feynman amplitudes. We focus here on some elementary aspects of iteration of subgraphs into each other, and will not clutter notation any further.

The core Hopf algebra is then obtained by relaxing the qualification on superficial divergence: we simply sum over all 1PI subgraphs.

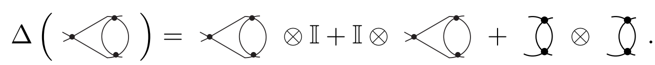
$$\Delta(\Gamma) = \Gamma \otimes \mathbb{I} + \mathbb{I} \otimes \Gamma + \sum_{\gamma = \cup \gamma_i} \gamma \otimes \Gamma/\gamma. \tag{2}$$

Note that this immediately implies that the only primitives are one-loop graphs. As an aside, we note that for the renormalization Hopf algebra of quantum gravity, the particular powercounting rules of gravity [**7**] ensure that for perturbative gravity, the renormalization Hopf algebra and the core algebra agree.

Let us give now an example for the core Hopf algebra in ϕ^4 theory.

(3) 

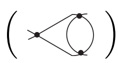
In the renormalization Hopf algebra we would simply have

(4) 

So why should we study the core Hopf algebra? Let us discuss the structure of the graph polynomial:

$$\phi(\Gamma) = \sum_{\text{spanning trees } T} \prod_{e \notin T} A_e, \tag{5}$$

accompanying this graph. Labeling the two straight edges on the left as A_1, A_2 and the other two as A_3, A_4, it reads

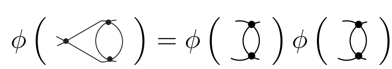
(6) ϕ  $= A_1A_3 + A_1A_4 + A_2A_3 + A_2A_4 + A_3A_4$

$$\begin{align} &= (A_1 + A_2)(A_3 + A_4) + A_3A_4, \tag{7}\\ &= (A_1 + A_2 + A_3)A_4 + (A_1 + A_2)A_3 \tag{8}\\ &= (A_1 + A_2 + A_4)A_3 + (A_1 + A_2)A_4 \tag{9}\end{align}$$

corresponding to the five spanning trees of the graph. We can find the coproduct of the renormalization as well as the core Hopf algebra from a factorization

$$\phi(\Gamma) = \phi(\Gamma/\gamma)\phi(\gamma) + r(\Gamma, \gamma) \tag{10}$$

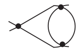
such that $r(\Gamma, \gamma)$ is of higher degree in the variables of $\phi(\gamma)$ than $\phi(\gamma)$ itself. For example, from (7)

(11)  $+ A_3A_4,$

where A_3A_4 is quadratic in the variables A_3, A_4 of the subgraph made up of edges $3, 4$ of the initial graph, while that subgraph γ itself, superficially divergent as $\omega(\gamma) = 0$, has graph polynomial

(12) $\phi\Big($  $\Big) = A_3 + A_4,$

while the cograph has

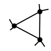
(13) $\phi\Big($  $/$  $\Big) = A_1 + A_2.$

Clearly, when A_3, A_4 tend to zero jointly, $r(\Gamma, \gamma)$ vanishes faster than $\phi(\Gamma/\gamma)\phi(\gamma)$ and hence we find a subdivergence with regard to the A_3, A_4 integration using the Feynman rules in parametric representation. The other factorizations (8,9) above have limits which remain integrable over the respective subgraph variables.

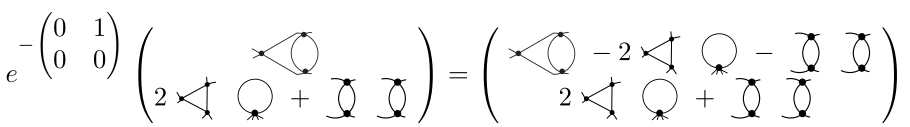
But any investigation of the algebro-geometric structure of periods assigned to graphs starts with the investigation of the graph hypersurface $X_\Gamma : \phi(\Gamma) = 0$, and the question of how that graph hypersurface meets the simplex $A_i > 0$. Integrability is a rather irrelevant criterion in this respect, as studied in detail in [**1**], and the two other factorizations (8,9)

$$\begin{aligned} A_1A_3 + A_1A_4 + A_2A_3 + A_2A_4 + A_3A_4 &= \underbrace{(A_1 + A_2 + A_3)A_4}_{\text{linear in } A_4} + \underbrace{(A_1 + A_2)A_3}_{\text{constant in A}_4} \\ &= \underbrace{(A_1 + A_2 + A_4)A_3}_{\text{linear in } A_3} + \underbrace{(A_1 + A_2)A_4}_{\text{constant in } A_3}, \end{aligned} \tag{14}$$

give the other two terms generated by the non-trivial part of Δ_c and are mandatory in order to study the situation from the perspective of a limiting mixed Hodge structure. Note that

(15) $\omega\Big($  $\Big) = +1.$

Actually, let us get an idea of how the connection between the Hopf algebra and limiting Hodge structures emerges. For details, the reader can refer to [**1**]. Consider



The entries of the column vector clearly originate from the coproduct on the graph. Replacing graphs by their graph polynomials and trying to integrate against the edge variables hits a non-integrable logarithmic singularity near the origin, parametrized say by a small cut-off t. Supplementing the exponential with a suitable $\ln t$ gives a rhs which allows for a limit $t \to 0$.

As an amusing side remark, let me mention that the famous problem of overlapping divergences in renormalization corresponds to precisely the coexistence of different factorizations in the above sense. While the above has three coexistent decompositions of the graph polynomial all contributing to the core coproduct, only a single term contributes to the renormalization coproduct, as this graph has no overlapping divergences with regard to renormalization.

For the renormalization Hopf algebra it has proved worthwhile to study its Hochschild cohomology, as this provides a prefered way to prove renormalizability of counterterms and illuminates the structure of Dyson–Schwinger equations (DSE). Let us see how the core Hopf algebra fares in this respect.

3. DSE in the core Hopf algebra

Let us stay for simplicity in the realm of massless ϕ^4 theory in four dimensions of space-time. In the renormalization Hopf algebra, we have to study two Green functions, one for the vertex function (four external legs), and one for the inverse propagator (two external legs).

Both are obtained from the evaluation by suitably renormalized Feynman rules ϕ_R of the series $X^4(g)$ and $X^2(g)$ of all 1PI graphs with the appropriate number of four or two external legs.

These series in the coupling g (series in g with coefficients in the Hopf algebra) are fixpoints of equations formed by studying the Hochschild cohomology [**8**], $bB_+^{j,m} = 0$, $m \in \{2,4\}$ of these Hopf algebras:

$$X^4(g) = \mathbb{I} + \sum_{j>0} g^j B_+^{j,4}\left(X^4(g)\left(\frac{X^4(g)}{(X^2(g))^2}\right)^j\right) \tag{16}$$

$$X^2(g) = \mathbb{I} - \sum_{j>0} g^j B_+^{j,2}\left(X^2(g)\left(\frac{X^4(g)}{(X^2(g))^2}\right)^j\right). \tag{17}$$

It is crucial that the $B_+^{j,m}$ are closed one-cocycles: it leads to a clean approach to non-perturbative aspects of local field theory and to an analysis of the structure of solution of Dyson–Schwinger equations in such theories [**9, 10, 11**].

In the above,

$$B_+^{j,m} = \sum_{|\gamma|=j,\Delta(\gamma)=\gamma\otimes\mathbb{I}+\mathbb{I}\otimes\gamma} \frac{1}{\mathrm{sym}(\gamma)} B_+^{\gamma}, \tag{18}$$

with γ having m external legs and

$$B_+^{\gamma}(X) = \sum_{\Gamma} \frac{\mathrm{bij}(\gamma, X, \Gamma)}{\mathrm{maxf}(\Gamma)[\gamma|X]|X|_{\wedge}} \Gamma. \tag{19}$$

The reader will have to consult [**12, 13**] for details. We just mention that $\mathrm{bij}(\gamma, X, \Gamma)$ counts the number of bijections between external edges of X and insertion places of γ so as to obtain Γ, maxf counts the number of ways to shrink 1PI subgraphs such that the cograph is primitive under the coproduct, $[\gamma|X]$ counts the number of insertion places for X in γ, and $|X|_{\wedge}$ gives the number of different graphs generated from permuting external edges.

Here is an illuminating example: First, from Hochschild closedness, $B_+^{\gamma}(\mathbb{I}) = \gamma, \forall\gamma$. Hence $X^4(g)$ starts as

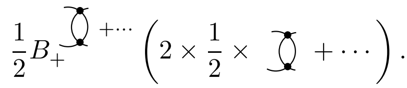
(20) $\mathbb{I} + \frac{1}{2}$  $+ \cdots + \mathcal{O}(g^2).$

Here, $+\cdots$ refers to the two other orientations of this graph (the s, t, u channels). Let us now look at

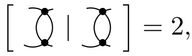
(21) $\frac{1}{2}B_+$  $^{+\cdots}\left(\left(\frac{X^4(g)}{X^2(g)}\right)^2\right)$

appearing on the rhs of (16).

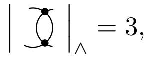
Let us Taylor expand the argument in g to first order and concentrate on the term coming from the expansion of the square of the vertex function. We find

(22) 

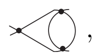
The number of insertion places is

(23) 

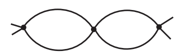
there are three orientations,

(24) 

giving two bijections leading to graphs of the form

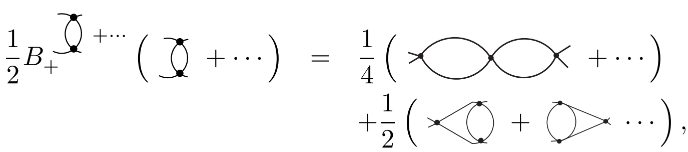
(25) 

(swapped or permuted possibly) each of which has a single maximal forest, maxf= 1, and one bijection leading to

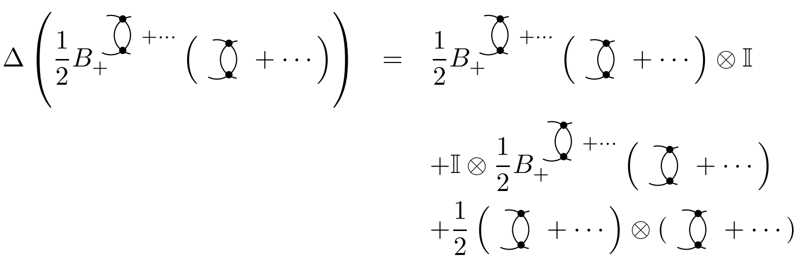
(26) 

with two maximal forests.

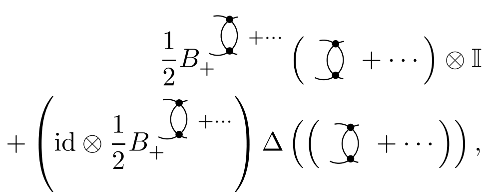
We hence find



with all the correct symmetry factors. Computing now the coproduct delivers



which agrees with



as required by Hochschild cohomology. Hochschild cohomology does us an enormous favor here, and it becomes even more impressive when one realizes how it conspires to give rhyme and reason to internal symmetries in a field theory [**12, 14, 15**].

So what changes if we try the same with the core Hopf algebra?

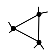
Let us first describe the primitives. We noted already that they are all one-loop graphs. Next, we observe that in the core Hopf algebra underlying the vertices and edges of ϕ^4_4 theory we must have vertices of arbitrary high but even valence.

Hence, the one-loop graphs can be described by partitions, where each entry j in the partition corresponds to a $(2j+2)$-valent vertex in the one-loop graph, the length of the partition gives the number of vertices on the one-loop graph (and equals the number of its edges), and the size of the partition gives the total number of external edges.

So for example

(27)  $\sim (1,1)$

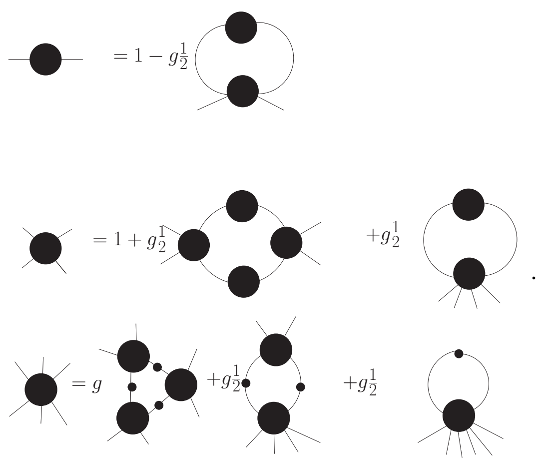
and

(28)  $\sim (1,1,1)$.

We are led to the following system:

$$
\begin{aligned}
(29)\quad X^2(g) &= \mathbb{I} - gB_+^{(1)}(X^4/X^2(g)),\\
(30)\quad X^4(g) &= \mathbb{I} + g\tfrac{1}{2}B_+^{(1,1)}((X^4)^2/(X^2(g))^2) + g\tfrac{1}{2}B_+^{(2)}(X^6(g)/X^2(g)),\\
(31)\quad X^6(g) &= gB_+^{(1,1,1)}((X^4(g))^3/(X^2(g))^3)\\
&\quad +g\tfrac{1}{2}B_+^{(2,1)}((X^6(g)X^4(g))/(X^2(g))^2)\\
&\quad +g\tfrac{1}{2}B_+^{(3)}(X^8(g)/X^2(g)),
\end{aligned}
$$

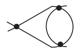
and so on, which is best understood graphically (we omit giving contributions obtained by swapping or permuting external edges):



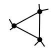
Note that we have only a finite number of one-loop primitives contributing to each fix-point equation, but we have an infinite set of equations to consider. Also, we emphasize that we maintain the B_+ operators to be closed one-cocycles in the Hochschild cohomology of the core Hopf algebra, and claim that the same definition (19) achieves precisely that.

The series X^4 and X^2 which are fixpoints of the above system are the same series as the one obtained in the Hochschild cohomology of the renormalization Hopf algebra above. This is a rather remarkable fact. We have actually done something very typical for the functional integral: we have traded a loop expansion for a leg expansion.

It is instructive to see in an example how this comes about. From $B_+(1,1)$ we get the same graphs as before, but

(32) 

now has three maximal forests (in the core Hopf algebra the number of maximal forests equals the number of non-self-intersecting closed paths we can draw on the graph). So this contribution gets an extra factor 1/3. The missing 2/3 is precisely provided from the same graph generated by insertions of

(33) 

into $B_+^{(2)}$, where the number of relevant bijections is two.

4. sub Hopf algebras and AdS/CFT

Now, for the renormalization Hopf algebra and its Hochschild cohomology we have learned a rather remarkable story: if we decompose a series of graphs by order,

$$X^s(g) = \sum_{j=0}^{\infty} c_j^s g^j, \tag{34}$$

with c_j^s Hopf algebra elements, these finite linear combinations of graphs provide a sub Hopf algebra. To achieve this in the presence of internal symmetries one has to divide by suitable ideals [**12, 14, 15**], and doing so, we finally can work with much simpler Hopf algebras. Combining with the structure of the renormalization group [**6, 9, 10, 11**] then fully exhibits the recursive structure of field theory parameterized by the periods underlying the motives coming with the graph hypersurfaces.

And for the core Hopf algebra? If we sum all graphs contributing to a chosen amplitude at a given loop order, do these form linear combinations the generators of a sub Hopf algebra? Certainly not as they stand, but what is the structure of the (co-) ideals such that we can obtain such a sub Hopf algebra when taking quotients?

Applying the techniques of [**12, 14, 15**] this is straightforward as we will report elsewhere [**16**]. The harder question is to study Feynman rules and see to what extent they respect such quotients.

Here, we note that the relations

$$X^{2k}/X^{2(k-1)} = X^{2(k+1)}/X^{2k} \tag{35}$$

determine a co-ideal such that we get the desired sub Hopf algebras. Similar relations will show up in the study of any core Hopf algebra for other quantum field theories.

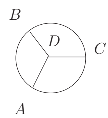
Two points deserve attention: if we had not considered ϕ^4 but perturbative gravity, these would be precisely the relations which, if tolerated by the Feynman rules, will render gravity renormalizable. Putting this together with the pecularities of powercounting of perturbative gravity as studied in [**7**] opens an avenue to a renormalizabe approach to gravity after all, which deserves much further study.

At tree level, the relations (35) have a recursive form very familiar from studying the now famous [**17**] on-shell recursion relations of tree (and actually one-loop) amplitudes. This also deserves much future work. Note in particular that one-loop recursion relations boil down here to relations between the Hochschild one-cocycles driving the equations of motion.

Indeed, in very recent work [**16**] we showed how co-ideals of the core Hopf algebra as above are in accordance with the celebrated BCFW [**3**] relations. Similar relations hold for the spin-2 graviton [**2**], possibly justifying the above hope for a renormalizable theory of gravity.

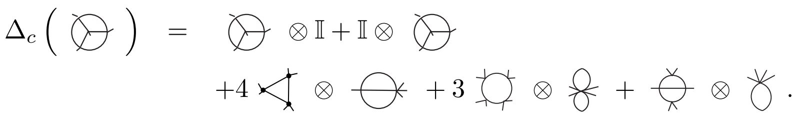
5. Unitarity of the S-matrix

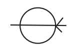
The main role which the core Hopf algebra has to play in the future is, I believe, in reconciling our understanding of renormalization with the unitarity of the S-matrix. The notion of a cut at a Feynman graph is compatible with the core coproduct. This again will be discussed elsewhere, but let us give us one example. Consider the wheel with three spokes

(36) 

.

We have labelled its vertices A, B, C, D. External edges are not drawn, but all vertices are supposed to be four-valent. We consider the graph as contributing to a $1 \to 3$ production amplitude, and consider the particle incoming at vertex A. The core Hopf algebra delivers the following coproduct for this graph:

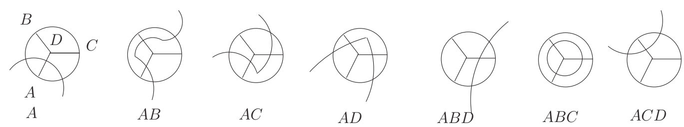


Note that from the terms on the rhs only  allows for cuts C separating incoming and outgoing particles.

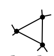
The other ones are too tadpole-ish to contribute:

(37) $C\Big($  $\Big) = C\Big($  $\Big) = 0.$

Now consider the cuts C determining the imaginary part.

(38) $C\Big($  $\Big) =$ 

.

We see four contributions which have an intact subgraph , and three contributions where no internal loop is left intact. We have labeled each cut by the set of vertices connected to vertex A.

If we now let CC (completely cut) be the operator which assigns the sum of all cuts to a graph such that no internal loop is left intact, then

(39) $(\mathrm{id} \otimes CC)\Delta_c\Big($  $\Big)$

is in one-to-one correspondence with $C($  $)$ and hence describes the structure of this imaginary part rather well.

This is the beginning of a mathematically beautifully approach to unitarity and the S-matrix based on the core Hopf algebra. I hope to report more on that in collaboration with Spencer Bloch, still continuing to celebrate a line of thought which started in [**18**] and first blossomed in my work with Alain.

References

[1] S. Bloch and D. Kreimer, *Mixed Hodge Structures and Renormalization in Phyiscs*, Comm. in Number Theory and Physics, Vol. 2.4, p.637-718; arXiv:0804.4399 [hep-th].

[2] C. Boucher-Veronneau. On-Shell recursion relations in general relativity. *Master Thesis, U. of Waterloo, Ontario, Canada 2007.*

[3] R. Britto, F. Cachazo, B. Feng and E. Witten. Direct proof of tree-level recursion relation in Yang-Mills theory. *Phys.Rev.Lett.* 94:181602,2005.

[4] A. Connes and D. Kreimer, *Hopf algebras, renormalization and noncommutative geometry,* Commun. Math. Phys. **199** (1998) 203 [arXiv:hep-th/9808042].

[5] A. Connes and D. Kreimer, *Renormalization in quantum field theory and the Riemann-Hilbert problem. I: The Hopf algebra structure of graphs and the main theorem,* Commun. Math. Phys. **210** (2000) 249 [arXiv:hep-th/9912092].

[6] A. Connes and D. Kreimer, *Renormalization in quantum field theory and the Riemann-Hilbert problem. II: The beta-function, diffeomorphisms and the renormalization group,* Commun. Math. Phys. **216** (2001) 215 [arXiv:hep-th/0003188].

[7] D. Kreimer, *A remark on quantum gravity,* Annals Phys. **323** (2008) 49 [arXiv:0705.3897 [hep-th]].

[8] C. Bergbauer and D. Kreimer, *Hopf algebras in renormalization theory: Locality and Dyson-Schwinger equations from Hochschild cohomology,* IRMA Lect. Math. Theor. Phys. **10** (2006) 133 [arXiv:hep-th/0506190].

[9] D. Kreimer and K. Yeats, *Recursion and growth estimates in renormalizable quantum field theory,* Commun. Math. Phys. **279** (2008) 401; [arXiv:hep-th/0612179].

[10] D. Kreimer and K. Yeats, *An etude in non-linear Dyson-Schwinger equations,* Nucl. Phys. Proc. Suppl. **160** (2006) 116 [arXiv:hep-th/0605096].

[11] G. van Baalen, D. Kreimer, D. Uminsky and K. Yeats, *The QED beta-function from global solutions to Dyson-Schwinger equations,* Annals Phys. **324**, 205 (2009) [arXiv:0805.0826 [hep-th]].

[12] D. Kreimer, *Anatomy of a gauge theory,* Annals Phys. **321** (2006) 2757 [arXiv:hep-th/0509135].

[13] Karen Yeats, *Growth Estimates for Dyson–Schwinger Equations,* Thesis, Boston U. 2008, arXiv:0810.2249.

[14] W. D. van Suijlekom, *Renormalization of gauge fields: A Hopf algebra approach,* Commun. Math. Phys. **276** (2007) 773 [arXiv:hep-th/0610137].

[15] W.D. van Suijlekom, *The structure of renormalization Hopf algebras for gauge theories I: Representing Feynman graphs on BV-algebras,* Commun. Math. Phys. **290** (2009) 291, arXiv:0807:0999.

[16] D. Kreimer, W.D. van Suijlekom, *Recursive relations in the core Hopf algebra,* arXiv:0903.2849 [hep-th], Nucl. Phys. **B** (2009), in press, DOI information: 10.1016/j.nuclphysb.2009.04.025.

[17] E. W. Nigel Glover and C. Williams, *One-Loop Gluonic Amplitudes from Single Unitarity Cuts,* JHEP **0812** (2008) 067 [arXiv:0810.2964 [hep-th]]; and references there.

[18] D. Kreimer, *On the Hopf algebra structure of perturbative quantum field theories,* Adv. Theor. Math. Phys. **2** (1998) 303 [arXiv:q-alg/9707029].

IHES, 35 rte. de Chartres, 91440 Bures-sur-Yvette, France, and Boston University, Boston, MA

E-mail address: `kreimer@ihes.fr`

URL: `http://math.bu.edu`

URL: `http://www.ihes.fr`

Clay Mathematics Proceedings
Volume 11, 2010

Propriété (T) Renforcée et Conjecture de Baum-Connes

Vincent Lafforgue

Cet article est dédié à Alain Connes
pour son soixantième anniversaire

Nous cherchons à comprendre pourquoi la propriété (T) de Kazhdan [**Kaz67, HV89, BHV08**], et plus particulièrement une forme renforcée de celle-ci introduite dans [**Laf08**], sont un obstacle à une démonstration de la surjectivité de l'application de Baum-Connes à coefficients arbitraires pour des groupes ayant un élément γ de Kasparov, à l'aide des méthodes connues.

Nous passons d'abord en revue trois méthodes pour montrer la surjectivité de l'application de Baum-Connes à coefficients pour des groupes ayant un élément γ de Kasparov (pour lesquels l'injectivité de l'application de Baum-Connes à coefficients est connue). La première méthode (due à Kasparov [**Kas88**]) consiste à montrer que $\gamma = 1$ dans $KK_G(\mathbb{C}, \mathbb{C})$. C'est la méthode qui a donné le plus de résultats positifs mais nous ne la mentionnons que brièvement car elle échoue pour les groupes non compacts ayant la propriété (T) pour des raisons évidentes alors que l'obstacle de la propriété (T) renforcée est plus subtil pour les autres méthodes. La deuxième méthode (d'abord proposée par Julg [**Jul97**]) consiste à construire une homotopie de 1 à γ en utilisant des représentations dans des espaces de Hilbert qui ne sont pas unitaires mais dont la croissance est contrôlée par une exponentielle arbitrairement petite. Nous justifions en détail, en nous appuyant sur des idées de Higson, le fait que, pour un groupe localement compact agissant de façon continue, isométrique et propre sur un espace de dimension asymptotique finie avec contrôle linéaire (et donc en particulier pour un groupe hyperbolique), l'existence d'une telle homotopie implique la surjectivité de l'application de Baum-Connes à coefficients arbitraires. La troisième méthode est la méthode banachique [**Laf02a**], qui fait intervenir des complétions inconditionnelles et dont le résultat dépend beaucoup des coefficients : elle ne montre la conjecture de Baum-Connes à coefficients arbitraires pour aucun groupe !

Nous proposons ensuite un cadre général (assez évident) englobant ces trois méthodes et nous en tirons une condition nécessaire pour qu'une méthode inscrite dans ce cadre général, c'est-à-dire utilisant la KK-théorie banachique et des arguments élémentaires de stabilité par calcul fonctionnel holomorphe, puisse montrer la surjectivité de l'application de Baum-Connes à coefficients arbitraires.

2010 *Mathematics Subject Classification.* Primary 19K35 ; Secondary 19K56, 46L80.

Nous expliquons ensuite pourquoi la propriété (T) renforcée [**Laf08**] est un obstacle à une démonstration de la surjectivité de l'application de Baum-Connes à coefficients arbitraires par une méthode de ce type.

Nous montrons enfin que $SL_3(\mathbb{Q}_p)$ a la propriété (T) renforcée (cela a été démontré dans [**Laf08**] mais comme nous donnons ici une définition différente de la propriété (T) renforcée, adaptée à la conjecture de Baum-Connes, nous devons reprendre la démonstration). En fait tout groupe algébrique presque simple sur un corps local non-archimédien ou archimédien, dont l'algèbre de Lie contient $\mathfrak{sl}_3$, possède la propriété (T) renforcée au sens de cet article (en modifiant un peu la définition dans le cas archimédien en raison de l'absence de sous-groupe compact ouvert) mais nous ne le montrons pas ici.

Bien sûr cela ne donne aucune indication qu'il puisse exister un contre-exemple à la conjecture de Baum-Connes à coefficients pour un tel groupe. Au contraire les idées de Jean-Benoît Bost sur le principe d'Oka [**Bos90**] restent intactes et font espérer que la conjecture de Baum-Connes à coefficients soit vraie pour tous les groupes de Lie réels ou p-adiques.

Quand Paul Baum et Alain Connes ont formulé leur conjecture la propriété (T) est apparue très vite comme un obstacle pour montrer la surjectivité de l'application de Baum-Connes. Bien que cet obstacle ait été contourné dans quelques cas, l'obstacle de la propriété (T) renforcée reste infranchissable actuellement.

C'est avec un très grand plaisir que je dédie cet article à Alain Connes, qui a inspiré tant de mathématiciens.

1. Rappels

Nous renvoyons à [**BCH94**] pour l'énoncé de la conjecture de Baum-Connes à coefficients, à [**Kas88, Laf02a**] pour la KK-théorie et la KK-théorie banachique, et à [**Ska99, Laf01, Laf02b**] pour des introductions bien adaptées à la suite.

Soit G un groupe localement compact et dg une mesure de Haar à gauche sur G. Soit A une C^*-algèbre munie d'une action continue de G. On note $C_c(G,A)$ l'algèbre des fonctions continues à support compact sur G à valeurs dans A munie du produit de convolution

$$(f_1.f_2)(g) = \int_{g_1 \in G} f_1(g_1) g_1(f_2(g_1^{-1}g)) dg_1$$

(cette formule est naturelle si on écrit les éléments de $C_c(G,A)$ sous la forme $\int_{g\in G} f(g)e_g dg$). On définit $L^1(G,A)$ comme la complétion de $C_c(G,A)$ pour la norme $\int ||f(g)||_A dg$ et $C^*_{\mathrm{red}}(G,A)$ comme la complétion pour la norme d'opérateur de la convolution à gauche sur le A-module hilbertien $L^2(G,A)$. On rappelle que $L^2(G,A)$ est le complété du A-module pré-hilbertien $C_c(G,A)$, dont la structure de A-module à droite est donnée par

$$(\int e_g a(g) dg) b = \int e_g a(g) b dg$$

et dont le produit hermitien est donné par

$$\langle \int e_g a(g) dg, \int e_g b(g) dg \rangle = \int a(g)^* b(g) dg.$$

Si on écrit les éléments du A-module pré-hilbertien $C_c(G,A)$ sous la forme $\int a(g) e_g dg$ les formules sont un peu plus compliquées.

On note $\underline{E}G$ un G-espace propre qui est final dans la catégorie dont les objets sont les G-espaces propres et dont les morphismes sont les G-morphismes à homotopie près. Pour toute G-C^*-algèbre A et pour $j \in \mathbb{Z}/2\mathbb{Z}$ on pose

$$K_j^{\mathrm{top}}(G,A) = \varinjlim KK_G^j(C_0(Y),A)$$

où la limite est prise suivant les parties G-invariantes et G-compactes Y de $\underline{E}G$. On a un morphisme de groupes abéliens, dit morphisme d'assemblage ou application de Baum-Connes,

$$\mu_{\mathrm{red}}^{G,A} : K_j^{\mathrm{top}}(G,A) \to K_j(C^*_{\mathrm{red}}(G,A)).$$

La conjecture de Baum-Connes à coefficients [**BCH94**] affirme que $\mu_{\mathrm{red}}^{G,A}$ est un isomorphisme de groupes abéliens. Higson, Skandalis et moi-même [**HLS02**] avons trouvé des contre-exemples à la surjectivité de l'application de Baum-Connes à coefficients pour certains groupes discrets construits par Gromov (pour lesquels on ne sait pas construire un élément γ au sens suivant).

On supposera toujours que G possède un élément γ de Kasparov. Dans [**Tu99a**], Jean-Louis Tu a axiomatisé les propriétés d'un élément γ, essentiellement sous la forme suivante. On appelle élément γ un élément de $KK_G(\mathbb{C},\mathbb{C})$ tel que $\gamma = \eta \otimes_B d$ avec B une G-C^*-algèbre propre, $\eta \in KK_G(\mathbb{C},B)$ et $d \in KK_G(B,\mathbb{C})$, et que pour toute partie G-compacte Y de $\underline{E}G$, $q^*(\gamma) = 1$ dans $KK_{G\ltimes Y}(C_0(Y),C_0(Y))$ où q est la projection de Y vers le point (on renvoie à [**Gal99**] pour la notation $KK_{G\ltimes Y}$). Ces conditions impliquent que $\gamma^2 = \gamma$ dans $KK_G(\mathbb{C},\mathbb{C})$.

Sous cette hypothèse, pour toute G-C^*-algèbre A, $\mu_{\mathrm{red}}^{G,A}$ est injective et son image est égale à celle de l'action sur $K_*(C^*_{\mathrm{red}}(G,A))$ de l'idempotent $j^G_{\mathrm{red}} \circ \sigma_A(\gamma)$. On rappelle que les morphismes d'anneaux

$$\sigma_A : KK_G(\mathbb{C},\mathbb{C}) \to KK_G(A,A)$$

$$\text{et} \qquad j^G_{\mathrm{red}} : KK_G(A,A) \to KK(C^*_{\mathrm{red}}(G,A), C^*_{\mathrm{red}}(G,A))$$

ont été construits par Kasparov [**Kas88**]. Donc la surjectivité de $\mu_{\mathrm{red}}^{G,A}$ équivaut au fait que $j^G_{\mathrm{red}} \circ \sigma_A(\gamma)$ agit par l'identité sur $K_*(C^*_{\mathrm{red}}(G,A))$.

D'après [**Kas88, KS91, KS94, KS03**] les groupes de Lie réels ou p-adiques et les groupes "boliques" (donc en particulier les groupes hyperboliques) possèdent un élément γ.

2. Quelques méthodes pour montrer la surjectivité de l'application de Baum-Connes à coefficients

Nous indiquons les principales méthodes pour montrer la surjectivité de $\mu_{\mathrm{red}}^{G,A}$ pour un groupe localement compact G possédant un élément γ de Kasparov. Elles s'appuient toutes sur la construction d'une certaine homotopie de 1 à γ.

2.1. Homotopie par des représentations unitaires. La première méthode pour montrer la conjecture de Baum-Connes à coefficients arbitraires pour G est due à Kasparov et consiste à montrer $\gamma = 1$ dans $KK_G(\mathbb{C},\mathbb{C})$ (voir [**Jul98**] pour un séminaire Bourbaki sur ce sujet). Cela a été fait pour $SO(n,1)$ [**Kas84**], les groupes agissant proprement sur des arbres [**JV84**] (voir aussi [**Pim86**]), $SU(n,1)$ [**JK95**], et enfin dans le cas des groupes de Haagerup qui contient tous les cas précédents [**HK01**]. Un groupe est dit de Haagerup ou encore a-T-menable s'il possède une action isométrique affine continue et propre sur un espace de Hilbert. Tous les groupes moyennables ont la propriété de Haagerup. Cependant on sait que $\gamma \neq 1$

dans $KK_G(\mathbb{C},\mathbb{C})$ si G a la propriété (T) de Kazhdan et n'est pas compact. Nous ne donnons pas plus de détails sur ces importants travaux car l'objet de cet article est l'obstacle de la propriété (T).

2.2. Homotopie par des représentations non unitaires dans des espaces de Hilbert. La rédaction de ce paragraphe a été très influencée par des discussions avec Guoliang Yu, qui m'a indiqué les références [**Mat07, Roe05**] et que je remercie.

Dans [**Jul97**], Julg a proposé d'utiliser des représentations bornées non unitaires dans des Hilbert pour montrer la conjecture de Baum-Connes à coefficients pour $Sp(n,1)$. En 1999, Higson, Julg et moi-même avons discuté de la possibilité d'utiliser des représentations à croissance exponentielle arbitrairement petite. Cette méthode permet de montrer la conjecture de Baum-Connes à coefficients pour $Sp(n,1)$ [**Jul02**] et pour les groupes hyperboliques [**Laf09**]. Cette méthode est explicitée dans le corollaire 2.12, qui résulte du théorème 2.3 et de la proposition 2.10.

Le théorème 2.3, dont l'idée est due à Nigel Higson, affirme que si un groupe localement compact G possède un élément γ de Kasparov et agit de façon continue, isométrique et propre sur un espace de dimension asymptotique finie avec contrôle linéaire, l'existence d'homotopies de 1 à γ, utilisant des représentations dans des espaces de Hilbert dont la croissance est contrôlée par une exponentielle arbitrairement petite, implique la surjectivité de l'application de Baum-Connes à coefficients. La notion de dimension asymptotique est due à Gromov [**Gro93**]. D'autre part la proposition 2.10 rappelle, d'après Gromov [**Gro93**] et Roe [**Roe05**], que tout espace métrique faiblement géodésique, hyperbolique et à géométrie grossière bornée est de dimension asymptotique finie avec contrôle linéaire.

Définition 2.1. *Soit $N \in \mathbb{N}$ et $(\mu_0,\mu_1) \in \mathbb{R}_+^2$. Un espace métrique (X,d) est de dimension asymptotique $\leq N$ avec contrôle linéaire (μ_0,μ_1) si pour tout $d \in \mathbb{R}_+$ il existe une partition $X = \bigcup_{i\in I} X_i$ et une application "couleur" $c : I \to \{0,1,\dots,N\}$ telles que*

- *pour tout $i \in I$, X_i est mesurable et* $\mathrm{diam}(X_i) \leq \mu_0 d + \mu_1$,
- *pour $i,j \in I$ vérifiant $c(i) = c(j)$ et $i \neq j$ on a $d(X_i,X_j) > d$, où l'on note $d(X_i,X_j) = \inf_{y\in X_i, z\in X_j} d(y,z)$.*

Remarque. La propriété de dimension asymptotique finie avec contrôle linéaire est invariante par quasi-isométrie.

Définition 2.2. *Soit G un groupe localement compact. On appelle longueur sur G une fonction continue $\ell : G \to \mathbb{R}_+$ vérifiant $\ell(g^{-1}) = \ell(g)$ et $\ell(g_1g_2) \leq \ell(g_1) + \ell(g_2)$ pour tous $g, g_1, g_2 \in G$.*

Soit G un groupe localement compact et ℓ une longueur sur G. Pour toutes G-C^*-algèbres A et B on définit $E_{G,\ell}(A,B)$ comme l'ensemble des classes d'isomorphisme de (E,π,T) où E est un (A,B)-bimodule hilbertien $\mathbb{Z}/2\mathbb{Z}$-gradué muni d'une action continue de G vérifiant $\|\pi(g)\| \leq e^{\ell(g)}$ pour tout $g \in G$, et d'un opérateur T borné impair tel que pour tout $a \in A$ les opérateurs $[a,T]$ et $a(T^2-1)$ soient compacts et que l'application $g \mapsto a(g(T)-T)$ soit une application normiquement continue de G dans $\mathcal{K}_B(E)$. On définit ensuite $KK_{G,\ell}(A,B)$ comme l'ensemble des classes d'homotopie dans $E_{G,\ell}(A,B)$: deux éléments sont homotopes si ils sont les évaluations en 0 et 1 d'un élément de $E_{G,\ell}(A,B[0,1])$. On rappelle que $B[0,1] = C([0,1], B)$ muni de la norme du supremum. On peut montrer que la somme directe munit $KK_{G,\ell}(A,B)$ d'une structure de groupe abélien.

En particulier $E_{G,\ell}(\mathbb{C},\mathbb{C})$ est l'ensemble des classes d'isomorphisme de (H,π,T) où H est un espace de Hilbert $\mathbb{Z}/2\mathbb{Z}$-gradué muni d'une action continue de G vérifiant $\|\pi(g)\| \leq e^{\ell(g)}$ pour tout $g \in G$, et d'un opérateur T borné impair tel que (T^2-1) soit compact et que l'application $g \mapsto g(T) - T$ soit une application normiquement continue de G dans $\mathcal{K}(H)$.

Le théorème suivant repose sur des idées de Nigel Higson.

Théorème 2.3. *Soit G un groupe localement compact agissant de façon isométrique et continue sur un espace métrique (X,d) de dimension asymptotique finie avec contrôle linéaire. Soit x_0 un point de X et ℓ la longueur sur G définie par $\ell(g) = d(x_0, gx_0)$. Soit $\gamma \in KK_G(\mathbb{C},\mathbb{C})$ tel que pour tout $s > 0$, il existe $C \in \mathbb{R}_+$ tel que l'image de $1-\gamma$ dans $KK_{G,s\ell+C}(\mathbb{C},\mathbb{C})$ soit nulle. Alors pour toute G-C^*-algèbre A, $j^G_{\mathrm{red}} \circ \sigma_A(\gamma)$ agit par l'identité sur $K_*(C^*_{\mathrm{red}}(G,A))$.*

Le théorème résultera de la conjonction des propositions 2.5 et 2.6.

Pour toute longueur ℓ sur G on note $\mathcal{E}_{G,\ell}$ la classe des représentations (H,π) de G dans un espace de Hilbert H telles que $\|\pi(g)\|_{\mathcal{L}(H)} \leq e^{\ell(g)}$ pour tout $g \in G$ et on définit l'algèbre de Banach $\mathcal{C}_\ell(G,A)$ comme la complétion de $C_c(G,A)$ pour la norme

$$\|f\|_{\mathcal{C}_\ell(G,A)} = \sup_{(H,\pi)\in\mathcal{E}_{G,\ell}} \|\alpha(f)\|_{\mathcal{L}_A(H\otimes L^2(G,A))}$$

où le morphisme $\alpha : C_c(G,A) \to \mathcal{L}_A(H \otimes L^2(G,A))$ est donné par la formule

$$\alpha(\int_G a(g)e_g dg)(h \otimes (\int_G b(g)e_g dg)) = \int_{G\times G} \pi(g_1)h \otimes a(g_1)g_1(b(g_2))e_{g_1g_2}dg_1dg_2.$$

Si ℓ et ℓ' sont deux longueurs avec $\ell'(g) \leq \ell(g)$ pour tout $g \in G$, on a un morphisme d'algèbres de Banach de $\mathcal{C}_\ell(G,A)$ dans $\mathcal{C}_{\ell'}(G,A)$.

On note que $\mathcal{E}_{G,0}$ est la classe des représentations unitaires de G. Le lemme suivant est un des ingrédients de la construction de la descente de Kasparov [**Kas88**] dans le cas particulier qui nous intéresse, c'est-à-dire

$$j^G_{\mathrm{red}} \circ \sigma_A : KK_G(\mathbb{C},\mathbb{C}) \to KK(C^*_{\mathrm{red}}(G,A), C^*_{\mathrm{red}}(G,A)).$$

Lemme 2.4. *On a $\mathcal{C}_0(G,A) = C^*_{\mathrm{red}}(G,A)$.*

Démonstration. Bien que ce lemme soit très connu, nous en rappelons la démonstration, car la démonstration du lemme 2.7 ci-dessous repose sur la même idée. Soit (H,π) une représentation unitaire de G. On doit montrer que α se prolonge en un morphisme de C^*-algèbres $\alpha : C^*_{\mathrm{red}}(G,A) \to \mathcal{L}_A(H \otimes L^2(G,A))$. Soit $g_0 \in G$. Notons θ_{g_0} l'opérateur unitaire sur le A-module hilbertien $H \otimes L^2(G,A)$ défini par

$$\theta_{g_0}(h \otimes \int_G a(g)e_g dg) = \int_G \pi(g_0g^{-1})h \otimes a(g)e_g dg.$$

D'autre part notons $\beta : C^*_{\mathrm{red}}(G,A) \to \mathcal{L}_A(H \otimes L^2(G,A))$ le morphisme défini par

$$\beta(\int_G a(g)e_g dg)(h \otimes (\int_G b(g)e_g dg)) = h \otimes \int_{G\times G} a(g_1)g_1(b(g_2))e_{g_1g_2}dg_1dg_2.$$

On a alors $\alpha(a) = \theta_{g_0}^{-1} \circ \beta(a) \circ \theta_{g_0}$ pour tout $a \in C_c(G,A)$, donc α se prolonge par continuité à $C^*_{\mathrm{red}}(G,A)$. Ceci achève la démonstration du lemme 2.4. □

Proposition 2.5. *Pour toute longueur ℓ sur G on a un morphisme de descente*

$$j_{\mathrm{red}}^{G,\ell,A} : KK_{G,\ell}(\mathbb{C},\mathbb{C}) \to KK^{\mathrm{ban}}(\mathcal{C}_\ell(G,A), C^*_{\mathrm{red}}(G,A))$$

qui coïncide avec $j^G_{\mathrm{red}} \circ \sigma_A$ si $\ell = 0$ et tel que si ℓ et ℓ' sont deux longueurs avec $\ell'(g) \leq \ell(g)$ pour tout $g \in G$, le diagramme suivant soit commutatif

$$\begin{array}{ccc} KK_{G,\ell'}(\mathbb{C},\mathbb{C}) & \xrightarrow{j_{\mathrm{red}}^{G,\ell',A}} & KK^{\mathrm{ban}}(\mathcal{C}_{\ell'}(G,A), C^*_{\mathrm{red}}(G,A)) \\ \downarrow & & \downarrow \\ KK_{G,\ell}(\mathbb{C},\mathbb{C}) & \xrightarrow{j_{\mathrm{red}}^{G,\ell,A}} & KK^{\mathrm{ban}}(\mathcal{C}_{\ell}(G,A), C^*_{\mathrm{red}}(G,A)) \end{array}$$

Remarque. Nous utiliserons le diagramme commutatif ci-dessus avec $\ell' = 0$.
Remarque. On pourrait aussi construire pour deux G-C^*-algèbres A et B un morphisme de descente

$$KK_{G,\ell}(A,B) \to KK^{\mathrm{ban}}(\mathcal{C}_{\ell,B}(G,A), C^*_{\mathrm{red}}(G,B))$$

en définissant $\mathcal{C}_{\ell,B}(G,A)$ comme le complété de $C_c(G,A)$ pour la norme d'opérateur sur les $C^*_{\mathrm{red}}(G,B)$-modules hilbertiens $C^*_{\mathrm{red}}(G,E)$ avec E un (A,B)-bimodule hilbertien muni d'une action continue de G vérifiant $\|\pi(g)\| \leq e^{\ell(g)}$ pour tout $g \in G$. Nous n'en aurons pas besoin.
Démonstration de la proposition 2.5. En suivant **[Kas88]**, si $(H,\pi,T) \in E_{G,\ell}(\mathbb{C},\mathbb{C})$, on construit

$$j_{\mathrm{red}}^{G,\ell,A}(H,\pi,T) \in E^{\mathrm{ban}}(\mathcal{C}_\ell(G,A), C^*_{\mathrm{red}}(G,A))$$

de la manière suivante. On considère le $C^*_{\mathrm{red}}(G,A)$-module hilbertien

$$E = H \otimes C^*_{\mathrm{red}}(G,A)$$

et on définit un morphisme $\alpha' : C_c(G,A) \to \mathcal{L}_{C^*_{\mathrm{red}}(G,A)}(E)$ par la formule évidente

$$\alpha'\Big(\int_G a(g)e_g dg\Big)\Big(h \otimes \Big(\int_G b(g)e_g dg\Big)\Big) = \int_{G\times G} \pi(g_1)h \otimes a(g_1)g_1(b(g_2))e_{g_1g_2}dg_1dg_2.$$

Comme $\mathcal{L}_{C^*_{\mathrm{red}}(G,A)}(E)$ s'injecte isométriquement dans $\mathcal{L}_A(H \otimes L^2(G,A))$, et que la composée de α' et de cette injection est

$$\alpha : C_c(G,A) \to \mathcal{L}_A(H \otimes L^2(G,A))$$

considéré précédemment, et par la définition même de $\mathcal{C}_\ell(G,A)$, α se prolonge par continuité en un morphisme $\mathcal{C}_\ell(G,A) \to \mathcal{L}_{C^*_{\mathrm{red}}(G,A)}(E)$. On définit d'autre part $\tilde{T} = T \otimes 1 \in \mathcal{L}_{C^*_{\mathrm{red}}(G,A)}(E)$ et on pose $j_{\mathrm{red}}^{G,\ell,A}(H,\pi,T) = (E,\tilde{T})$. □

Proposition 2.6. *Soient $N \in \mathbb{N}$ et $(\mu_0,\mu_1) \in \mathbb{R}_+^2$. Soit G un groupe localement compact agissant de façon isométrique et continue sur un espace métrique (X,d) de dimension asymptotique finie $\leq N$ avec contrôle linéaire (μ_0,μ_1). Soit x_0 un point de X et ℓ la longueur sur G définie par $\ell(g) = d(x_0, gx_0)$. Soit A une G-C^*-algèbre. Pour tout $r \in \mathbb{R}_+$ on note B_r l'ensemble des éléments de G qui vérifient $\ell(g) \leq r$. Pour $f \in C_c(G,A)$ on note* $\mathrm{supp}(f)$ *l'adhérence dans G du sous-ensemble des g tels que $f(g) \neq 0$.*

Alors pour tous $s > 0$, $C \in \mathbb{R}_+$, $r \in \mathbb{R}_+^*$ et $f \in C_c(G,A)$ avec $\mathrm{supp}(f) \subset B_r$, *on a*

$$\|f\|_{\mathcal{C}_{s\ell+C}(G,A)} \leq (N+1)e^{(4\mu_0+1)sr+(2\mu_1 s+2C)}\|f\|_{C^*_{\mathrm{red}}(G,A)}.$$

Démonstration du théorème 2.3 en admettant la proposition 2.6. On désigne par ρ le rayon spectral. Il résulte immédiatement de la proposition 2.6 que pour tous $s \in \mathbb{R}_+^*$, $C \in \mathbb{R}_+$, $r \in \mathbb{R}_+$ et pour toute fonction $f \in C_c(G, A)$ telle que $\mathrm{supp}(f) \subset B_r$ on a

$$(1) \qquad \rho_{\mathcal{C}_{s\ell+C}(G,A)}(f) \leq e^{(4\mu_0+1)sr} \rho_{C^*_{\mathrm{red}}(G,A)}(f).$$

Pour tout $i \in \mathbb{N}^*$ soit $C_i \in \mathbb{R}_+$ tel que l'image de $1 - \gamma$ dans $KK_{G, \frac{1}{i}\ell+C_i}(\mathbb{C}, \mathbb{C})$ soit nulle. Par la proposition 2.5, $j^G_{\mathrm{red}} \circ \sigma_A(\gamma)$ agit par l'identité sur l'image de $K_*(\mathcal{C}_{\frac{1}{i}\ell+C_i}(G, A))$ dans $K_*(C^*_{\mathrm{red}}(G, A))$. Il résulte alors de l'inégalité (1), appliquée à $A \otimes M_k(\mathbb{C})$ pour tout entier $k \in \mathbb{N}^*$, et aux longueurs $\frac{1}{i}\ell + C_i$, que la suite d'algèbres de Banach $(\mathcal{C}_{\frac{1}{i}\ell+C_i}(G, A))_{i\in\mathbb{N}^*}$ vérifie les hypothèses du lemme 1.7.2 de [**Laf02a**]. Par conséquent $K_*(C^*_{\mathrm{red}}(G, A))$ est égal à la réunion des images de $K_*(\mathcal{C}_{\frac{1}{i}\ell+C_i}(G, A))$ et on a montré le théorème 2.3 en admettant la proposition 2.6. $\square$

Voici maintenant la démonstration de la proposition 2.6, dont l'idée est due à Nigel Higson.

On plonge G dans X par $\kappa : g \mapsto g^{-1}x_0$. On munit G de la distance invariante à droite d définie par $d(g, g') = d(\kappa(g), \kappa(g')) = \ell(g'g^{-1})$.

Soient $s > 0$, $C \in \mathbb{R}_+$ et (H, π) dans $\mathcal{E}_{G,s\ell+C}$.

Lemme 2.7. *Soit $m, r \in \mathbb{R}_+$, $f \in C_c(G, A)$ avec $\mathrm{supp}(f) \subset B_r$, et $w \in H \otimes L^2(G, A)$ tel que $\mathrm{supp}(w)$ a un diamètre inférieur ou égal à m pour la distance ci-dessus (c'est-à-dire de façon équivalente $\kappa(\mathrm{supp}(w))$ a un diamètre inférieur ou égal à m). Alors*

$$\|\alpha(f)w\|_{H\otimes L^2(G,A)} \leq e^{s(2m+r)+2C} \|f\|_{C^*_{\mathrm{red}}(G,A)} \|w\|_{H\otimes L^2(G,A)}.$$

Démonstration. Soit $g_0 \in \mathrm{supp}(w)$. On a $\alpha(f)w = \theta_{g_0}^{-1}(\beta(f)\theta_{g_0}(w))$. Pour $w' \in H \otimes L^2(G, A)$ tel que $\mathrm{supp}(w')$ est inclus dans la boule fermée de centre g_0 et de rayon R, on a $\|\theta_{g_0}(w')\|_{H\otimes L^2(G,A)} \leq e^{sR+C}\|w'\|_{H\otimes L^2(G,A)}$ et de même pour $\theta_{g_0}^{-1}$, ce qui démontre le lemme, car $\mathrm{supp}(w)$ est inclus dans la boule fermée de centre g_0 et de rayon m et $\mathrm{supp}(\beta(f)\theta_{g_0}(w))$ est inclus dans la boule fermée de centre g_0 et de rayon $m + r$. $\square$

Démonstration de la proposition 2.6. Soient $r \in \mathbb{R}_+^*$ et $f \in C_c(G, A)$ avec $\mathrm{supp}(f) \subset B_r$. Soit $X = \bigcup_{i\in I} X_i$ une partition et $c : I \to \{0, 1, \ldots, N\}$ une application couleur, satisfaisant les conditions de la définition 2.1 avec $d = 2r$. Pour tout $i \in I$, on note p_i le projecteur orthogonal de $H \otimes L^2(G, A)$ sur $H \otimes L^2(\kappa^{-1}(X_i), A)$. Pour tout $j \in \{0, 1, \ldots, N\}$ on note q_j le projecteur orthogonal de $H \otimes L^2(G, A)$ sur $H \otimes L^2(\kappa^{-1}(\bigcup_{i\in I, c(i)=j} X_i), A)$. On a $\alpha(f) = \sum_{j=0}^N \alpha(f)q_j$ donc il suffit de montrer, pour tout $j \in \{0, 1, \ldots, N\}$,

$$\|\alpha(f)q_j\|_{\mathcal{L}(H\otimes L^2(G,A))} \leq e^{(4\mu_0+1)sr+(2\mu_1 s+2C)} \|f\|_{C^*_{\mathrm{red}}(G,A)}.$$

Pour $i \in I$ et $w \in H \otimes L^2(\kappa^{-1}(X_i), A)$, $\alpha(f)w$ est supporté par

$$\kappa^{-1}(\{x \in X, d(x, X_i) \leq r\}).$$

Soit $j \in \{0, 1, \ldots, N\}$. Par hypothèse les parties $\{x \in X, d(x, X_i) \leq r\}$ de X, pour i vérifiant $c(i) = j$, sont deux à deux disjointes. Donc

$$\|\alpha(f)q_j\|_{\mathcal{L}(H\otimes L^2(G,A))} = \sup_{i\in I, c(i)=j} \|\alpha(f)p_i\|_{\mathcal{L}(H\otimes L^2(G,A))}.$$

Enfin pour tout $i \in I$ et $w \in H \otimes L^2(G, A)$ on a

$$\|\alpha(f)p_i w\|_{H\otimes L^2(G,A)} \leq e^{(4\mu_0+1)sr+(2\mu_1 s+2C)}\|f\|_{C^*_{\mathrm{red}}(G,A)}\|w\|_{H\otimes L^2(G,A)}$$

grâce au lemme 2.7 puisque $p_i w$ est supporté sur $\kappa^{-1}(X_i)$ dont le diamètre est inférieur ou égal à $2\mu_0 r + \mu_1$. $\square$

DÉFINITION 2.8. *Soit $\delta \geq 0$ et (X, d) un espace métrique. On dit que (X, d) est δ-hyperbolique si pour tout quadruplet (x, y, z, t) de points de X on a*

$$d(x, z) + d(y, t) \leq \max\big(d(x, t) + d(y, z), d(x, y) + d(z, t)\big) + \delta.$$

On dit que (X, d) est faiblement δ-géodésique si pour tous $x, y \in X$ et pour tout $s \in [0, d(x, y) + \delta]$ il existe $z \in X$ tel que $d(x, z) \leq s$ et $d(z, y) \leq d(x, y) - s + \delta$.

Un espace métrique (X, d) est dit hyperbolique (resp. faiblement géodésique) s'il existe $\delta \geq 0$ tel que (X, d) soit δ-hyperbolique (resp. faiblement δ-géodésique).

DÉFINITION 2.9. *Un espace métrique (X, d) est dit à géométrie grossière bornée s'il existe $\Delta > 0$ tel que pour tout $R > 0$ il existe un entier N tel que dans toute boule fermée de rayon R le nombre maximal de points dont les distances mutuelles sont supérieures ou égales à Δ est inférieur ou égal à N.*

C'est la définition 3.1 de [**KS03**].

La proposition suivante est due à Gromov ([**Gro93**] page 31) et Roe [**Roe05**]. Nous rappelons la démonstration parce que nos hypothèses sont légèrement différentes de celles de [**Roe05**].

PROPOSITION 2.10. *Tout espace métrique faiblement géodésique, hyperbolique et à géométrie grossière bornée est de dimension asymptotique finie avec contrôle linéaire.*

Plus précisément soit (X, d) un espace métrique, $\delta, \Delta \in \mathbb{R}_+^*$ et $M \in \mathbb{N}^*$ tels que (X, d) soit faiblement δ-géodésique et δ-hyperbolique et que toute boule fermée de rayon $4\delta + 2\Delta$ dans X contienne au plus M points dont les distances mutuelles sont supérieures ou égales à Δ. Alors (X, d) est de dimension asymptotique $\leq 2M - 1$ avec contrôle linéaire $(3, 5\delta + 2\Delta)$.

LEMME 2.11. *Soit $\delta, \epsilon > 0$ et (X, d) un espace métrique δ-hyperbolique. Soient $x_0, x, x', y, y' \in X$ tels que*

$$d(x_0, x') + d(x', x) \leq d(x_0, x) + \epsilon \quad \text{et} \quad d(x_0, y') + d(y', y) \leq d(x_0, y) + \epsilon. \tag{2}$$

Alors

$$d(x', y') \leq \max(|d(x_0, x') - d(x_0, y')| + \epsilon + 2\delta, d(x, y) - d(x, x') - d(y, y') + 2\epsilon + 2\delta).$$

Démonstration. La propriété d'hyperbolicité pour x, x_0, y', y donne

$$d(x, y') \leq \max(d(x, y) + d(x_0, y') - d(x_0, y), d(x, x_0) + d(y, y') - d(x_0, y)) + \delta$$

d'où l'on déduit grâce à la deuxième partie de (2),

$$d(x, y') \leq \max(d(x, y) - d(y, y'), d(x, x_0) - d(x_0, y')) + \epsilon + \delta. \tag{3}$$

La propriété d'hyperbolicité pour y', x_0, x', x donne

$$d(y', x') \leq \max(d(y', x) + d(x_0, x') - d(x_0, x), d(y', x_0) + d(x, x') - d(x_0, x)) + \delta.$$

Grâce à (3), on a

$$d(y', x) + d(x_0, x') - d(x_0, x)$$

$$\leq \max(d(x,y) - d(y,y') + d(x_0,x') - d(x_0,x), d(x_0,x') - d(x_0,y')) + \epsilon + \delta.$$

Par la première moitié de (2),

$$d(x,y) - d(y,y') + d(x_0,x') - d(x_0,x) \leq d(x,y) - d(y,y') - d(x,x') + \epsilon.$$

Enfin par la première moitié de (2),

$$d(y',x_0) + d(x,x') - d(x_0,x) \leq d(x_0,y') - d(x_0,x') + \epsilon.$$

□

Démonstration de la proposition 2.10. D'abord nous choisissons une partie $Y \subset X$ telle que pour tous $y, y' \in Y$ avec $y \neq y'$ on ait $d(y,y') \geq \Delta$, et que Y soit maximale pour cette propriété. Alors tout point de X est à distance $< \Delta$ de Y. On munit Y de la distance induite par X, si bien que toute boule fermée de rayon $4\delta + 2\Delta$ dans Y contient au plus M points. On fixe une application $c_Y : Y \to \{0, \ldots, M-1\}$ telle que pour $y, y' \in Y$ vérifiant $c_Y(y) = c_Y(y')$ et $y \neq y'$ on ait $d(y,y') > 4\delta + 2\Delta$. Une telle application existe par le lemme de Zorn car si elle est définie sur $Z \subset Y$, pour tout $z \in Y \setminus Z$ on peut la prolonger à $Z \cup \{z\}$ grâce à la propriété précédente, et par maximalité elle peut donc être définie sur Y.

Soit $d \in \mathbb{R}_+$. Pour tout $k \in \mathbb{N}$, on pose

$$A_k = \{x \in X, d(x_0,x) \in [k(d+\delta), (k+1)(d+\delta)[\}.$$

On va définir pour tout $k \in \mathbb{N}$ un ensemble fini I_k muni d'une application $c_k : I_k \to \{0, \ldots, M-1\}$, et une partition $A_k = \bigcup_{i \in I_k} X_{k,i}$ avec $X_{k,i}$ mesurable et de diamètre $\leq 3d + 5\delta + 2\Delta$, de telle sorte que la partition $X = \bigcup_{(k,i) \in I} X_{k,i}$, paramétrée par la réunion disjointe $I = \{(k,i), k \in \mathbb{N}, i \in I_k\}$ munie de l'application couleur $c : I \to \{0, \ldots, 2M-1\}$ définie par $c(k,i) = c_k(i)$ pour $i \in I_k$ avec k pair et $c(k,i) = M + c_k(i)$ pour $i \in I_k$ avec k impair, vérifie les conditions de la définition 2.1.

D'abord on pose $I_0 = \{x_0\}$ et $X_{0,x_0} = A_0$ dont le diamètre est $\leq 2d + 2\delta$. On définit $c_0 : I_0 \to \{0, \ldots, M-1\}$ en posant $c_0(x_0) = 0$. Pour $k \in \mathbb{N}^*$ on choisit une application mesurable $\mu_k : A_k \to Y$ telle que pour tout $x \in A_k$ il existe $x' \in X$ vérifiant

$$(4) \qquad d(x_0,x') \leq (k - \frac{1}{2})(d+\delta) \quad \text{et} \quad d(x',x) \leq d(x_0,x) - (k - \frac{1}{2})(d+\delta) + \delta$$

et tel que $d(x', \mu_k(x)) \leq \Delta$. Une telle application existe car pour tout $x \in X$ un tel x' existe puisque (X,d) est faiblement δ-géodésique, et $d(x', Y) < \Delta$ par construction de Y. On pose alors $I_k = \mu_k(A_k)$, on note

$$c_k : I_k \to \{0, \ldots, M-1\}$$

la restriction de c_Y à I_k et pour $i \in I_k$ on note $X_{k,i} = \mu_k^{-1}(i)$. Pour tout $i \in I_k$, le diamètre de $X_{k,i}$ est $\leq 3d + 5\delta + 2\Delta$. En effet pour x, x' comme ci-dessus on a

$$d(x,x') \leq (k+1)(d+\delta) - (k - \frac{1}{2})(d+\delta) + \delta = \frac{3}{2}d + \frac{5}{2}\delta$$

et $d(x', \mu_k(x)) \leq \Delta$, donc pour tout $i \in I_k$, $X_{k,i}$ est inclus dans la boule fermée de centre i et de rayon $\frac{3}{2}d + \frac{5}{2}\delta + \Delta$. Montrons maintenant que pour (k,i) et (l,j) des éléments distincts de I tels que $c(k,i) = c(l,j)$ on a $d(X_{k,i}, X_{l,j}) \geq d + \delta$. C'est clair si $k \neq l$ car l'hypothèse $c(k,i) = c(l,j)$ implique alors $|k - l| \geq 2$. Pour terminer il suffit donc de montrer que pour $k \in \mathbb{N}^*$ et $x, y \in A_k$ tels que $d(x,y) \leq d + \delta$, on a $d(\mu_k(x), \mu_k(y)) \leq 4\delta + 2\Delta$. Soient $x', y' \in X$ vérifiant (4) ainsi que la même condition pour y, y' et tels que $d(x', \mu_k(x)) \leq \Delta$ et $d(y', \mu_k(y)) \leq \Delta$.

Alors $d(x_0,x')$ et $d(x_0,y')$ appartiennent à $[(k-\frac{1}{2})(d+\delta)-\delta,(k-\frac{1}{2})(d+\delta)]$, donc $|d(x_0,x')-d(x_0,y')|\leq\delta$. De plus $d(x,x')\geq d(x_0,x)-d(x_0,x')\geq\frac{d+\delta}{2}$ et de même $d(y,y')\geq\frac{d+\delta}{2}$. En appliquant le lemme 2.11 à x_0,x,x',y,y' avec $\epsilon=\delta$ on en déduit $d(x',y')\leq 4\delta$ et donc

$$d(\mu_k(x),\mu_k(y))\leq 4\delta+2\Delta.$$

□

Corollaire 2.12. *Soit G un groupe localement compact agissant de façon isométrique, continue et propre sur un espace métrique (X,d) hyperbolique, faiblement géodésique et à géométrie grossière bornée. Soit x_0 un point de X et ℓ la longueur sur G définie par $\ell(g)=d(x_0,gx_0)$. Soit $\gamma\in KK_G(\mathbb{C},\mathbb{C})$ l'élément défini sous ces hypothèses par Kasparov et Skandalis* [**KS03**]. *Supposons que pour tout $s>0$ il existe $C\in\mathbb{R}_+$ tel que l'image de $1-\gamma$ dans $KK_{G,s\ell+C}(\mathbb{C},\mathbb{C})$ soit nulle. Alors G vérifie la conjecture de Baum-Connes à coefficients, c'est-à-dire que pour toute G-C^*-algèbre A, $\mu_{\mathrm{red}}^{G,A}:K_*^{\mathrm{top}}(G,A)\to K_*(C^*_{\mathrm{red}}(G,A))$ est une bijection.*

Démonstration. D'après [**KS03**] $\mu_{\mathrm{red}}^{G,A}$ est injective et son image est égale à l'image de l'idempotent de $\mathrm{End}(K_*(C^*_{\mathrm{red}}(G,A)))$ qui est associé à $j^G_{\mathrm{red}}(\sigma_A(\gamma))$. Il suffit donc de montrer que $j^G_{\mathrm{red}}(\sigma_A(\gamma))$ agit par l'identité sur $K_*(C^*_{\mathrm{red}}(G,A))$. Cela résulte de la conjonction du théorème 2.3 et de la proposition 2.10. □

Remarque. Comme la propriété de dimension asymptotique finie avec contrôle linéaire passe aux produits, le corollaire 2.12 est encore vrai si (X,d) est un produit fini d'espaces métriques hyperboliques, faiblement géodésiques et à géométrie grossière bornée.

Remarque. D'après [**Mat07, BD06, CG04**] les espaces symétriques et les immeubles affines sont de dimension asymptotique finie. Par les mêmes arguments on vérifie facilement qu'ils sont de dimension asymptotique finie avec contrôle linéaire. Donnons l'idée, pour la commodité du lecteur. Comme la propriété de dimension asymptotique finie avec contrôle linéaire passe aux sous-espaces, il suffit de le montrer pour G/K, avec $G=SL_n(F)$, F un corps local, et K un sous-groupe compact maximal de G. Si B est le sous-groupe de Borel des matrices triangulaires supérieures, on a $G/K=B/B\cap K$ et B est un produit semi-direct itéré de groupes abéliens. Or la propriété de dimension asymptotique finie avec contrôle linéaire passe aux produits semi-directs et est vraie pour ces groupes abéliens.

Donc par exemple pour montrer la conjecture de Baum-Connes à coefficient pour $SL_3(\mathbb{Q}_p)$ il suffirait de construire des homotopies reliant 1 à l'élément γ de [**KS91**], comme dans les hypothèses du théorème 2.3. Malheureusement nous verrons plus loin qu'à cause de la propriété (T) renforcée, quelle que soit la longueur ℓ sur $G=SL_3(\mathbb{Q}_p)$, il existe $s>0$ tel que pour tout $C\in\mathbb{R}_+$, $\gamma\neq 1$ dans $KK_{G,s\ell+C}(\mathbb{C},\mathbb{C})$. La propriété (T) renforcée est satisfaite par tous les groupes presque simples sur un corps local dont l'algèbre de Lie contient $\mathfrak{sl}_3$ et on s'attend à ce qu'elle soit satisfaite par tous les groupes presque simples sur un corps local dont le rang déployé est ≥ 2.

2.3. Méthode banachique. Une complétion $\mathcal{A}(G)$ de $C_c(G)$ est dite inconditionnelle si $\|f\|$ ne dépend que $g\mapsto|f(g)|$. Pour toute G-C^*-algèbre A on définit alors $\mathcal{A}(G,A)$ comme la complétion de $C_c(G,A)$ pour la norme $\left\|g\mapsto\|f(g)\|_A\right\|_{\mathcal{A}(G)}$. On construit dans [**Laf02a**] (juste avant la proposition 1.7.4) une application d'assemblage $\mu_{\mathcal{A}}^{G,A}:K_j^{\mathrm{top}}(G,A)\to K_j(\mathcal{A}(G,A))$. Si on a $\|f\|_{C^*_{\mathrm{red}}(G)}\leq\|f\|_{\mathcal{A}(G)}=$

$\|f^*\|_{\mathcal{A}(G)}$ pour tout $f \in C_c(G)$, d'après les propositions 1.6.4 et 1.7.6 de [**Laf02a**], l'identité de $C_c(G,A)$ s'étend en un morphisme $i : \mathcal{A}(G,A) \to C^*_{\mathrm{red}}(G,A)$ et on a un diagramme commutatif

$$\begin{tikzcd} K_j^{\mathrm{top}}(G,A) \arrow[r, "\mu_{\mathcal{A}}^{G,A}"] \arrow[dr, "\mu_{\mathrm{red}}^{G,A}"'] & K_j(\mathcal{A}(G,A)) \arrow[d, "i_*"] \\ & K_j(C^*_{\mathrm{red}}(G,A)) \end{tikzcd}$$

La conjecture de Bost (initialement formulée pour $\mathcal{A} = L^1$) affirme que $\mu_{\mathcal{A}}^{G,A}$ est un isomorphisme pour tout groupe G, toute complétion inconditionnelle $\mathcal{A}$ et toute G-C^*-algèbre A. On ne connait pas de contre-exemples à cette conjecture. Nous ne nous intéressons ici qu'à la surjectivité de $\mu_{\mathcal{A}}^{G,A}$ et $\mu_{\mathrm{red}}^{G,A}$. Dans [**Laf02a**] on a montré que si G a un élément γ et si

(C1) *il existe une longueur ℓ sur G telle que, pour tout $s > 0$, on ait $\gamma = 1$ dans $KK^{\mathrm{ban}}_{G,s\ell}(\mathbb{C},\mathbb{C})$,*

alors pour toute complétion inconditionnelle $\mathcal{A}(G)$ et pour toute G-C^*-algèbre A, $\mu_{\mathcal{A}}^{G,A}$ est une surjection. En effet on note $\mathcal{A}_{s\ell}(G,A)$ le complété de $C_c(G,A)$ pour la norme $\|g \mapsto e^{s\ell(g)}\|f(g)\|_A\|_{\mathcal{A}(G)}$. On rappelle que l'on a introduit dans [**Laf02a**] une descente

$$KK^{\mathrm{ban}}_{G,s\ell}(\mathbb{C},\mathbb{C}) \to KK^{\mathrm{ban}}(\mathcal{A}_{s\ell}(G,A),\mathcal{A}(G,A)).$$

Comme $\mathcal{A}(G,A)$ est la limite inductive des $\mathcal{A}_{s\ell}(G,A)$ quand s tend vers 0, $K_*(\mathcal{A}(G,A))$ est la réunion des images des $K_*(\mathcal{A}_{s\ell}(G,A))$. D'autre part (C1) est vrai pour tous les groupes de Lie réels ou p-adiques d'après [**Laf02a**] et pour les groupes hyperboliques d'après [**Laf02a, MY02**].

Remarque. Pour montrer la surjectivité de $\mu_{\mathcal{A}}^{G,A}$ pour toute complétion inconditionnelle $\mathcal{A}(G)$ et pour toute G-C^*-algèbre A, on voit grâce au lemme 1.7.2 de [**Laf02a**] qu'il suffit d'avoir la condition plus faible suivante

(C1′) *il existe une longueur ℓ sur G telle que pour tout $s > 0$, il existe $C \in \mathbb{R}_+$ tel que l'on ait $\gamma = 1$ dans $KK^{\mathrm{ban}}_{G,s\ell+C}(\mathbb{C},\mathbb{C})$.*

Cependant on ne connaît pas de cas où (C1′) soit réalisée et pas (C1).

Soit A une G-C^*-algèbre. Faisons l'hypothèse suivante.

(C2) *Il existe une complétion inconditionnelle $\mathcal{A}(G)$ telle que $\mathcal{A}(G,A)$ soit stable par calcul fonctionnel holomorphe dans $C^*_{\mathrm{red}}(G,A)$.*

On rappelle qu'un morphisme injectif d'algèbres de Banach dont l'image est dense et stable par calcul fonctionnel holomorphe induit un isomorphisme en K-théorie (voir l'appendice de [**Bos90**]). Par conséquent si G a un élément γ et si (C1) et (C2) sont vraies, le diagramme commutatif ci-dessus montre la surjectivité de $\mu_{\mathrm{red}}^{G,A}$.

Malheureusement la condition (C2) n'est pratiquement jamais vraie pour A arbitraire et pour $A = \mathbb{C}$ elle n'est montrée que pour quelques groupes : les groupes de Lie semi-simples réels ou p-adiques [**Laf02a**], et les groupes discrets ayant la propriété (RD), en particulier les groupes hyperboliques et d'après [**RRS98, Laf00, Cha03**] les réseaux cocompacts dans des produits de $SL_3(F)$ avec F corps local, $SL_3(\mathbb{H})$ et $E_{6(-26)}$.

Remarque. Non seulement $SL_3(\mathbb{Z})$, qui est un réseau non cocompact de $SL_3(\mathbb{R})$, n'a pas la propriété (RD), mais la condition (C2) elle-même est fausse pour $G =$

$SL_3(\mathbb{Z})$ avec $A = \mathbb{C}$. En effet soit

$$H = \mathbb{Z} \ltimes \mathbb{Z}^2 \text{ où } \mathbb{Z} \text{ agit sur } \mathbb{Z}^2 \text{ par } n \mapsto \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}^n.$$

On sait que H s'identifie à un sous-groupe de G par le morphisme

$$(n, \begin{pmatrix} a \\ b \end{pmatrix}) \mapsto \begin{pmatrix} \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}^n & \begin{pmatrix} a \\ b \end{pmatrix} \\ 0 & 1 \end{pmatrix}.$$

Soit $\mathcal{A}(G)$ une complétion inconditionnelle de $\mathbb{C}G$ incluse dans $C^*_{\mathrm{red}}(G)$. Soit $\mathcal{A}(H)$ la complétion de $\mathbb{C}H$ pour la norme de $\mathcal{A}(G)$ (c'est-à-dire l'intersection de $C^*_{\mathrm{red}}(H)$ avec $\mathcal{A}(G)$). Alors $\mathcal{A}(H)$ est une complétion inconditionnelle de $\mathbb{C}H$ incluse dans $C^*_{\mathrm{red}}(H)$. Comme H est moyennable, pour toute fonction positive à support fini f sur H on a $\|f\|_{C^*_{\mathrm{red}}(H)} = \|f\|_{\ell^1(H)}$, donc $\mathcal{A}(H)$ est incluse dans $\ell^1(H)$. Or $\ell^1(H)$ n'est pas stable par calcul fonctionnel holomorphe dans $C^*_{\mathrm{red}}(H)$. La preuve que nous allons donner est très proche de [**Jen68**]. Posons

$$x_i = (1, \begin{pmatrix} i \\ 0 \end{pmatrix}) \in H = \mathbb{Z} \ltimes \mathbb{Z}^2 \text{ pour } i = 0, 1, 2.$$

Alors x_0, x_1, x_2 engendrent un semi-groupe libre dans H car, si on note

$$\begin{pmatrix} a_k \\ b_k \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 2 & 1 \end{pmatrix}^k \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

pour $n \in \mathbb{N}$ et $i_0, \ldots, i_{n-1} \in \{0, 1, 2\}^n$ on a

$$x_{i_0} \ldots x_{i_{n-1}} = (n, \begin{pmatrix} \sum_{k=0}^{n-1} i_k a_k \\ \sum_{k=0}^{n-1} i_k b_k \end{pmatrix})$$

et comme $a_{k+1} \geq 3a_k > 0$ pour tout $k \in \mathbb{N}$ la connaissance de $\sum_{k=0}^{n-1} i_k a_k$ détermine $i_0, \ldots, i_{n-1}$. Donc $e_{x_0} + e_{x_1} - e_{x_2}$ a pour rayon spectral 3 dans $\ell^1(H)$, alors que sa norme dans $C^*_{\mathrm{red}}(H)$ est $\sup_{z \in \mathbb{C}, |z|=1} |1 + z - z^2| < 3$.

Remarque. La conjecture de Baum-Connes à coefficients peut aussi être énoncée pour des groupoïdes [**Tu99a, Tu99b**] en utilisant [**Gal99**]. La notion de complétion inconditionnelle existe aussi dans ce cadre et a été utilisée dans [**Laf07**] pour montrer la conjecture de Baum-Connes sans coefficients pour certains "groupoïdes hyperboliques", en particulier les produits croisés de groupes hyperboliques par des espaces localement compacts. On montre ainsi dans [**Laf07**] la conjecture de Baum-Connes à coefficients commutatifs pour les groupes hyperboliques.

Remarque. Dans tous les cas où une homotopie (E, π, T) de 1 à γ a été construite dans $E^{\mathrm{ban}}_{G,\ell}(\mathbb{C}, \mathbb{C}[0,1])$ pour une certaine longueur ℓ, la $\mathbb{C}[0,1]$-paire E est à dualité isométrique au sens de la définition suivante.

Définition 2.13. *Soit B une algèbre de Banach. Une B-paire E est dite à dualité isométrique si les applications B-linéaires $E^> \to \mathcal{L}_B(E^<, B)$ et $E^< \to \mathcal{L}_B(E^>, B)$ sont des injections isométriques.*

Notons qu'à toute B-paire E on peut associer une B-paire $\hat{E}$ à dualité isométrique de la façon suivante : on note $\hat{E}^>$ le B-module de Banach complété de $E^>$ pour la norme

$$\|x\|_{\hat{E}^>} = \sup_{\xi \in E^<, \|\xi\|_{E^<} \leq 1} \|\langle \xi, x \rangle\|_B.$$

On définit alors $\hat{E}^<$ comme le B-module de Banach complété de $E^<$ pour la norme

$$\|\xi\|_{\hat{E}^<} = \sup_{x \in \hat{E}^>, \|x\|_{\hat{E}^>} \leq 1} \|\langle \xi, x \rangle\|_B.$$

On vérifie que $\hat{E}$ est une B-paire à dualité isométrique. Le prolongement par continuité donne un morphisme $\mathcal{L}_B(E) \to \mathcal{L}_B(\hat{E})$ qui envoie $\mathcal{K}_B(E)$ dans $\mathcal{K}_B(\hat{E})$ et que l'on note $T \mapsto \hat{T}$. En particulier si G est un groupe localement compact, ℓ une longueur sur G et A et B des G-algèbres de Banach, et si (E, π, T) appartient à $E_{G,\ell}^{\mathrm{ban}}(A, B)$, $(\hat{E}, \hat{\pi}, \hat{T})$ appartient aussi à $E_{G,\ell}^{\mathrm{ban}}(A, B)$ et en construisant un cône on montre que les deux sont homotopes. Plus bas, dans la condition (D4), nous demanderons que E soit à dualité isométrique, car nous ne savons pas montrer le lemme 4.4 sans cette hypothèse.

3. Un cadre général englobant ces méthodes

Soit G un groupe localement compact possédant un élément γ, et ℓ une longueur sur G. Soit A une G-C^*-algèbre. Toutes les méthodes présentées dans le paragraphe précédent pour montrer la surjectivité de

$$\mu_{\mathrm{red}}^{G,A} : K_*^{\mathrm{top}}(G, A) \to K_*(C_{\mathrm{red}}^*(G, A))$$

(sauf **[Laf07]** mentionné dans l'avant-dernière remarque) obéissent au schéma suivant :

(D) *Pour tout $s > 0$ trouver $C \in \mathbb{R}_+$, et une sous-algèbre de Banach $\mathcal{B}$ de $C_{\mathrm{red}}^*(G, A)$ contenant $C_c(G, A)$, telle que*

– **(D1)** *pour tout $n \in \mathbb{N}^*$*

$$\sup_{f \in C_c(G,A) \text{ supporté dans la boule fermée de rayon } n} \frac{\|f\|_{\mathcal{B}}}{\|f\|_{C_{\mathrm{red}}^*(G,A)}} \leq Ce^{sn}.$$

– **(D2)** *on possède une homotopie (E, π, T) de 1 à γ dans $E_{G,?}^{\mathrm{ban}}(\mathbb{C}, \mathbb{C}[0,1])$ (où ? indique que la longueur n'est pas précisée, c'est-à-dire qu'il n'y a pas de condition de norme sur l'action de G) qui fournit par descente un élément $(\tilde{E}, \tilde{\pi}, \tilde{T})$ de $E^{\mathrm{ban}}(\mathcal{B}, C_{\mathrm{red}}^*(G, A)[0,1])$, tel que $(\tilde{E}^<, \tilde{E}^>)$ soit une complétion de $(C_c(G, A \otimes^{\mathrm{alg}} E^<), C_c(G, A \otimes^{\mathrm{alg}} E^>))$ pour certaines normes.*

Précisions. On rappelle que pour toute algèbre de Banach B, on note $B[0,1] = C([0,1], B)$ muni de la norme du supremum. On ne s'attend pas à avoir en général un morphisme de descente

$$KK_{G,?}^{\mathrm{ban}}(\mathbb{C}, \mathbb{C}) \to KK^{\mathrm{ban}}(\mathcal{B}, C_{\mathrm{red}}^*(G, A))$$

(on n'a pas supposé $\mathcal{B}$ de la forme $\mathcal{A}(G, A)$ où $\mathcal{A}(G)$ est une complétion inconditionnelle car avec cette restriction la condition (D1) serait impossible à réaliser dans la plupart des cas). C'est seulement pour l'homotopie particulière (E, π, T) de (D2) que l'on demande une descente, c'est-à-dire l'existence de $(\tilde{E}, \tilde{\pi}, \tilde{T})$ comme ci-dessus, afin de montrer l'égalité entre les images de γ et 1 dans $KK^{\mathrm{ban}}(\mathcal{B}, C_{\mathrm{red}}^*(G, A))$.

Montrons que (D) implique la surjectivité de $\mu_{\mathrm{red}}^{G,A}$. On note $\mathcal{B}_m$ associée à $s = 1/m$. La condition (D1) implique que pour $m, n \in \mathbb{N}^*$ et $f \in C_c(G, A)$ supportée dans la boule fermée de rayon n on a $\rho_{\mathcal{B}_m}(f) \leq e^{\frac{n}{m}} \rho_{C_{\mathrm{red}}^*(G,A)}(f)$ (et la même condition en remplaçant A par $M_k(A)$ pour tout $k \in \mathbb{N}^*$) et donc le lemme 1.7.2 de **[Laf02a]** montre que $K_*(C_{\mathrm{red}}^*(G, A))$ est la réunion des images des $K_*(\mathcal{B}_m)$. Or par la condition (D2), $j^G_{\text{red}} \circ \sigma_A(\gamma)$ agit par l'identité sur l'image de $K_*(\mathcal{B}_m)$ dans $K_*(C^*_{\text{red}}(G,A))$, pour tout $m \in \mathbb{N}^*$.

Montrons maintenant que (D) englobe les méthodes du paragraphe précédent. D'abord si $\gamma = 1$ dans $KK_G(\mathbb{C},\mathbb{C})$, on prend $\mathcal{B} = C^*_{\text{red}}(G,A)$ et la condition (D2) résulte de la descente de Kasparov

$$j^G_{\text{red}} \circ \sigma_A : E_G(\mathbb{C}, \mathbb{C}[0,1]) \to E(C^*_{\text{red}}(G,A), C^*_{\text{red}}(G,A)[0,1]).$$

Pour la méthode du sous-paragraphe 2.2, avec $s > 0$ et $C \in \mathbb{R}_+$ comme dans le théorème 2.3, on prend $\mathcal{B} = \mathcal{C}_{s\ell+C}(G,A)$ et les conditions (D1) et (D2) sont assurées par les propositions 2.6 et 2.5. Enfin pour la méthode banachique décrite dans le sous-paragraphe 2.3, où l'on a $A = \mathbb{C}$, on prend, pour $s > 0$ et $\mathcal{A}(G)$ comme dans les conditions (C1) et (C2), $\mathcal{B} = \mathcal{A}_{s\ell}(G)$.

Remarque. La preuve de la conjecture de Baum-Connes à coefficients commutatifs pour les groupes hyperboliques donnée dans [**Laf07**] et mentionnée à la fin du paragraphe précédent rentre pratiquement dans le cadre de (D) en prenant $A = C_0(Y)$, où Y est un espace localement compact muni d'une action d'un groupe hyperbolique Γ et $\mathcal{B}$ est une certaine complétion inconditionnelle de $C_c(\mathcal{G})$ en notant $\mathcal{G}$ le groupoïde produit croisé de Y par Γ. La condition (D2) est assurée par la descente en KK-théorie banachique pour les groupoïdes (introduite dans [**Laf07**]) mais la condition (D1) n'est pas vérifiée par $\mathcal{B}$ car la stabilité par calcul fonctionnel holomorphe de $\mathcal{B}$ dans $C^*_{\text{red}}(\mathcal{G}) = C^*_{\text{red}}(\Gamma, C_0(Y))$ est démontrée par une méthode ad hoc. Cette méthode ne se généralise pas à d'autres cas (par exemple dans [**Laf07**] on ne montre pas la conjecture de Baum-Connes à coefficients commutatifs pour un produit de deux groupes hyperboliques).

4. Un obstacle à $(\widetilde{D})$ pour certains groupes

Dans ce paragraphe on se donne un sous-groupe compact ouvert $K \subset G$. On suppose $\int_K dg = 1$ et on note $e_K = \int_K e_g dg \in C_c(G)$. Nous étudierons dans le paragraphe suivant le cas où $G = SL_3(\mathbb{Q}_p)$ et $K = SL_3(\mathbb{Z}_p)$.

On note $(\widetilde{D})$ la réunion de $(D) = (D1) + (D2)$ et de deux conditions techniques (D3) et (D4) que nous expliciterons plus loin. Nous allons définir une propriété (T) renforcée pour G (relativement à K) qui empêche que $(\widetilde{D})$ soit vraie pour toute G-C^*-algèbre A (ou même pour toute G-C^*-algèbre commutative A). Nous noterons $(\mathrm{T}_{\mathrm{Schur}})$ cette propriété car sa définition, qui est adaptée à la conjecture de Baum-Connes, fait intervenir des produits de Schur (voir la remarque après le lemme 4.3). On renvoie à [**CH85, Haa86, CH89, Dor93, CDSW05**] pour des travaux sur l'approximation de la fonction constante égale à 1 par des éléments de l'algèbre de Fourier dont les normes de Schur sont bornées (ce problème est lié à l'étude des représentations uniformément bornées, alors que nous sommes intéressés dans cet article par les représentations dont la croissance est contrôlée par une exponentielle assez petite). Dans le paragraphe suivant nous montrerons que $SL_3(\mathbb{Q}_p)$ possède $(\mathrm{T}_{\mathrm{Schur}})$ relativement à $SL_3(\mathbb{Z}_p)$. Dans la définition suivante, la C^*-algèbre $C_0(G)$ est munie de l'action de G par translations à droite.

Définition 4.1. *On dit que G a la propriété* $(\mathrm{T}_{\mathrm{Schur}})$ *relativement à K si pour toute longueur ℓ sur G invariante à gauche et à droite par K, il existe $s > 0$ et une fonction $\phi : G \to \mathbb{R}_+$ invariante à gauche et à droite par K et tendant vers* 0 *à l'infini tels que la propriété suivante soit vraie :*

si $c : G \to \mathbb{C}$ est une fonction K-invariante à gauche et à droite telle que pour tout $n \in \mathbb{N}$ et pour tout $f \in e_K C_c(G, C_0(G))e_K$ supporté dans $\{g \in G, \ell(g) \leq n\}$ on ait

$$\|g \mapsto c(g)f(g)\|_{C^*_{\mathrm{red}}(G,C_0(G))} \leq e^{sn}\|f\|_{C^*_{\mathrm{red}}(G,C_0(G))}$$

alors c admet une limite à l'infini $c_\infty \in \mathbb{C}$ et on a $|c(g) - c_\infty| \leq \phi(g)$ pour tout $g \in G$.

Remarque. La propriété ($\mathrm{T}_{\mathrm{Schur}}$) que nous venons de définir diffère un peu de la propriété (T) renforcée de la définition 0.1 de [**Laf08**]. Expliquons le rapport logique entre les deux, en nous restreignant au cas où $G = SL_3(\mathbb{Q}_p)$ et $K = SL_3(\mathbb{Z}_p)$ pour pouvoir citer [**Laf08**] (mais ce qui suit reste vrai quels que soient G et K). Le théorème 3.2 de [**Laf08**] montre la propriété (T) renforcée (au sens de la définition 0.1 de [**Laf08**], en oubliant ici la généralisation à certains espaces de Banach) pour $SL_3(\mathbb{Q}_p)$ à l'aide des propositions 3.3 et 3.4 de [**Laf08**] qui donnent des renseignements sur les coefficients de matrice entre vecteurs $SL_3(\mathbb{Z}_p)$-invariants (resp. d'un autre type sous l'action de $SL_3(\mathbb{Z}_p)$). Le rapport logique entre ($\mathrm{T}_{\mathrm{Schur}}$) et (T) renforcé est le suivant : la propriété ($\mathrm{T}_{\mathrm{Schur}}$) pour $G = SL_3(\mathbb{Q}_p)$ relativement à $K = SL_3(\mathbb{Z}_p)$ implique facilement l'énoncé de la proposition 3.3 de [**Laf08**] mais pas celui de la proposition 3.4 de [**Laf08**]. Néanmoins pour tout groupe localement compact G et tout sous-groupe compact ouvert K, si G a la propriété ($\mathrm{T}_{\mathrm{Schur}}$) relativement à K, G a la propriété (T) de Kazhdan (adapter le lemme 3.5 de [**Laf08**] et l'argument qui le précède, qui montrent que la proposition 3.3 de [**Laf08**] implique la propriété (T) de Kazhdan pour $SL_3(\mathbb{Q}_p)$).

Si un groupe localement compact G a la propriété (T) renforcée au sens de la définition 0.1 de [**Laf08**], c'est-à-dire si la représentation triviale est isolée parmi les représentations dans des espaces de Hilbert à croissance exponentielle suffisamment petite, il est clair que la méthode proposée au paragraphe 2.2 ne peut s'appliquer à G. Néanmoins on pourrait imaginer qu'une méthode hybride comme (D) permette de montrer la conjecture de Baum-Connes à coefficients pour G. La proposition 4.2 montre que si G possède la propriété ($\mathrm{T}_{\mathrm{Schur}}$) relativement à un sous-groupe ouvert compact K, la méthode (D) échoue elle aussi pour des coefficients arbitraires. Cependant pour montrer la proposition 4.2 nous devons ajouter à (D) les deux conditions techniques suivantes, qui complètent (D2).

- **(D3)** *La $\mathbb{C}[0,1]$-paire E est à dualité isométrique.*
- **(D4)** *La C^*-algèbre A est unifère, l'action de K sur E est isométrique et le $\mathcal{B}$-$C^*_{\mathrm{red}}(G,A)[0,1]$-bimodule $(\tilde{E}^<, \tilde{E}^>)$ qui est une complétion de*

  $$(C_c(G, A \otimes^{\mathrm{alg}} E^<), C_c(G, A \otimes^{\mathrm{alg}} E^>))$$

  vérifie l'estimée suivante. On note χ_K la fonction caractéristique de K. Pour $x \in E^>$ et $\xi \in E^<$ des éléments K-invariants, on note

  $$e_K \otimes 1_A \otimes x \in C_c(G, A \otimes^{\mathrm{alg}} E^>) \text{ la fonction } g \mapsto \chi_K(g)1_A \otimes x$$

  $$\text{et } \ e_K \otimes 1_A \otimes \xi \in C_c(G, A \otimes^{\mathrm{alg}} E^<) \ \text{ la fonction } \ g \mapsto \chi_K(g)1_A \otimes \xi.$$

  Alors on demande

  $$\|e_K \otimes 1_A \otimes x\|_{\tilde{E}^>} \leq \|x\|_{E^>} \text{ et } \ \|e_K \otimes 1_A \otimes \xi\|_{\tilde{E}^<} \leq \|\xi\|_{E^<}.$$

PROPOSITION 4.2. *Si G n'est pas compact et a la propriété ($\mathrm{T}_{\mathrm{Schur}}$) relativement à K, la condition $(\widetilde{D}) = (D1) + (D2) + (D3) + (D4)$ n'est satisfaite pour aucune G-C^*-algèbre commutative unifère A contenant $C_0(G)$ comme sous-G-C^*-algèbre.*

La proposition 4.2 résultera des lemmes 4.3 et 4.4.

Le lemme 4.3 donne des estimées sur les produits de Schur par les coefficients de matrice des représentations intervenant dans une homotopie de 1 à γ, qui sont nécessaires pour que les conditions (D1), (D2) et (D4) soient satisfaites.

Lemme 4.3. *Soit A une G-C^*-algèbre commutative unifère, $s > 0$, $C \in \mathbb{R}_+$. Soit $(E, \pi, T) \in E^{\mathrm{ban}}_{G,?}(\mathbb{C}, \mathbb{C}[0,1])$ une homotopie de 1 à γ telle que, pour une certaine sous-algèbre de Banach $\mathcal{B}$ de $C^*_{\mathrm{red}}(G, A)$ contenant $C_c(G, A)$ les conditions (D1), (D2) et (D4) soient satisfaites.*

Alors pour tout $t \in [0, 1]$, pour $x \in E^>_t$ et $\xi \in E^<_t$ des éléments K-invariants, pour tout $n \in \mathbb{N}$ et pour tout $f \in e_K C_c(G, A)e_K$ supporté dans $\{g \in G, \ell(g) \le n\}$ on a, en notant $c(g) = \langle \xi, \pi_t(g)x\rangle$,

$$\|g \mapsto c(g)f(g)\|_{C^*_{\mathrm{red}}(G,A)} \le Ce^{sn}\|f\|_{C^*_{\mathrm{red}}(G,A)}\|x\|_{E^>_t}\|\xi\|_{E^<_t}.$$

Précisons que dans ce lemme, pour $t \in [0, 1]$, E_t désigne la $\mathbb{C}$-paire image directe de E par le morphisme $\mathbb{C}[0, 1] \to \mathbb{C}$ d'évaluation en t. En particulier $E^>_t = E^> \otimes^{\pi}_{\mathbb{C}[0,1]} \mathbb{C}$ et $E^<_t = \mathbb{C} \otimes^{\pi}_{\mathbb{C}[0,1]} E^<$.

Démonstration du lemme 4.3. Soit $t \in [0, 1]$, et $x \in E^>_t$ et $\xi \in E^<_t$ des éléments K-invariants. On note

$$e_K \otimes 1_A \otimes x \in C_c(G, A \otimes^{\mathrm{alg}} E^>_t) \text{ la fonction } g \mapsto \chi_K(g)1_A \otimes x$$

où χ_K désigne la fonction caractéristique de K. De même on note

$$e_K \otimes 1_A \otimes \xi \in C_c(G, A \otimes^{\mathrm{alg}} E^<_t) \text{ la fonction } g \mapsto \chi_K(g)1_A \otimes \xi.$$

Dans la condition (D) apparaît un $\mathcal{B}$-$C^*_{\mathrm{red}}(G, A)[0, 1]$-bimodule $(\tilde{E}^<, \tilde{E}^>)$ qui est une complétion de

$$(C_c(G, A \otimes^{\mathrm{alg}} E^<), C_c(G, A \otimes^{\mathrm{alg}} E^>)).$$

Donc pour tout $t \in [0, 1]$, le $\mathcal{B}$-$C^*_{\mathrm{red}}(G, A)$-bimodule $(\tilde{E}^<_t, \tilde{E}^>_t)$ est une complétion de

$$(C_c(G, A \otimes^{\mathrm{alg}} E^<_t), C_c(G, A \otimes^{\mathrm{alg}} E^>_t)).$$

Grâce à la condition (D4), pour tous t, x, ξ comme ci-dessus on a

$$\|e_K \otimes 1_A \otimes x\|_{\tilde{E}^>_t} \le \|x\|_{E^>_t} \text{ et } \|e_K \otimes 1_A \otimes \xi\|_{\tilde{E}^<_t} \le \|\xi\|_{E^<_t}.$$

Voici maintenant le calcul fondamental. On pose $c(g) = \langle \xi, \pi_t(g)x\rangle$. Pour tout $f \in e_K C_c(G, A)e_K$ on a

$$\langle e_K \otimes 1_A \otimes \xi, f(e_K \otimes 1_A \otimes x)\rangle = (g \mapsto c(g)f(g)) \in C_c(G, A).$$

Or comme $\tilde{E}_t$ doit être un $\mathcal{B}$-$C^*_{\mathrm{red}}(G, A)$-bimodule, pour $X \in \tilde{E}^>_t$, $\Xi \in \tilde{E}^<_t$ et $f \in C_c(G, A)$ on doit avoir

$$\|\langle \Xi, fX\rangle\|_{C^*_{\mathrm{red}}(G,A)} \le \|\Xi\|_{\tilde{E}^<_t}\|X\|_{\tilde{E}^>_t}\|f\|_{\mathcal{B}}.$$

On en déduit que pour tout $f \in e_K C_c(G, A)e_K$ on doit avoir

$$\|g \mapsto c(g)f(g)\|_{C^*_{\mathrm{red}}(G,A)} \le \|\xi\|_{E^<_t}\|x\|_{E^>_t}\|f\|_{\mathcal{B}}. \tag{5}$$

Grâce à la condition (D1) le lemme 4.3 est alors démontré. ☐

Remarque. Pour toute fonction $c : K\backslash G/K \to \mathbb{C}$, on peut noter

$$\mathrm{Schur}_c : e_K C_c(G, A)e_K \to e_K C_c(G, A)e_K$$

le produit de Schur, qui à $g \mapsto f(g)$ associe $g \mapsto c(g)f(g)$. Alors on peut réexprimer (5) en disant que Schur_c s'étend en une application continue de $e_K \mathcal{B} e_K$ dans $e_K C^*_{\mathrm{red}}(G,A)e_K$ et que

$$\left\|\mathrm{Schur}_c\right\|_{\mathcal{L}(e_K\mathcal{B}e_K, e_K C^*_{\mathrm{red}}(G,A)e_K)} \leq \|\xi\|_{E_t^<}\|x\|_{E_t^>}.$$

LEMME 4.4. *On suppose que G n'est pas compact. Soit $\psi : G \to \mathbb{R}_+$ invariante à gauche et à droite par K et tendant vers 0 à l'infini. Il n'existe pas d'homotopie (E,π,T) de 1 à γ dans $E^{\mathrm{ban}}_{G,?}(\mathbb{C},\mathbb{C}[0,1])$ (où ? indique que la longueur n'est pas précisée), telle que K agisse isométriquement sur E, que E soit à dualité isométrique, et que pour tout $t \in [0,1]$, pour $x \in E_t^>$ et $\xi \in E_t^<$ des éléments K-invariants de norme 1, le coefficient de matrice $c : K\backslash G/K \to \mathbb{C}, g \mapsto \langle \xi, \pi_t(g)x\rangle$, admette une limite à l'infini c_∞ et vérifie $|c(g) - c_\infty| \leq \psi(g)$ pour tout $g \in G$.*

Démonstration. Supposons par l'absurde qu'une telle homotopie (E,π,T) existe. Pour $g \in G$, on pose $P^g = e_K e_g e_K \in \mathcal{L}_{\mathbb{C}[0,1]}(E)$. Grâce à l'hypothèse que E est à dualité isométrique, pour $g,g' \in G$ on a $\|P^g - P^{g'}\| \leq \psi(g) + \psi(g')$. Donc P^g est une suite de Cauchy dans $\mathcal{L}_{\mathbb{C}[0,1]}(E)$ (quand g tend vers l'infini) et admet une limite $P \in \mathcal{L}_{\mathbb{C}[0,1]}(E)$. Comme $P^g P^{g'} = \int_K P^{gkg'}dk$ on montre facilement que P est un idempotent dans $\mathcal{L}_{\mathbb{C}[0,1]}(E)$. De plus pour $t = 0,1$, P_t est le projecteur orthogonal sur les vecteurs G-invariants (E_0 et E_1 sont des espaces de Hilbert munis de représentations unitaires de G).

Pour tout $g \in G$, $[P^g,T] \in \mathcal{K}_{\mathbb{C}[0,1]}(E)$, d'où en passant à la limite $[P,T] \in \mathcal{K}_{\mathbb{C}[0,1]}(E)$. On note $\mathrm{Im}\,P$ la $\mathbb{C}[0,1]$-paire formée des images de $P^>$ et $P^<$ dans $E^>$ et $E^<$. Alors $PTP \in \mathcal{L}_{\mathbb{C}[0,1]}(\mathrm{Im}\,P)$ et $(PTP)^2 - 1 \in \mathcal{K}_{\mathbb{C}[0,1]}(\mathrm{Im}\,P)$ donc $(\mathrm{Im}\,P, PTP) \in E^{\mathrm{ban}}(\mathbb{C},\mathbb{C}[0,1])$. Or $(PTP)_t$ est d'indice 1 pour $t = 0$ et d'indice 0 pour $t = 1$ (en effet E_0 représente 1 dans $KK_G(\mathbb{C},\mathbb{C})$ et comme G n'est pas compact, E_1, qui représente γ, fait intervenir des représentations unitaires de G dans des espaces de Hilbert qui n'admettent pas de vecteurs G-invariants). Cette contradiction achève la démonstration du lemme 4.4. ☐

Démonstration de la proposition 4.2. On raisonne par l'absurde. On suppose que G n'est pas compact et vérifie ($\mathrm{T}_{\mathrm{Schur}}$) relativement à K. Soit ℓ la longueur fixée avant l'énoncé de (D). Soit A une G-C^*-algèbre commutative unifère contenant $C_0(G)$ comme sous-G-C^*-algèbre, telle que G vérifie la condition $(\widetilde{D})$ pour A. On fixe s et ϕ comme dans ($\mathrm{T}_{\mathrm{Schur}}$). Soit C associé à s dans (D). On applique le lemme 4.3 à A, s et C, puis ($\mathrm{T}_{\mathrm{Schur}}$), puis le lemme 4.4 avec $\psi = C\phi$, et on arrive à une contradiction. ☐

5. Démonstration de la propriété (T) renforcée pour $SL_3(\mathbb{Q}_p)$

Le but de ce paragraphe est de montrer le théorème suivant.

THÉORÈME 5.1. *Pour tout nombre premier p, $SL_3(\mathbb{Q}_p)$ a la propriété* ($\mathrm{T}_{\mathrm{Schur}}$) *relativement à $SL_3(\mathbb{Z}_p)$, au sens de la définition 4.1.*

On note $G = SL_3(\mathbb{Q}_p)$ et $K = SL_3(\mathbb{Z}_p)$. On rappelle que $C_0(G)$ est muni de l'action de G par translations à droite. On a

$$C^*_{\mathrm{red}}(G, C_0(G)) = \mathcal{K}(L^2(G)).$$

La sous-algèbre $e_K C^*_{\mathrm{red}}(G,C_0(G))e_K$ de $C^*_{\mathrm{red}}(G,C_0(G))$ s'identifie à $\mathcal{K}(\ell^2(G/K))$.

Soit $\Lambda = \{(i,j) \in \mathbb{N}^2, i-j = 0 \text{ modulo } 3\}$. L'application

$$(i,j) \mapsto K\Big(p^{\frac{i+2j}{3}} \begin{pmatrix} p^{-(i+j)} & 0 & 0 \\ 0 & p^{-j} & 0 \\ 0 & & 1 \end{pmatrix}\Big)K$$

est une bijection de Λ vers $K\backslash G/K$. On munit G de la longueur ℓ définie par

$$\ell\Big(k\Big(p^{\frac{i+2j}{3}} \begin{pmatrix} p^{-(i+j)} & 0 & 0 \\ 0 & p^{-j} & 0 \\ 0 & & 1 \end{pmatrix}\Big)k'\Big) = i+j$$

pour $k, k' \in K$ et $(i,j) \in \Lambda$.

On note B l'immeuble de G. On rappelle que les sommets de B sont identifiés aux réseaux de $\mathbb{Q}_p^3$, à homothétie près par $\mathbb{Q}_p^*$. Pour tout réseau M on note $[M]$ sa classe d'équivalence. Etant donnés $x, y \in B$ il existe un unique couple $(i,j) \in \mathbb{N}^2$ tel que dans une certaine base (v_1, v_2, v_3) de $\mathbb{Q}_p^3$ on ait

$$x = [\mathbb{Z}_p v_1 + \mathbb{Z}_p v_2 + \mathbb{Z}_p v_3] \text{ et } y = [\mathbb{Z}_p p^{-i-j} v_1 + \mathbb{Z}_p p^{-j} v_2 + \mathbb{Z}_p v_3].$$

On écrit $\sigma(x,y) = (i,j)$. On a alors $\sigma(y,x) = (j,i)$. On munit B de la distance d définie par $d(x,y) = i+j$ si $\sigma(x,y) = (i,j)$. Le déterminant d'une classe d'équivalence de réseaux (pour la relation d'homothétie) est bien déterminé dans $\mathbb{Q}_p^*/\mathbb{Z}_p^* p^{3\mathbb{Z}} = \mathbb{Z}/3\mathbb{Z}$ (où $p^{-1}\mathbb{Z}_p^* p^{3\mathbb{Z}}$ correspond à $1 \in \mathbb{Z}/3\mathbb{Z}$) et on appelle type : $B \to \mathbb{Z}/3\mathbb{Z}$ la fonction correspondante. On note B^0 l'ensemble des points de B de type 0 et on note $x_0 \in B^0$ la classe d'équivalence du réseau $\mathbb{Z}_p^3$. Comme $SL_3(\mathbb{Q}_p)$ agit transitivement sur B^0 et que le stabilisateur de x_0 est $SL_3(\mathbb{Z}_p)$, on a une bijection $G/K \to B^0$ qui à gK associe gx_0. Pour $x, y \in B^0$ on a $\sigma(x,y) \in \Lambda$. Pour $x, y, x', y' \in B^0$ on a $\sigma(x,y) = \sigma(x',y')$ si et seulement si il existe un élément de $SL_3(\mathbb{Q}_p)$ transportant x sur x' et y sur y'.

Démonstration du théorème 5.1 en admettant la proposition 5.2. Pour tout $n \in \mathbb{N}$, $f \in e_K C^*_{\mathrm{red}}(G, C_0(G))e_K$ est supporté sur $B_n = \{g \in G, \ell(g) \le n\}$ si et seulement si l'opérateur $T \in \mathcal{K}(\ell^2(B^0))$ correspondant vérifie $T_{x,y} = 0$ lorsque $d(x,y) > n$. On voit donc que le théorème 5.1 résulte de la proposition suivante. □

Proposition 5.2. *Soient $s \in [0, \frac{1}{4}[$ et $c : \Lambda \to \mathbb{C}$ une fonction telle que pour tout $n \in \mathbb{N}$*

$$\sup_T \frac{\|(T_{xy} c(\sigma(x,y)))_{x,y \in B^0}\|_{\mathcal{K}(\ell^2(B^0))}}{\|T\|_{\mathcal{K}(\ell^2(B^0))}} \le p^{sn} \tag{6}$$

où le supremum est pris sur les matrices T indexées par B^0, ayant un nombre fini de coefficients non nuls, et vérifiant $T_{xy} = 0$ si $d(x,y) > n$. Alors c admet une limite c_∞ à l'infini, et

$$|c(i,j) - c_\infty| \le \Big(\frac{2p^s}{p^{\frac{1}{2}+s} - 1} + \frac{2(p^s + p^{2s})}{1 - p^{\frac{4s-1}{2}}}\Big) p^{\frac{4s-1}{6}(i+j+\max(i,j))}$$

pour tout $(i,j) \in \Lambda$.

La fin de ce paragraphe est consacrée à la démonstration de la proposition 5.2.

Lemme 5.3. *Soit $(m, m') \in \mathbb{N}^2$ avec $m \le m'$. Soit (e_1, e_2, e_3) une base de $\mathbb{Q}_p^3$. Soient $x_1, x_2, x_3 \in \mathbb{Z}_p$ tel que l'un d'entre eux appartienne à $\mathbb{Z}_p^*$. Soient $\xi_1, \xi_2, \xi_3 \in \mathbb{Z}_p$ tel que l'un d'entre eux appartienne à $\mathbb{Z}_p^*$. Soit*

$$M = \mathbb{Z}_p p^{-m}(x_1 e_1 + x_2 e_2 + x_3 e_3) + \mathbb{Z}_p e_1 + \mathbb{Z}_p e_2 + \mathbb{Z}_p e_3 \quad \textit{et}$$

$$N = \{u_1e_1 + u_2e_2 + u_3e_3 \in \mathbb{Z}_pe_1 + \mathbb{Z}_pe_2 + \mathbb{Z}_pe_3, \xi_1u_1 + \xi_2u_2 + \xi_3u_3 \in p^{m'}\mathbb{Z}_p\}.$$

Alors $\sigma([M],[N]) = (i, m + m' - 2i)$ *où* i *est le plus grand entier de* $\{0, ..., m\}$ *tel que* $\xi_1x_1 + \xi_2x_2 + \xi_3x_3 \equiv 0$ *modulo* p^i.

Démonstration. Comme l'énoncé est invariant par l'action de $SL_3(\mathbb{Z}_p)$ sur le vecteur $\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ et le covecteur (ξ_1, ξ_2, ξ_3) on peut supposer $\xi_2 = \xi_3 = 0$, $\xi_1 = 1$ et $x_3 = 0$. Si $i < m$ la valuation p-adique de x_1 est égale à i et on a

$$M = \mathbb{Z}_pp^{-m}(x_1e_1 + x_2e_2) + \mathbb{Z}_pp^{-i}e_2 + \mathbb{Z}_pe_3$$

$$\text{et} \qquad N = \mathbb{Z}_pp^{m'-i}(x_1e_1 + x_2e_2) + \mathbb{Z}_pe_2 + \mathbb{Z}_pe_3.$$

Si $i = m$ on a $M = \mathbb{Z}_pe_1 + \mathbb{Z}_pp^{-m}e_2 + \mathbb{Z}_pe_3$ et $N = \mathbb{Z}_pp^{m'}e_1 + \mathbb{Z}_pe_2 + \mathbb{Z}_pe_3$. □

Remarque. On a toujours

$$\sigma([\mathbb{Z}_pe_1 + \mathbb{Z}_pe_2 + \mathbb{Z}_pe_3], [M]) = (m, 0) \text{ et } \sigma([\mathbb{Z}_pe_1 + \mathbb{Z}_pe_2 + \mathbb{Z}_pe_3], [N]) = (0, m').$$

Début de la démonstration de la proposition 5.2. Fixons $(m, m') \in \mathbb{N}^2$ avec $m \leq m'$ et $m + m' \in 3\mathbb{N}$. Soit (e_1, e_2, e_3) une base de $\mathbb{Q}_p^3$ telle que le type de $[\mathbb{Z}_pe_1 + \mathbb{Z}_pe_2 + \mathbb{Z}_pe_3]$ soit m' modulo 3. Nous définissons deux applications injectives $\alpha : (\mathbb{Z}/p^m\mathbb{Z})^2 \to B^0$ et $\beta : (\mathbb{Z}/p^{m'}\mathbb{Z})^2 \to B^0$ de la façon suivante. Pour $x = (x_2, x_3) \in (\mathbb{Z}/p^m\mathbb{Z})^2$,

$$\alpha(x) = [\mathbb{Z}_pp^{-m}(e_1 + x_2e_2 + x_3e_3) + \mathbb{Z}_pe_2 + \mathbb{Z}_pe_3],$$

et pour $y = (y_1, y_2) \in (\mathbb{Z}/p^{m'}\mathbb{Z})^2$,

$$\beta(y) = [\{u = u_1e_1 + u_2e_2 + u_3e_3 \in \mathbb{Z}_pe_1 + \mathbb{Z}_pe_2 + \mathbb{Z}_pe_3, u_1y_1 + u_2y_2 + u_3 \in p^{m'}\mathbb{Z}_p\}].$$

Par abus nous avons supposé dans les formules ci-dessus que x_2, x_3, y_1, y_2 étaient relevés en des éléments de $\mathbb{Z}_p$, mais $\alpha(x)$ et $\beta(y)$ ne dépendent pas des relèvements.

Lemme 5.4. *Soient* $x = (x_2, x_3) \in (\mathbb{Z}/p^m\mathbb{Z})^2$ *et* $y = (y_1, y_2)$ *dans* $(\mathbb{Z}/p^{m'}\mathbb{Z})^2$. *Soit* i *le plus grand entier de* $\{0, ..., m\}$ *tel que* $y_1 + x_2y_2 + x_3 \equiv 0$ *modulo* p^i. *Alors* $\sigma(\alpha(x), \beta(y)) = (i, m + m' - 2i)$.

Démonstration. Ce lemme est une conséquence immédiate du lemme 5.3. □

Lemme 5.5. *Soient* k, k' *des entiers supérieurs ou égaux à* 2 *avec* k' *multiple de* k, l_1, l_2 *deux éléments distincts de* $\mathbb{Z}/k\mathbb{Z}$ *et* $a_1, a_2 \in \mathbb{R}$. *Soit*

$$(T_{x,y})_{x \in (\mathbb{Z}/k\mathbb{Z})^2, y \in (\mathbb{Z}/k'\mathbb{Z})^2}$$

la matrice définie par

$$\begin{aligned} T_{(x_2,x_3),(y_1,y_2)} &= a_1 \text{ si } y_1 + x_2y_2 + x_3 = l_1 \text{ dans } \mathbb{Z}/k\mathbb{Z}, \\ &= a_2 \text{ si } y_1 + x_2y_2 + x_3 = l_2 \text{ dans } \mathbb{Z}/k\mathbb{Z}, \\ &= 0 \text{ sinon.} \end{aligned}$$

Alors

$$\begin{aligned} \|T\|_{M_{k^2,k'^2}(\mathbb{C})} &= \frac{k'}{k} \sup_{q \in \mathbb{Z}/k\mathbb{Z}} \left| a_1e^{i2\pi ql_1/k} + a_2e^{i2\pi ql_2/k} \right| \left\| (e^{-i2\pi qx_2y_2/k}) \right\|_{M_k(\mathbb{C})} \\ &= \frac{k'}{k} \sup_{q \in \mathbb{Z}/k\mathbb{Z}} \sqrt{k \times \mathrm{pgcd}(k,q)} \left| a_1e^{i2\pi ql_1/k} + a_2e^{i2\pi ql_2/k} \right|. \end{aligned}$$

Démonstration. Comme $T_{x,y}$ ne dépend que de l'image de y dans $(\mathbb{Z}/k\mathbb{Z})^2$, on se ramène facilement au cas où $k' = k$. Le lemme découle alors de la diagonalisation des matrices circulantes, car lorsqu'on fixe x_2 et y_2 la matrice $(T_{(x_2,x_3),(y_1,y_2)})_{x_3,y_1}$ est une matrice circulante, que l'on diagonalise par une transformation de Fourier en x_3 et y_1. $\square$

Fin de la démonstration de la proposition 5.2. Pour $i \in \{0,...,m\}$ on note $((T_i)_{x,y})_{x\in(\mathbb{Z}/p^m\mathbb{Z})^2, y\in(\mathbb{Z}/p^{m'}\mathbb{Z})^2}$ la matrice définie par

$$\begin{aligned}(T_i)_{(x_2,x_3),(y_1,y_2)} &= p^{-m'} \text{ si } y_1 + x_2y_2 + x_3 = p^i \ \text{ dans } \ \mathbb{Z}/p^m\mathbb{Z},\\ &= 0 \text{ sinon}\end{aligned}$$

On note que T_i est normalisée pour être de norme 1 dans $M_{p^{2m},p^{2m'}}(\mathbb{C})$.

Soit $i \in \{1,...,m\}$. Alors

$$\|T_i - T_{i-1}\| \leq 2p^{-\frac{i}{2}}$$

En effet par le lemme 5.5 on a

$$\|T_i - T_{i-1}\| = \sup_{q\in\mathbb{Z}/p^m\mathbb{Z}} p^{-\frac{m}{2}} \sqrt{\operatorname{pgcd}(p^m,q)} \left|1 - e^{i2\pi qp^{i-1}(p-1)/p^m}\right|$$

et $1 - e^{i2\pi qp^{i-1}(p-1)/p^m}$ s'annule si p^{m-i+1} divise q et

$$p^{-\frac{m}{2}} \sqrt{\operatorname{pgcd}(p^m,q)} \left|1 - e^{i2\pi qp^{i-1}(p-1)/p^m}\right| \leq 2p^{-\frac{m}{2}} \sqrt{\operatorname{pgcd}(p^m,q)} \leq 2p^{-\frac{i}{2}}$$

si p^{m-i+1} ne divise pas q.

Comme le vecteur de norme 1 de $\mathbb{C}^{(\mathbb{Z}/p^{m'}\mathbb{Z})^2}$ dont toutes les coordonnées sont égales à $p^{-m'}$ a pour image par T_i et T_{i-1} le vecteur de norme 1 de $\mathbb{C}^{(\mathbb{Z}/p^m\mathbb{Z})^2}$ dont toutes les coordonnées sont égales à p^{-m} , on a, pour $a, b \in \mathbb{C}$, $\|aT_i+bT_{i-1}\| \geq |a+b|$.

Soit $(m, m') \in \mathbb{N}^2$ avec $m \leq m'$. Pour tout $i \in \{0, \ldots, m\}$ on note $\tilde{T}_i \in \mathcal{K}(\ell^2(B^0))$ la matrice (qui a un nombre fini de coefficients non nuls) telle que

$$(\tilde{T}_i)_{x,y} = 0 \text{ si } \ x \notin \operatorname{Im}(\alpha) \text{ ou } \ y \notin \operatorname{Im}(\beta)$$

$$\text{et } (\tilde{T}_i)_{\alpha(x),\beta(y)} = (T_i)_{x,y} \text{ pour } \ x \in (\mathbb{Z}/p^m\mathbb{Z})^2 \text{ et } y \in (\mathbb{Z}/p^{m'}\mathbb{Z})^2.$$

Soit $i \in \{1, \ldots, m\}$. En appliquant (6) à $T = \tilde{T}_i - \tilde{T}_{i-1}$ et $n = m + m' - i + 1$, on obtient

$$\begin{aligned}&\Big|c(i, m + m' - 2i) - c(i - 1, m + m' - 2i + 2))\Big|\\ &\leq \Big\|c(i, m + m' - 2i)T_i - c(i - 1, m + m' - 2i + 2)T_{i-1}\Big\|\\ &\leq \|T_i - T_{i-1}\|p^{s(m+m'-i+1)} \leq 2p^{-\frac{i}{2}}p^{s(m+m'-i+1)}.\end{aligned}$$

En appliquant l'inégalité précédente avec $m' = m$ ou $m' = m + 1$ (c'est-à-dire $m = i + [\frac{j}{2}] \geq i$) on trouve que pour tous $(i, j) \in \Lambda$ avec $i > 0$,

$$|c(i, j) - c(i - 1, j + 2)| \leq 2p^{-\frac{i}{2}+s(i+j+1)}.$$

En faisant agir l'automorphisme

$$g \mapsto \begin{pmatrix}0 & 0 & 1\\0 & 1 & 0\\1 & 0 & 0\end{pmatrix} {}^tg^{-1} \begin{pmatrix}0 & 0 & 1\\0 & 1 & 0\\1 & 0 & 0\end{pmatrix}$$

de G, qui stabilise K, et envoie

$$p^{\frac{i+2j}{3}}\begin{pmatrix} p^{-(i+j)} & 0 & 0 \\ 0 & p^{-j} & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ sur } p^{\frac{2i+j}{3}}\begin{pmatrix} p^{-(i+j)} & 0 & 0 \\ 0 & p^{-i} & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

on obtient, pour $(i,j) \in \Lambda$ avec $j > 0$,

$$|c(i,j) - c(i+2, j-1)| \leq 2p^{-\frac{j}{2}+s(i+j+1)}.$$

Donc on a, pour $(i,j) \in \Lambda$, avec $i \geq j$,

$$\begin{aligned} &\Big|c(i,j) - c(\frac{2i+j}{3}, \frac{2i+j}{3})\Big| \\ &\leq 2\Big(\big(p^{-\frac{1}{2}-s}\big) + \big(p^{-\frac{1}{2}-s}\big)^2 + \cdots + \big(p^{-\frac{1}{2}-s}\big)^{\frac{i-j}{3}}\Big)p^{\frac{2i+j}{6}(4s-1)+s} \qquad (7) \\ &\leq \frac{2p^s}{p^{\frac{1}{2}+s} - 1}p^{\frac{4s-1}{6}(2i+j)}, \end{aligned}$$

ainsi que la même inégalité en permutant i et j. Pour $i \geq 1$ on trouve

$$\begin{aligned} |c(i,i) - c(i+1,i+1)| &\leq |c(i,i) - c(i-1,i+2)| + |c(i-1,i+2) - c(i+1,i+1)| \\ &\leq 2\big(p^s + p^{2s-1}\big)p^{\frac{4s-1}{2}i}. \end{aligned}$$

En prenant $x_1 \in B^0$ tel que $\sigma(x_0, x_1) = (1,1)$, et en appliquant (6) à la matrice T indexée par B^0 telle que $T_{x,y} = 1$ pour $x, y \in \{x_0, x_1\}$ et 0 sinon, on obtient

$$|c(0,0) - c(1,1)| \leq \left\| \begin{pmatrix} c(0,0) & c(1,1) \\ c(1,1) & c(0,0) \end{pmatrix} \right\| \leq p^{2s}\left\| \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right\| = 2p^{2s}.$$

On a donc pour tout $i \in \mathbb{N}$,

$$|c(i,i) - c(i+1,i+1)| \leq 2\big(p^s + p^{2s}\big)p^{\frac{4s-1}{2}i}.$$

On en déduit que $c(i,i)$ admet une limite c_∞ à l'infini et que

$$\Big|c(i,i) - c_\infty\Big| \leq 2\frac{p^s + p^{2s}}{1 - p^{\frac{4s-1}{2}}}p^{\frac{4s-1}{2}i}$$

pour tout $i \in \mathbb{N}$. En utilisant (7) et l'inégalité obtenue en permutant i et j on a alors, pour $(i,j) \in \Lambda$,

$$\Big|c(i,j) - c_\infty\Big| \leq \Big(\frac{2p^s}{p^{\frac{1}{2}+s} - 1} + 2\frac{p^s + p^{2s}}{1 - p^{\frac{4s-1}{2}}}\Big)p^{\frac{4s-1}{6}(i+j+\max(i,j))}.$$

La proposition 5.2 est démontrée. □

Références

[BC82] P. Baum and A. Connes. Geometric K-theory for Lie groups and foliations (preprint de 1982). *Enseign. Math. (2)*, 46 :3–42, 2000.

[BCH94] P. Baum and A. Connes and N. Higson. Classifying space for proper actions and K-theory of group C^*-algebras. *C^*-algebras : 1943–1993* (San Antonio, TX, 1993), Contemp. Math., **167**, Amer. Math. Soc. : 240–291, 1994.

[BHV08] B. Bekka, P. de la Harpe et A. Valette. Kazhdan's property (T). Cambridge University Press, 2008.

[BD06] G. Bell et A. Dranishnikov A Hurewicz-type theorem for asymptotic dimension and applications to geometric group theory. *Trans. Amer. Math. Soc.* 358(11) :4749–4764, 2006.

[Bos90] J.-B. Bost. Principe d'Oka, K-théorie et systèmes dynamiques non commutatifs. *Invent. Math.* 101 :261–333, 1990.

[CH85] J. De Cannière et U. Haagerup. Multipliers of the Fourier algebras of some simple Lie groups and their discrete subgroups. *Amer. J. Math.* 107(2) :455–500, 1985.

[CG04] G. Carlsson et B. Goldfarb. On homological coherence of discrete groups. *J. Algebra* 276(2) :502–514, 2004.

[Cha03] I. Chatterji. Property (RD) for cocompact lattices in a finite product of rank one Lie groups with some rank two Lie groups. *Geometriae Dedicata* 96 : 161–177, 2003.

[CDSW05] M. Cowling, B. Dorofaeff, A. Seeger et J. Wright. A family of singular oscillatory integral operators and failure of weak amenability. *Duke Math. J.* 127(3) : 429–486, 2005.

[CH89] M. Cowling et U. Haagerup. Completely bounded multipliers of the Fourier algebra of a simple Lie group of real rank one. *Invent. Math.* 96(3) :507–549, 1989.

[Dor93] B. Dorofaeff. The Fourier algebra of $SL(2,\mathbb{R}) \rtimes \mathbb{R}^n$, $n \geq 2$, has no multiplier bounded approximate unit. *Math. Ann.* 297(4) : 707–724, 1993.

[Gal99] P.Y. Le Gall. Théorie de Kasparov équivariante et groupoïdes. I. *K-Theory*, 16(4) :361–390, 1999.

[Gro93] M. Gromov. Asymptotic invariants for infinite groups. Geometric Group Theory, LMS Lect. Notes 182, 1993.

[Haa86] U. Haagerup. Group C^*-algebras without the completely bounded approximation property, Preprint, 1986.

[HV89] P. de la Harpe and A. Valette. La propriété (T) de Kazhdan pour les groupes localement compacts. Astérisque 175, 1989.

[HK01] N. Higson and G. Kasparov. E-theory and KK-theory for groups which act properly and isometrically on Hilbert space. *Invent. Math.* 144(1) : 23–74, 2001.

[HLS02] N. Higson, V. Lafforgue and G. Skandalis. Counterexamples to the Baum-Connes conjecture. *Geom. Funct. Anal.* 12(2) :330–354, 2002.

[Jen68] J. Jenkins. An amenable group with a nonsymmetric group algebra. *Bull. Amer. Math. Soc.* 75 :357–360, 1969.

[JV84] P. Julg and A. Valette. K-theoretic amenability for $SL_2(\mathbb{Q}_p)$, and the action on the associated tree. *J. Funct. Anal.* 58 :194–215, 1984.

[JK95] P. Julg and G. Kasparov. Operator K-theory for the group $SU(n,1)$. *J. Reine Angew. Math.*, 463 :99–152, 1995.

[Jul97] P. Julg. Remarks on the Baum-Connes conjecture and Kazhdan's property T. Operator algebras and their applications, Waterloo (1994/1995), Fields Inst. Commun., Amer. Math. Soc. **13**, 145–153, 1997.

[Jul98] P. Julg. Travaux de N. Higson et G. Kasparov sur la conjecture de Baum-Connes. Séminaire Bourbaki, 40 (1997-1998), Exp. No. 841.

[Jul02] P. Julg. La conjecture de Baum-Connes à coefficients pour le groupe $Sp(n,1)$. *C. R. Acad. Sci.* 334(7) :533–538, 2002.

[Kas84] G. G. Kasparov. Lorentz groups : K-theory of unitary representations and crossed products. *Dokl. Akad. Nauk SSSR*, 275(3) :541–545, 1984.

[Kas88] G.G. Kasparov. Equivariant KK-theory and the Novikov conjecture. *Invent. Math.*, 91 :147–201, 1988.

[KS91] G.G. Kasparov and G. Skandalis. Groups acting on buildings, operator K-theory, and Novikov's conjecture. *K-Theory*, 4(4) :303–337, 1991.

[KS94] G. Kasparov and G. Skandalis. Groupes boliques et conjecture de Novikov. *C. R. Acad. Sci. Paris Sér. I Math.*, 319(4) :815–820, 1994.

[KS03] G. Kasparov and G. Skandalis. Groups acting properly on "bolic" spaces and the Novikov conjecture. *Ann. of Math. (2)*, 158(1) :165–206, 2003.

[Kaz67] D. Kazhdan. Connection of the dual space of a group with the structure of its closed subgroups. *Funct. Anal. Appl.* 1 :63–65, 1967.

[Laf00] V. Lafforgue. *A proof of property* (RD) *for cocompact lattices of* $\mathrm{SL}(3,\mathbb{R})$ *and* $\mathrm{SL}(3,\mathbb{C})$. *J. Lie Theory* 10(2) :255–267, 2000.

[Laf01] V. Lafforgue. Banach KK-theory and the Baum-Connes conjecture. European Congress of Mathematics, Vol. II, (Barcelone,2000), Progr. Math. 202 :31–46, Birkhäuser, 2001.

[Laf02a] V. Lafforgue. K-théorie bivariante pour les algèbres de Banach et conjecture de Baum-Connes. *Invent. Math.*, 149(1) :1–95, 2002.

[Laf02b] V. Lafforgue. Banach KK-theory and the Baum-Connes conjecture. Proceedings of the ICM 2002 (Beijing), vol II, 795–812.

[Laf07] V. Lafforgue (avec un appendice de Hervé Oyono-Oyono). K-théorie bivariante pour les algèbres de Banach, groupoïdes et conjecture de Baum-Connes. *J. Inst. Math. Jussieu* 6(3) : 415–451, 2007.

[Laf08] V. Lafforgue. Un renforcement de la propriété (T). *Duke Math. J.*, 143(3) :559–602, 2008.

[Laf09] V. Lafforgue. La conjecture de Baum-Connes à coefficients pour les groupes hyperboliques. Preprint 2009, disponible à l'adresse http ://people.math.jussieu.fr/~vlafforg/

[Mat07] D. Matsnev. The Baum-Connes conjecture and proper group actions on affine buildings. arXiv :math/0703923

[MY02] I. Mineyev and G. Yu. The Baum-Connes conjecture for hyperbolic groups. *Invent. Math.*, 149(1) :97–122, 2002.

[Pim86] M. Pimsner. KK-groups of crossed products by groups acting on trees. *Invent. Math.* 86 : 603–634, 1986.

[RRS98] J. Ramagge, G. Robertson, T. Steger. A Haagerup inequality for $\widetilde{A}_1 \times \widetilde{A}_1$ and $\widetilde{A}_2$ buildings. *Geom. Funct. Anal.* 8 : 702–731, 1998.

[Roe05] J. Roe. Hyperbolic groups have finite asymptotic dimension. *Proc. Amer. Math. Soc.* 133(9) :2489–2490, 2005.

[Ska91] G. Skandalis. Kasparov's bivariant K-theory and applications. *Expositiones Math.* 9 : 193–250, 1991.

[Ska99] G. Skandalis. Progrès récents sur la conjecture de Baum-Connes, contribution de Vincent Lafforgue. Séminaire Bourbaki, No. 869 (novembre 1999).

[Tu99a] J.-L. Tu. La conjecture de Novikov pour les feuilletages hyperboliques. *K-Theory*, 16(2) :129–184, 1999.

[Tu99b] J.L. Tu. La conjecture de Novikov pour les feuilletages moyennables. *K-Theory* 17(3) : 215–264, 1999.

Institut de mathématiques de Jussieu, 175 rue du Chevaleret, 75013 Paris
E-mail address: `vlafforg@math.jussieu.fr`

Clay Mathematics Proceedings
Volume 11, 2010

Combinatorial Hopf algebras

Jean-Louis Loday and María Ronco

Ce travail est dédié à Alain Connes en témoignage d'estime et d'une longue amitié

Abstract. We define a "combinatorial Hopf algebra" as a Hopf algebra which is free (or cofree) and equipped with a given isomorphism to the free algebra over the indecomposables (resp. the cofree coalgebra over the primitives). We show that the choice of such an isomorphism implies the existence of a finer algebraic structure on the Hopf algebra and on the indecomposables (resp. the primitives). For instance a cofree-cocommutative right-sided combinatorial Hopf algebra is completely determined by its primitive part which is a pre-Lie algebra. The key example is the Connes-Kreimer Hopf algebra. The study of all these combinatorial Hopf algebra types gives rise to several good triples of operads. It involves the operads: dendriform, pre-Lie, brace, and variations of them.

Introduction

Many recent papers are devoted to some infinite dimensional Hopf algebras called collectively "combinatorial Hopf algebras". The Connes-Kreimer Hopf algebra is one of them [**13**]. Among other examples we find: the Faà di Bruno algebra, variations of the symmetric functions: $Sym, NSym, FSym, QSym, FQSym$, $PQSym$, (Aguiar-Sottile [**2, 3**], Bergeron-Hohlweg [**4**], Chapoton-Livernet [**12**], Livernet [**33**], Malvenuto-Reutenauer [**44**], Hivert-Novelli-Thibon [**25**], Palacios-Ronco [**48**]), examples related to knot theory (Turaev [**61**]), to quantum field theory (Brouder [**5**], Figueroa-Gracia-Bondía [**16**], Brouder-Frabetti-Krattenthaler [**6, 7**], van Suijlekom [**59**]), to foliations (Connes-Moscovici [**14**]), to K-theory (Gangl-Goncharov-Levin [**21**], Loday-Ronco [**39**], Lam-Pylyavskyy [**32**]). The aim of this paper is to make precise the meaning of combinatorial Hopf algebras and to unravel their fine algebraic structure.

Here is a typical example of a combinatorial Hopf algebra. The Connes-Kreimer algebra $\mathcal{H}_{CK}$ is a free-commutative Hopf algebra over some families of trees t. The

---

2010 *Mathematics Subject Classification.* Primary 16T30; Secondary 18D50.

Key words and phrases. Tree, Hopf algebra, combinatorial, dendriform, dipterous, brace, multibrace, pre-Lie.

This work has been partially supported by the FONDECYT Project 1085004, the "Ecos-Sud" project E01C06, and an "Agence Nationale pour la Recherche" (ANR) contract.

coalgebra structure is of the form

$$\Delta(t) = \sum_c P_c(t) \otimes R_c(t),$$

where $P_c(t)$ is a polynomial of trees and $R_c(t)$ is a tree (the sum is over admissible cuts c, cf. 5.7). In this description there are two key points. First, we observe that, not only is $\mathcal{H}_{CK}$ free as a commutative algebra, but we are given a precise isomorphism with the polynomial algebra over the indecomposables (spanned by the trees). This is why we call it a "combinatorial" Hopf algebra. Second, the coproduct has a special form since its second component is linear (a tree instead of a polynomial of trees). This is called the "right-sided condition". A priori the space of indecomposables is a Lie coalgebra. But it is known, in this case, that it has a finer structure: it is a *pre-Lie* coalgebra. Our claim is that the existence of this pre-Lie structure is not special to the Connes-Kreimer algebra, but is a general fact for combinatorial Hopf algebras satisfying the right-sided condition.

This case (free-commutative Hopf algebra) can be dualized (in the graded sense) to give rise to the following statement: a Hopf algebra structure on the cofree-cocommutative coalgebra $S^c(V)$, which satisfies some "right-sided condition" (cf. 3.1), induces a pre-Lie algebra structure on V (cf. Theorem 5.3). The Lie product on the primitive part V is the anti-symmetrization of this pre-Lie product.

We will see that there are several contexts for combinatorial Hopf algebras, depending on the following choice of options: free or cofree, associative or commutative, general or right-sided. Since the free and cofree cases are dual to one another under graded dualization we will only study one of them in detail, namely the cofree case. Some generalizations are evoked in the last section.

In the cofree-coassociative general context a combinatorial Hopf algebra is a Hopf algebra structure on the tensor coalgebra $T^c(R)$ for a certain vector space R. We show that, for such a combinatorial Hopf algebra, the primitive part R is a *multibrace algebra* (MB-algebra) and that any multibrace algebra R gives rise to a Hopf structure on $T^c(R)$. A multibrace algebra is determined by $(p+q)$-ary operations M_{pq} satisfying some relations analogous to the relations satisfied by the brace algebras (see 2.3). Moreover $T^c(R)$ inherits a finer algebraic structure: it is a *dipterous algebra* (see 2.7). So, in the cofree-coassociative context, the classification of the combinatorial Hopf algebras involves the operad $\mathcal{P} = MB$ governing the primitive part and the operad $\mathcal{A} = Dipt$ governing the algebra structure. We show that the combinatorial Hopf algebra associated to a free multibrace algebra is in fact a free dipterous algebra:

$$Dipt(V) \cong T^c(MB(V)).$$

This is part of a more general result which says that there is a "good triple of operads" (in the sense of [**38**]):

$$(As, Dipt, MB),$$

where As is the operad of associative algebras.

Our aim is to handle similarly the three other cases (cofree-coassociative right-sided, cofree-cocommutative general, cofree-cocommutative right-sided) and to determine the operad $\mathcal{P}$ and the operad $\mathcal{A}$ when possible. We show that these two operads are strongly related since there is a good triple of operads

$$(\mathcal{C}, \mathcal{A}, \mathcal{P})$$

where $\mathcal{C} = As$ or Com (governing commutative algebras) depending on the context we are working in. We show that the operads $\mathcal{P}$ and $\mathcal{A}$ are as follows:

	$\mathcal{C}$	$\mathcal{A}$	$\mathcal{P}$
general	As	$Dipt$ or $2as$	MB
right-sided	As	$Dend$ or $\mathcal{Y} := 2as/\sim$	$Brace$
general	Com	$ComAs$	SMB
right-sided	Com	$\mathcal{X} :=?$	$SBrace = PreLie$

In the last case we conjecture the existence of an operad structure on $Com \circ preLie$, that we denote by $\mathcal{X}$. Some of these operads are known, the other ones are variations of known types:

- MB : multibrace algebra, also called $\mathbf{B}_\infty$-algebra in [**42**], non-differential B_∞ in [**19**], Hirsch algebra in [**30**], LR-algebra in [**45**], see 2.5,
- $Brace$: brace algebra [**19, 20**], see 3.3,
- SMB : symmetric multibrace algebra, a symmetric variation of MB, see 4.5,
- $PreLie$: pre-Lie algebra, also called Vinberg algebra, or right-symmetric algebra, see 5.1,
- $Dipt$: dipterous algebra [**40**], see 2.7,
- $Dend$: dendriform algebra [**35**], see 3.6,
- $2as$: 2-associative algebra (two associative products) [**42**], see 2.19,
- $\mathcal{Y} := 2as/\{M_{pq} = 0 \mid p \geq 2\}$, see 3.17,
- $ComAs$: commutative-associative algebra (a commutative product and an associative product), see 4.7.

The dimension of the space of n-ary operations is given in the following tableau:

	$\mathcal{C}$	$\mathcal{A}$	$\mathcal{P}$
general, As	$n!$	$2C_{n-1} \times n!$	$C_n \times n!$
right-sided, As	$n!$	$c_n \times n!$	$c_{n-1} \times n!$
general, Com	1	d_n	f_n
right-sided, Com	1	$(n+1)^{n-1}$	n^{n-1}

where c_n is the Catalan number, C_n is the super-Catalan number, d_n is the number of "labeled series-parallel posets with n points" and f_n is the number of "connected labeled series-parallel posets with n nodes" (cf. [**57, 58**] problem 5.39 and 4.14):

n	1	2	3	4	5	$\cdots$
c_n	1	2	5	14	42	$\cdots$
C_n	1	3	11	45	197	$\cdots$
$\dim ComAs(n) = d_n$	1	3	19	195	2791	$\cdots$
$\dim SMB(n) = f_n$	1	2	12	122	1740	$\cdots$

We observe that some of these operads are rather simple since they are binary and quadratic: $PreLie, Dipt, Dend, 2as, ComAs$. The others are more complicated since they involve n-ary generating operations for any n: $MB, Brace, SMB$, or nonquadratic relations: $\mathcal{Y}$.

Summarizing our results, we have shown the existence of the following good triples of operads:

$$\begin{array}{ccc}(As, Dipt, MB), & (As, 2as, MB), & (Com, ComAs, SMB), \\ (As, Dend, Brace), & (As, \mathcal{Y}, Brace), & (Com, \mathcal{X}, PreLie),\end{array}$$

except for the last one which is conjectural.

In most examples the space of primitives R comes endowed with a basis of "combinatorial objects" like permutations, trees, graphs, tableaux. Therefore the Hopf algebra is made up of polynomials on these combinatorial objects. For instance in every context the free algebra $\mathcal{P}(V)$ can be described by means of trees. Changing the basis of R does not change the $\mathcal{P}$-algebra structure of R. However, if the coalgebra isomorphism $\mathcal{H} \cong T^c(R)$ (resp. $S^c(R)$) is modified, then the $\mathcal{P}$-algebra structure of R is modified accordingly. For instance, in the cofree-cocommutative right-sided context, the structure of pre-Lie algebra of R is modified (but not its structure of Lie algebra). In 6.4 we make explicit two different combinatorial Hopf algebra structures on the Malvenuto-Reutenauer Hopf algebra. This importance of the basis had been envisioned by Joni and Rota in their seminal paper [**29**] where they say: *"It must be stressed that the coalgebras of combinatorics come equipped with a distinguished basis, and many an interesting combinatorial problem can be formulated algebraically as that of transforming this basis into another basis with more desirable properties. Thus, a mere structure theory of coalgebras—or Hopf algebras—will hardly be sufficient for combinatorial purposes."*

The plan of the paper is as follows. In the first section we recall the basic notion of Hopf algebra and the results on triples of operads that are going to be used in the proofs. The other four sections are devoted to the following four cases:

- section 2: cofree-coassociative CHA and MB-algebras,
- section 3: cofree-coassociative right-sided CHA and $Brace$-algebras,
- section 4: cofree-cocommutative CHA and SMB-algebras,
- section 5: cofree-cocommutative right-sided CHA and $PreLie$-algebras.

The plan of each of these four sections is as follows:

- definition of the CHAs involved,
- the algebraic structure of the primitives, the operad $\mathcal{P}$, the equivalence,
- the algebraic structure of the Hopf algebra, the operad $\mathcal{A}$,
- comparison of $\mathcal{P}$ and $\mathcal{A}$, the good triple $(\mathcal{C}, \mathcal{A}, \mathcal{P})$, where $\mathcal{C} = As$ or Com,
- combinatorial description of the free algebras $\mathcal{P}(V)$ and $\mathcal{A}(V)$ when available,
- the dual case (free-associative or free-commutative),
- variation (in the associative context)

The first case (cofree-coassociative general CHA) is treated in details. When the proofs in the other cases are analogous they are, most of the time, omitted. In the last section we list several examples from the literature.

Finally, let us say a word about the terminology. In the literature the term "combinatorial Hopf algebras" is used to call Hopf algebras based on combinatorial objects (cf. for instance [**25**]), without a precise definition about what is a combinatorial object. In all cases these Hopf algebras are free (or cofree) and the combinatorial objects provide a basis, whence an isomorphism as required in our definition. The only exception is the use of combinatorial Hopf algebra in the paper [**3**] where an extra piece of information is required (namely a character map).

Acknowledgement. Thanks to Emily Burgunder, Vladimir Dotsenko, Ralf Holtkamp, Muriel Livernet and Bruno Vallette for useful comments on earlier versions of this paper. Thanks also to the referee whose comments were very helpful in clarifying a few points.

Notation. In this paper $\mathbb{K}$ is a field and all vector spaces are over $\mathbb{K}$. Its unit is denoted by 1. The vector space spanned by the elements of a set X is denoted by $\mathbb{K}[X]$. The tensor product of vector spaces over $\mathbb{K}$ is denoted by $\otimes$. The tensor product of n copies of the space V is denoted by $V^{\otimes n}$. For $v_i \in V$ the element $v_1 \otimes \cdots \otimes v_n$ of $V^{\otimes n}$ is denoted by the concatenation of the elements: $v_1 \cdots v_n$. The *tensor module* over V is the direct sum

$$T(V) := \mathbb{K}1 \oplus V \oplus V^{\otimes 2} \oplus \cdots \oplus V^{\otimes n} \oplus \cdots$$

and the *reduced tensor module* is $\overline{T}(V) := T(V)/\mathbb{K}1$. The *symmetric module* over V is the direct sum

$$S(V) := \mathbb{K}1 \oplus V \oplus S^2(V) \oplus \cdots \oplus S^n(V) \oplus \cdots$$

and the *reduced symmetric module* is $\bar{S}(V) := S(V)/\mathbb{K}1$, where $S^n(V) = (V^{\otimes n})_{S_n}$ is the quotient of $V^{\otimes n}$ by the action of the symmetric group. We still denote by $v_1 \cdots v_n$ the image in $S^n(V)$ of $v_1 \cdots v_n \in V^{\otimes n}$. If V is generated by $x_1, \ldots, x_n$, then $S(V)$ (resp. $T(V)$) can be identified with the polynomials (resp. noncommutative polynomials) in n variables.

If the set X is a basis of $V = \mathbb{K}[X]$, then we write $T(X)$ (resp. $S(X)$) in place of $T(V)$ (resp. $S(V)$).

1. Prerequisites: Hopf algebras and operads

1.1. Hopf algebras. Recall that a *bialgebra* is a vector space $\mathcal{H}$ equipped with an associative and unital algebra structure $(\mathcal{H}, *, u)$ and a coassociative counital coalgebra structure $(\mathcal{H}, \Delta, \epsilon)$, which satisfy the *Hopf compatibility relation*, that is, $\Delta : \mathcal{H} \to \mathcal{H} \otimes \mathcal{H}$ and $\epsilon : \mathcal{H} \to \mathbb{K}$ are morphisms of associative unital algebras.

It is helpful to introduce the augmentation ideal $\overline{\mathcal{H}} := \operatorname{Ker} \epsilon$ and the reduced comultiplication $\overline{\Delta} : \overline{\mathcal{H}} \to \overline{\mathcal{H}} \otimes \overline{\mathcal{H}}$ defined by the formula $\Delta(x) = x \otimes 1 + 1 \otimes x + \overline{\Delta}(x)$. The iteration of the map $\overline{\Delta}$ gives rise to n-ary cooperations $\overline{\Delta}^n : \overline{\mathcal{H}} \to \overline{\mathcal{H}}^{\otimes n}$. The filtration of $\overline{\mathcal{H}}$ is defined by:

$$F_r\overline{\mathcal{H}} := \{x \in \overline{\mathcal{H}} \mid \overline{\Delta}^n(x) = 0 \text{ for any } n > r\}.$$

By definition $\mathcal{H}$ is said to be *conilpotent* if $\overline{\mathcal{H}} = \bigcup_{r \geq 1} F_r\overline{\mathcal{H}}$. Observe that the first piece of the filtration is the space of primitives of the bialgebra: $F_1\overline{\mathcal{H}} = \operatorname{Prim} \mathcal{H}$. Since any conilpotent bialgebra can be equipped with an antipode, there is an equivalence between conilpotent Hopf algebras and conilpotent bialgebras.

1.2. Cofree bialgebras. We say that a conilpotent bialgebra is *cofree-coassociative* if, as a conilpotent coalgebra, it is a cofree object. Denoting by R the space of primitives, it means that there exists an isomorphism of coalgebras $\mathcal{H} \cong T^cR$. Recall that the cofree coalgebra T^cR is the tensor module as a graded vector space, and that the coproduct is the *deconcatenation*:

$$\Delta(x_1 \cdots x_n) = \sum_{i=0}^{i=n} x_1 \cdots x_i \otimes x_{i+1} \cdots x_n.$$

As a result the following universal property holds: any conilpotent coalgebra homomorphism $C \to T^c(R)$ is completely determined by the composite $C \to T^c(R) \to R$. On the other hand, any linear map $\varphi : C \to R$ which maps 1 to 0 determines a unique coalgebra homomorphism $C \to T^c(R)$. The component in $R^{\otimes n}$ of the image of $c \in C$ is $\sum \varphi(c_{(1)}) \otimes \cdots \otimes \varphi(c_{(n)})$, where $\Delta^n(c) = \sum c_{(1)} \otimes \cdots \otimes c_{(n)}$.

Similar statements hold in the cocommutative case with $S^c(R)$ in place of $T^c(R)$.

1.3. Operads. For an operad $\mathcal{P}$ the free $\mathcal{P}$-algebra over the vector space V is denoted by $\mathcal{P}(V)$. It is of the form $\mathcal{P}(V) = \bigoplus_n \mathcal{P}(n) \otimes_{S_n} V^{\otimes n}$ where $\mathcal{P}(n)$ is the space of n-ary operations considered as a right module. The left module structure of $V^{\otimes n}$ is given by $\sigma(v_1 \cdots v_n) = v_{\sigma^{-1}(1)} \cdots v_{\sigma^{-1}(n)}$. If $\mathcal{P}(n)$ is free as a representation of S_n, then we write it $\mathcal{P}(n) = \mathcal{P}_n \otimes \mathbb{K}[S_n]$ and $\mathcal{P}_n$ is called the space of nonsymmetric n-ary operations. In this case we obtain $\mathcal{P}(V) = \bigoplus_n \mathcal{P}_n \otimes V^{\otimes n}$.

The generating series of the operad $\mathcal{P}$ is

$$f^{\mathcal{P}}(t) := \sum_{n \geq 1} \dim \frac{\mathcal{P}(n)}{n!} t^n.$$

The operads governing the associative algebras, commutative algebras, Lie algebras, pre-Lie algebras are denoted *As*, *Com*, *Lie*, *PreLie* respectively. For more on operads, see for instance [**43**].

1.4. Triple of operads. Recall briefly from [**38**] what it means for $(\mathcal{C}, \mathcal{A}, \mathcal{P})$ to be a good triple of operads. First, we assume that there is a well-defined notion of $\mathcal{C}^c$-$\mathcal{A}$-bialgebra, $\mathcal{C}$ governing the coalgebra structure and $\mathcal{A}$ governing the algebra structure. The operations and the cooperations are assumed to be related by *compatibility relations*. In our cases these compatibility relations are always distributive, that is, the composite of an operation with a cooperation can be written as the composite of a cooperation followed by an operation (or an algebraic sum of them). Then,

– the primitive part of any $\mathcal{C}^c$-$\mathcal{A}$-bialgebra is a $\mathcal{P}$-algebra,

– there is a pair of adjoint functors $(U \dashv F)$: $\mathcal{A}\text{-alg} \underset{F}{\overset{U}{\leftrightarrows}} \mathcal{P}\text{-alg}$,

– the following structure theorem holds:
for any $\mathcal{C}^c$-$\mathcal{A}$-bialgebra $\mathcal{H}$ the following are equivalent:

(a) $\mathcal{H}$ is conilpotent,
(b) $\mathcal{H}$ is isomorphic to $U(\operatorname{Prim}\mathcal{H})$,
(c) $\mathcal{H}$ is cofree over its primitive part, i.e. isomorphic to $\mathcal{C}^c(\operatorname{Prim}\mathcal{H})$.

An important consequence of this theorem is an isomorphism of vector spaces $\mathcal{A}(V) \cong \mathcal{C}^c(\mathcal{P}(V))$, which is functorial in V.

In [**38**] we found a criterion which ensures that the notion of $\mathcal{C}^c$-$\mathcal{A}$-bialgebra gives rise to a good triple of operads $(\mathcal{C}, \mathcal{A}, \mathcal{P})$, where the operad $\mathcal{P} := \mathrm{Prim}_{\mathcal{C}}\mathcal{A}$ is made up of the primitive operations. There are three conditions:
(H0) The compatibility relations are distributive.
(H1) The free $\mathcal{A}$-algebra $\mathcal{A}(V)$ is naturally a $\mathcal{C}^c$-$\mathcal{A}$-bialgebra.
(H2epi) The $\mathcal{C}$-coalgebra map $\mathcal{A}(V) \to \mathcal{C}^c(V)$ (deduced from H1) admits a splitting.

The isomorphism $\mathcal{A}(V) \cong \mathcal{C}^c(\mathcal{P}(V))$, which is a consequence of these hypotheses, depends on the choice of the splitting. It implies $f^{\mathcal{A}}(t) = f^{\mathcal{C}}(f^{\mathcal{P}}(t))$.

In the cases that we are looking at in this paper $\mathcal{C}$ is either the operad As or the operad Com. The condition (H0) is immediate to verify by direct inspection. So the first task (once $\mathcal{A}$ is discovered) consists in verifying the hypotheses (H1) and (H2epi). The second task consists in unravelling the operad $\mathcal{P}$ by providing a presentation by generators and relations.

2. (Co)free-(co)associative CHA

In this section we study the cofree-coassociative CHAs. Some of the results of this section were announced in [**40**], where the notion of dipterous algebras was introduced for the first time. Examples will be given in section 6. They include the Malvenuto-Reutenauer algebra (6.4), the incidence Hopf algebras (6.8) and the Hopf algebra on a tensor module (6.5).

2.1. Definition. A *cofree-coassociative combinatorial Hopf algebra $\mathcal{H}$* (cofree-coassociative CHA for short) is a cofree bialgebra with a prescribed isomorphism with $T^c(R)$ (here $R := \mathrm{Prim}\,\mathcal{H}$). Equivalently, it is a vector space R with a product $*$ on $T^c(R)$ which makes $(T^c(R), *, \Delta)$ (Δ = deconcatenation) into a Hopf algebra.

It is important to notice that the isomorphism $\mathcal{H} \cong T^c(R)$ is part of the structure. Any change of this isomorphism changes the product $*$. Such an isomorphism is completely determined by a splitting of the inclusion $R = \mathrm{Prim}\,\mathcal{H} \rightarrowtail \mathcal{H}$ (analogue of a Lie idempotent).

A morphism of cofree-coassociative CHAs is a linear map $\varphi : R \to R'$ whose extension $T^c\varphi : T^c(R) \to T^c(R')$ is a morphism of Hopf algebras.

2.2. Multibraces. Let $\mathcal{H} = T^c(R)$ be a cofree-coassociative CHA. Since it is cofree, the multiplication $* : T^c(R) \otimes T^c(R) \to T^c(R)$ is completely determined by its projection onto R. So, there are well-defined maps $M_{pq} : R^{\otimes p} \otimes R^{\otimes q} \to R$, called *multibrace operations*, given by the following composite:

$$\begin{tikzcd} T^c(R) \otimes T^c(R) \arrow[r, "*"] & T^c(R) \arrow[d, "\mathrm{proj}"] \\ R^{\otimes p} \otimes R^{\otimes q} \arrow[u, "\mathrm{inc}"] \arrow[r, "M_{pq}"] & R \end{tikzcd}$$

It is sometimes helpful to write

$$\{x_1, \ldots, x_p; y_1, \ldots, y_q\} := M_{pq}(x_1 \cdots x_p; y_1 \cdots , y_q)$$

One also finds the notation $\{x_1, \ldots, x_p\}\{y_1, \ldots, y_q\}$ in the literature [**19, 1**].

2.3. Proposition. *Let M_{pq} be $(p+q)$-ary operations on R ($p \geq 0, q \geq 0$), and let $*$ be the unique binary operation on $T^c(R)$ which is a coalgebra morphism and* *whose projection onto R coincides with the M_{pq}'s. The following assertions are equivalent:*

(a) the operation $*$ is associative and unital,

(b) the operations M_{pq} satisfy the following relations $\mathcal{R}$:

$$M_{00} = 0, M_{01} = \mathrm{id} = M_{10}, M_{0q} = 0 = M_{p0},\ \textit{for}\ p > 1, q > 1,$$

and, for any integers i, j, k greater than or equal to 1,

$(\mathcal{R}_{ijk})$:

$$\sum_{1\leq l\leq i+j} M_{lk} \circ (M_{i_1 j_1} \cdots M_{i_l j_l}; \mathrm{id}^{\otimes k}) = \sum_{1\leq m\leq j+k} M_{im} \circ (\mathrm{id}^{\otimes i}; M_{j_1 k_1} \cdots M_{j_m k_m})$$

where the left sum is extended to all sets of indices $i_1, \cdots, i_l; j_1, \cdots, j_l$ such that $i_1 + \cdots + i_l = i; j_1 + \cdots + j_l = j$, and the right sum is extended to all sets of indices $j_1, \cdots, j_m; k_1, \cdots, k_m$ such that $j_1 + \cdots + j_m = j; k_1 + \cdots + k_m = k$.

In the formula the concatenation of operations $M_{pq}M_{rs}$ has the following meaning:

$$M_{pq}M_{rs}(x_1 \cdots x_{p+r}; y_1 \cdots y_{q+s}) = M_{pq}(x_1 \cdots x_p; y_1 \cdots y_q) M_{rs}(x_{p+1} \cdots x_{p+r}; y_{q+1} \cdots y_{q+s}).$$

Proof. See [**40**] and Proposition 1.6 of [**42**]. □

2.4. Example. The first nontrivial relation is $(\mathcal{R}_{111})$ which reads

$$M_{21}(uv + vu, w) + M_{11}(M_{11}(u, v), w) = M_{11}(u, M_{11}(v, w)) + M_{12}(u, vw + wv)$$

for any $u, v, w \in R$.

2.5. Multibrace algebra. By definition a *multibrace algebra* (or *MB*-algebra) is a vector space R equipped with multibrace operations M_{pq}, that is $(p + q)$-ary operations $M_{pq} : R^{\otimes p} \otimes R^{\otimes q} \to R$, which satisfy all the conditions $\mathcal{R}$ of Proposition 2.3.

This notion was called $\mathbf{B}_\infty$-algebra in [**42**] and first appeared in the differential graded framework as B_∞-algebra in [**19, 20**]. It appears in [**45**] as *LR*-algebras.

2.6. Theorem. *There is an equivalence of categories between the cofree-coassociative combinatorial Hopf algebras and the multibrace algebras.*

Proof. Taking the primitives gives a functor Prim : *coAs-CHA* $\to$ *MB*-alg. In the other direction the functor *MB*-alg $\to$ *coAs-CHA*, $R \mapsto T^c(R)$, is a consequence of Proposition 2.3. These two functors are inverse to each other. □

2.7. Dipterous structures. A *dipterous algebra* (cf. [**40**]) is a vector space A equipped with two binary operations $*$ and $\succ$ which satisfy the two relations

$$(x * y) * z = x * (y * z), \quad \text{and} \quad (x * y) \succ z = x \succ (y \succ z)$$

for any $x, y, z \in A$. In most examples the associative algebras that we are working with are unital and augmented. Therefore it is helpful to introduce the notion of *unital dipterous algebra.* It is an augmented algebra $A = \mathbb{K}\,1 \oplus \overline{A}$ such that $\overline{A}$ is a dipterous algebra, i.e. equipped with a right operation. We ask that the following relations hold for any $a \in \overline{A}$:

$$a \succ 1 = 0, \quad \text{and} \quad 1 \succ a = a.$$

Observe that $1 \succ 1$ is not defined, but 1 is a unit for $*$.

If A and B are two dipterous algebras, then there is a way to construct a dipterous structure on the tensor product $A\otimes B$ as follows. The associative product is as usual:

$$(a\otimes b)*(a'\otimes b')=(a*a')\otimes(b*b').$$

The right product is given by

$$(a\otimes b)\succ(a'\otimes b')=(a*a')\otimes(b\succ b').$$

This formula makes sense provided that we do not have $b=1$ and $b'=1$ simultaneously. In that case we put

$$(a\otimes 1)\succ(a'\otimes 1)=(a\succ a')\otimes 1\ .$$

Again this formula makes sense provided that we do not have $a=1$ and $a'=1$ simultaneously. But, in that case, both elements are the unit of $A\otimes B$ for which we do not need to define the left product with itself.

It is straightforward to show that, equipped with these two products, $A\otimes B$ becomes a unital dipterous algebra (cf. [**40**]).

By definition a *dipterous bialgebra* is a unital dipterous algebra $(\mathcal{H},*,\succ)$ equipped with a coassociative counital coproduct $\Delta:\mathcal{H}\to\mathcal{H}\otimes\mathcal{H}$ which is a morphism of unital dipterous algebras. This last condition means that the following two compatibility relations hold:

$$\Delta(x*y)=\Delta(x)*\Delta(y)\quad\text{and}\quad\Delta(x\succ y)=\Delta(x)\succ\Delta(y)=\sum x_{(1)}*y_{(1)}\otimes x_{(2)}\succ y_{(2)}.$$

2.8. Proposition. *Any cofree-coassociative CHA has a natural structure of unital dipterous bialgebra.*

Proof. We denote by $*$ the associative product on $T^c(R)$ induced by the *MB*-structure of R. Define a new binary operation on $T^c(R)$ by:

$$(u_1\cdots u_k)\succ(v_1\cdots v_l):=((u_1\cdots u_k)*(v_1\cdots v_{l-1}))\succ v_l,$$

for $u_1,\dots,u_k,v_1,\dots,v_l\in R$ for $l\geq 2$ and for $l=1$:

$$(u_1\cdots u_k)\succ v_1=u_1\cdots u_k v_1.$$

In particular, we get:

$$(((u_1\succ u_2)\succ u_3)\succ\cdots)\succ u_k:=u_1u_2u_3\cdots u_k.$$

It is immediate to verify that the two binary operations $*$ and $\succ$ satisfy the dipterous relation.

The compatibility relation with the deconcatenation Δ is also immediate by induction. □

2.9. Proposition. [**40**] *The free unital dipterous algebra on the vector space V, denoted $Dipt(V)$, is naturally equipped with a dipterous bialgebra structure, and, a fortiori, with a Hopf algebra structure.*

Proof. We follow the general procedure given in [**36**]. Consider the map $V\to Dipt(V)\otimes Dipt(V), v\mapsto v\otimes 1+1\otimes v$. Since $Dipt(V)$ is free, there exists a unique extension of this map as a dipterous morphism

$$\Delta: Dipt(V)\to Dipt(V)\otimes Dipt(V).$$

Using again the universal property of $Dipt(V)$ it is immediate to check that Δ is coassociative. Hence $(Dipt(V),*,\succ,\Delta)$ is a dipterous bialgebra and this structure is functorial in V. □

2.10. Remark. Let $\mathcal{H}$ be an *As*^{c}-*Dipt*-bialgebra which is conilpotent. By results of [**40**] there exists an isomorphism of coassociative coalgebras $T^c(\operatorname{Prim}\mathcal{H}) \cong \mathcal{H}$. Any choice of such an isomorphism makes $\mathcal{H}$ into a cofree-coassociative CHA.

2.11. From dipterous algebras to multibrace algebras. Let $A = (A, \succ, *)$ be a dipterous algebra. For elements $u_1, \ldots, u_p$ in A we define :

$$\omega^{\succ}(u_1 \cdots u_p) := (((u_1 \succ u_2) \succ u_3) \succ \cdots) \succ u_p.$$

We construct $(p+q)$-ary operations on A inductively as follows:

$$M_{00} = 0,\ M_{10} = \mathrm{id}_A = M_{01},\ \text{and}\ M_{n0} = 0 = M_{0n}\ \text{for}\ n \geq 2,$$

and

$$M_{pq} :=\ \left(\omega^{\succ}\right) * \left(\omega^{\succ}\right) - \sum_{k\geq 2} \sum \omega^{\succ}\left(M_{i_1 j_1} M_{i_2 j_2} \cdots M_{i_k j_k}\right)$$

where the second sum (for which $k \geq 2$ is fixed) is extended to all sets of indices $(i_1, \ldots, i_k; j_1, \ldots, j_k)$ such that $i_1 + \cdots + i_k = p$ and $j_1 + \cdots + j_k = q$.

For instance:

$$M_{11}(u; v) = u * v - u \succ v - v \succ u\ ,$$

$$\begin{aligned} M_{21}(uv; w) = {} & (u \succ v) * w - u \succ M_{11}(v; w) - M_{11}(u; w) \succ v \\ & - (u \succ v) \succ w - (u \succ w) \succ v - (w \succ u) \succ v\ , \\ = {} & (u \succ v) * w - (u \succ v) \succ w - u \succ (v * w) + u \succ (v \succ w), \end{aligned}$$

$$\begin{aligned} M_{12}(u; vw) = {} & u * (v \succ w) - M_{11}(u; v) \succ w - v \succ M_{11}(u; w) \\ & - (u \succ v) \succ w - (v \succ u) \succ w - (v \succ w) \succ u\ . \\ = {} & u * (v \succ w) - (u * v) \succ w - v \succ (u * w - u \succ w - w \succ u) - (v \succ w) \succ u. \end{aligned}$$

Let us remark that, if we define $x \prec y := x * y - x \succ y$, then M_{21} becomes:

$$M_{21}(uv; w) = (u \succ v) \prec w - u \succ (v \prec w).$$

2.12. Proposition. *There is a (forgetful) functor F from the category of dipterous algebras to the category of multibrace algebras given by*

$$(A, *, \succ) \mapsto (A, \{M_{pq}\}_{p\geq 0, q\geq 0}).$$

The functor F has a left adjoint denoted U. For a brace algebra R, the dipterous algebra $U(R)$ is the quotient of the free dipterous algebra on the vector space R modulo the relations which identify the MB-structure of R with the MB-structure coming from the dipterous structure. Moreover $U(R)$ is a dipterous bialgebra.

Proof. The proof follows immediately from the definition of a multibrace algebra, using that $\omega^{\succ}(u_1 \cdots u_p) = u_1 \otimes \cdots \otimes u_p$ and that $*$ is associative.

Since F is a forgetful functor, it is immediate that it has a left adjoint, which is precisely U as described in the statement.

Since the free dipterous algebra is a dipterous bialgebra by Proposition 2.9, it follows that its quotient $U(R)$ is also a dipterous bialgebra. $\square$

2.13. THEOREM. *There is a good triple of operads*

$$(As, Dipt, MB).$$

For any vector space V there is a natural isomorphism of CHAs:

$$Dipt(V) \cong T^c(MB(V)).$$

Proof. Let $\mathcal{P} = \mathrm{Prim}_{As} Dipt$ denote the operad made up of the primitive operations for As^c-$Dipt$-bialgebras. The hypothesis (H1), recalled in 1.4, is fullfilled thanks to Proposition 2.9. The hypothesis (H2epi) is also fullfilled because a splitting of the coalgebra map $Dipt(V) \to As(V) = T^c(V)$ is given by

$$v_1 \cdots v_n \mapsto \omega^{\succ}(v_1 \cdots v_n).$$

The coalgebra morphism property is proved by induction on n. Hence, by [**38**], there is a good triple of operads

$$(As, Dipt, \mathcal{P}).$$

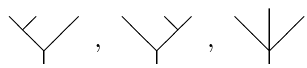
It remains to show that we have $\mathcal{P} = MB$. Since the operations M_{pq} are primitive by construction there is a natural map $MB(V) \to \mathcal{P}(V)$, with extension $T^c(MB(V)) \to T^c(\mathcal{P}(V))$. By Theorem 2.6 $T^c(MB(V))$ is a dipterous algebra, therefore there is a natural dipterous morphism from $Dipt(V)$ to it. Since $(As, Dipt, \mathcal{P})$ is a good triple of operads, the composite

$$Dipt(V) \to T^c(MB(V)) \to T^c(\mathcal{P}(V))$$

is an isomorphism. As a consequence the map $MB(n) \to \mathcal{P}(n)$ on the spaces of n-ary operations is surjective. It suffices to show that they have the same (finite) dimension to deduce that it is an isomorphism.

The dimension of $MB(n)$ is known to be $C_n \times n!$, where C_n is the super-Catalan number (number of planar rooted trees with n leaves), cf. [**42**]. The dimension of $Dipt(n)$ is known to be $2 \times C_n \times n!$, see 2.15. From the functional equation relating the generating series of the two functors $Dipt$ and $\mathcal{P}$ deduced from $Dipt = As \circ \mathcal{P}$ and from the functional equation satisfied by the generating series of the super-Catalan numbers (cf. 2.14) we conclude that $\dim \mathcal{P}(n) = C_n \times n!$. Hence $\dim MB(n) = \dim \mathcal{P}(n)$ and $MB = \mathcal{P}$ as expected. □

2.14. Planar trees. In order to describe the structure of the free dipterous algebra we introduce the combinatorial objects named planar trees. We denote by PT_n the set of *planar trees* with n leaves, $n \geq 1$, which are reduced (every vertex has at least two inputs) and rooted. Here are the first of them:

$PT_1 = \{|\}$, $PT_2 = \{$  $\}$, $PT_3 = \Big\{$  $\Big\}$.

The lowest vertex is called the *root vertex*. Observe that the tree $|$ has no vertex.

We define $PT_\infty := \bigcup_{n \geq 1} PT_n$. The number of elements in PT_n is the so-called *super Catalan number* C_n.

By definition the *grafting* of k planar trees $\{x^{(1)}, \ldots, x^{(k)}\}$ is a planar tree denoted $\bigvee(x^{(1)}, \ldots, x^{(k)})$ obtained by joining the k roots to a new vertex and adding a new root. For $k = 2$, sometimes we shall write $x^{(1)} \vee x^{(2)}$ instead of $\bigvee(x^{(1)}, x^{(2)})$. Observe that the grafting operation is not associative. In fact the three trees $(x \vee y) \vee z, x \vee (y \vee z), \bigvee(x, y, z)$ are all different. Any planar tree x can be uniquely obtained as $x = \bigvee(x^{(1)}, \ldots, x^{(k)})$, where k is the number of inputs of the root vertex.

The grafting operation gives a bijection between the set of k-tuples of planar trees, for any $k \geq 2$, and $PT_\infty \backslash \{|\}$. Indeed, the inverse is the map

$$\bigvee(x^{(1)}, \ldots, x^{(k)}) \mapsto \{x^{(1)}, \ldots, x^{(k)}\}.$$

As a consequence the generating series $C(t) = \sum_{n \geq 0} C_n t^n$ satisfies the equation $2tCf(t)^2 - (1+t)C(t) + 1 = 0$.

2.15. Free dipterous algebra. Since, in the relations defining the notion of a dipterous algebra, the variables stay in the same order, we only need to understand the free dipterous algebra on one generator (i.e. over $\mathbb{K}$). In other words the operad *Dipt* is a nonsymmetric operad (see [**43**]).

Let $T(PT_\infty)$ be the free unital associative algebra over the vector space $\mathbb{K}[PT_\infty]$ spanned by the set PT_∞. So we write $T(PT_\infty)$ instead of $T(\mathbb{K}[PT_\infty])$. A set of linear generators of $T(PT_\infty)$ is made up of the monomials $t_1 t_2 \cdots t_k$ where the t_i's are planar trees. We define a right product on the augmentation ideal as follows:

$$(t_1 t_2 \cdots t_k) \succ (s_1 s_2 \cdots s_l) := \Big(t_1 \vee \big(t_2 \vee (\cdots \vee \bigvee(t_k, s_1, \ldots, s_l) \cdot)\big)\Big).$$

Observe that the right product of two monomials is always a tree. As before we extend the right operation by $1 \succ \omega = \omega$ and $\omega \succ 1 = 0$ for ω in the augmentation ideal.

2.16. Proposition. *The associative algebra $T(PT_\infty)$ equipped with the right product defined above is the free unital dipterous algebra on one generator.*

Proof. Let us verify the dipterous relation. On one hand we have

$$\begin{aligned}((t_1 \cdots t_k) * (s_1 \cdots s_l)) \succ (r_1 \cdots r_m) &= (t_1 \cdots t_k s_1 \cdots s_l) \succ (r_1 \cdots r_m) \\ &= \Big(t_1 \vee \big(\cdots \vee (t_k \bigvee(s_1, \ldots, s_{l-1}, \bigvee(s_l, r_1, \ldots, r_m) \cdot)) \cdot\big)\Big)\end{aligned}$$

On the other hand we have

$$\begin{aligned}(t_1 \cdots t_k) \succ \big((s_1 \cdots s_l) \succ (r_1 \cdots r_m)\big) &= (t_1 \cdots t_k) \succ \Big(s_1 \vee (\cdots \vee (s_l \vee r_1 \cdots \vee r_m) \cdot)\Big) \\ &= \Big(t_1 \vee \big(\cdots \vee (t_k \bigvee(s_1, \ldots, s_{l-1}, \bigvee(s_l, r_1, \ldots, r_m) \cdot)) \cdot\big)\Big) .\end{aligned}$$

And so $(T(PT_\infty, *, \succ)$ is a dipterous algebra. Since any linear generator can be obtained from the tree $|$ by combining the two operations $*$ and $\succ$, it is a quotient of $Dipt(\mathbb{K})$. In order to show that it is free on one generator it is sufficient to show that, for any dipterous algebra A and any element $a \in A$, there is a unique dipterous morphism $T(PT_\infty) \to A$ mapping $|$ to a. We already know how to construct a map and we know that it is unique. To prove that it is a dipterous morphism we proceed by induction (on the degree of the trees) like in the proof of the dendriform case performed in [**35**] Proposition 5.7. $\square$

2.17. Remark. As in the case of the operad $2as$ governing the 2-associative algebras (cf. [**42**]) the operad *Dipt* can be described in terms of two copies of the set PT_∞.

2.18. Free-associative CHA and co-multibraces. A *free-associative combinatorial Hopf algebra* $\mathcal{H} = \mathbb{K}1 \oplus \overline{\mathcal{H}}$, free-associative CHA for short, is a Hopf algebra structure on the tensor algebra $T(C)$. Here C is the space of indecomposables of $\mathcal{H}$, that is, $C := \overline{\mathcal{H}}/\overline{\mathcal{H}}^2$. This structure is encoded into the coproduct map Δ. The preceding results admit an obvious dualization. So, C is a *multibrace coalgebra* and this comultibrace structure determines completely the Hopf algebra $(T(C), *, \Delta)$. The space $T(C)$ inherits a *dipterous coalgebra* structure which makes it into a $Dipt^c$-As-bialgebra.

For any vector space V there is an isomorphism of CHAs

$$T(MB^c(V)) \cong Dipt^c(V).$$

These spaces are spanned by trees and the (co)operations can be made explicit through "admissible cuts" (the dual notion of grafting).

2.19. 2-associative bialgebras. A 2-associative algebra A (or $2as$-algebra for short) is a vector space equipped with two associative products. Here we assume that they are unital, that is, A contains an element 1 which is a unit for the two associative products $\cdot$ and $*$. In [**42**] we introduced the notion of 2-associative bialgebra (more precisely As^c-$2as$-bialgebra) which is a $2as$-algebra equipped with a coassociative (and counital) coproduct satisfying the Hopf compatibility relation with $*$ and the unital infinitesimal compatibility relation with $\cdot$ (see loc.cit. for details).

Let $\mathcal{H} = (T^c(R), *)$ be a cofree-coassociative CHA. In [**42**] we have shown that, if we consider the concatenation product $\cdot$ on the tensor module $T^c(R)$, then $(T^c(R), *, \cdot, \Delta)$ is a 2-associative bialgebra. The same results as before are valid with dipterous replaced by 2-associative. They are the subject of [**42**] where it is proved that

$$(As, 2as, MB)$$

is a good triple of operads.

Let us recall from loc. cit. that there is a functor

$$F : 2as\text{-alg} \to MB\text{-alg}$$

which provides operations M_{pq} out of the two associative operations $\cdot$ and $*$. For instance:

$$M_{11}(x; y) = x * y - x \cdot y - y \cdot x\ ,$$
$$M_{21}(xy; z) = (x \cdot y) * z - x \cdot (y * z) - (x * z) \cdot y + x \cdot y \cdot z,$$
$$M_{12}(x; yz) = x * (y \cdot z) - (x * y) \cdot z - y \cdot (x * z) + y \cdot x \cdot z.$$

The choice of 2-associative bialgebras behaves well with the symmetrization procedure that we deal with in Section 4. The choice of dipterous bialgebras behaves well with the right-sided hypothesis, dealt with in Section 3.

2.20. From dipterous to 2-associative. Let λ be a formal parameter (or an element in $\mathbb{K}$). We consider algebras having two binary operations $*$ and $\succ$ such that $*$ is associative and the following relation holds:

$$(\lambda x * y + (1 - \lambda)x \succ y) \succ z = x \succ (y \succ z).$$

Let us denote by $\mathcal{Z}_\lambda$ the associated operad. It is clear that $\mathcal{Z}_0 = 2as$ and $\mathcal{Z}_1 = Dipt$. So we have a homotopy between the operads $2as$ and $Dipt$. It would be interesting to know if we always have a good triple of operads

$$(As, \mathcal{Z}_\lambda, MB)$$

such that any cofree-coassociative CHA is a $\mathcal{Z}_\lambda$-bialgebra.

3. (Co)free-(co)associative right-sided CHA

In this section we study the cofree-coassociative CHAs (and then free-associative CHAs) which satisfy the right-sided property (r-s). We will see that they are dendriform bialgebras which have been thoroughly studied in [**35, 39, 52, 53, 54**]. The space of primitives inherits a brace algebra structure. In the last part of the section we study the alternative structure, denoted $\mathcal{Y}$, which is close to the 2-associative structure. Examples will be given in section 6. They include the Hopf algebra of quasi-symmetric functions (6.2), the Malvenuto-Reutenauer algebra (6.4), the Solomon-Tits algebra (6.2), the algebra of parking functions (6.2) and Hopf algebras constructed out of operads (6.7).

3.1. Right-sided condition on combinatorial Hopf algebras. Let $\mathcal{H} = (T^c(R), *)$ be a cofree-coassociative CHA. There is a natural grading $\mathcal{H} = \oplus_n \mathcal{H}^n$ given by $\mathcal{H}^n := R^{\otimes n}$. We study the cofree-coassociative CHAs satisfying the following condition.

Right-sided condition:
(r-s) for any integer q the subspace $\bigoplus_{n\geq q} \mathcal{H}^n$ is a right-sided ideal of $\mathcal{H}$.

Since $\mathcal{H}$ is a cofree-coassociative CHA, by Proposition 2.3 the associative product $*$ is given by a family $\{M_{pq}\}_{p,q\geq 0}$ of $(p+q)$-ary operations, satisfying the relations $\mathcal{R}$.

3.2. Proposition. *The combinatorial Hopf algebra $\mathcal{H}$ is right-sided if and only if $M_{pq} = 0$ for $p \geq 2$.*

Proof. Recall that $\mathcal{H} = T^c(V)$ is right-sided if the subspace $\mathbb{F}_r = \bigoplus_{n\geq r} V^{\otimes n}$ is a right ideal of $T^c(V)$.

If $\mathbb{F}_r$ is a right-sided ideal, then

$$(x_1 \cdots x_p) * (y_1 \cdots y_q) = M_{pq}(x_1 \cdots x_p; y_1 \cdots y_q) + z,$$

where $z \in \mathbb{F}_2$. But, for $p \geq 2$, $(x_1 \cdots x_p) * (y_1 \cdots y_q)$ must belong to $\mathbb{F}_p$ and $M_{pq}(x_1 \cdots x_p; y_1 \cdots y_q) \in \mathbb{F}_1 \setminus \mathbb{F}_2$, which implies that $M_{pq} = 0$ for all q and $p \geq 2$.

Conversely, suppose that $M_{pq} = 0$ for all m and $n \geq 2$. The formula for $*$ is then given by:

$$(x_1 \cdots x_p) * (y_1 \cdots y_q) =$$
$$\sum_r \Big(\sum_{\substack{\underline{k}\vdash q \\ |k|=p+2}} y_1 \cdots y_{k_0} M_{1k_1}(x_1; y_{k_0+1}, \cdots, y_{k_0+j_1}) \cdots y_{k_1} M_{1j_2}(x_2; \cdots) \cdots y_{k_1+j_2}) \cdots$$
$$M_{1j_p}(x_p; \cdots) \cdots y_q)\Big),$$

where $0 \leq k_0 \leq k_0 + j_1 \leq \cdots \leq k_{p-1} + j_p \leq q$. So, the element $(x_1 \cdots x_p) * (y_1 \cdots y_q)$ belongs to $\mathbb{F}_p$, which ends the proof. $\square$

3.3. Brace algebras. A *brace algebra* is a vector space R equipped with a $(1+q)$-ary operation denoted $\{-;-,\dots,-\}$ for any $q\geq 1$ satisfying the following formulas:

$$\{\{x;y_1,\dots,y_n\};z_1,\dots,z_m\}=\sum\{x;\dots,\{y_1;\dots\},\dots,\{y_n;\dots,\},\dots\}.$$

On the right-hand side the dots are filled with the variables z_i's (in order) with the convention $\{y_k;\emptyset\}=y_k$. This notion appears in the work of Gerstenhaber and Voronov in [**19**].

3.4. LEMMA. *A multibrace algebra for which $M_{pq}=0$ when $p\geq 2$ is a brace algebra and vice-versa.*

Proof. It is straightforward to check that, under the identification

$$M_{1q}(x;y_1\cdots y_q)=\{x;y_1,\dots,y_q\},$$

the relations $\mathcal{R}_{ijk}$ are exactly the brace relations. □

3.5. THEOREM. *The primitive part of a cofree-coassociative r-s CHA is a brace algebra. There is an equivalence of categories between cofree-coassociative r-s CHAs and brace algebras.*

Proof. By Proposition 2.3 the primitive part of a cofree bialgebra is a MB-algebra. By Lemma 3.2 property (r-s) implies that $M_{pq}=0$ when $p\geq 2$. Therefore, by Lemma 3.4 this primitive part is a brace algebra. The two functors $R\mapsto T^c(R)$ and $\mathcal{H}=T^c(R)\mapsto R$ are obviously inverse to each other. □

3.6. Dendriform algebras [35]. A *dendriform algebra* is a vector space A equipped with two binary operations denoted $\prec$ and $\succ$ which satisfy the conditions:

$$\begin{cases} (x\prec y)\prec z &=& x\prec(y*z),\\ (x\succ y)\prec z &=& x\succ(y\prec z),\\ (x*y)\succ z &=& x\succ(y\succ z).\end{cases}$$

where $x*y:=x\prec y+x\succ y$.

As in the case of dipterous algebras, there is a notion of *unital dendriform algebra*. It is a vector space $A=\mathbb{K}\,1\oplus\overline{A}$ such that $\overline{A}$ is a dendriform algebra and the two binary operations are extended as follows:

$$\begin{cases} 1\prec x &=& 0 &,& x\prec 1 &=& x,\\ 1\succ x &=& x &,& x\succ 1 &=& 0.\end{cases}$$

Observe that one has $1*x=x=x*1$ as expected, but $1\prec 1$ and $1\succ 1$ are *not* defined.

3.7. PROPOSITION. *Any cofree-coassociative CHA which is right-sided is a dendriform bialgebra.*

Proof. By Proposition 2.8 the Hopf algebra $\mathcal{H}$ is a dipterous algebra. Let us define the operation $\prec$ by the equality $x*y=x\prec y+x\succ y$.

Condition (r-s) implies that the operations $\prec$ and $\succ$ satisfy the relation

$$(x\succ y)\prec z=x\succ(y\prec z).$$

It is clear that the axioms for a dipterous algebra satisfying this extra condition are equivalent to the axioms for dendriform algebras. □

3.8. From dendriform algebras to brace algebras. The forgetful functor from dipterous algebras to multibrace algebras sends a dendriform algebra to a brace algebra with the same underlying space. In fact its restriction is precisely the functor constructed in [**52**]. Explicitly it is given by

$$M_{1q}(v; v_1, \ldots, v_q) = \sum_{i=0}^{i=q} (-1)^i \omega_{\prec}(v_1 \cdots v_i) \succ v \prec \omega^{\succ}(v_{i+1} \cdots v_q),$$

where

$$\omega_{\prec}(v_1 \cdots v_n) := v_1 \prec (v_2 \prec (\cdots \prec (v_{n-1} \prec v_n)))$$

and where

$$\omega^{\succ}(v_1 \cdots v_n) := ((v_1 \succ v_2) \succ \cdots) \succ v_n,$$

is the element defined in Subsection 2.9. The proofs of the following results are similar to those of section 2 and can be found in [**52, 53, 54**].

3.9. Theorem. *There is a good triple of operads*

$$(As, Dend, Brace).$$

For any vector space V there is a natural isomorphism of cofree dendriform bialgebras, and, a fortiori, of cofree-coassociative right-sided CHAs:

$$Dend(V) \cong T^c(Brace(V)).$$

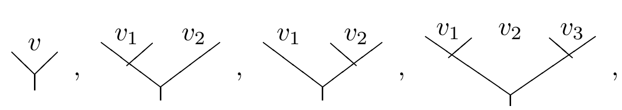
3.10. Planar binary trees and free dendriform algebra. The set of planar binary trees with n leaves, denoted PBT_n, is made up of the planar trees (cf. 2.14) with exactly two inputs at each vertex. The number of elements of PBT_n is the Catalan number c_n. For any $n \geq 1$ a tree $t \in PBT_n$ can be uniquely written as a grafting $t = t^l \vee t^r$. It is shown in [**35**] that the free dendriform algebra on one generator $Dend(\mathbb{K})$ is the Hopf algebra $\mathcal{H}_{LR} := \oplus_{n \geq 1} \mathbb{K}[PBT_n]$ equipped with the following right and left product:

$$t \prec s := t^l \vee (t^r * s), \quad t \succ s := (t * s^l) \vee s^r.$$

It was first introduced in our paper [**39**]. Notice that the tree $|$ is the unit and the tree $y :=$  is the generator. In order to compute in $\mathcal{H}_{LR}$ the following formula proves to be helpful:

$$t \succ y \prec s = t \vee s.$$

In order to work with the free dendriform algebra on a vector space V (resp. on a set X) it suffices to decorate the trees by putting elements of V (resp. X) in between the leaves. For instance:



represent respectively

$$v, \quad v_1 \succ v_2, \quad v_1 \prec v_2, \quad v_1 \succ v_2 \prec v_3.$$

3.11. Planar binary trees and the primitives. The primitive operations can be described by planar binary trees as follows. The following results, stated without proofs, are taken from [**54**].

The dendriform analog of the Eulerian idempotent acting on the dendriform bialgebra $Dend(\mathbb{K})$ is defined by the formula

$$e := \sum_{n\geq 1}(-1)^{n+1}\omega_{\succ}^{n}\circ\overline{\Delta}^{n},$$

where $\overline{\Delta}$ is the reduced coproduct of $Dend(\mathbb{K})$, and where

$$\omega_{\succ}^{n}(x_1,\dots,x_n) := x_1\succ(x_2\succ(\cdots\succ(x_{n-1}\succ x_n))),\ \text{for}\ x_i\in Dend(\mathbb{K}).$$

This operator has the following properties:

(1) the element $e(x)$ is a primitive element, for any $x\in Dend(\mathbb{K})$,
(2) $e(x\succ y)=0$, for any pair of elements x and y of $Dend(\mathbb{K})$,
(3) $e(x)=x$ for any $x\in \mathrm{Prim}(Dend(\mathbb{K}))$.

Moreover, if $t = |\vee x$ is a tree in PBT_n, with $x\in PBT_{n-1}$, then $e(t) = t+\sum_i t_1^i\vee t_2^i$, for some $t_1^i\in PBT_{n_i}$ with $n_i\geq 2$.

An easy argument on the dimensions of $Dend(\mathbb{K})_n$ and $T(\bigoplus_{m\geq 0}\mathbb{K}[PBT_m])_n$ shows that the elements $e(|\vee t)$, with $t\in PBT_n$, form a basis of $\mathrm{Prim}(Dend(\mathbb{K}))_n$, for all $n\geq 1$. From these results it follows that any element of $Dend(\mathbb{K})$ is a sum of elements of type $((x_1\succ x_2)\succ\dots)\succ x_r$, with $x_i\in\mathrm{Prim}(Dend(\mathbb{K}))$ for $1\leq i\leq r$.

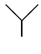
Moreover, the coproduct Δ of $Dend(\mathbb{K})$, satisfies that:

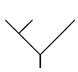
$$\Delta(((x_1\succ x_2)\succ\cdots)\succ x_r) = \sum_{i=0}^{r}((x_1\succ x_2)\succ\cdots)\succ x_i\otimes((x_{i+1}\succ x_{i+2})\succ\cdots)\succ x_r,$$

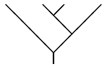
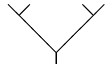
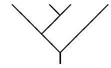
for any $x_1,\dots,x_r\in\mathrm{Prim}(Dend(\mathbb{K}))$. Therefore there is a bijection

$$\mathbb{K}[PBT_{n-1}\times S_n]\cong(\mathrm{Prim}_{As}Dend)(n)$$

given by $(t;x_1,\dots,x_n)\mapsto(e(|\vee t);x_1,\dots,x_n)$. For instance, in the free dendriform algebra $Dend(V)$ over the vector space V we get the following formulas in low dimension for any $v_i\in V$:

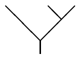
$(|;v_1)\mapsto v_1,\quad ($  $;v_1v_2)\mapsto v_1\prec v_2 - v_2\succ v_1 = M_{11}(v_1;v_2),$

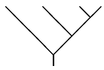
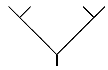
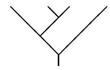
$\Big($  $;v_1v_2v_3\Big)\mapsto$

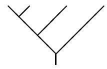
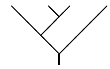
$\Big($  $;v_1v_2v_3\Big)-\Big($  $;v_2v_1v_3\Big)+\Big($  $;v_2v_3v_1\Big)$

$= v_1\prec(v_2\succ v_3) - v_2\succ v_1\prec v_3) + (v_2\prec v_3)\succ v_1)$

$= M_{12}(v_1;v_2v_3),$

( $; v_1v_2v_3\Big) \mapsto$

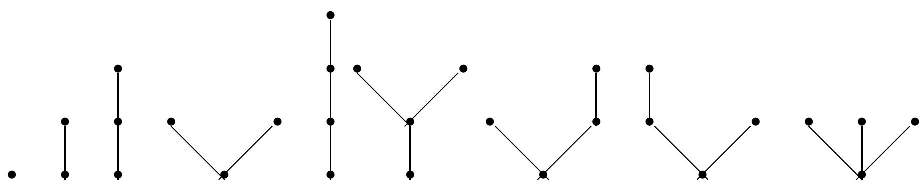
( $; v_1v_2v_3\Big) - \Big($  $; v_3v_1v_2\Big) - \Big($  $; v_2v_3v_1\Big)$

$+\Big($  $; v_3v_2v_1\Big) + \Big($  $; v_3v_2v_1\Big)$

$$= v_1 \prec (v_2 \prec v_3) - v_3 \succ v_1 \prec v_2 - (v_2 \prec v_3) \succ v_1 + (v_3 \succ v_2) \succ v_1 + (v_3 \prec v_2) \succ v_1$$
$$= M_{11}(M_{11}(v_1; v_2); v_3) - M_{12}(v_1; v_2v_3).$$

Various basis of $Dend(\mathbb{K})$ and comparison between them have been studied in [**2, 18, 27**].

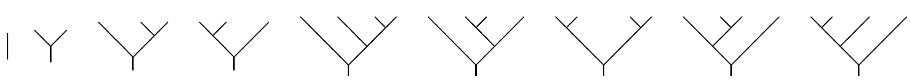
3.12. Unreduced planar trees. We denote the set of *planar unreduced trees* with n vertices by PUT_n. In low dimensions we have, for $n = 1$ to 4,



The number of elements of PUT_n is also the Catalan number c_n. Let us define a grafting operation on PUT_∞ as follows. The unreduced planar tree $x \vee y$ is obtained by putting the tree y on the right-hand side of x and joining the root of x to the root of y by a new edge. The root-vertex of $x \vee y$ is taken to be the root-vertex of x.

3.13. Lemma. [P. Palacios, thesis, unpublished]. *The map $\varphi : PBT_n \to PUT_n$ given by $\varphi(t \vee s) = \varphi(t) \vee \varphi(s)$ and $\varphi(|) = \bullet$ is a bijection.*

The following list of planar binary trees gives under φ the list of planar unreduced trees cited above:



3.14. Labelled planar unreduced trees and the operad *Brace*. We recall the explicit description of the operad *Brace* in terms of planar unreduced trees (see for instance [**10**]). We first construct the S_n-module $Brace(n)$ and then we give the formula for the partial composition operation $\circ_i$.

The S_n-module $Brace(n)$ is the following regular representation:

$$Brace(n) = \mathbb{K}[PUT_n \times S_n].$$

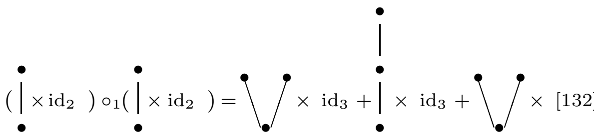
The generating operation M_{1n} corresponds to $c_n \times \mathrm{id}_n$ where c_n is the corolla. In order to describe $\mu \circ_i \nu$ for $\mu \in Brace(m)$ and $\nu \in Brace(n)$ it suffices to describe it for $\mu = t \times \mathrm{id}_m$ and $\nu = r \times \mathrm{id}_n$ (t and r are trees). The composite

$$(t \times \mathrm{id}_m) \circ_1 (r \times \mathrm{id}_n)$$

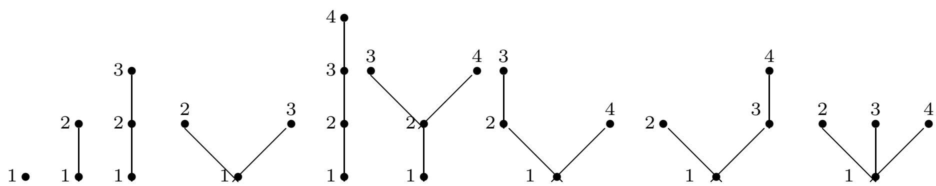
is obtained through the brace relation 3.3. For instance, letting $n = m = 2$, the relation

$$\{\{x_1; x_2\}; x_3\} = \{x_1; x_2, x_3\} + \{x_1; \{x_2; x_3\}\} + \{x_1; x_3; x_2\}$$

gives:



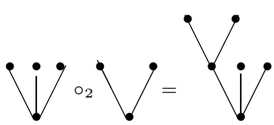
Observe that the action of the symmetric group is involved in this formula. In order to describe the $\circ_i$ composition operation we label the vertices of a planar unreduced tree as follows. The root-vertex is labelled by 1. The other vertices are labelled according to the following rule: for $t = x \vee y$ the integers labelling the vertices of x are less than the the integers labelling the vertices of y. Here are the first of them:



For $i = 2, \ldots, m$ we have

$$(t \times \mathrm{id}_m) \circ_i (r \times \mathrm{id}_n) = (t \circ_i r \times \mathrm{id}_{m+n-1})$$

where the $t \circ_i r$ is obtained by grafting the tree r to the tree t by adjoining an edge from the ith vertex of t to the root of r. For instance:



3.15. Comparison of $\mathrm{Prim}\,_{As}Dend$ **and** *Brace***.** Putting together the results of the preceding paragraphs we see that there is a bijection

$$\mathbb{K}[PBT_n \times S_n] \xrightarrow{\varphi} \mathbb{K}[PUT_n \times S_n] \cong Brace(n) = \mathrm{Prim}\,_{As}Dend(n) \cong \mathbb{K}[PBT_n \times S_n],$$

where the last isomorphism is the inverse of $t \mapsto e(| \vee t)$, cf. 3.11. This composite is not the identity.

3.16. Free-associative right-sided CHA. In the dual framework, that is, for an associative CHA $\mathcal{H} = T(C)$, where C is a multibrace coalgebra, the right-sided condition reads as follows:
(r-s) the coproduct Δ on $T(C)$, given by $\Delta(v) = v \otimes 1 + 1 \otimes v + \sum v_{(1)} \otimes v_{(2)}$, is such that

$$v_{(2)} \in C \text{ for any } v \in C.$$

This condition implies that C is in fact a brace coalgebra and that $\mathcal{H}$ is a dendriform coalgebra.

All the results of this section can be dualized: C is a brace coalgebra and $T(C)$ is a $Dend^c$-As-bialgebra. For instance the free-associative CHA $Dend^c(\mathbb{K})$ can be seen as a noncommutative variation of the Connes-Kreimer algebra. The coproduct can be described on planar binary trees and hence on planar unreduced trees by means of admissible cuts. Some of these results have been worked out explicitly in [**18**].

3.17. $\mathcal{Y}$-algebras. If, instead of looking at the dipterous structure of a cofree-coassociative **r-s** CHA, we look its 2-associative structure (cf. [**42**] and 2.19), then the relevant quotient structure is more complicated: it is a $\mathcal{Y}$-algebra, where $\mathcal{Y}$ is the quotient of the operad $2as$ by the operadic ideal generated by the operations M_{pq} for $p \geq 2$:

$$\mathcal{Y} := 2as/\{M_{pq} \mid p \geq 2\}.$$

Since the operations M_{pq} are primitive for $2as$-bialgebras, the notion of $\mathcal{Y}$-bialgebra is well-defined. Let $\mathcal{H}$ be a cofree-coassociative CHA which is right-sided. By Theorem 4.2 of [**42**] it is a $2as$-bialgebra. By Proposition 3.2 the operations M_{pq} are 0 on $\mathcal{H}$ for $p \geq 2$. Therefore $\mathcal{H}$ is a $\mathcal{Y}$-algebra, and even a $\mathcal{Y}$-bialgebra.

3.18. THEOREM. *There is a good triple of operads*

$$(As, \mathcal{Y}, Brace).$$

Proof. Since the operations M_{pq} are primitive, by Proposition 3.1.1 of [**38**] there is a good triple of operads

$$(As, \mathcal{Y}, \mathrm{Prim}_{As}\mathcal{Y}).$$

Let us consider the free $\mathcal{Y}$-algebra over V, denoted $\mathcal{Y}(V)$. Its primitive part is $\mathrm{Prim}\,\mathcal{Y}(V) = (\mathrm{Prim}_{As}\mathcal{Y})(V)$. It is a multibrace algebra for which $M_{pq} = 0$ for $p \geq 2$, therefore it is a brace algebra. Since these structures are functorial in V we get a surjective map of operads

$$Brace \twoheadrightarrow \mathrm{Prim}_{As}\mathcal{Y}.$$

In order to show that this is an isomorphism, it suffices to show that, for any n, the spaces $Brace(n)$ and $(\mathrm{Prim}_{As}\mathcal{Y})(n)$ have the same dimension. Since we have isomorphisms $Dend \cong As^c \circ Brace$ and $\mathcal{Y} \cong As^c \circ \mathrm{Prim}_{As}\mathcal{Y}$, it suffices to show that $\dim Dend(n) = \dim \mathcal{Y}(n)$. We claim that the isomorphism of CHAs $Dipt(V) \cong T^c(MB(V)) \cong 2as(V)$ induces an isomorphism $Dend(V) \cong \mathcal{Y}(V)$ which implies the expected equality. This last isomorphism is a consequence of the following Proposition. $\square$

3.19. PROPOSITION. *The quotient of the operad Dipt by the relations $M_{pq} = 0$ for $p \geq 2$ is the operad Dend.*

Proof. We know that the operad $Dend$ is the quotient of $Dipt$ by the relation $M_{21} = 0$ (cf. last line of section 2.11). Therefore it suffices to show that, in $Dipt$, $M_{21} = 0$ implies $M_{pq} = 0$ for any $p \geq 2$. This is an immediate consequence of the following Lemma. $\square$

3.20. LEMMA. *In the free dipterous algebra $Dipt(V)$ the following relations hold:*

$$M_{n1}(x_1 \cdots x_n; y) = M_{(n-1)1}((x_1 \succ x_2)x_3 \cdots x_n; y) - x_1 \succ M_{(n-1)1}(x_2 x_3 \cdots x_n; y)$$

for $n \geq 3$,

$$M_{nm}(x_1 \cdots x_n; y_1 \cdots y_m) = M_{n(m-1)}(x_1 \cdots x_n; (y_1 \succ y_2)y_3 \cdots y_m) \\ - y_1 \succ M_{n(m-1)}(x_1 \cdots x_n; y_2)y_3 \cdots y_m)$$

for $m > 1$.

Proof. We prove the first equality applying a recursive argument. The result may be checked for $n = 3$ by a straightforward calculation.
Let $n > 3$ and suppose that the equality holds for all M_{h1}, with $h < n$. We get:

$$(1)\ \omega^{\succ}(x_1 \cdots x_n) * y = \sum_{r=0}^{n} \Big(\sum_{j=0}^{n-r} \omega^{\succ}(x_1 \cdots x_r M_{j1}(\cdots x_{r+j}; y) \cdots x_n) \Big).$$

$$(2)\ \omega^{\succ}(x_1 \cdots x_n) * y = \\ \omega^{\succ}((x_1 \succ x_2) \cdots x_n) * y = \\ \sum_{r=2}^{n} \Big(\sum_{j=0}^{n-r} \omega^{\succ}(x_1 \cdots x_r M_{j1}(\cdots x_{r+j}; y) \cdots x_n) \Big) + \\ \sum_{j=2}^{n} \omega^{\succ}(M_{(j-1)1}((x_1 \succ x_2) \cdots x_j; y) \cdots x_n) + \omega^{\succ}(y(x_1 \succ x_2) \cdots x_n).$$

Equalities (1) and (2) imply:

$$(3)\ \sum_{j=0}^{n} \omega^{\succ}(M_{j1}(x_1 \cdots x_j; y) \cdots x_n) + \sum_{j=0}^{n-1} \omega^{\succ}(x_1 M_{j1}(\cdots x_{j+1}; y) \cdots x_n) = \\ \sum_{j=2}^{n} \omega^{\succ}(M_{(j-1)1}((x_1 \succ x_2) \cdots x_j; y) \cdots x_n) + \omega^{\succ}((y * x_1)x_2 \cdots x_n).$$

The recursive hypothesis states that:

$$M_{j1}(x_1 \cdots x_j; y) + x_1 \succ M_{(j-1)1}(x_2 \cdots x_j; y) = M_{(j-1)1}((x_1 \succ x_2)x_3 \cdots x_j; y),$$

for $3 \leq j \leq n - 1$.

Applying it, (3) becomes:

$$M_{n1}(x_1 \cdots x_n; y) + x_1 \succ M_{(n-1)1}(x_2 \cdots x_n; y) + \omega^{\succ}(yx_1 \cdots x_n) + \\ \omega^{\succ}(M_{11}(y; x_1)x_2 \cdots x_n) + \omega^{\succ}(M_{11}(x_1 \succ x_2; y)x_3 \cdots x_n) + \omega^{\succ}(x_1 y x_2 \cdots x_n) = \\ \omega^{\succ}(M_{11}(x_1 \succ x_2; y)x_3 \cdots x_n) + M_{(n-1)1}((x_1 \succ x_2)x_3 \cdots x_n; y) + \omega^{\succ}((y * x_1)x_2 \cdots x_n).$$

But, since $M_{11}(y; x_1) = y * x_1 - y \succ x_1 - x_1 \succ y$, we get:

$$M_{n1}(x_1 \cdots x_n; y) + x_1 \succ M_{(n-1)1}(x_2 \cdots x_n; y) = M_{(n-1)1}((x_1 \succ x_2)x_3 \cdots x_n; y),$$

which ends the proof of the first equality.

For the second statement, the result is easy to verify for all $n \geq 0$ and $m = 2$. For $m > 2$, suppose that the equality holds for all $1 \leq k \leq n$ and $1 \leq h < m$.

We get:

$$(4)\ \omega^{\succ}(x_1\cdots x_n)*\omega^{\succ}(y_1\cdots y_m)=\omega^{\succ}(x_1\cdots x_n)*\omega^{\succ}((y_1\succ y_2)y_3\cdots y_m)=$$
$$\sum\omega^{\succ}(x_1\cdots x_hM_{j(k-1)}(\cdots x_{h+j};(y_1\succ y_2)\cdots y_k)M_{r_1s_1}\cdots M_{r_ls_l}(\cdots x_n\cdots y_m))$$
$$+M_{n(m-1)}(x_1\cdots x_n;(y_1\succ y_2)\cdots y_m),$$

where the sum is taken over $0\le h\le n, 2\le k\le m, 0\le j\le n-h, j<n$ and the compositions $(r_1,\dots,r_l)$ of $n-h-j$, and $(s_1,\dots,s_l)$ of $m-k$. Note that if $j=0$, then $k=2$.

Applying the formula recursively, we get:

(1) if $h>0$, then

$$\omega^{\succ}(x_1\cdots x_hM_{j(k-1)}(\cdots x_{h+j};(y_1\succ y_2)\cdots y_k))=$$
$$\omega^{\succ}(x_1\cdots x_hM_{jk}(\cdots x_{h+j};y_1\cdots y_k))+$$
$$\omega^{\succ}((\omega^{\succ}(x_1\cdots x_h)*y_1)M_{j(k-1)}(\cdots x_{h+j};y_2\cdots y_k))$$
$$=\sum_{0\le l\le s\le h}\omega^{\succ}(x_1\cdots x_lM_{(s-l)1}(\cdots x_s;y_1)\cdots x_hM_{j(k-1)}(\cdots x_{h+j};y_2\cdots y_k)).$$

(2) if $h=0$, then

$$M_{j(k-1)}(x_1\cdots x_j;(y_1\succ y_2)\cdots y_k)=$$
$$M_{jk}(x_1\cdots x_j;y_1\cdots y_k)+y_1\succ M_{j(k-1)}(x_1\cdots x_j;y_2\cdots y_k).$$

The formulas above imply the equality:

$$\sum\omega^{\succ}(x_1\cdots x_hM_{j(k-1)}(\cdots x_{h+j};(y_1\succ y_2)\cdots y_k)M_{r_1s_1}\cdots M_{r_ls_l}(\cdots x_n\cdots y_m))=$$
$$\sum\omega^{\succ}(M_{i_1j_1}\cdots M_{i_rj_r})(x_1\cdots x_ny_1\cdots y_m)-y_1\succ M_{n(m-1)}(x_1\cdots x_n;y_2\cdots y_m).$$

where the first sum is taken over $0\le h<n, 2\le k\le m, 0\le j\le n-h$ and the compositions $(r_1,\dots,r_l),(s_1,\dots,s_l)$, and the second sum is taken over $k\ge 2, i_1+\cdots+i_r=n, j_1+\cdots+j_r=m$. Replacing in (4), we get the equality:

$$\sum_{k\ge 2}\Big(\sum_{\substack{i_1+\cdots+i_r=n\\ j_1+\cdots+j_r=m}}\omega^{\succ}(M_{i_1j_1}\cdots M_{i_rj_r})(x_1\cdots x_ny_1\cdots y_m)+M_{nm}(x_1\cdots x_n;y_1\cdots y_m)=$$
$$\omega^{\succ}(x_1\cdots x_n)*\omega^{\succ}(y_1\cdots y_m)=$$
$$\sum_{k\ge 2}\Big(\sum_{\substack{i_1+\cdots+i_r=n\\ j_1+\cdots+j_r=m}}\omega^{\succ}(M_{i_1j_1})\cdots M_{i_rj_r})(x_1\cdots x_ny_1\cdots y_m)-$$
$$y_1\succ M_{n(m-1)}(x_1\cdots x_n;y_2\cdots y_m)+M_{n(m-1)}(x_1\cdots x_n;(y_1\succ y_2)\cdots y_m),$$

which implies:

$$M_{nm}(x_1\cdots x_n;y_1\cdots y_m)=$$
$$y_1\succ M_{n(m-1)}(x_1\cdots x_n;y_2\cdots y_m)+M_{n(m-1)}(x_1\cdots x_n;(y_1\succ y_2)\cdots y_m),$$

as expected. □

3.21. On the operad $\mathcal{Y}$. As a corollary of Theorem 3.18 we deduce an isomorphism of S-modules:

$$\mathcal{Y} \cong T^c(Brace).$$

Hence $\mathcal{Y}(n)$ has the same dimension as $Dend(n)$, that is,

$$\dim \mathcal{Y}(n) = c_n \times n!.$$

4. (Co)free-(co)commutative CHA

In this section we study the cofree-cocommutative CHAs. We will see that they are $ComAs$-bialgebras. The space of primitives inherits a symmetric multibrace (SMB) algebra structure. We omit the proofs, which are completely analogous to the non-symmetric case. In this section $\mathbb{K}$ is a characteristic zero field.

4.1. Definition. A *cofree-cocommutative combinatorial Hopf algebra $\mathcal{H}$* (cofree-cocommutative CHA for short) is a cofree-cocommutative bialgebra with a prescribed isomorphism with $S^c(R)$ (here $R := \mathrm{Prim}\,\mathcal{H}$). Equivalently, it is a vector space R with a product $*$ on $S^c(R)$ which makes $(S^c(R), *, \Delta)$ (Δ = deconcatenation) into a Hopf algebra.

4.2. Symmetric multibraces. Let $\mathcal{H} = S^c(R)$ be a cofree-coassociative CHA. Since it is cofree, the multiplication $* : S^c(R) \otimes S^c(R) \to S^c(R)$ is completely determined by its projection onto R. So, there are well-defined maps $M_{pq} : S^p R \otimes S^q R \to R$, called *symmetric multibrace operations*, given by the following composite:

$$\begin{tikzcd} S^c(R) \otimes S^c(R) \arrow[r, "*"] & S^c(R) \arrow[d, "\mathrm{proj}"] \\ S^p R \otimes S^q R \arrow[u, "\mathrm{inc}"] \arrow[r, "M_{pq}"] & R \end{tikzcd}$$

In the next proposition we make explicit all the relations satisfied by the symmetric multibrace operations. We use the following notation. First $Sh(p,q)$ denotes the set of permutations made up of (p,q)-shuffles. Second, for any integer m, a *composition* of m is an ordered sequence $\underline{k} = (k_1, \dots, k_r)$ of nonnegative integers k_i such that $\sum_{i=1}^{r} k_i = m$. We write $(k_1, \dots, k_r) \vdash m$ when $\underline{k}$ is a composition of m.

4.3. Proposition. *Let $M_{pq} : S^p R \otimes S^q R \to R$ be (p,q)-symmetric operations on the vector space R. Let $*$ be the unique binary operation on $S^c R$ which is a cocommutative coalgebra morphism and whose projection onto R coincides with the M_{pq}'s. The following assertions are equivalent:*

(a) the operation $*$ is associative and unital,

(b) the (p,q)-symmetric operations M_{pq} satisfy the following relations $\mathcal{SR}$:

$$M_{00} = 0, M_{01} = \mathrm{id} = M_{10}, M_{0q} = 0 = M_{p0}, \ \textit{for } p > 1, q > 1,$$

and, for any integers i, j, k greater than or equal to 1 and any elements $x_1, \dots, x_n, y_1, \dots, y_m, z_1, \dots, z_r \in R$:

$(\mathcal{SR}_{ijk})$:

$$\sum_{r\geq 1}\Big(\sum_{\substack{(n_1,\dots,n_k)\vdash n\\(m_1,\dots,m_k)\vdash m}}\frac{1}{k!}\Big(\sum_{\substack{\sigma\in Sh(n_1,\dots,n_k)\\ \tau\in Sh(m_1,\dots,m_k)}} M_{kr}(M_{n_1m_1}(\underline{x}_1^{\sigma};\underline{y}_1^{\tau})\cdots M_{n_km_k}(\underline{x}_k^{\sigma};\underline{y}_k^{\tau});z_1,\dots,z_r)\Big)\Big)$$
$$=\sum_{l\geq 1}\Big(\sum_{\substack{(m_1,\dots,m_l)\vdash m\\(r_1,\dots,r_l)\vdash l}}\frac{1}{l!}\Big(\sum_{\substack{\gamma\in Sh(m_1,\dots,m_l)\\ \delta\in Sh(r_1,\dots,r_l)}} M_{nl}(x_1\cdots x_n;M_{m_1r_1}(\underline{y}_1^{\gamma};\underline{z}_1^{\delta}),\dots,M_{m_lr_l}(\underline{y}_l^{\gamma};\underline{z}_l^{\delta}))\Big)\Big),$$

where the first sum is taken over all pairs of compositions $(n_1,\dots,n_k)\vdash n$ *and* $(m_1,\dots,m_k)\vdash m$ *such that if* $n_i=0$ *then* $m_i\neq 0$, *for* $1\leq i\leq k$, *while the second one is taken over all pairs of compositions* $(m_1,\dots,m_l)\vdash m$ *and* $(r_1,\dots,r_l)\vdash r$ *such that if* $m_j=0$ *then* $r_j\neq 0$, *for* $1\leq j\leq l$. *The elements* $\underline{x}_i^{\sigma}$, $\underline{y}_i^{\tau}$, $\underline{y}_j^{\gamma}$ *and* $\underline{z}_j^{\delta}$ *are defined by:*

$$w_i^{\omega}:=w_{\omega(p_1+\cdots+p_{i-1}+1)}\otimes\cdots\otimes w_{\omega(p_1+\cdots+p_i)},$$

for $w\in R^{\otimes p}$, $(p_1,\dots,p_s)\vdash p$, $\omega\in Sh(p_1,\dots,p_s)$, *and* $1\leq i\leq s$.

Proof. From the unitality and counitality property of the Hopf algebra $S^c(R)$ we deduce that

$$M_{00}=0,\ M_{10}=\mathrm{id}_V=M_{01},\ \text{and}\ M_{n0}=0=M_{0n}\ \text{for}\ n\geq 2.$$

The operation $*$ is related to the operations M_{pq} by the following formula:

$$u_1\cdots u_p*v_1\cdots v_q=\sum_{k=1}^{p+q}(u_1\cdots u_p*v_1\cdots v_q)_k$$

where the component $(u_1\cdots u_p*v_1\cdots v_q)_k\in S^k(R)$ is given by:

$$\begin{aligned}&(u_1\cdots u_p*v_1\cdots v_q)_k=\\&\sum M_{i_1j_1}(u_1\cdots u_{i_1},v_1\cdots v_{j_1})M_{i_2j_2}(u_{i_1+1}\cdots u_{i_1+i_2},v_{j_1+1}\cdots v_{j_1+j_2})\cdots\\&\qquad\qquad\cdots M_{i_kj_k}(\cdots u_p,\cdots v_q)\in V^{\otimes k}\ .\end{aligned}$$

where the sum is extended to all the sequences of indices satisfying the following conditions:

– each sequence of integers $\{1,\dots i_1\},\{i_1+1,\dots,i_1+i_2\},\dots$ is ordered
– $1\leq i_1+1\leq i_1+i_2+1\leq\dots$,
– each sequence of integers $\{1,\dots j_1\},\{j_1+1,\dots,j_1+j_2\},\dots$ is ordered
– $1\leq j_1+1\leq j_1+j_2+1\leq\dots$.

This formula, which holds in the cocommutative case, is analogous to the formula given in the coassociative case in section 1.4 of [**42**] (cf. Proposition 2.3). The difference (restriction on the sequences of indices) is due to the fact that the coproduct is cocommutative, hence the map $\Delta: S(V)\to S(V)\otimes S(V)$ is viewed as a composite $\Delta: S(V)\to S^2(S(V)\otimes S(V))\rightarrowtail S(V)\otimes S(V)$.

For instance we get

$$(u_1\cdots u_p*v_1\cdots v_q)_1=M_{pq}(u_1\cdots u_p;v_1\cdots v_q),$$
$$(u_1\cdots u_p*v_1\cdots v_q)_{p+q}=u_1\cdots u_pv_1\cdots v_q,$$
$$(u_1u_2*v_1)_2=u_1M_{11}(u_2;v_1)+M_{11}(u_1;v_1)u_2.$$

By computing the component in R of the two elements:

$$(x_1\cdots x_i*y_1\cdots y_j)*z_1\cdots z_k=x_1\cdots x_i*(y_1\cdots y_j*z_1\cdots z_k)$$

we get the relation $(\mathcal{SR}_{ijk})$. The rest of the proof is as in Proposition 2.3. $\square$

4.4. Example. The first nontrivial relation is $(\mathcal{SR}_{111})$, which reads

$$M_{21}(uv;w)+M_{11}(M_{11}(u;v);w)=M_{11}(u;M_{11}(v;w))+M_{12}(u;vw)$$

for any $u,v,w\in R$. Remember that we are working in the symmetric framework, so $uv=vu$. Here are the next ones:

$(\mathcal{SR}_{112})$

$$\begin{aligned}&M_{22}(uv;wx)+M_{12}(M_{11}(u;v);wx)=\\&\quad M_{13}(u;vwx)+M_{12}(u;M_{11}(v;w)x)+M_{12}(u;wM_{11}(v;x))+M_{11}(u;M_{12}(v;wx)),\end{aligned}$$

$(\mathcal{SR}_{121})$

$$\begin{aligned}&M_{31}(uvw;x)+M_{21}(M_{11}(u;v)w;x)+M_{21}(vM_{11}(u;w);x)+M_{11}(M_{12}(u;vw);x)=\\&\quad M_{13}(u;vwx)+M_{12}(u;vM_{11}(w;x))+M_{12}(u;M_{11}(v;x)w)+M_{11}(u;M_{21}(vw;x)),\end{aligned}$$

$(\mathcal{SR}_{211})$

$$\begin{aligned}&M_{31}(uvw;x)+M_{21}(uM_{11}(v;w);x)+M_{21}(M_{11}(u;wv);x)+M_{11}(M_{21}(uv;w);x)=\\&\qquad\qquad M_{22}(uv;wx)+M_{21}(uv;M_{11}(w;x)),\end{aligned}$$

4.5. Symmetric multibrace algebra. By definition a *symmetric multibrace algebra* (or SMB-algebra) is a vector space R equipped with symmetric multibrace operations M_{pq}, that is $(p+q)$-ary operations $M_{pq}:S^pR\otimes S^qR\to R$, which satisfy all the conditions $\mathcal{SR}$ of Proposition 4.3.

4.6. Theorem. *There is an equivalence of categories between the cofree-cocommutative CHAs and the SMB-algebras.*

4.7. Commutative-associative bialgebra. A *$ComAs$-algebra* is a vector space equipped with two associative operations, one of them being commutative. We denote by $x\cdot y$ the commutative one and by $x*y$ the associative one. A $ComAs$-algebra is unital if there is an element 1 which is a unit for both operations.

A *$ComAs$-bialgebra* (i.e. a Com^c-$ComAs$-bialgebra) is a unital $ComAs$-algebra equipped with a counital cocommutative coproduct which satisfies the Hopf compatibility relation for both operations.

4.8. Proposition. *Any cofree-cocommutative CHA is a $ComAs$-bialgebra.*

Proof. Let $\mathcal{H}=S^c(V)$ be a cofree-cocommutative CHA. We equip $S^c(V)=S(V)$ with the standard commutative product, denoted $\cdot$, so that $(S(V),\cdot,\Delta)$ is the classical polynomial algebra. Since $\mathcal{H}$ is endowed with another associative product, denoted $*$, we have a $ComAs$-bialgebra $\mathcal{H}=(S(V),\cdot,*,\Delta)$. □

4.9. Proposition. *The free $ComAs$ algebra over V is a natural $ComAs$-bialgebra.*

Proof. The tensor product of two $ComAs$-algebras is still a $ComAs$-algebra. So we can apply the method of [**36**] to construct the expected coproduct. □

4.10. From $ComAs$-algebras to SMB-algebras. Let us recall that a conilpotent commutative and cocommutative Hopf algebra $(\mathcal{H}, \cdot, \Delta)$ is isomorphic to $S^c(\mathrm{Prim}(\mathcal{H}))$. We choose the Eulerian idempotent $e^{(1)} : \mathcal{H} \longrightarrow \mathcal{H}$ (cf. [**51, 34**]), to construct this isomorphism.

If $\mathcal{H}$ is a Com^c-$ComAs$-bialgebra, then, a fortiori, it is a Com^c-Com-bialgebra and we identify $\mathcal{H}$ with $S^c(\mathrm{Prim}(\mathcal{H}))$ as above. We identify the dot product $x{\cdot}y$ with the polynomial product xy. The associative product $*$ induces a SMB structure on $\mathrm{Prim}(\mathcal{H})$. As in the case of 2-associative bialgebras, the relationship between the associative product, the commutative product and the multibraces is given by:

$$(x_1 \cdots x_n) * (y_1 \cdots y_m) = \\ \sum_r \frac{1}{r!} \Big(\sum_{\substack{(k_1,\ldots,k_r)\vdash n \\ (h_1,\ldots,h_r)\vdash m}} \Big(\sum_{\substack{\sigma\in Sh(k_1,\ldots,k_r) \\ \tau\in Sh(h_1,\ldots,h_r)}} M_{n_1 m_1}(\underline{x}_1^{\sigma}, \underline{y}_1^{\tau}) \cdots M_{n_r m_r}(\underline{x}_r^{\sigma}, \underline{y}_r^{\tau}) \Big) \Big),$$

where the sum is taken over all pairs of compositions

$$(k_1, \ldots, k_r) \vdash n \qquad \text{and} \qquad (h_1, \ldots, h_r) \vdash m$$

such that if $k_i = 0$ then $h_i \neq 0$.

Note that the formula for M_{nm} may be obtained easily in terms of $*$, $\cdot$ and the M_{ij}'s, for $1 \leq i \leq n$, $1 \leq j \leq m$ and $i + j < n + m$. For instance:

$$\begin{aligned} M_{11}(x_1; y_1) &= x_1 * y_1 - x_1 y_1, \\ M_{12}(x_1; y_1 y_2) &= x_1 * (y_1 y_2) - x_1 y_1 y_2 - M_{11}(x_1; y_1) y_2 - y_1 M_{11}(x_1; y_2) \\ &= x * (yz) - (x * y)z - y(x * z) + xyz. \end{aligned}$$

4.11. Proposition. *There is a forgetful functor F from the category of $ComAs$-algebras to the category of SMB-algebras.*

4.12. Theorem. *There is a good triple of operads*

$$(Com, ComAs, SMB).$$

For any V there is a canonical isomorphism

$$ComAs(V) \cong S^c(SMB(V)).$$

4.13. The operad $ComAs$. The operad $ComAs$ is a set-theoretic operad. Its underlying set was described by R. Stanley in [**57**]. The dimension of $ComAs(n)$ is the number d_n of "labeled series-parallel posets with n points":

$$\{1, 3, 19, 195, 2791, 51303, \ldots \}.$$

4.14. The operad SMB. The operad SMB can be described by using the "connected labeled series-parallel posets with n nodes", cf. [**57**]. Let us mention the first dimensions $f_n = \dim SMB(n)$ for $n \geq 1$:

$$\{1, 2, 12, 122, 1740, 31922, \ldots \}.$$

From the presentation of the operad SMB it is clear that we can modify it by reducing the number of generators. For instance the relation $\mathcal{SR}_{111}$ shows that we can get rid of one of the two generators M_{12}, M_{21}. More symmetrically we can replace the two generators M_{12} and M_{21} by $M_{12} + M_{21}$. This discussion is postponed to a further paper.

4.15. From Lie algebras to SMB-algebras. Let $\mathfrak{g}$ be a Lie algebra and let $U(\mathfrak{g})$ be its universal enveloping algebra. By the Cartier-Milnor-Moore theorem we know that $U(\mathfrak{g})$ is isomorphic to $S^c(\mathfrak{g})$ as a coalgebra. Let us choose the Eulerian idempotents to realize this isomorphism explicitly, cf. [**51, 34**]:

$$e : U(\mathfrak{g}) \cong S^c(\mathfrak{g}).$$

Recall from [**34**] that the component $e^{(1)}$, whose image lies in $\mathfrak{g} = \operatorname{Prim} U(\mathfrak{g}) \subset U(\mathfrak{g})$ is obtained by

$$e^{(1)} = \log^\star(\mathrm{Id}),$$

where $\log^\star$ is the convolution log. The other components are given by

$$e^{(i)} = \frac{(e^{(1)})^{\star i}}{i!}.$$

Once this isomorphism is chosen, Theorem 4.6 implies that there is an SMB-algebra structure on $\mathfrak{g}$. Hence we have constructed a forgetful functor

$$Lie\text{-alg} \to SMB\text{-alg}.$$

In low dimension we get the following formulas:

$$M_{11}(x;y) = \frac{1}{2}[x,y],$$

$$M_{12}(x;yz) = \frac{1}{6}[[x,y],z] - \frac{1}{12}[x,[y,z]],$$

$$M_{21}(xy;z) = -\frac{1}{12}[[x,y],z] + \frac{1}{6}[x,[y,z]].$$

4.16. Free-commutative CHA. As in the preceding examples there is a dual version of the results of this section. We leave it to the interested reader to phrase them in detail.

5. (Co)free-(co)commutative right-sided CHA

In this section we study the cofree-cocommutative CHAs which satisfy the right-sided property **(r-s)**. The space of primitives inherits a symmetric brace algebra structure, which turns out to be the same as a pre-Lie algebra structure. As shown in [**11**] the free pre-Lie algebra on one generator gives rise to a cofree-cocommutative CHA which is the Grossman-Larson algebra [**24**]. In the dual framework, the cofree pre-Lie coalgebra on one generator gives rise to a free-commutative CHA which is the Connes-Kreimer Hopf algebra [**13**]. In this section $\mathbb{K}$ is a characteristic zero field. Examples will be given in section Examples will be given in section 6. They include the dual of the Faà di Bruno algebra (6.1) and the symmetric functions algebra (6.3).

5.1. Right-sided cofree-cocommutative CHA. Let $\mathcal{H}$ be a cofree-cocommutative CHA. By definition $\mathcal{H} = (S^c(R), *, \Delta)$ is said to be *right-sided* if the following condition holds:

(r-s) for any integer q the subspace $\bigoplus_{n\geq q} \mathcal{H}^n$ is a right-sided ideal of $\mathcal{H}$,

where $\mathcal{H}^n := S^n(R)$. We have seen that the associative product $*$ is given by a family $\{M_{pq}\}_{p,q\geq 0}$ of symmetric multibrace operations (cf. Proposition 4.3). If $\mathcal{H}$ is right-sided, then, as in 3.2, the operations M_{pq} are 0 for any $p \geq 2$. Hence we get a *symmetric brace algebra*, *SBrace*-algebra for short, that is, a vector space R equipped with operations M_{1q} verifying:

$$M_{1m}(M_{1n}(x;y_1\cdots y_n);z_1\cdots z_m) =$$
$$\sum_{(m_1,\dots,m_{n+1}\vdash m}\Big(\sum_{\sigma\in Sh(m_1,\dots,m_{n+1})} M_{1(n+m_{n+1})}(x;M_{1m_1}(y_1,\underline{z}_1^\sigma)\cdots M_{1m_n}(y_n;\underline{z}_n^\sigma)\underline{z}_{n+1}^\sigma)\Big),$$

where $\underline{z}_i^\sigma := z_{\sigma(m_1+\cdots+m_{i-1}+1)}\otimes\cdots\otimes z_{\sigma(m_1+\cdots+m_i)}$.

We remark that the first relation ($n=1$) implies that the operation M_{11} is pre-Lie. Recall that an operation $\{-,-\}$ is *pre-Lie* if

$$\{\{x,y\},z\}-\{x,\{y,z\}\}=\{\{x,z\},y\}-\{x,\{z,y\}\}.$$

Moreover it follows from $\mathcal{SR}_{111}$ (cf. 4.4) that the operation M_{12} is completely determined by the operation M_{11} since $M_{21}=0$. More generally, all the operations M_{1n} are determined by M_{11} and the following result holds.

5.2. Proposition (Guin-Oudom [**22, 23**]). *A symmetric brace algebra is equivalent to a pre-Lie algebra.*

Proof. A proof can be found in [**23**], see also [**31**]. The operations M_{1n} are obtained from M_{11} recursively by the formulas:

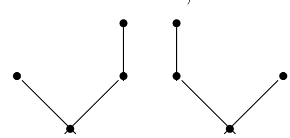
$$\begin{array}{rcc} M_{11}(x;y) & := & \{x,y\}, \\ M_{1n}(x;y_1\cdots y_n) & := & M_{11}(M_{1(n-1)}(x;y_1\cdots y_{n-1});y_n) - \\ & & \displaystyle\sum_{1\le i\le n-1} M_{1(n-1)}(x;y_1\cdots M_{11}(y_i;y_n)\cdots y_{n-1}), \end{array}$$

for $x,y_1,\dots,y_n\in V$. □

5.3. Theorem. *The category of cofree-cocommutative right-sided CHAs is equivalent to the category of pre-Lie algebras.*

Proof. Taking the primitives gives a functor from cofree-cocommutative r-s CHAs to pre-Lie algebras. Since any pre-Lie algebra gives rise to a symmetric brace algebra, we can apply the reconstruction functor to get a cofree-cocommutative CHA (cf. 4.6). Since we started with a symmetric brace algebra rather than a general SMB-algebra, the Hopf algebra is right-sided, hence the existence of a functor from pre-Lie algebras to cofree-cocommutative r-s CHAs. Obviously these two functors are inverse to each other. □

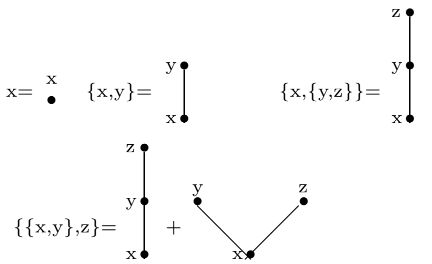
5.4. Unreduced trees and free pre-Lie algebra. In 3.12 we defined the planar unreduced trees. The quotient by the obvious relation gives the notion of (non-planar) *unreduced trees*. For instance, in this setting, the two drawings



define the same element. The set of unreduced trees with n vertices is denoted UT_n. The free pre-Lie algebra on V has been described by Livernet and Chapoton in [**11**] in terms of unreduced trees labeled by elements of V:

$$PreLie(V)=\sum_n \mathbb{K}[UT_n]\otimes V^{\otimes n}.$$

The pre-Lie product is given by the following rule. If ω and ω' are two labelled unreduced trees, then the pre-Lie product $\{\omega, \omega'\}$ is the sum of all the trees obtained from ω and ω' by drawing an edge from a vertex of ω to the root of ω'. The root of this new tree is the root of ω. For instance:



5.5. The conjectural operad $\mathcal{X}$. The results of the three other cases lead us naturally to conjecture the existence of an operad structure $\mathcal{X}$ on $Com \circ preLie$ compatible with the operad structure of Com and $preLie$ (extension of operads, see 3.4.2 in [**38**]). Observe that

$$\dim \mathcal{X}(n) = (n+1)^{n-1}.$$

If so, then we conjecture that any cofree-cocommutative right-sided CHA is a $\mathcal{X}$-algebra.

5.6. Grossman-Larson Hopf algebra. In [**24**] Grossman and Larson constructed a cocommutative Hopf algebra on some trees. It turns out that this is exactly the combinatorial Hopf algebra $S^c(preLie(\mathbb{K}))$.

5.7. Free-commutative right-sided CHAs and Connes-Kreimer algebra. As before the results of this section can be linearly dualized. It should be said that many examples in the literature appear as free-commutative (rather than cofree-cocommutative). For instance the dual of $S^c(preLie(\mathbb{K}))$, that is, the algebra $S(copreLie(\mathbb{K}))$, is the Connes-Kreimer Hopf algebra (cf. [**49, 26**]). It can be constructed either directly as in [**13**], or by means of the cofree pre-Lie coalgebra on one generator.

The direct construction consists in taking the free-commutative algebra over the set of unreduced trees. The coproduct is given by the formula

$$\Delta(t) = \sum_c P_c(t) \otimes R_c(t),$$

where $P_c(t)$ is a polynomial of trees and $R_c(t)$ is a tree, the sum is over admissible cuts c. Let us recall that a cut of the tree t is admissible if there is one and only one cut on any path from the root to a leaf. Among the pieces, one of them contains the root, this is the tree $R_c(t)$. The other ones assemble to give a polynomial, which is $P_c(t)$. Observe that there are two extreme cuts: under the root (it gives the element $t \otimes 1$), above the leaves (it gives the element $1 \otimes t$).

We end this section with the dual version of Theorem 5.3.

5.8. Theorem. *The category of free-commutative right-sided CHAs is equivalent to the category of pre-Lie coalgebras.*

5.9. Cofree-cocommutative CHAs being right and left-sided. Let $\mathcal{H}$ be a cofree-cocommutative CHA which is both right-sided and left-sided. As a consequence its primitive part is such that $M_{nm} = 0$ except possibly for M_{11}. By formula $\mathcal{R}_{111}$ it follows that M_{11} is associative.

This case has been studied in [**37**] where it is proved that there is a good triple of operads

$$(Com, CTD, As)$$

where As is the operad of nonunital associative algebras, and CTD is the operad of *commutative tridendriform algebras*. Recall from [**41**] that a tridendriform algebra is determined by three binary operations $x \prec y, x \succ y, x \cdot y$ satisfying 7 relations. It is said to be *commutative* whenever $x \prec y = y \succ x$ and $x \cdot y = y \cdot x$ for any x and y. Then, the 7 relations come down to the following 4 relations:

$$\begin{gathered}(x \prec y) \prec z = x \prec (y \prec z + z \prec y + y \cdot z),\\ (x \prec y) \cdot z = x \cdot (z \prec y),\\ (x \cdot y) \prec z = x \cdot (y \prec z),\\ (x \cdot y) \cdot z = x \cdot (y \cdot z),\end{gathered}$$

The second relation has been, unfortunately, mistakenly omitted in [**37**].

There is an equivalence of categories between the right-and-left-sided cofree-cocommutative CHAs and the nonunital associative algebras.

6. Examples

There are many examples of combinatorial Hopf algebras in the literature. We already described some of them: the free algebras of type *Dipt*, $2as$, *ComAs*, *Dend*, $\mathcal{Y}$ and their dual. In these cases the combinatorial objects at work are trees (of various types). The examples coming from quantum field theory are usually free-associative or free-commutative (for instance Connes-Kreimer algebra). In this section we list some specific examples and refer to the literature for more information. In 6.4 we give an example of two different CHA structures on a given cofree-coassociative Hopf algebra.

6.1. Faà di Bruno algebra. On $F = \bigoplus_{n\geq 1} \mathbb{K}x_n$ the operation

$$\{x_p, x_q\} := -px_{p+q}$$

is a pre-Lie product since

$$\{\{x_p, x_q\}, x_r\} - \{x_p, \{x_q, x_r\}\} = p^2 x_{p+q+r},$$

is symmetric in q and r.

Up to a minor modification, this is the pre-Lie algebra coming from the nonsymmetric operad As, cf. 6.7. By Theorem 5.3 the coalgebra $S^c(F)$ is a combinatorial Hopf algebra which is cofree-commutative and right-sided as an algebra. This is exactly the graded dual of the Faà di Bruno Hopf algebra, since the Lie bracket is given by $[x_p, x_q] = (-p+q)x_{p+q}$, see for instance [**29, 16, 17**]. More precisely, let us introduce the dual basis $a_n := \frac{1}{n!}x^*_{n-1}$. The linear dual $S^c(F)^*$ is the polynomial algebra in $a_2, \ldots, a_n, \ldots$ and the coproduct is given by

$$\Delta(a_n) = \sum_{k=1}^{n} \sum_{\lambda} \binom{n}{\lambda; k} a_1^{\lambda_1} \cdots a_n^{\lambda_n} \otimes a_k$$

where $\binom{n}{\lambda;k} = \frac{n!}{\lambda_1!\cdots\lambda_n!(1!)^{\lambda_1}\cdots(n!)^{\lambda_n}}$ for $\lambda_1+\cdots+\lambda_n = k$ and $\lambda_1+2\lambda_2+\cdots+n\lambda_n = n$. We see immediately from this formula that the Faà di Bruno Hopf algebra is a right-sided coalgebra.

This Hopf algebra proves helpful in studying the higher derivatives of a composition of formal power series (Faà di Bruno's formula) because it is the Hopf algebra of functions on the group of invertible formal power series in one variable.

It is immediate to check that the pre-Lie algebra F is generated (in characteristic zero) by one element, namely x_1. Hence F is a quotient of $PreLie(\mathbb{K})$. As a consequence the Faà di Bruno Hopf algebra gets identified with a subalgebra of the Connes-Kreimer algebra.

There is an analogue of the Faà di Bruno algebra in the noncommutative framework due to Brouder, Frabetti and Menous, cf. [**8**]. Since F is a pre-Lie algebra, a fortiori it is a brace algebra. Therefore, taking the cofree-coassociative coalgebra gives a combinatorial Hopf algebra. The graded dual is a noncommutative analogue of the Faà di Bruno algebra.

6.2. The algebra of quasi-symmetric functions QSym. (see [**25**]). Let $\mathbb{K}[\xi_1,\ldots,\xi_n,\ldots]$ be the algebra of all polynomial functions on an infinite family of variables $\{\xi_n\}_{n\geq 1}$. The algebra of *quasi-symmetric functions* **QSym** is the subalgebra of all polynomial functions $f \in \mathbb{K}[\xi_1,\ldots,\xi_n,\ldots]$ such that $f = \sum_{n_1\cdots n_r} f_{n_1\cdots n_r}\big(\sum_{i_1<\cdots<i_r} \xi_{i_1}^{n_1}\cdots\xi_{i_r}^{n_r}\big)$, for certain coefficients $f_{i_1\cdots i_r} \in \mathbb{K}$. The product $*$ of **QSym** is the usual product of p olynomials.

The subspace of homogeneous elements of degree n of **QSym** has a natural basis $\{x_{n_1\cdots n_r}\}_{n_1+\cdots+n_r=n}$, with:

$$x_{n_1\cdots n_r} := \sum_{i_1<\cdots<i_r} \xi_{i_1}^{n_1}\cdots\xi_{i_r}^{n_r}.$$

The coproduct on **QSym** is defined by:

$$\Delta(x_{n_1\cdots n_r}) := \sum_{0\leq j\leq r} x_{n_1\cdots n_j}\otimes x_{n_{j+1}\cdots n_r}.$$

It is clear that the subspace of primitive elements of **QSym** is the vector space $\bigoplus_{n\geq 1}\mathbb{K}x_n$. Moreover, the coalgebra isomorphism with $T^c(\bigoplus_{n\geq 1}\mathbb{K}x_n)$ sends the element $x_{n_1\cdots n_r}$ to $x_{n_1}\otimes\cdots\otimes x_{n_r}$.

Note that **QSym** is cofree-coassociative and its product is associative and commutative. The multibrace structure of $\bigoplus_{n\geq 1}\mathbb{K}x_n$ associated to the product $*$ is given by:

$$M_{nm}(x_{i_1}\cdots x_{i_n};x_{j_1}\cdots x_{j_m}) = \begin{cases} x_{i_1+j_1}, & \text{for } n=m=1,\\ 0, & \text{otherwise.}\end{cases}$$

The graded dual of **QSym** is the Solomon descent algebra **NSym**. It is the free algebra over the space $\bigoplus_{n\geq 1}\mathbb{K}x_n$, with the coassociative cocommutative coproduct Δ^* given by:

$$\Delta^*(x_n) := \sum_{i=0}^{n} x_i\otimes x_{n-i}.$$

The Solomon descent algebra is a cofree-cocommutative CHA, isomorphic as a coalgebra to $S^c\big(Lie(\bigoplus_{n\geq 1}\mathbb{K}x_n)\big)$. This is a particular example of the case discussed in 4.15.

6.3. The Hopf algebra Sym. The Hopf algebra **Sym** is a cofree-cocommutative CHA isomorphic as a coalgebra to $S^c\big(\bigoplus_{n\geq 1}\mathbb{K}x_n\big)$. This is a particular example of the case discussed in 4.15.

6.4. The Malvenuto-Reutenauer algebra $\mathcal{H}_{MR}$. (See [**44**]). The underlying space of $\mathcal{H}_{MR}$ is the graded space $\bigoplus_{n\geq 0}\mathbb{K}[S_n]$, where S_n denotes the group of permutations of n elements. The product of $\mathcal{H}_{MR}$ is the shuffle product, given by:

$$\sigma * \tau := \sum_{\delta\in Sh(n,m)} (\sigma\times\tau)\cdot\delta^{-1},$$

for $\sigma\in S_n$ and $\tau\in S_m$; where $Sh(n,m)$ denotes the set of all *(n,m)*-shuffles and $\times$ denotes the concatenation of permutations.

Given a permutation $\sigma\in S_n$ and an integer $0\leq i\leq n$, let $\sigma^i_{(1)}$ be the map from $\{1,\dots,i\}$ to $\{1,\dots,n\}$ whose image is $\sigma^i_{(1)} = (\sigma(1),\dots,\sigma(i))$ and $\sigma^i_{(2)}$ be the map whose image is $\sigma^i_{(2)} = (\sigma(i+1),\dots,\sigma(i))$.

The coproduct is defined as follows:

$$\Delta(\sigma) := \sum_{i=0}^{n} std(\sigma^i_{(1)})\otimes std(\sigma^i_{(2)}),$$

where $std(\tau)$ is the unique permutation in S_n such that $std(\tau)(i) < std(\tau)(j)$ if and only if $\tau(i)<\tau(j)$, for any injective map $\tau:\{1,\dots,k\}\longrightarrow\{1,\dots,n\}$. In degree two the primitive space is 1-dimensional generated by $12-21$. In degree three it is 3-dimensional generated by the elements:

$$u := 213 - 312,$$
$$v := 231 - 132,$$
$$w := 321 - 132 - 213 + 123.$$

An element $\sigma\in S_n$ is called *irreducible* if $\sigma\notin\bigcup_{1\leq i\leq n-1} S_i\times S_{n-i}$. We denote by Irr_n the subset of irreducible elements of S_n. In low dimension we get $Irr_2 = \{21\}, Irr_3=\{231,312,321\}$.

We will give two different structures of combinatorial Hopf algebra on $\mathcal{H}_{MR}$. For the first one it appears as a cofree-coassociative general CHA and for the second one it appears as a cofree-coassociative right-sided CHA.

1) Define an isomorphism φ from $T^c\big(\bigoplus_n \mathbb{K}[Irr_n]\big)$ to $\mathcal{H}_{MR}$ by

$$\varphi(x) = \begin{cases} e^{(1)}_{inf}(\sigma), & \text{for } x=\sigma\in\bigcup_n Irr_n \\ e^{(1)}_{inf}(\sigma_1)\times\cdots\times e^{(1)}_{inf}(\sigma_m), & \text{for } x=\sigma_1\otimes\cdots\otimes\sigma_m, \end{cases}$$

where

$$e^{(1)}_{inf}(\sigma) := \sum_{i\geq 1}(-1)^{i+1}\times^i\circ\overline{\Delta}^i(\sigma) = \sigma - \sigma_{(1)}\times\sigma_{(2)} + \cdots$$

is the (infinitesimal) idempotent defined in [**42**]. For instance, in low dimension, we get:

$$\varphi(21) = 21 - 12,$$
$$\varphi(231) = 231 - 132 - 123 + 123 = 231 - 132,$$
$$\varphi(312) = 312 - 123 - 213 + 123 = 312 - 213,$$
$$\varphi(321) = 321 - 132 - 213 + 123.$$

Under this construction $(\mathcal{H}_{MR}, \phi)$ become a cofree-coassociative general CHA. Hence the space $\bigoplus_{n\geq 1}\mathbb{K}[Irr_n]$ inherits a structure of MB-algebra (see [**55**] for explicit formulas). The $2as$-bialgebra structure is given by the products $\times$ and $*$, and the coproduct Δ.

2) Define an isomorphism $\theta : T^c\big(\bigoplus_n \mathbb{K}[Irr_n]\big) \longrightarrow \mathcal{H}_{MR}$ as follows:

$$\theta(x) = \begin{cases} e(\sigma), & \text{for } x = \sigma \in \bigcup_n Irr_n \\ ((e(\sigma_1) \succ e(\sigma_2)) \dots) \succ e(\sigma_m), & \text{for } x = \sigma_1 \otimes \cdots \otimes \sigma_m, \end{cases}$$

where $e(\sigma) := \sum_{i\geq 1}(-1)^{i+1}\omega_{\succ}^i \circ \overline{\Delta}^i(\sigma)$ is the idempotent defined in section 3.11. So $(\mathcal{H}_{MR}, \theta)$ is a cofree-coassociative right-sided CHA. Indeed, one can check that the multi-brace operations satisfy $M_{pq} = 0$ for $p \geq 2$, so there is a brace structure on $\mathbb{K}[Irr_n]$. The dendriform algebra structure on $\mathcal{H}_{MR}$ is given by

$$\sigma \succ \tau := \sum_{\substack{\delta\in Sh(n,m)\\ \delta(n+m)=n+m}} (\sigma \times \tau)\cdot \delta^{-1},$$
$$\sigma \prec \tau := \sum_{\substack{\delta\in Sh(n,m)\\ \delta(n)=n+m}} (\sigma \times \tau)\cdot \delta^{-1}.$$

Observe that, because of the second condition under the summation sign, the sum is only over half-shuffles. In low dimension ($n \leq 3$) θ coincides with φ, but for $n \geq 4$ it is different.

The Hopf algebra of Solomon-Tits and the algebra **FPQSym** of parking functions may be described as CHA in a similar way, cf. [**47, 48**].

In [**50**] the authors show that there exists a subalgebra of $\mathcal{H}_{MR}$ which is spanned by the image of the standard Young tableaux. They construct a coalgebra structure for which the "connected tableaux" form a basis of the primitive part. Hence there is a MB-algebra structure on this latter space.

6.5. Tensor module as indecomposables. In the literature there are several examples of combinatorial Hopf algebras whose space of indecomposables is the tensor module.

- in [**9**] Brouder and Schmitt construct a Hopf algebra structure on $T(\overline{T}(H))$, $S(\overline{T}(H))$ and on $S(\overline{S}(H))$ where H is a non-unital bialgebra. These constructions generalize a construction of G. Pinter related to renormalisation in perturbative quantum field theory.

- In [**61**] Turaev constructs a Hopf algebra structure on $T(T(V))$ (resp. $T(\overline{T}(V))$) which is free-associative right-sided. As we know, there is a structure of pre-Lie coalgebra on $T(V)$ (resp. $\overline{T}(V)$) which is studied in detail. It is part of a more general structure: a brace coalgebra structure, not studied in loc.cit.

6.6. Quantum field theory. The Connes-Kreimer algebra is an example of the combinatorial Hopf algebras which appear in quantum field theory. They are based on graphs (Feynman graphs) and the rule is always the same: a subgraph is singled out and put on the right side. What is left is used to construct the left-hand side of the coproduct. See for instance [**14, 16, 6, 5, 60**].

6.7. Operads. For any nonsymmetric operad $\mathcal{P}$, the space $\bigoplus_n \mathcal{P}_n$ has a natural structure of brace algebra, therefore $T^c(\bigoplus_n \mathcal{P}_n)$ is a cofree-coassociative r-s CHA (cf. for instance [**28, 46**]). Similarly, if $\mathcal{P}$ is a properad, then it can be shown that $\oplus_{n,m}\mathcal{P}(n,m)$ is naturally equipped with a structure of MB-algebra, cf. [**45**]. Therefore $T^c\big(\bigoplus_{n,m}\mathcal{P}(n,m)\big)$ is a cofree-coassociative CHA.

6.8. Incidence Hopf algebras. Continuing the paper of Joni and Rota [**29**] W. Schmitt studied in [**56**] the notion of incidence Hopf algebras. In the last sections of the work he described incidence Hopf algebras related to families of graphs closed under formation of induced subgraphs and sums. His examples include the the Faà di Bruno algebra. The incidence Hopf algebras on graphs defined in [**56**] are either free-commutative or free-associative, so they enter in the examples studied in this paper.

Consider for instance the free associative algebra spanned by the set $\tilde{\mathcal{G}}_0$ of all the isomorphism classes of connected simple graphs (i.e. having neither loops nor multiple edges) with linearly ordered vertex set, and let $H_l(\mathcal{G})$ be the free associative algebra generated by $\tilde{\mathcal{G}}_0$. Note that as a vector space $H_l(\mathcal{G})$ is spanned by the isomorphism classes of simple graphs with linearly ordered vertex set.

For any simple connected graph G with the set of vertices $V(G)$ linearly ordered and any subset of $U \subseteq V(G)$ the induced graph $G|U$ is the graph whose set of vertices is U, with the edge set formed by all the edges of G which have both end-vertices in U. Note that $G|U$ is simple, and that U inherits the linear order of $V(G)$. So, the isomorphism class of $G|U$ belongs to $H_l(\mathcal{G})$. The coalgebra structure on $H_l(\mathcal{G})$ is given by

$$\Delta(\langle G\rangle) := \sum_{U\subseteq V(G)} \langle G|U\rangle \otimes \langle G|(V(G)-U)\rangle,$$

where $\langle -,-\rangle$ denotes the isomorphism class.

Clearly $H_l(\mathcal{G})$ is free-associative and is cofree-cocommutative. Its graded dual $H_l(\mathcal{G})^*$ is cofree-coassociative isomorphic to $T^C(\mathbb{K}[\tilde{\mathcal{G}}_0])$. The multibrace structure on $\mathbb{K}[\tilde{\mathcal{G}}_0]$ is given by:

$M_{nm}\big(\langle G_1\rangle\cdots\langle G_n\rangle;\langle H_1\rangle\cdots\langle H_m\rangle\big)$ = the sum of the isomorphism classes of all simple connected graphs G such that:

(1) the set of vertices $V(G)$ is the disjoint union

$$\big(\bigcup_{i=1}^{n} V(G_i)\big)\bigcup\big(\bigcup_{j=1}^{m} V(H_j)\big),$$

(2) If $i \neq k$, there does not exist any edge of G between a vertex of G_i and a vertex of G_k,

(3) If $j \neq l$, there does not exist any edge of G between a vertex of H_j and a vertex of H_l,

(4) the linear order on $V(G)$ is such that for $1 \leq i < k \leq n$, the minimal vertex of G_i is smaller than the minimal vertex of G_k; and for $1 \leq j < k \leq m$, the minimal vertex of H_j is smaller than the minimal vertex of H_l. Moreover, the induced orders on $G|G_i$ and $G|H_j$ coincide with the orders of G_i and H_j, respectively.

6.9. Relationship with algebraic K-theory. The free dendriform algebra on one generator is spanned by the planar binary trees (cf. 3.10 and [**35**]). If, in each dimension, we take the sum of all the trees, then it spans a one-dimensional space, and the sum over all dimensions form a sub-Hopf algebra of the Loday-Ronco algebra $Dend(\mathbb{K})$. The same procedure applied to the free dendriform algebra over some decoration set X provides a combinatorial Hopf algebra which is dual to the Hopf algebra constructed by Gangl, Goncharov and Levin in [**21**]. The importance of this combinatorial Hopf algebra lies in its close relationship with computation in algebraic K-theory (cf. loc.cit.).

7. Variations

Another interesting case consists in assuming both right-sidedness and left-sidedness on a CHA. It is treated in the first paragraph of this section.

In this paper our object of study is conilpotent Hopf algebras. As we know it involves an associative product and a coassociative coproduct. But there exist more general types of bialgebras involving various kinds of operads. We briefly mention the pattern of a similar theory to generalized bialgebras in the second paragraph.

7.1. Cofree-associative right and left-sided CHAs. Let us assume that the CHA $\mathcal{H}$ is both right-sided and left-sided. Then it turns out that the multibrace operations are all trivial with the exception of M_{11}. From the formula 2.4 we deduce that M_{11} is an associative operation. So the primitive part is simply an associative algebra. This example is well-documented in the literature, it is called "quasi-shuffle algebra", cf. [**15, 37**].

7.2. Generalized bialgebras. In [**38**] we introduced the notion of *generalized bialgebras*, more specifically $\mathcal{C}^c$-$\mathcal{A}$-bialgebras, where $\mathcal{C}$ governs the coalgebra structure and $\mathcal{A}$ governs the algebra structure. Under some hypothesis a $\mathcal{C}^c$-$\mathcal{A}$-bialgebra $\mathcal{H}$ is cofree and its primitive part is governed by an operad $\mathcal{P}$. We assume that the triple of operads $(\mathcal{C}, \mathcal{A}, \mathcal{P})$ is good (in particular $\mathcal{A} \cong \mathcal{C} \circ \mathcal{P}$ as $\mathbf{S}$-modules). As a consequence there is an isomorphism $\mathcal{H} \cong \mathcal{C}^c(\mathrm{Prim}\,\mathcal{H})$. Once such an isomorphism is chosen, we call this data a *combinatorial $\mathcal{C}^c$-$\mathcal{A}$-bialgebra*. In this framework one can generalize the results of this paper as follows. There exists an operad $\mathcal{Q}$ and a morphism of operads $\mathcal{P} \to \mathcal{Q}$ such that the $\mathcal{P}$-algebra of $\mathrm{Prim}\,\mathcal{H}$ can be lifted to a $\mathcal{Q}$-algebra structure.

References

[1] *F. Akman*, Multibraces on the Hochschild space J. Pure Appl. Algebra 167 (2002), no. 2-3, 129–163.

[2] *M. Aguiar, F. Sottile*, Structure of the Loday-Ronco Hopf algebra of trees. J. Algebra 295 (2006), no. 2, 473–511.

[3] *M. Aguiar, N. Bergeron, F. Sottile*, Combinatorial Hopf algebras and generalized Dehn-Sommerville relations. Compos. Math. 142 (2006), no. 1, 1–30.

[4] *N. Bergeron, C. Hohlweg*, Coloured peak algebras and Hopf algebras. J. Algebraic Combin. 24 (2006), no. 3, 299–330.

[5] *A. Brouder*, Quantum field theory meets Hopf algebra, Math. Nachr. 282 (2009), no. 12, 1664–1690.

[6] *A. Brouder, A. Frabetti*, QED Hopf algebras on planar binary trees. J. Algebra 267 (2003), no. 1, 298–322.

[7] *A. Brouder, A. Frabetti, C. Krattenthaler*, Non-commutative Hopf algebra of formal diffeomorphisms. Adv. Math. 200 (2006), no. 2, 479–524.

[8] *A. Brouder, A. Frabetti, F. Menous*, Combinatorial Hopf algebras from renormalization, to appear in Journal of Algebraic Combinatorics, arXiv:09093362
[9] *A. Brouder, W. Schmitt*, Renormalization as a functor on bialgebras. J. Pure Appl. Algebra 209 (2007), no. 2, 477–495.
[10] *F. Chapoton*, Un théorème de Cartier-Milnor-Moore-Quillen pour les bigèbres dendriformes et les algèbres braces. J. Pure Appl. Algebra 168 (2002), no. 1, 1–18.
[11] *F. Chapoton, M. Livernet*, Pre-Lie algebras and the rooted trees operad. Internat. Math. Res. Notices 2001, no. 8, 395–408.
[12] *F. Chapoton, M. Livernet*, Relating two Hopf algebras built from an operad. Int. Math. Res. Not. IMRN 2007, no. 24, Art. ID rnm131, 27 pp.
[13] *A. Connes, D. Kreimer*, Hopf algebras, renormalization and noncommutative geometry Comm. Math. Phys. 199 (1998), no. 1, 203–242.
[14] *A. Connes, H. Moscovici*, Hopf algebras, cyclic cohomology and the transverse index theorem. Comm. Math. Phys. 198 (1998), no. 1, 199–246.
[15] *K. Ebrahimi-Fard, L. Guo*, Mixable shuffles, quasi-shuffles and Hopf algebras. J. Algebraic Combin. 24 (2006), no. 1, 83–101.
[16] *H. Figueroa, J.M. Gracia-Bondía*, Combinatorial Hopf algebras in quantum field theory. I. Rev. Math. Phys. 17 (2005), no. 8, 881–976.
[17] *H. Figueroa, J.M. Gracia-Bondía, J.C. Várilly*, Faà di Bruno Hopf algebras, preprint (2005), 7pp. arXiv:math/0508337
[18] *L. Foissy*, Les algèbres de Hopf des arbres enracinés décorés. II. Bull. Sci. Math. 126 (2002), no. 4, 249–288.
[19] *M. Gerstenhaber, A. Voronov*, Homotopy G-algebras and moduli space operad. Internat. Math. Res. Notices (1995), no. 3, 141–153.
[20] *E. Getzler, J.D.S. Jones*, Operads, homotopy algebra, and iterated integrals for double loop spaces. `arXiv:hep-th/9403055`
[21] *H. Gangl, A.B. Goncharov, A. Levin*, Multiple logarithms, algebraic cycles and trees. Frontiers in number theory, physics, and geometry. II, 759–774, Springer, Berlin, 2007.
[22] *D. Guin, J.-M. Oudom*, Sur l'algèbre enveloppante d'une algèbre pré-Lie. C. R. Math. Acad. Sci. Paris 340 (2005), no. 5, 331–336.
[23] *D. Guin, J.-M. Oudom*, On the Lie enveloping algebra of a pre-Lie algebra. J. K-Theory 2 (2008), no. 1, 147–167.
[24] *R. Grossman, R.G. Larson*, Hopf-algebraic structure of families of trees. J. Algebra 126 (1989), no. 1, 184–210.
[25] *F. Hivert, J.-C. Novelli, J.-Y. Thibon*, Commutative combinatorial Hopf algebras, J. Algebraic Combin. 28 (2008), no. 1, 65–95.
[26] *M. Hoffman*, Combinatorics of rooted trees and Hopf algebras. Trans. Amer. Math. Soc. 355 (2003), no. 9, 3795–3811 (electronic).
[27] *R. Holtkamp*, Comparison of Hopf algebras on trees, Arch. Math. (Basel) 80 (2003), no. 4, 368–383.
[28] *R. Holtkamp*, On Hopf algebra structures over free operads, Adv. Math. 207 (2006), 544–565.
[29] *S.A. Joni, G.-C. Rota*, Coalgebras and bialgebras in combinatorics. Stud. Appl. Math. 61 (1979), no. 2, 93–139. Also printed in Umbral calculus and Hopf algebras (Norman, Okla., 1978), pp. 1–47, Contemp. Math., 6, Amer. Math. Soc., Providence, R.I., 1982.
[30] *T. Kadeishvili*, Measuring the noncommutativity of DG-algebras. Topology and noncommutative geometry. J. Math. Sci. (N. Y.) 119 (2004), no. 4, 494–512.
[31] *T. Lada, M. Markl*, Symmetric brace algebras. Appl. Categ. Structures 13 (2005), no. 4, 351–370.
[32] *T. Lam, P. Pylyavskyy*, Combinatorial Hopf algebras and K-homology of Grassmannians. Int. Math. Res. Not. IMRN 2007, no. 24, Art. ID rnm125, 48 pp.
[33] *M. Livernet*, From left modules to algebras over an operad: application to combinatorial Hopf algebras, Annales Mathématiques Blaise Pascal 17, no. 1 (2010), 47–96.
[34] *J.-L. Loday*, Série de Hausdorff, idempotents eulériens et algèbres de Hopf. Exposition. Math. 12 (1994), no. 2, 165–178.
[35] *J.-L. Loday*, Dialgebras, in "Dialgebras and related operads", Springer Lecture Notes in Math. 1763 (2001), 7-66.
[36] *J.-L. Loday*, Scindement d'associativité et algèbres de Hopf, Proceedings of the Conference in honor of Jean Leray, Nantes 2002, Séminaire et Congrès (SMF) 9 (2004), 155–172.

[37] *J.-L. Loday*, On the algebra of quasi-shuffles. Manuscripta Mathematica 123, no1, (2007), 79–93.

[38] *J.-L. Loday*, **Generalized bialgebras and triples of operads**, Astérisque 320 (2008), x+116 pp.

[39] *J.-L. Loday, M. Ronco*, Hopf algebra of the planar binary trees. Adv. in Maths 139 (1998), 293–309.

[40] *J.-L. Loday, M. Ronco*, Algèbres de Hopf colibres. C. R. Math. Acad. Sci. Paris 337 (2003), no. 3, 153–158.

[41] *J.-L. Loday, M. Ronco*, Trialgebras and families of polytopes. Homotopy theory: relations with algebraic geometry, group cohomology, and algebraic K-theory, 369–398, Contemp. Math., 346, Amer. Math. Soc., Providence, RI, 2004.

[42] *J.-L. Loday, M. Ronco*, On the structure of cofree Hopf algebras, J. reine angew. Math. 592 (2006) 123–155.

[43] *J.-L. Loday, B. Vallette*, **Algebraic Operads**, book in progress.

[44] *C. Malvenuto, C. Reutenauer*, Duality between quasi-symmetric functions and the Solomon descent algebra. J. Algebra 177 (1995), 967-982.

[45] *S. Merkulov, B. Vallette*, Deformation theory of representations of prop(erad)s. I. J. Reine Angew. Math. 634 (2009), 51–106.

[46] *I. Moerdijk*, On the Connes-Kreimer construction of Hopf algebras. Homotopy methods in algebraic topology (Boulder, CO, 1999), 311–321, Contemp. Math., 271, Amer. Math. Soc., Providence, RI, 2001.

[47] *J.-C. Novelli, J.-Y. Thibon*, Hopf algebras and dendriform structures arising from parking functions. Fund. Math. 193 (2007), no. 3, 189–241.

[48] *P. Palacios, M. Ronco*, Weak Bruhat order on the set of faces of the permutohedron and the associahedron. J. Algebra 299 (2006), no. 2, 648–678.

[49] *F. Panaite*, Relating the Connes-Kreimer and Grossman-Larson Hopf algebras built on rooted trees. Lett. Math. Phys. 51 (2000), no. 3, 211–219.

[50] *S. Poirier, C. Reutenauer*, Algèbres de Hopf de tableaux. Ann. Sci. Math. Québec 19 (1995), no. 1, 79–90.

[51] *C. Reutenauer*, Theorem of Poincaré-Birkhoff-Witt, logarithm and symmetric group representations of degrees equal to Stirling numbers. Combinatoire énumérative (Montreal, Que., 1985/Quebec, Que., 1985), 267–284, Lecture Notes in Math., 1234, Springer, Berlin, 1986.

[52] *M. Ronco*, Primitive elements in a free dendriform algebra, New trends in Hopf algebra theory (La Falda, 1999), 245–263, Contemp. Math., 267, Amer. Math. Soc., Providence, RI, 2000.

[53] *M. Ronco*, A Milnor-Moore theorem for dendriform Hopf algebras. C. R. Acad. Sci. Paris Sr. I Math. 332 (2001), no. 2, 109–114.

[54] *M. Ronco*, Eulerian idempotents and Milnor-Moore theorem for certain non-cocommutative Hopf algebras, J. Algebra 254 (1) (2002), 152-172.

[55] *M. Ronco*, Shuffle bialgebras, to appear in Ann. Inst. Fourier, arXiv:math/0703437

[56] *W. Schmitt*, Incidence Hopf algebras. J. Pure Appl. Algebra 96 (1994), no. 3, 299–330.

[57] *R.P. Stanley*, Enumeration of posets generated by disjoint unions and ordinal sums. Proc. Amer. Math. Soc. 45 (1974), 295-299.

[58] *R.P. Stanley*, **Enumerative Combinatorics**, vol. 2, Cambridge Studies in Advanced Mathematics, 62. Cambridge University Press, Cambridge, 1999. xii+581 pp.

[59] *W.D. van Suijlekom*, Renormalization of gauge fields: a Hopf algebra approach. Comm. Math. Phys. 276 (2007), no. 3, 773–798.

[60] *A. Tanasa, F. Vignes-Tourneret*, Hopf algebra of non-commutative field theory, J. Noncommutative Geometry 2 (2008), no 1, 125–139.

[61] *V. Turaev*, Coalgebras of words and phrases. J. Algebra 314 (2007), no. 1, 303–323.

Institut de Recherche Mathématique Avancée, CNRS et Université de Strasbourg, 7 rue R. Descartes, 67084 Strasbourg Cedex, France

E-mail address: loday@math.unistra.fr

Depto. de Estadística. CIMFAV, Facultad de Ciencias-Universidad de Valparaiso, Avda. Gran Bretana 1091, Valparaiso, Chile

E-mail address: maria.ronco@uv.cl

Clay Mathematics Proceedings
Volume 11, 2010

Cyclotomy and Analytic Geometry over $\mathbb{F}_1$

Yuri I. Manin

To Alain Connes, for his sixtieth anniversary

Abstract. Geometry over non–existent “field with one element” $\mathbb{F}_1$ conceived by Jacques Tits [Ti] half a century ago recently found an incarnation, in several related but different guises. In this paper I analyze the crucial role of roots of unity in this geometry and propose a version of the notion of “analytic functions” over F_1. The paper combines a focused survey of various approaches with some new constructions.

Introduction: many faces of cyclotomy

0.1. Roots of unity and the field with one element. The basics of algebraic geometry over an elusive “field with one element F_1” were laid down recently in [So], [De1], [De2], [TV], fifty years after a seminal remark by J. Tits [Ti]. There are many motivations to look for F_1; a hope to imitate Weil's proof for Riemann's zeta is one of them, cf. [CCMa3], [Ku], [Ma1].

An important role in the formalization of F_1–geometry was played by the suggestion made in [KS] that one should simultaneously consider all the “finite extensions” F_{1^n}. This resulted in the approach of [So], where a geometric object, say a scheme, V over F_1, acquired flesh after a base extension to $\mathbf{Z}$, and the F_1– geometry of V was reflected in (and in fact, formally defined in terms of) the geometry of “cyclotomic” points of an appropriate ordinary scheme $V_{\mathbf{Z}}$. In [De1] and [TV], schemes over F_1 are defined in categorical terms independently of cyclotomy, but the latter reappears soon: see the Definition 1.1 below and the following discussion.

All these ideas are interrelated but lead to somewhat different versions of basic definitions, and develop the initial intuition in different directions, so that their divergence can be fruitfully exploited. With this goal in mind, I have chosen the topics to be discussed in sec. 1, where four approaches to the definition of F_1–geometry are sketched and compared.

Of course, roots of unity appear naturally in many different geometric contexts, not motivated by geometry over F_1: some of these contexts are reviewed below in the subsections 0.2–0.6 of this Introduction. I have compiled a sample of them with

2010 *Mathematics Subject Classification.* Primary 14G15; Secondary 11M55, 14A15, 32P05.

an explicit goal: to guess how the insights gained within these contexts could help develop F_1–geometry.

Seemingly, a similar desire moved the authors of [CCMa2] to put the theory of Bost–Connes in the framework of F_1–geometry.

I show in Sec. 2 and Sec. 3 that the results of [Ha1], the preparatory part to K. Habiro's work on Witten–Reshetikhin–Turaev invariants of homology spheres [Ha2], and discoveries about these invariants made in [Law] and [LawZ], can be naturally viewed as a contribution to the rudiments of analytic geometry over F_1.

Finally, in Sec. 4 I discuss Witt vectors and a series of F_1–models of moduli spaces.

Acknowledgements. V. Golyshev's note [Go] prompted me to think about cyclotomy in the F_1 context. K. Habiro read a preliminary version of this paper and suggested several complements and simplifications. H. Lenstra kindly referred me to [van D] and other useful sources on profinite numbers. A. Connes and C. Consani sent me a copy of their new paper [CC] which was being written during the same weeks as the first version this article.

When this first version appeared on the arXiv on Sept. 09, 2008, I received several messages commenting upon and developing the framework involving roots of unity in F_1–geometry.

Matilde Marcolli used the multivariable Habiro ring in [Marc] in order to generalize the Bost–Connes system.

James Borger drew my attention to the fact that my treatment of the cyclotomic coordinates on Witt schemes perfectly matches his remarkable basic idea that "a lambda–ring structure (in the sense of Grothendieck–Riemann–Roch) on a ring R should be thought of as descent data for R from $\mathbf{Z}$ to F_1" (message of Sept. 11, 2008). Borger's approach promises to be a significant breakthrough in our understanding of F_1–geometry, and I have added a brief discussion of it in this new version.

Finally, a totally anonymous referee provided a list of useful remarks and suggestions.

I deeply appreciate their interest and help.

0.2. Roots of unity and Morse–Smale diffeomorphisms. This aspect of cyclotomy is described by D. Grayson in [Gr].

Let M be a compact smooth manifold, f a diffeomorphism of M. It is called *Morse–Smale* if it is structurally stable, and only a finite number of points x are non–wandering. (A point x is called non–wandering if, for any neighborhood U of x, we have $U \cap f^n(U) \neq \emptyset$ for some $n > 0$).

Assume that all eigenvalues of the action of f on the integral cohomology of M are roots of unity and pose the question: *When f is isotopic to a Morse–Smale map?*

There is an obstruction to this, lying in the group $SK_1(\mathcal{R})$, where $\mathcal{R}$ is the ring obtained by localizing $\mathbf{Z}[q]$ with respect to $\Phi_0(q) := q$ and all cyclotomic polynomials

$$\Phi_n(q) := \prod_\eta (q - \eta)$$

where η runs over all primitive roots of unity of degree $n \geq 1$.

This ring turns out to be a principal ideal domain. The reason for this is that each closed point (a prime ideal of depth two) of the "arithmetical plane" $Spec\,\mathbf{Z}[q]$ is situated on an arithmetic curve $\Phi_n(q) = 0, n \geq 0$, because all finite fields consist of roots of unity and zero.

Localization cuts all these curves off, and all closed points go with them. The remaining prime ideals are of height one, and they are principal.

The same effect can be achieved by localizing with respect to all primes $p \in \mathbf{Z}$, thus getting the principal ideal domain $\mathbf{Q}[q]$. This localization excises the closed fibers of the projection $Spec\,\mathbf{Z}[q] \to Spec\,\mathbf{Z}$, and all the closed points with them.

This suggests that the union of cyclotomic arithmetic curves $\Phi_n(q) = 0$ can be imagined as the union of closed fibers of the projection $Spec\,\mathbf{Z}[q] \to Spec\,F_1[q]$, and the arithmetic plane itself as the product of two coordinate axes, the arithmetic one $Spec\,\mathbf{Z}$ and the geometric one, $Spec\,F_1[q]$, over the "absolute point" $Spec\,F_1$.

In Sec. 1 below, I review several versions of algebraic F_1–geometry where this intuition can be made precise.

0.2.1. *Question.* Is there a context in which diffeomorphisms f, acting on the integral cohomology of M with eigenvalues roots of unity, could be interpreted as "Frobenius maps in characteristic 1", and their fixed (or non–wandering) points in a Morse–Smale situation as F_{1^n}–points of an appropriate variety?

0.3. Roots of unity and the Witten–Reshetikhin–Turaev invariants. An apparently totally different line of thought led to the consideration of completions of $\mathbf{Z}[q]$ with respect to various linear topologies generated by the cyclotomic polynomials $\Phi_n(q)$. Namely, it turned out that the invariants of 3–dimensional homology spheres, introduced first by E. Witten by means of path integrals, and mathematically constructed by Reshetikhin and Turaev, can be unified into objects lying in completions of the kind described above.

0.3.1. *Question.* Can these completions be interpreted in a framework of F_1–geometry?

We try to answer this question affirmatively in Sec. 2 and Sec. 3 below.

(Similar completions along the arithmetical axis produce for example direct products of p–adic integers $\prod \mathbf{Z}_{p_i}$ and the ring $\varprojlim_N \mathbf{Z}/(p_1 \dots p_N)$, in which $\mathbf{Z}$ can be embedded.)

We suggest two interpretations, one in the framework of Soulé's axiomatics, and another more in the spirit of Toën–Vaquié and Deitmar's definitions. Here are brief explanations.

Soulé's definition of an F_1–scheme X involves, besides $X_{\mathbf{Z}}$, a $\mathbf{C}$–algebra $\mathcal{A}_X$, and each cyclotomic point of $X_{\mathbf{Z}}$ coming from X must assign "values" to the elements of $\mathcal{A}_X$. His choice of $\mathcal{A}_X$ for the multiplicative group G_{m,F_1} is that of continuous functions on the unit circle in $\mathbf{C}$ (cf. [So], 5.2.2). For the affine line he uses holomorphic functions in the open unit circle continuous on the boundary.

We suggest to consider respectively *the ring of Habiro's analytic functions* and *the ring of Habiro's functions admitting an analytic continuation in the open unit* *disc.* The first one consists of formal series

$$f(q) = a_0 + \sum_{n=1}^{\infty} a_n (1-q)\cdots(1-q^n). \tag{0.1}$$

where the a_n are polynomials in q of degree $\le n-1$. At any root of unity, only a finite number of terms do not vanish, so f is a well defined function on cyclotomic points.

The second option consists in considering holomorphic functions $\varphi(q)$ in the unit circle, such that for any root of unity ζ, a radial limit $\lim_{r\to 1_-} r\zeta$ exists, and the family of such limits can be given by the series (0.1).

Versions of this choice might involve functions holomorphic inside variable rings with outer boundary $|q| = 1$ admitting radial limits at roots of unity, or even pairs of functions φ_-, resp. φ_+, holomorphic inside narrow rings with outer, resp. inner, boundary unit circle, which restrict to a C^∞–function on all small radial intervals $(1-\varepsilon_\zeta, 1+\varepsilon_\zeta)\cdot\zeta$ containing roots of unity ζ. In particular, they must satisfy

$$\lim_{r\to 1_-} \varphi_-\,(r\zeta) = \lim_{r\to 1_+} \varphi_+(r\zeta)$$

The limit values should admit the representation (0.1).

The fact that there exist highly nontrivial and interesting examples of such functions was discovered in the theory of Witten–Reshetikhin–Turaev invariants: cf. [Law], [LawZ]. Don Zagier says that $\varphi_\pm$ "leak through" roots of unity.

On the other hand, if $\mathcal{A}_X$ is not a part of the definition of an F_1–scheme, as in the versions of [TV] and [De1], one can still imagine that a ring of the type discussed above would form a part of the structure of analytic F_1–varieties when F_1–geometry becomes mature enough to include analytic geometry.

[CCMa2] also suggests that time is ripe for such generalizations.

0.4. Roots of unity and the Bost–Connes system. In the paper [BoCo] roots of unity appear in the following setting. Consider the Hecke algebra $\mathcal{H}$ with involution over $\mathbf{Q}$ given by the following presentation. The generators are denoted μ_n, $n \in \mathbf{Z}_+$, and $e(\gamma)$, $\gamma \in \mathbf{Q}/\mathbf{Z}$. The relations are

$$\mu_n^*\mu_n = 1, \quad \mu_{mn} = \mu_m\mu_n, \quad \mu_m^*\mu_n = \mu_n\mu_m^* \text{ for } (m,n) = 1;$$

$$e(\gamma)^* = e(-\gamma), \quad e(\gamma_1+\gamma_2) = e(\gamma_1)\,e(\gamma_2);$$

$$e(\gamma)\mu_n = \mu_n e(n\gamma), \quad \mu_n e(\gamma)\mu_n^* = \frac{1}{n}\sum_{n\delta=\gamma} e(\delta).$$

The idèle class group $\widehat{\mathbf{Z}}^*$ of $\mathbf{Q}$ acts upon $\mathcal{H}$ in a very explicit and simple way: on the $e(\gamma)$'s the action is induced by the multiplication $\widehat{\mathbf{Z}}^* \times \mathbf{Q}/\mathbf{Z} \to \mathbf{Q}/\mathbf{Z}$, whereas on the μ_n's it is the identity.

The algebra $\mathcal{H}$ admits an involutive representation ρ in $l^2(\mathbf{Z}_+)$: denoting by $\{\epsilon_k\}$ the standard basis of this space, we have

$$\rho(\mu_n)\,\epsilon_k = \epsilon_{nk}, \quad \rho(e(\gamma))\epsilon_k = e^{2\pi i k\gamma}\epsilon_k.$$

From this, one can produce the whole Gal$\,(\mathbf{Q}^{ab}/\mathbf{Q})$–orbit $\{\rho_g\}$ of such representations, applying $g \in \mathrm{Gal}\,(\mathbf{Q}^{ab}/\mathbf{Q})$ to all roots of unity occuring on the right hand sides of the expressions for $\rho(e(\gamma))\epsilon_k$. All these representations can be canonically extended to the C^*–algebra completion C of $\mathcal{H}$ constructed from the regular representation of $\mathcal{H}$. Let us denote them by the same symbol ρ_g.

To formulate the main theorem of [BoCo], we need some more explanations. The algebra C admits a canonical action of $\mathbf{R}$, which can be interpreted as *time evolution* represented on the algebra of observables. This is a general (and deep) fact in the theory of C^*–algebras, but for C the action of $\mathbf{R}$ can be quite explicitly described on the generators. Let us denote by σ_t the action of $t \in \mathbf{R}$. *A* KMS$_\beta$ *state at inverse temperature* β on (C, σ_t) is defined as a state φ on C such that for any $x, y \in C$ there exists a bounded holomorphic function $F_{x,y}(z)$ defined in the strip $0 \le \operatorname{Im} z \le \beta$ and continuous on the boundary, satisfying

$$\varphi(x\sigma_t(y)) = F_{x,y}(t),\ \varphi(\sigma_t(y)x) = F_{x,y}(t + i\beta).$$

Now denote by H the positive operator on $l^2(\mathbf{Z}_+)$: $H\epsilon_k = (\log k)\,\epsilon_k$. Then for any $\beta > 1$, $g \in \operatorname{Gal}(\mathbf{Q}^{ab}/\mathbf{Q})$ one can define a KMS$_\beta$ state $\varphi_{\beta,g}$ on (C, σ_t) by the following formula:

$$\varphi_{\beta,g}(x) := \zeta(\beta)^{-1}\operatorname{Trace}(\rho_g(x)\,e^{-\beta H}), \quad x \in C$$

where ζ is the Riemann zeta–function. The map $g \mapsto \varphi_{\beta,g}(x)$ is a homeomorphism of $\operatorname{Gal}(\mathbf{Q}^{ab}/\mathbf{Q})$ with the space of extreme points of the Choquet simplex of all KMS$_\beta$ states.

On the contrary, for $\beta < 1$ there is a unique KMS$_\beta$ state. This is a remarkable "arithmetical symmetry breaking" phenomenon.

The description of the Hecke algebra above involves denominators in the last relation. In [CCMa2], the authors construct $\mathbf{Z}$–models of finite layers of this object and natural morphisms between them, and show that the resulting system is a lift to $\mathbf{Z}$ of an F_1–tower.

This picture is generalized to the multivariable case in [Marc].

0.5. Witt vectors. It is desirable to consider the arithmetical axis *Spec* $\mathbf{Z}$ as an F_1–space as well, but in the current framework it is certainly not a scheme of finite type. In fact, its base extension to $\mathbf{Z}$ is elusive, being precisely what we would like to see as the spectrum of $\mathbf{Z} \otimes \mathbf{Z}$.

Nevertheless, in a certain sense primes can be considered as cyclotomic points of *Spec* $\mathbf{Z}$, at which the "cyclotomic coordinates", all integers, take values that are roots of unity or zero.

In fact, roots of unity of degree $q = p^n - 1$ (and zero), considered together with their embedding into a fixed unramified extension $\mathbf{Z}_p^{nr}$ of $\mathbf{Z}_p$ rather than $\mathbf{C}$, appear as natural coefficients of p–adic expansions discovered by Teichmüller and Witt. Namely, each residue class in $\mathbf{Z}_p^{nr}/(p)$ has a unique (Teichmüller) representative ζ which is either a root of unity or 0 in $\mathbf{Z}_p^{nr}$, so that an element of such an extension can be written as a well-defined series $\sum_{i=0}^{\infty} \zeta_i p^i$. Moreover, coefficients of a sum or a product of two such series are given by Witt's universal polynomials in the coefficients of the summands/factors in the following sense: one must reduce Witt's coefficients modulo p, apply these polynomials (which are defined over $\mathbf{Z}$), and lift the results back to roots of unity.

This can be generalized to the so called "big Witt ring" and interpreted in the following way. On affine spaces $A^k_{\mathbf{Z}} = \mathit{Spec}\,\mathbf{Z}\,[u_1, u_2, \dots, u_k]$ there exists a natural system of “cyclotomic coordinates” (in the p–adic context sometimes called “ghost coordinates”). In terms of these coordinates, one can define an F_1–gadget à la Soulé, requiring that, in the subfunctor of points, these coordinates take cyclotomic values (including zero). However, Witt’s addition/multiplication becomes well defined only after extension to $\mathbf{Z}$, unless the notion of morphism over F_1 is drastically extended.

To me, this looks like a strong argument for considering options for such an extension.

We supply some more details in Sec. 4.

0.6. Roots of unity and the analogy between Hilbert polynomials and zeta functions. There were several suggestions that Hilbert polynomials $H(n)$, say, of graded commutative rings, behave like toy zeta functions.

Rather precise recent observations by V. Golyshev in [Go] can be summarized as follows.

a) The comparison to zetas becomes most striking if one restricts oneself to the following Hilbert polynomials of projective smooth manifolds X:

(i) If X is of general type or Fano: consider $H_{-K_X}(n) := \chi(-nK_X)$.

(ii) If X is a Calabi–Yau manifold embedded as an anticanonical section in a Fano manifold: consider the Euler characteristic of the powers of the induced anticanonical sheaf.

b) With this normalization, Serre duality leads to a functional equation for the Hilbert polynomial of the $s \mapsto -1-s$ type.

c) In many cases, the well known inequalities for Chern numbers of X imply that all roots of $H(s)$ lie in the critical strip $-1 < \mathrm{Re}\, s < 0$, and sometimes even more precise statements. For example, Yau’s inequality $c_1^3 \geq 8/3c_1c_2$ for Fano threefolds shows this fact for them.

0.6.1. *Question.* Is there a systematic relationship between Hilbert polynomials and zeta functions of schemes (or more general spaces) over F_1?

The existence of such zeta functions and their structure in certain cases was heuristically suggested in [Ma1] (cf. also [Ku]). They make precise sense for some specimens in Soulé’s category, and are indeed polynomials; see also [CC] for essential complements. An obvious attack on question 0.6.1 might start with comparing the counting of F_{1^n}–points with counting of monomials in cyclotomic coordinates.

Roots of unity appear in this context via the following beautiful observation due to F. Rodriguez–Villegas [RV].

Consider first an arbitrary polynomial $H(q) \in \mathbf{Z}[q]$. Define another polynomial $P(t)$ such that

$$\sum_{n=0}^{\infty} H(n)t^n = \frac{P(t)}{(1-t)^d}$$

where $P(1) \neq 0$. Let $d := \deg H + 1,\ e := \deg P$.

Rodriguez–Villegas proves that *if all roots of P lie on the unit circle, then $H(q)$ has simple roots at $q = -1, \dots, e+1-d$ and possible additional roots at the middle of this critical strip* $\mathrm{Re}\, q = \dfrac{e-d}{2}$.

This result can be applied to the case when $H(n) = \dim A_n$ is the Hilbert polynomial of a graded algebra $\oplus_{n=0}^{\infty} A_n$ generated by A_1, of Krull dimension d, a complete intersection of polydegree $(n_1, \dots, n_s)$.

It turns out that $e = n_1 + \dots + n_s$ and $P(t) = \prod_{j=1}^{s}(1 + t + \dots + t^{n_j})$, so that all roots of P are in fact roots of unity.

The critical strip in [Go] has width 1, because Golyshev differently normalizes the grading via $-K_X$. However, the Rodriguez–Villegas grading agrees with the motivic philosophy involving weights and Tate's motives over F_1, see [Ma1].

0.7. Summary. We are guided by the following heuristics. Each time that roots of unity appear in a certain context, we try to interpret the functions whose values are these roots of unity as cyclotomic coordinates on a relevant F_1–scheme, in the sense of the Definition 1.1 below, or a version thereof.

An appropriate version of (big) Witt vectors must furnish the basic F_1–analytic (or formal) approximation to the arithmetic line $Spec\,\mathbf{Z}$.

1. Geometry over F_1: generalities

This section sketches and compares four approaches to the definition of F_1–geometry. Preparing a colloquium talk in Paris, I have succumbed to the temptation to associate them with some dominant trends in the history of art.

1.1. Affine schemes over F_1 according to Toën and Vaquié. *(Abstract Expressionism).* Affine schemes over F_1 arise in the most straightforward (and allowing vast generalizations) manner in the framework of [TV], according to which algebraic geometry over F_1 is a special case of algebraic geometry relative to a monoidal symmetric category $(C, \otimes, \mathbf{1})$, which is assumed to be complete, cocomplete, and to admit internal Hom's.

Such a category C gives rise to the category of commutative, associative and unitary monoids $Comm\,(C)$ which serves as a substitute for the category of ordinary commutative rings. Each object A of $Comm\,(C)$ determines the category of A–modules A–Mod consisting of pairs (M, μ) where M is an object of C together with an action $\mu : A \otimes M \to M$ and satisfying the usual formalism.

The opposite category $Aff_C := Comm\,(C)^{opp}$ is called the category of C–affine schemes, and the tautological functor $Comm\,(C) \to Aff_C$ is called *Spec*.

Florian Marty in [Mart2] defines and studies the notion of smoothness in the Toën–Vaquié geometry. This requires passing to the homotopical algebra in appropriate simplicial categories.

According to [TV], we obtain F_1–geometry as the geometry relative to the monoidal category of sets and direct products $(Ens, \times, *)$.

Commutative rings relative to $(Ens, \times, *)$ are just the ordinary commutative (associative, unital) monoids written multiplicatively: this explains the popular motto that to do F_1–algebra one must forget the additive structure: cf. [Har]. This structure is restored when one applies the functor "base change" $\otimes_{F_1}\mathbf{Z}$: a monoid M turns into the commutative associative unital ring $\mathbf{Z}[M]$. The opposite to the monoidal category will be denoted by Aff_{F_1}.

More generally, any commutative ring R determines the base extension functor

$$\otimes_{F_1} R: \ Aff_{F_1} \to Aff_R, \ M \mapsto R[M],$$

from affine schemes over F_1 to affine schemes over $Spec\, R$.

Elements of monoids M will be called *cyclotomic coordinates* on the respective affine scheme. The same term will refer to their images in $R[M]$. On more general schemes, we may speak about local cyclotomic coordinates.

1.2. Deitmar's affine schemes. *(Minimalism).* A. Deitmar in [De1] adopts the same definition of the category Aff_{F_1}. Moreover, he associates to a monoid M a topological space which we will denote $spec\, M$ (to distinguish it from the spectrum of prime ideals of a ring $Spec\,(*)$), and which is endowed with a structure sheaf. Points of this space are prime ideals $P \subset M$: submonoids such that $xy \in P$ implies $x \in P$ or $y \in P$. Basic open sets and the structure (pre)sheaf of monoids are determined via localization, just as in the classical case of commutative rings. Moreover, Deitmar characterizes morphisms in Aff_{F_1} in terms of appropriate morphisms of topological spaces *spec* with structure sheaves.

1.3. Examples. *(i) Affine F_1–schemes associated to abelian groups.* Let M be an abelian group considered as a monoid in *Ens*. We have

$$spec\, M \otimes_{F_1} \mathbf{Z} = Spec\, \mathbf{Z}[M].$$

In particular, [TV] define $\mathbb{F}_{1^n}$ as the monoid (group) $\mathbf{Z}/n\mathbf{Z}$, and its spectrum after lifting to $\mathbf{Z}$ becomes

$$spec\, F_{1^n} \otimes_{F_1} \mathbf{Z} := Spec\, \mathbf{Z}[q]/(q^n - 1) = Spec\, \mathbf{Z}[q]/\{n\}_q. \tag{1.1}$$

In [So], the study of $\mathbf{Z}[q]/\{n\}_q$–points of an a priori given ordinary scheme X gives clues to finding its F_1–forms identified with certain subfunctors of F_{1^n}–points.

In our paper, formula (1.1) motivates the introduction of analytic functions on (certain) F_1–schemes via Habiro's formalism: morally, they are functions that are defined *at all F_{1^n}–points, but nowhere else.* (In fact, the latter stricture should not be taken too literally: some functions have very interesting p–adic, and sometimes complex, arguments and values as well).

(ii). Affine scheme G_{m,F_1}. Over F_1, it is represented by the spectrum of the infinite cyclic group $\mathbf{Z}$. Lifted to $\mathbf{Z}$, it becomes the ordinary $\mathbb{G}_m = Spec\, \mathbf{Z}[q, q^{-1}]$.

1.4. Affine spaces. The affine line $A^1_{F_1}$ is the spectrum of the infinite cyclic monoid $\mathbf{N}$. Its lift to $\mathbf{Z}$ is $A^1_{\mathbf{Z}} := Spec\, \mathbf{Z}[q]$. Similarly, $A^k_{F_1}$ is the spectrum of $\mathbf{N}^k$, $k \geq 1$.

The space $spec\, \mathbf{N}$ consists of one closed point (q) and one generic point.

One can also consider $\mathbf{N}^{\times}$, that is, the free monoid freely generated by all primes. Its lift to $\mathbf{Z}$ is the ring of polynomials in infinitely many variables indexed by primes.

1.5. Affine scheme $GL(n)_{F_1}$. According to [TV], Proposition 4.1, the natural sheaf (in the Grothendieck flat topology, see [TV] and 1.6 below) of automorphisms of a free module of rank n is represented, after lifting to $\mathbf{Z}$, by the semidirect product of $\mathbb{G}_m^n$ and $\mathbf{S}_n$ (the symmetric group).

In the more down–to–earth language of [De1], sec. 5, this is expressed as follows. Let A be a commutative monoid. Define the (set–theoretic) group of "A–valued points of $GL(n)_{F_1}$" as

$$GL(n)_{F_1}(A) := \mathrm{Aut}_A(A^n).$$

This can be identified with the group of (n,n) matrices with entries in $A \subset \mathbf{Z}[A]$, having exactly one non–zero element in each row and each column. This is precisely the description of [TV] quoted above.

The reader should be warned that, contrary to what happened with A^k and $\mathbb{G}_m$, after lifting $GL(n)_{F_1}$ to $\mathbf{Z}$ we *do not* get the usual $GL(n)_{\mathbf{Z}}$. This caused a difficulty in the framework of [So], where it was not obvious how to choose "cyclotomic points of $GL(n)_{\mathbf{Z}}$." In fact, according to [TV], Proposition 4.1, $GL(n)_{\mathbf{Z}}$ for $n > 1$ is not a lift of an F_1–scheme in their sense.

1.6. General schemes over F_1. Glueing general schemes from affine ones is defined differently in [TV] and [De1], respectively.

For Deitmar, an F_1–scheme is a topological space with a sheaf of monoids that is everywhere locally affine, that is, locally isomorphic to some *spec* M.

Toën and Vaquié endow the category Aff_C with a natural Grothendieck topology, which is called *the flat topology.* Using it, one can defined general *schemes* relative to C, as functors that can be obtained from disjoint unions of affine schemes X by taking the quotient with respect to an equivalence relation $R \subset X \times X$ such that projections $R \to X$ are local Zariski isomorphisms. Such schemes form a category denoted $Sch\,(C)$.

Florian Marty in [Mart1] presents a thorough study of the Zariski topology on the category of commutative monoids in C and applies it to the comparison of Deitmar's schemes with Toën–Vaquié's ones.

1.7. Schemes over F_1 à la Soulé. *(Critical Realism).* The idea of Soulé's definitions in [So] can be succinctly formulated as the project of direct reconstruction of F_1–schemas X *of finite type* from certain schemes $X_{\mathbf{Z}}$ over $\mathbf{Z}$ endowed with some kind of *descent data* from $\mathbf{Z}$ to F_1.

However, more than only descent data to F_1 is required: Soulé's spaces come with an additional data $\mathcal{A}_X$ which is a $\mathbf{C}$–algebra, morally an algebra of functions on the "∞–adic completion" of X.

This latter structure embeds F_1–geometry into a wider context, potentially containing also rich structures of Arakelov, or ∞–adic geometry. Some hints that this should be necessary and possible can be glimpsed in the remark made in [Ma1], 1.7. Namely, in [Ma1] it was suggested that the zeta function of $\mathbf{P}^k_{F_1}$ must be $(2\pi)^{-(k+1)}s(s-1)\cdots(s-k)$.

Combining this with Deninger's representation of the basic Euler Γ–factor at arithmetical infinity as a regularized product

$$\Gamma_{\mathbf{C}}(s)^{-1} := \frac{(2\pi)^s}{\Gamma(s)} = \prod_{n\geq 0} \frac{s+n}{2\pi}$$

we see that this gamma–factor should be understood as the zeta–function of the (motivic dual of) an infinite dimensional projective space over F_1.

However, the existing framework is too narrow to make sense of this statement: although the zeta of $\mathbf{P}^k_{F_1}$ is now defined in [So] and agrees with expectations of [Ma1] (up to a power of 2π), the infinite–dimensional case and its connections with ∞–adic geometry still elude us. A promising approach extensively elaborated in the thesis by N. Durov (cf. [Du]) might pave the way to this unification. The treatment of the Bost–Connes dynamical system in [CCMa2] provides another bridge between F_1–geometry and the Archimedean world.

Returning to [So], we will now sketch his version of F_1–schemes of finite type.

The data defining such a scheme X consist of:

(i) A $\mathbf{Z}$*–scheme of finite type* $X_{\mathbf{Z}}$.

(ii) A subfunctor $X(R)$ *of the functor of points of* $X_{\mathbf{Z}}$ *from a category of rings to the category of sets:*

$$X_{\mathbf{Z}}(R) := Hom(Spec\,R, X_{\mathbf{Z}}).$$

Here R runs over rings that are direct summands of $\otimes_i \mathbf{Z}[q]/(\{n_i\}_q)$, and each $X(R)$ is required to be a finite set. We will call elements of $X(R)$ "cyclotomic points".

(iii) A $\mathbf{C}$*–algebra* $\mathcal{A}_X$*, and an assignment of complex values to each element* $f \in \mathcal{A}_X$ *at each pair consisting of a point of* $X(R)$ *and a ring homomorphism* $R \to \mathbf{C}$.

We will not spell out here the compatibility requirements between these data, which are pretty straightforward.

Morphisms of schemes over F_1 are pairs consisting of functor morphisms of cyclotomic points and contravariant homomorphisms of function algebras, compatible with the rest of the data.

It is natural to call an F_1–scheme X affine if $X_{\mathbf{Z}}$ is affine. But, without further restrictions, one would get many schemes over $\mathbf{Z}$ into which the cyclotomic points could be embedded as a subfunctor. The restriction that restores the uniqueness of $X_{\mathbf{Z}}$ once X is known declares that $X_{\mathbf{Z}}$ must be the initial object in the category of such embeddings (see [So], Sec. 4, Definition 3, for a precise statement). A similar universality requirement defines general F_1–schemes (loc. cit., Definition 5).

We will now formally define the notion of *cyclotomic coordinates* on Soulé's F_1–schemes. Let X be an affine F_1–scheme, $X_{\mathbf{Z}} = Spec\,A$.

1.7.1. *Definition.*

DEFINITION 1.1. *A cyclotomic coordinate on the affine* F_1*–scheme* X *in the sense of Soulé is any element* $f \in A$ *whose values at all cyclotomic points* $X(R)$ *are either 0 or roots of unity.*

Clearly, cyclotomic coordinates in this sense form a commutative monoid with unit. If the scheme X is not affine, local cyclotomic coordinates can be defined, forming a (pre)sheaf of commutative monoids.

Recall that in the framework of [TV] and [De1], where $\mathbf{Z}$–lifts of F_1–spaces are patched from spectra of monoid ring $\mathbf{Z}\,[S]$, the elements of S themselves were called cyclotomic coordinates. However, since these versions of F_1–schemes are not equivalent to Soulé's version, we should use this term being aware of its context.

Notice also that

a) To reconstruct cyclotomic coordinates in the sense 1.7.1, it is sufficient to know $X_{\mathbf{Z}}$ and the functor $R \to X(R) \subset X_{\mathbf{Z}}(R)$. This is a part of the structure of *a gadget*, as Soulé's *truc* was translated in [CCMa2]).

b) The rings R used by Soulé to probe schemes over F_1 are essentially group rings of finite abelian groups.

Conceivably, one could replace finite abelian groups by finite commutative unital monoids, thus narrowing the gap between [So] and [TV], [De].

c) Moreover, one could sketch rudiments of *supergeometry* over F_1, by requiring Z_2–grading of our monoids, a structure subgroup $\{\pm 1\}$, and the anticommutation rule for odd elements.

The following example from [So] serves as a good illustration of similarities and differences between affine schemes in the sense of [So] and [TV] respectively, and of relationships of F_1 to Arakelov geometry.

1.7.2. *Arakelov vector bundles over Spec* $\mathbf{Z}$ *as affine* F_1*–schemes.* An Arakelov vector bundle $\bar{\Lambda}$ over *Spec* $\mathbf{Z}$ is defined as a pair consisting of a free abelian group Λ of finite rank and a Hermitian norm $||\cdot||$ on $\Lambda_{\mathbf{C}} := \Lambda \otimes \mathbf{C}$, "integral structure at arithmetical infinity". The global sections of $\bar{\Lambda}$ over the "compactification" *Spec* $\mathbf{Z} \cup \infty$ are defined as $B \cap \Lambda$, where $B := \{x \in \Lambda_{\mathbf{C}} \,|\, ||x|| \le 1\}$.

In order to produce a Soulé affine scheme $X(\bar{\Lambda})$ out of $\bar{\Lambda}$, make an additional choice (on which the final product will not depend): choose a finite subset $\Phi \subset B \cap \Lambda \setminus \{0\}$ such that if $v \in B \cap \Lambda \setminus \{0\}$, then exactly one element of the pair $\{v, -v\}$ belongs to Φ. Let Λ_0 be the sublattice of Λ generated by Φ, Λ_0^t the dual lattice.

Now we can define the structure data.

(i) $X(\bar{\Lambda})_{\mathbf{Z}} := \mathbf{Z}[\Lambda_0^t]$.

(ii) The points of $X(\bar{\Lambda})(R)$ are given by the following prescription:

$$X(\bar{\Lambda})(R) := \{x = \sum_{v \in \Phi} v \otimes \zeta_v \,|\, x \in \Lambda \otimes_{\mathbf{Z}} R,\, \zeta_v \in \mu(R) \cup \{0\}\,\}.$$

Equivalently, coefficients of $v \in \Phi$ are cyclotomic coordinates.

(iii) $\mathcal{A}_{X(\bar{\Lambda})}$ is defined as the algebra of holomorphic functions which are continuous on the boundary of the following domain:

$$C := \{x \in \Lambda_0 \otimes \mathbf{C} \,|\, ||x|| \le \mathrm{card}\,\Phi\}.$$

Given a homomorphism $\sigma :\, R \to \mathbf{C}$, a cyclotomic R–valued point x and a function $f \in \mathbf{Z}[\Lambda_0^t]$, we get its value at (x, σ) in an obvious way.

Comparing this example to the definitions of [TV] and [De1], we see that the algebra $X(\bar{\Lambda})_{\mathbf{Z}} := \mathbf{Z}[\Lambda_0^t]$ fits in their framework, but that other elements of the structure significantly change morphisms and points.

1.7.3. *A non–affine case: toric varieties.* The treatment of this case in [TV], [De1] and [So] leads to essentially the same object (although Soulé produces his $\mathbf{C}$–algebra only in the smooth case, i.e. for regular fans).

Let Δ be a fan. Each element $\sigma \in \Delta$ determines the dual cone σ^*. Let M_σ be the commutative monoid of integer points of σ^*, U_σ is its spectrum as F_1–scheme. If τ is a face of σ, we get a morphism of monoids $M_\sigma \to M_\tau$. The respective morphism of schemes $U_\tau \to U_\sigma$ is Zariski open. Put $X := \coprod_{\sigma\in\Delta} U_\sigma$. According to [TV], 4.2, the quotient of $X \times X$ modulo equivalence relation $R := \coprod_{\sigma,\tau\in\Delta} U_{\sigma\cap\tau}$ defines an F_1–scheme $X(\Delta)_{F_1}$. Lifting it to $\mathbf{Z}$, we get the classical toric scheme $X(\Delta)$.

In [De1], the same quotient is straightforwardly interpreted as a glueing of monoid spectra. In [So] the picture is enhanced by an appropriate $\mathbf{C}$–algebra.

1.8. Tits's problem and Connes–Consani schemes. Tits remarked that one can substitute $q = 1$ in the classical formulas for the number of $\mathbb{F}_q$ points of a projective space $\mathbf{P}^{n-1}$ (resp. Grassmannian $\mathrm{Gr}\,(n, j)$) and get formulas for the cardinality of $\{1, \dots, n\}$ (resp. of the set of subsets of cardinality j in it). Thus we get a version of classical combinatorial projective geometry, in which each line has two points, each plane has three points, etc. Tits asked in [Ti] how to extend this to Chevalley groups and the respective homogeneous spaces: it would be a version of geometry of homogeneous spaces "over a field of characteristic 1" as he put it then.

This project was realized only in 2008, when A. Connes and C. Consani adapted Soulé's definition to this problem in [CC]. Their main innovation consists in considering the functor of cyclotomic points $X(R)$ as taking values in the category of *graded sets*. Only components of degree zero are taken into account in various point counting contexts. After clarifying this issue they find out that Chevalley schemes have F_{1^2} as a natural field of definition, rather than F_1.

1.9. Lambda–rings and Borger's project. *(Futurism).* As I have already mentioned, the key idea of James Borger consists in a totally new conception of $\mathbf{Z}$–to–F_1 descent data: namely, *a restricted λ–ring structure* in the sense of Grothendieck.

According to [BorS], one can think about such a structure on a ring without additive torsion R as a family $\psi_p : R \to R$ of commuting ring endomorphisms indexed by primes such that $\psi_p(x) - x \in pR$ for all x, p.

More generally, as is sketched in [Bor2], we may consider the category of "spaces" $\mathrm{Sp}_{\mathbf{Z}}$, defined as sheaves of sets on the category of affine schemes with the étale topology. It is endowed with the endofunctor W^* of infinite big Witt vectors (cf. the definition in 4.1 below). This endofunctor carries a canonical monad structure. A Λ–structure on a space X is defined as an action of W^* on X. Λ–spaces with W^*–equivariant morphisms form a category $\mathrm{Sp}_{\mathbf{Z}/\Lambda}$. The functor forgetting the Λ–structure is called $v^* : \mathrm{Sp}_{\mathbf{Z}/\Lambda} \to \mathrm{Sp}_{\mathbf{Z}}$. It admits a left adjoint $v_!$ and a right one v_*. The first one must be thought of as (geometric) forgetting the base, and the second one as Weil's restriction of scalars functor.

Using the general topos formalism, Borger looks at the *algebraic geometry of Λ–rings* as a lifted *algebraic geometry over* $\mathbb{F}_1$, represented by the big etale topos over $\mathbf{F}_1$.

In particular, the ring $W(\mathbf{Z})$ of big Witt vectors with entries in $\mathbf{Z}$ should be thought of as (a completed version of) $\mathbf{Z} \otimes_{\mathbf{F}_1} \mathbf{Z}$.

Varieties of finite type over $\mathbf{F}_1$ (in this sense) are very *rigid, combinatorial objects.* They are essentially quotients of toric varieties by toric equivalence relations. In particular, only Tate motives descend to $\mathbf{F}_1$.

Non–finite–type schemes over $\mathbf{F}_1$ are more interesting. The big de Rham–Witt cohomology of X "is" the de Rham cohomology of X "viewed as an $\mathbf{F}_1$–scheme". It should contain the full information of the motive of X and is probably a concrete universal Weil cohomology theory.

The Weil restriction of scalars from $\mathbf{Z}$ to $\mathbf{F}_1$ is an arithmetically global version of Buium's p–jet space.

1.10. A summary. Deitmar's definition of the category of schemes over F_1 is, as he himself stresses in the opening paragraph of [De2], a minimalistic one. It is quite transparent, but obviously does not allow one to treat some more sophisticated situations, such as Soulé's scheme 1.7.2. In fact, Theorem 4.1 of [De2] shows that if X is a connected integral F_1–scheme of finite type, then its lift to $\mathbf{C}$, $X_{\mathbf{C}}$, consists of a finite union of mutually isomorphic toric varieties.

The richness of Toën and Vaquié's definition becomes apparent when it is applied to other basic symmetric monoidal categories. Especially remarkable is the extension of F_1–geometry $\mathbf{S}_1$-*Sch* which is the category of schemes relative to the category $(SEns, \times, *)$ of simplicial sets with direct product. There is a canonical functor "base extension" $\mathbf{S}_1\text{-}Sch \to F_1\text{-}Sch$, so that this geometry lies "below" F_1–geometry, in the same sense as F_1–geometry lies below $\mathbf{Z}$–geometry. Another extension with great promise is the algebraic geometry over "brave new rings".

One outstanding problem is to extend cyclotomy to the homotopical framework.

This is an appropriate place to stress that in a wider context of [TV], or eventually in noncommutative F_1–geometry, the spectrum of F_1 loses its privileged position of a final object of a geometric category. For example, in noncommutative geometry, or in an appropriate category of stacks, the quotient of this spectrum modulo the trivial action of a group must lie below this spectrum.

Soulé's algebras $\mathcal{A}_X$ are a very important element of the structure, in particular, because they form a bridge to Arakelov geometry. Soulé uses concrete choices of them in order to produce a "just right" supply of morphisms, without *a priori* constraining these choices formally.

However, these algebras appear as an *ad hoc* and somewhat arbitrary supplement to the natural F_1–algebraic objects. Perhaps a way to think about them is to imagine *a possible definition of 1–adic numbers.*

Borger's context might lead to progress in this direction.

2. Habiro's analytic functions of many variables

statements of results

2.1. Notations. *Rings* in this and the next sections are associative, commutative and unital, unless the context suggests otherwise. *Ring homomorphisms* are unital. Letters $R, R_0, R_1, \dots$ denote rings, $q, q_0, q_1, \dots$ are independent commuting variables.

Let R be a ring, $\mathcal{I} = \{I_\alpha\}$ a family of ideals filtered by inclusion. The ring projective limit $\varprojlim_\alpha R/I_\alpha$ is called the completion of R with respect to $\mathcal{I}$ and denoted $\widehat{R}_{\mathcal{I}}$ or some version of this notation. When $\mathcal{I}$ is (cofinal to) the family of powers of one ideal I, the respective limit is called the I–adic completion.

We say that R is $\mathcal{I}$– (resp. I–adically) *separated* if $\cap_\alpha I_\alpha = \emptyset$. Equivalently, the canonical homomorphism $R \to \widehat{R}_{\mathcal{I}}$ is injective. Example: $R = \mathbf{Z}$, $\mathcal{I}$ any infinite filtering system.

When q is considered as a "quantization parameter", our quantized (Gaussian) versions of integers and factorials are, as in [Ha2],

$$\{N\}_q := q^N - 1,\ \{N\}_q! := \{N\}_q\{N-1\}_q \dots \{1\}_q. \tag{2.1}$$

Fix an integral domain R_0 of characteristic zero and put $R_n := R_0[q_0, \dots, q_n]$, with natural embeddings $R_0 \subset R_1 \subset R_2 \subset \cdots$

Denote by $I_{n,N} \subset R_n$ the ideal $(\{N\}_{q_1}!, \dots, \{N\}_{q_n}!)$, $N \geq 1$. Clearly, $I_{n,N} \subset I_{n,N+1}$ so that the rings $R_n^{(N)} := R_n/I_{n,N}$, $n \geq 1$ being fixed, form an inverse system.

2.2. Definition.

Definition 2.1. *The ring of Habiro's analytic functions of n variables over R_0 is defined as*

$$\widehat{R}_n := \varprojlim_N R_n^{(N)}.$$

2.3. Taylor series of analytic functions. Choose a vector of roots of unity $\zeta = (\zeta_1, \dots, \zeta_n)$ such that all ζ_i are in R_0. For any integer $M > 0$, there exists $N_0 = N_0(\zeta, M)$ such that $I_{n,N} \subset (q_1 - \zeta_1, \dots, q_n - \zeta_n)^M$ for all $N \geq N_0$. In fact, $\{N\}_{q_i}!$ is divisible by any fixed monomial $(q_i - \zeta)^M, \zeta \in \mu$, if N is large enough.

The completion $\varprojlim_M R_n/(q_1 - \zeta_1, \dots, q_n - \zeta_n)^M$ is $R[[q_1 - \zeta_1, \dots, q_n - \zeta_n]]$.

Therefore we obtain a ring homomorphism "Taylor expansion at the point ζ":

$$T_n(\zeta) : \widehat{R}_n \to R_0[[q_1 - \zeta_1, \dots, q_n - \zeta_n]].$$

2.3.1. *Theorem.*

THEOREM 2.2. *If R_0 is an integral domain, p–adically separated for all primes p, then the same is true for $\widehat{R}_n$, and the Habiro–Taylor homomorphism $T_n(\zeta)$ is injective.*

More generally, let $F = \{\mathbb{F}_1, \dots, \mathbb{F}_n\} \in \mathbf{Z}[q]$ be a family of monic polynomials in $R_0[q]$ all of whose roots are roots of unity. Denote by (F) the ideal generated by $F_1(q_1), \dots, F_n(q_n)$ in R_n. In place of the formal series ring above, we can consider the completion

$$\widehat{R}_F := \varprojlim_M R_n/(F)^M$$

and the respective Taylor expansion homomorphism:

$$T_n(F) : \widehat{R}_n \to \widehat{R}_F.$$

2.3.2. *Theorem.*

THEOREM 2.3. *If R_0 is an integral domain, p–adically separated for all p, $R[[F]]$ is also p–adically separated, and the homomorphism $T_n(F)$ is injective.*

K. Habiro proved these results, as well as their generalizations, for $n = 1$, and we build upon his proof.

2.3.3. *Differential calculus.* Divided powers of partial derivatives with respect to q_k are continuous in the linear topologies generated by $I_{n,N}$, resp. by all $(q_1 - \zeta_1, \dots, q_n - \zeta_n)^M$. Hence these derivatives make sense in $\widehat{R}_n$, and their values at $(\zeta_1, \dots, \zeta_n)$ are the Taylor coefficients of the respective series.

(In order to check the continuity with respect to $I_{n,N}$ it suffices to notice that as N tends to infinity, $\{N\}_q!$ as a polynomial of q vanishes at a growing set of roots of unity with infinitely growing multiplicity at each root of unity. Taking a derivative of such a sequence of polynomials does not destroy this property).

Thus we can develop for $\widehat{R}_n$ the conventional formalism of tangent and cotangent modules, differential forms, etc.

2.4. Elements of $\widehat{R}_n$ as functions on roots of unity. Let $R_0' \supset R_0$ be an integral domain flat over R_0 and containing all roots of unity (that is, all cyclotomic polynomials $q^n - 1$ completely split in R_0'). Denote by μ the set of all roots of unity in R_0'. Choose $\zeta := (\zeta_1, \dots, \zeta_n) \in \mu^n$. Any element of R_n, being a polynomial in $(q_1, \dots, q_n)$, takes a certain value at ζ belonging to R_0'. If $N \geq N_0(\zeta)$, all elements of $I_{n,N}$ vanish at ζ. Hence any element $f \in \widehat{R}_n$ defines a map $\bar{f} : \mu^n \to R_0'$. This map is R_0–linear and compatible with pointwise addition and multiplication of functions.

Besides assuming that R_0 is p–adically separated for all primes p, impose the following separatedness condition: *for any infinite sequence of pairwise distinct primes $p_1, \dots, p_k, \dots,$ we have*

$$\cap_{m=1}^{\infty} R p_1 \cdots p_m = \{0\}. \tag{2.2}$$

2.4.1. *Theorem.*

THEOREM 2.4. *Under these assumptions, the map $f \mapsto \bar{f}$ is injective.*

One can also formulate this statement without adjoining to R_0 roots of unity.

2.4.2. *Theorem.*

THEOREM 2.5. *The natural map $\widehat{R}_n \to \prod_{m=1}^{\infty} \widehat{R}_n \bmod (\Phi_m(q_1), \dots, \Phi_m(q_n))$ is injective.*

For $n = 1$, these results were established by K. Habiro. He has also shown that vanishing of $\bar{f}$ on certain sufficiently large subsets of μ suffices to establish the vanishing of f.

More precisely, *Habiro's topology* on the set μ of all roots of unity is defined as follows (cf. [Ha2], 1.2).

Two roots of unity ξ, η are called *adjacent* if $\xi\eta^{-1}$ is of order p^m, $m \in \mathbf{Z}$, p a prime; or equivalently, if $\xi - \eta$ is not a unit (as an algebraic number). Clearly, the action of $\mathrm{Gal}\,(\overline{\mathbf{Q}}/\mathbf{Q})$ preserves adjacency.

2.4.3. *Definition.*

DEFINITION 2.6. *A subset $U \subset \mu$ is called open if, for any point $\xi \in U$, all except for finitely many $\eta \in \mu$ adjacent to ξ belong to U.*

The Galois action is continuous in this topology, in marked contrast to the topology induced from $\mathbf{C}$.

Let now μ' be an infinite set of roots of unity. A point $\xi \in \mu'$ is a limit point of μ' if for any open neighborhood U of ξ we have $\mu' \cap (U \setminus \xi) \neq \emptyset$. In Habiro's topology, this means that μ' contains infinitely many points adjacent to ξ.

2.4.4. *Theorem.*

THEOREM 2.7. *Under the notations and assumptions of Theorem 2.2, let $\nu = \nu_1 \times \cdots \times \nu_n \subset \mu^n$ be a set, such that each $\nu_i \subset \mu$ has a limit point. Let $f \in \widehat{R}_n$. If the restriction $\bar{f}|_\nu$ is identically zero, then $f = 0$.*

In the next section, we will prove this last result. Theorems 2.4 and 2.5 follow from it.

2.5. Analogs of Habiro's functions on the arithmetic axis and analytic continuation. The Habiro ring of one variable $\varprojlim_N \mathbf{Z}[q]/(\{N\}_q!)$ "is" the lift to $\mathbf{Z}$ of an imaginary ring $\varprojlim_N F_1[q]/(\{N\}_q!)$.

Along the arithmetic axis, the straightforward analog of the latter exists: this is the topological ring of profinite integers $\widehat{\mathbf{Z}} := \varprojlim_N \mathbf{Z}/(N!)$. Its elements can be uniquely represented by infinite series $\sum_{n=1}^{\infty} c_n n!$, where the c_n are integers with $0 \leq c_n \leq n$, cf. [van D].

H. Lenstra in [Le] discusses profinite Fibonacci numbers: continuous extrapolation to $n \in \widehat{\mathbf{Z}}$ of the Fibonacci function $n \mapsto u_n$.

An analog of the profinite number $1 + \sum_{n=1}^{\infty}(-1)^n n!$ is the remarkable example of Habiro function of one variable

$$1 + \sum_{n=1}^{\infty}(-1)^n \{n\}_q! = 1 + \sum_{n=1}^{\infty}(1-q)\cdots(1-q^n).$$

As a function on roots of unity, it emerged in a work of M. Kontsevich on Feynman integrals (talk at MPIM, 1997). Don Zagier in [Za] proved that its values, as well as values of its derivatives, are radial limits of the function (resp. its derivatives) holomorphic in the unit circle

$$\frac{1}{2}\sum_{n=1}^{\infty} n\chi(n)q^{(n^2-1)/24},$$

where χ is the quadratic character of conductor 12.

2.6. Habiro's functions on F_1–schemes. Let X be an F_1–scheme in the sense of one of the definitions from Sec. 1. Let $(x_1,\dots,x_n)$ be a finite family of local cyclotomic coordinates on X. For any ring R as in 1.7. (ii), denote by $U(R)\subset X(R)$ the set of cyclotomic points at which all x_i are defined and take non–zero values.

Consider an analytic function $f\in\widehat{R}_n$ in the sense of Habiro. This function then defines a map

$$f_R: U(R)\to R,\ f_R(r):=\bar{f}(x_1(r),\dots,x_n(r)),$$

with evident functorial properties.

In an appropriate setting such functions must be local sections of a global sheaf. I hope to return to this problem in another paper. Here I will restrict myself to the following observations.

(i) We have to exclude zero values, because q_1 is invertible in $\widehat{R}_1$, and hence each monomial $q_1^{m_1}\cdots q_n^{m_n}$ is invertible in $\widehat{R}_n$. In fact,

$$q^{-1}=1+\sum_{n=1}^{\infty}(-1)^n q^n\{n\}_q!,$$

see [Ha1], Proposition 7.1.

(ii) From the perspective of this paper, it seems quite natural to consider localizations with respect to functions such as $q_1^{m_1}\cdots q_n^{m_n}-1$, deleting sets of roots of unity closed in Habiro's topology. However, such functions are generally not cyclotomic coordinates. This runs counter to the spirit of Toën–Vaqué's definitions, and requires rethinking of their framework.

3. Habiro's analytic functions of many variables

Proofs and generalizations

3.1. The case $n=1$. Assuming that a ring R is $\mathcal{I}$–separated for each member $\mathcal{I}$ of some set of filters $\mathcal{S}_R$, we can deduce that the ring $R[q]$, and certain its completions, are separated with respect to the members of another set of filters, say $\mathcal{S}_{R[T]}$. Results of this type are collected and proved in [Ha1]. They will allow us to perform inductive steps, passing from n to $n+1$,

3.2. Proof of the Theorem 2.2. We will perform induction on n, using Habiro's theorem for $n = 1$ ([Ha1], Theorem 5.2) as the basis of induction.

Assuming the theorem proved for $\widehat{R}_n$, we will proceed by decomposing the Taylor series map $\widehat{R}_{n+1} \to R_0[[q_1 - \zeta_1, \dots, q_n - \zeta_n, q_{n+1} - \zeta_{n+1}]]$ into the product of two ring homomorphisms and checking injectivity of each one:

$$\widehat{R}_{n+1} \xrightarrow{\alpha} \widehat{R}_n[[q_{n+1} - \zeta_{n+1}]] \xrightarrow{\beta} R_0[[q_1 - \zeta_1, \dots, q_n - \zeta_n]]\,[[q_{n+1} - z_{n+1}]]$$

Now we will define the arrows α, β and check their properties.

The arrow β is continuous in $(q_{n+1} - \zeta_{n+1})$–adic topology, acts identically on $q_{n+1} - \zeta_{n+1}$, and sends each element of $\widehat{R}_n$ to its Taylor series at $(\zeta_1, \dots, \zeta_n)$. In view of the inductive assumption, β is injective.

To define α, consider an element $g \in \widehat{R}_{n+1}$. It can be represented as the limit of a sequence of polynomials $g_1, g_2, \dots, g_N, \dots,$ where $g_i \in R_0[q_1, \dots, q_{n+1}]$ such that $g_{N+1} \equiv g_N \bmod I_{n+1,N}$.

From the definition it follows that

$$I_{n+1,N} = I_{n,N}[q_{n+1}] + R_{n+1} \cdot \{N\}_{q_{n+1}}!$$

Therefore,

$$g_{N+1} = g_N + i_N + r_N \cdot \{N\}_{q_{n+1}}!, \tag{3.1}$$

where

$$i_N \in I_{n,N}[q_{n+1}],\ r_N \in R_{n+1}.$$

Now consider a point $(\zeta_1, \dots, \zeta_{n+1})$ as above. Clearly,

$$I_{n,N}[q_{n+1}] = I_{n,N}[q_{n+1} - \zeta_{n+1}].$$

Write $g_N, i_N, \{N\}_{q_{n+1}}!$ as polynomials in $q_{n+1} - \zeta_{n+1}$ with coefficients in R_n. When N becomes large enough, $\{N\}_{q_{n+1}}!$ starts with an arbitrarily large power of $q_{n+1} - \zeta_{n+1}$. Therefore for any given M, the coefficient at $(q_{n+1} - \zeta_{n+1})^M$ in g_{N+1} is the same as in $g_N + i_N$ if $N \geq N_1(M, \zeta)$. Hence the sequence of these coefficients (M being fixed and N growing) converges to a certain element $a_M \in \widehat{R}_n$.

Put $\alpha(g) := \sum_{M=0}^{\infty} a_M (q_{n+1} - \zeta_{n+1})^M$. One can routinely check that $\alpha(g)$ depends only on $g \in \widehat{R}_{n+1}$ and not on the system (g_N) chosen to represent g. Moreover, we get a ring homomorphism

$$\alpha : \widehat{R}_{n+1} \to \widehat{R}_n[[q_{n+1} - \zeta_{n+1}]]. \tag{3.2}$$

Let us check that α is injective. In fact, take a nonzero element $g = \lim g_N$. Then there exist arbitrarily large N such that $g_N \notin I_{n+1,N}$. Representing g_N as a polynomial in $q_{n+1} - \zeta_{n+1}$ with coefficients in R_n, we can find in this polynomial a coefficient, not belonging to $I_{n,N}$. In the limit, it will produce a nonvanishing a_M.

Finally, $\beta \circ \alpha = T_{n+1}(\zeta)$ by construction.

3.3. Proof of the Theorem 2.7. We first remark that the case $n = 1$ is essentially covered by Theorem 6.1 of [Ha1], if one weakens the assumption $R_0 \subset \bar{\mathbf{Q}}$ in the statement of this Theorem. In fact, this assumption is used only at the end of the proof, in order to ensure the validity of the separatedness condition (2.2). Instead, we will simply postulate (2.2). for R_0, and then deduce it for each $\widehat{R}_n$ using the Taylor embedding of R_n into $R_0[[q_1 - \zeta_1, \dots, q_{n+1} - \zeta_{n+1}]]$.

To pass from n to $n+1$, I will start with the following remarks.

Let R be a ring endowed with a filtering family of ideals $\mathcal{I} = \{I_\alpha\}$. Consider the following two families of ideals in the polynomial ring $R[q]$:

(i) $\mathcal{I}_1 := \{I_\alpha[q] + (\{N\}_q!) \,|\, \alpha, N \text{ arbitrary}\}$.

(ii) $\mathcal{I}_2 := \{(\{N\}_q!) \,|\, N \text{ arbitrary}\}$.

Denote by $R[q]^{\widetilde{}}$ (resp. $R[q]^{\widehat{}}$) the completion of $R[q]$ with respect to $\mathcal{I}_1$ (resp. $\mathcal{I}_2$.

For any N and α, we have natural surjections

$$R[q]/(\{N\}_q!) \to R[q]/(I_\alpha[q] + (\{N\}_q!)).$$

Passing to the limit, we get a canonical surjection

$$\varphi : R[q]^{\widehat{}} \to R[q]^{\widetilde{}}.$$

3.3.1. *Lemma.*

LEMMA 3.1. *Consider the case $R = \widehat{R}_n$, $\mathcal{I} = \{\widehat{I}_{n,N})$ where*

$$\widehat{I}_{n,N} := (\{N\}_{q_1}!, \dots, \{N\}_{q_n}!) \subset \widehat{R}_n.$$

Then the homomorphism

$$\varphi : \widehat{R}_n[q_{n+1}]^{\widehat{}} \to \widehat{R}_n[q_{n+1}]^{\widetilde{}} = \widehat{R}_{n+1}$$

is an isomorphism.

PROOF. It suffices to check that $\operatorname{Ker} \varphi = \{0\}$. In fact, as in 3.2, we have an injection

$$\alpha : \widehat{R}_n[q_{n+1}]^{\widetilde{}} \to \widehat{R}_n[[q_{n+1} - 1]]$$

and a one–variable Taylor series injection

$$T : \widehat{R}_n[q_{n+1}]^{\widehat{}} \to \widehat{R}_n[[q_{n+1} - 1]].$$

By construction, $\alpha \circ \varphi = T$, hence φ is an injection as well. □

3.3.2. *End of proof of Theorem 2.7.* Suppose now that $g \in \widehat{R}_{n+1}$ vanishes at all points $(\zeta_1, \dots, \zeta_{n+1})$, $\zeta_i \in \nu_i \subset \mu$, each of ν_i having a limit point. To simplify notation, assume that all roots of unity are in R_0.

The evaluation of g at $(\zeta_1, \dots, \zeta_{n+1})$ can be decomposed into the composition of two arrows:

$$\mathrm{ev}_{(\zeta_1,\dots,\zeta_n)} \circ \mathrm{ev}_{\zeta_{n+1}} : \widehat{R}_{n+1} \to \widehat{R}_n \to R_0,$$

where the first arrow $\mathrm{ev}_{\zeta_{n+1}}$ is obtained by taking the constant term in $\alpha(g)$, (3.2), and the second one is the evaluation at $(\zeta_1, \dots, \zeta_n)$.

First, fix $(\zeta_1, \dots, \zeta_n)$ and vary $\zeta_{n+1} \in \nu_{n+1}$. We have already identified $\widehat{R}_{n+1}$ with $\widehat{R}_n[q_{n+1}]^{\widehat{}}$ in a way which is clearly compatible with evaluation maps.

From the Habiro Theorem 6.1, [Ha1], we obtain that

$$\mathrm{ev}_{(\zeta_1,\dots,\zeta_n)}(g) = 0$$

for all

$$(\zeta_1,\dots,\zeta_n) \in \nu_1 \times \dots \times \nu_n.$$

By the inductive assumption, $g = 0$. This finishes the proof.

3.4. General monoids, coordinate independence, and functorality. Let M be a commutative monoid with unit.

We can consider the completion $R_0'[M]$ of $R_0[M]$ with respect to the system of ideals I_N, where I_N is generated by all elements $\{N\}_m! := (m^n - 1)\cdots(m-1)$ for $m \in M$.

Obviously, any morphism $\psi : M \to N$ induces the respective morphism of the completed rings. In particular, the diagonal morphism $M \to M \times M$ produces a structure of Hopf algebra on $R_0[M]$ and its completed version on $R_0'[M]$.

As K. Habiro noticed in a message to the author (Aug. 23, 2008), applying this construction to $M = \mathbf{Z}^n$, we get precisely $\widehat{R}_n$ (if q_i corresponds to the basic vector $(0,\dots,1,0,\dots,0)$, with 1 at the i–th place.) Since the q_i are invertible in $\widehat{R}_n$, we could as well start with $R_0[q_1, q_1^{-1},\dots,q_n,q_n^{-1}]$, but it seemed more natural to me to deduce the invertibility at the end of the construction.

4. Schemes with natural cyclotomic coordinates

Witt vectors and moduli spaces

In this section we treat two disjoint constructions.

4.1. Witt functors. The (big) Witt ring scheme W can be defined as an infinite dimensional affine space $Spec\,\mathbf{Z}[u_1,u_2,u_3,\dots]$, whose polynomial algebra of functions A is endowed with two homomorphisms $A \to A \otimes A$, "coaddition" α and "comultiplication" μ.

The functor of its R–points, for a variable commutative ring R, set theoretically is $W(R) = \prod_{k=1}^{\infty} R$ where the k–th coordinate of the product is the value of u_k at the respective R–point. The maps α and μ induce on $W(R)$ the structure of commutative ring, functorial in R. This structure can be described quite explicitly, if we use in place of $\{u_k\}$ the "ghost coordinates"

$$q_n := \sum_{d|n} d u_d^{n/d}.$$

In these coordinates, α and μ induce respectively componentwise addition and multiplication (cf. [Haz], sec. 9 and 14, in particular (14.3)).

The N–truncated Witt scheme $W^{(N)}$ is obtained if we apply this to the subring $\mathbf{Z}[u_1,\dots,u_N]$ with induced α_N and μ_N. For a prime p, the scheme W_p is obtained by taking the subring generated by all u_{p^k}, $k \geq 0$. The truncated version $W_p^{(N)}$ jumps only at powers of p as well. In this way we get quotient functors of the Witt functor, valued in commutative algebras.

In place of subrings, one can consider quotients by the ideals generated by the complementary coordinates.

4.1.1. *Definition.*

Definition 4.1. *The (truncated) Witt gadget $\mathcal{W}^{(N)}$ is defined by the following data:*

(i) $\mathcal{W}^{(N)}_{\mathbf{Z}} := W^{(N)}$.

(ii) *For a ring R as in 1.7, (ii), the subfunctor of cyclotomic points $\mathcal{W}^{(N)}_{\mathbf{Z}}(R)$ of $W^{(N)}(R)$ is defined as consisting of points whose ghost coordinates are 0 or roots of unity.*

Thus, ghost coordinates are cyclotomic coordinates in the sense of Definition 1.1

4.2. Moduli spaces $\overline{L}_{0;2,B}$. In an ideal world, not only schemes allowing "finite combinatorial" description ([So], p. 217) must be extensions of objects over F_1, but perhaps "all" rigid structures as well. An obvious challenge is presented by $\overline{M}_{0,n}$, moduli spaces of stable curves of genus zero with n marked points forming the basic operad of quantum cohomology.

As the first approximation, we look in this subsection to some moduli spaces introduced in [LoMa1] and studied further in [LoMa2] and [Ma2]. Generally, they parametrize curves of genus g, with marked points, a part of which (carrying "black" labels) being allowed to merge between them, although not with singular or "white labeled" points. There is an appropriate notion of stability and a representability theorem. (Both were vastly generalized in the study [BaMa].)

Here we will focus on the case of genus zero, two white points and arbitrary (≥ 1) number of black points. The resulting moduli spaces turn out to be toric, based upon permutohedral fans. Therefore they are certainly lifts to $\mathbf{Z}$ of toric F_1–schemes. We discuss which of the canonical morphisms between them descend to F_1.

It is convenient to label the black points by elements of a finite set B, carrying no additional structure (rather than, say, by $\{1,\dots,n\}$, which suggests a complete order on labels).

Below we will give a toric description of the respective moduli space that we will now denote by $\overline{L}_B$. For proofs, see [LoMa1].

4.3. Partitions. A partition $\{\sigma\}$ of a finite set B is *a totally ordered set of non–empty subsets of B whose union is B and whose pairwise intersections are empty.* If a partition consists of N subsets, it is called an N–partition. If its components are denoted $\sigma_1,\dots,\sigma_N$, or otherwise listed, this means that they are listed in their structure order.

Let τ be an $(N+1)$–partition of B. If $N\geq 1$, it determines a well ordered family of N 2–partitions $\sigma^{(a)}$:

$$\sigma^{(a)}_1 := \tau_1\cup\dots\cup\tau_a,\ \sigma^{(a)}_2 := \tau_{a+1}\cup\dots\cup\tau_{N+1},\ a=1,\dots,N\,. \tag{4.1}$$

In the reverse direction, call a family of 2–partitions $(\sigma^{(i)})$ *good* if for any $i\neq j$ we have $\sigma^{(i)}\neq\sigma^{(j)}$ and either $\sigma^{(i)}_1\subset\sigma^{(j)}_1$ or $\sigma^{(j)}_1\subset\sigma^{(i)}_1$. Any good family is naturally well–ordered by the relation $\sigma^{(i)}_1\subset\sigma^{(j)}_1$, and we will consider this ordering as a part of the structure. If a good family of 2–partitions consists of N members, we will usually choose superscripts $1, \dots, N$ to number these partitions in such a way that $\sigma_1^{(i)} \subset \sigma_1^{(j)}$ for $i < j$.

Such a good family produces one $(N+1)$–partition τ:

$$\tau_1 := \sigma_1^{(1)},\ \tau_2 := \sigma_1^{(2)} \setminus \sigma_1^{(1)},\ \dots,\ \tau_N := \sigma_1^{(N)} \setminus \sigma_1^{(N-1)},\ \tau_{N+1} = \sigma_2^{(N)}. \tag{4.2}$$

This correspondence between good N–element families of 2–partitions and $(N+1)$–partitions is one–to–one, because clearly $\sigma_1^{(i)} = \tau_1 \cup \dots \cup \tau_i$ for $1 \le i \le N$.

4.4. The fan F_B. Now we will describe a fan F_B in the space $N_B \otimes \mathbf{R}$, where $N_B := \mathrm{Hom}\,(\mathbf{G}_m, T_B)$, $T_B := \mathbf{G}_m^B/\mathbf{G}_m$. Clearly, N_B can be canonically identified with $\mathbf{Z}^B/\mathbf{Z}$, the latter subgroup being embedded diagonally. Similarly, $N_B \otimes \mathbf{R} = \mathbf{R}^B/\mathbf{R}$. We will write the vectors of this space (resp. lattice) as functions $B \to \mathbf{R}$ (resp. $B \to \mathbf{Z}$) considered modulo constant functions. For a subset $\beta \subset B$, let χ_β be the function equal 1 on β and 0 elsewhere.

4.4.1. *Definition.*

DEFINITION 4.2. *The fan F_B consists of the following l–dimensional cones $C(\tau)$ labeled by $(l+1)$–partitions τ of B.*

If τ is the trivial 1–partition, $C(\tau) = \{0\}$.

If σ is a 2–partition, $C(\sigma)$ is generated by χ_{σ_1}, or, equivalently, $-\chi_{\sigma_2}$, modulo constants.

Generally, let τ be an $(l+1)$–partition, and $\sigma^{(i)}$, $i = 1, \dots, l$, the respective good family of 2–partitions (4.1)*. Then $C(\tau)$ as a cone is generated by all $C(\sigma^{(i)})$.*

4.5. Toric varieties $\overline{L}_B$ and forgetful morphisms. We denote by $\overline{L}_B$ the variety associated with the fan F_B. It is smooth and proper, in fact projective.

Assume that $B \subset B'$. Then we have the projection morphism $\mathbf{Z}^{B'} \to \mathbf{Z}^B$ which induces the morphism $f^{B',B} : N_{B'} \to N_B$. It satisfies the following property: for each cone $C(\tau') \in F_{B'}$, there exists a cone $C(\tau) \in F_B$ such that $f^{B',B}(C(\tau')) \subset C(\tau)$. In fact, τ is obtained from τ' by deleting elements of $B' \setminus B$ and then deleting the empty subsets of the resulting partition of B.

Therefore, we have a morphism $f_*^{B',B} : \overline{\mathcal{L}}_{B'} \to \overline{\mathcal{L}}_B$ which we will call the *forgetful morphism* (it forgets elements of $B' \setminus B$).The forgetful morphism is flat, because locally in toric coordinates it is described as adjoining variables and localization.

4.6. $\overline{L}_B$ as families of curves with two white and B black points. This structure can be defined in terms of forgetful morphisms forgetting just one point B. Let $B \subset B'$, $\mathrm{card}\, B' \setminus B = 1$.

We start with describing structure sections.

In order to define the two white sections of the forgetful morphism, consider two partitions $(B' \setminus B, B)$ and $(B, B' \setminus B)$ of B' and the respective closed strata. The forgetful morphism restricted to these strata identifies them with $\overline{\mathcal{L}}_B$. We will call them x_0 and x_∞ respectively.

Finally, to define the j–th black section, $j \in B$, consider the morphism of lattices $s_j : N_B \to N_{B'}$ which extends a function χ on B to the function $s_j(\chi)$ on B' taking the value $\chi(j)$ at the forgotten point. This morphism satisfies the following condition: each cone $C(\tau)$ from F_B lands in an appropriate cone $C(\tau')$ from $F_{B'}$. Hence we have the induced morphisms $s_{j*} : \overline{\mathcal{L}}_B \to \overline{\mathcal{L}}_{B'}$ which obviously are sections. Moreover, they do not intersect x_0 and x_∞.

4.6.1. *Proposition.*

Proposition 4.3. *With the notations and assumptions above, the forgetful morphism is a universal family of (painted stable) marked curves of genus zero with two white points and B black points.*

In order to see the structure of fibers of the forgetful morphism, one should notice that the inverse image of any point $x \in \mathcal{L}_\tau$ is acted upon by the multiplicative group $\mathbf{G}_m = \mathrm{Ker}\,(T_{B'} \to T_B)$. This action breaks the fiber into a finite number of orbits which coincide with the intersections of this fiber with various $\mathcal{L}_{\tau'}$ described above. When τ' is obtained by adding the forgotten point to one of the parts, this intersection is a torsor over the kernel, otherwise it is a point. As a result, we get that the fiber is a chain of $\mathbf{P}^1$'s, whose components are labeled by the components of τ and singular points by the neighboring pairs of components.

4.7. Clutching morphisms. They are morphisms of the type $\overline{L}_{B_1} \times \overline{L}_{B_2} \to \overline{L}_{B_1 \coprod B_2}$ whose fiberwise description is this: glue ∞ of the first curve to 0 of the second curve. They admit an obvious toric description.

About their operadic role, see [Ma2].

4.8. Proposition.

Proposition 4.4. *Forgetful and clutching morphisms descend to the F_1–models of the toric varieties $\overline{L}_B$.*

References

[BaMa] A. Bayer, Yu. Manin. *Stability conditions, wall–crossing and weighted Gromov–Witten invariants.* Moscow Math. J., 9 (2009), 3–32. e-print math.AG/0607580.

[Bor1] J. Borger. *The basic geometry of Witt vectors.* e-print math.AG/0801.1691

[Bor2] J. Borger. *Λ–rings and the field with one element. Preliminary version.* Preprint, 2009.

[BorS] J. Borger, B. de Smit. *Galois theory and integral models of Λ–rings.* Bull. Lond. Math. Soc., 40 (2008), 439–446. e–print arxiv:0801.2352

[BoCo] J. Bost, A. Connes. *Hecke algebras, type III factors and phase transitions with spontaneous symmetry breaking in number theory.* Selecta Math., 3 (1995), 411–457.

[CC] A. Connes, C. Consani. *On the notion of geometry over $\mathbb{F}_1$.* e–print arxiv:0809.2926

[CCMa1] A. Connes, C. Consani, M. Marcolli. *Noncommutative geometry and motives: the thermodynamics of endomotives.* Advances in Math., 214 (2) (2007), 761–831.

[CCMa2] A. Connes, C. Consani, M. Marcolli. *Fun with $\mathbb{F}_1$.* J. Number Theory, 129 (2009), 1532-1561. e–print math.AG/0806.2401

[CCMa3] A. Connes, C. Consani, M. Marcolli. *The Weil proof and the geometry of the adèles class space.* in “Algebra, arithmetic and Geometry – Manin Festschrift”, Progress in Math., Birkhäuser, 2009, 339–405. e-print math/0703392

[De1] A. Deitmar. *Schemes over $\mathbb{F}_1$.* In: Number Fields and Function Fields – Two Parallel Worlds, 87–104. Ed. by G. van der Geer, B. Moonen, R. Schoof. Progr. in Math, vol. 239, 2005 . e-print math.NT/0404185

[De2] A. Deitmar. *$\mathbb{F}_1$–schemes and toric varieties.* Beiträge Algebra Geom., 49 (2008), 517–525. e-print math.NT/0608179

[De3] A. Deitmar. *Remarks on zeta functions and K–theory over $\mathbb{F}_1$.* Proc. Jap. Ac. Sci, ser. A, vol. 82:8 (2007), 141–146.

[Du] N. Durov. *New approach to Arakelov geometry.* e-print math/0704.2030

[Go] V. Golyshev. *On the canonical strip I.* Uspekhi Mat. Nauk., 64 (2009), 139–140.

[Gr] D. Grayson. *SK_1 of an interesting principal ideal domain.* Journ. of Pure and Appl. Algebra, 20 (1981), 157–163.

[Ha1] K. Habiro. *Cyclotomic completions of polynomial rings.* Publ. RIMS, Kyoto Univ., 40 (2004), 1127–1146.

[Ha2] K. Habiro. *A unified Witten–Reshetikhin–Turaev invariant for integral homology spheres.* Inv. Math., 171 (2008), 1–81.

[Har] S. Haran. *Non–additive geometry.* Comp. Math., 143 (2007), 618–688.

[Haz] M. Hazewinkel. *Witt vectors. Part I.* Handbook of algebra, Vol. 6, 319–470, Elsevier. e-print math/0804.3888

[Ka] M. Kapranov. *Some conjectures on the absolute direct image.* Letter, May 1995.

[KaS] M. Kapranov, A. Smirnov. *Cohomology determinants and reciprocity laws: number field case.* Unpublished manuscript, 15 pp.

[Ku] N. Kurokawa. *Zeta functions over F_1.* Proc. Jap. Ac. Sci., Ser. A, vol. 81 (2005), 180–184.

[Law] R. J. Lawrence. *Witten–Reshetikhin–Turaev invariants of 3–manifolds as holomorphic functions.* In: Geometry and Physics (Aarhus, 1995), LN in Pure and Appl. Math., 184 (1997), Dekker, NY.

[LawZ] R. J. Lawrence, D. Zagier. *Modular forms and quantum invariants of 3–manifolds.* Asian J. Math., 3 (1999), 93–107.

[Le] H. Lenstra. *Profinite Fibonacci numbers.* Nieuw Arch. Wiskd. (5) 6 (2005), no. 4, 297–300.

[LoMa1] A. Losev, Y. Manin. *New moduli spaces of pointed curves and pencils of flat connections.* Fulton's Festschrift, Michigan Journ. of Math., vol. 48, 2000, 443–472. e-print math.AG/0001003

[LoMa2] A. Losev, Y. Manin. *Extended modular operad.* In: Frobenius Manifolds, ed. by C. Hertling and M. Marcolli, Vieweg & Sohn Verlag, Wiesbaden, 2004, 181–211. e-print math.AG/0301003

[Ma1] Yu. Manin. *Lectures on zeta functions and motives (according to Deninger and Kurokawa).* Astérisque 228:4 (1995), 121–163.

[Ma2] Yu. Manin. *Moduli stacks $\overline{L}_{g,S}$.* Moscow Math. Journal, 4:1 (2004), 181–198. e–print math.AG/0206123

[Marc] M. Marcolli. *Cyclotomy and endomotives.* p-adic numbers Ultrametric Anal. Appl., 1 (2009), 217–263. e–print arxiv:0901.3167

[Mart1] F. Marty. *Relative Zariski open objects.* Math. Aq. 0712–3676.

[Mart2] F. Marty. *Smoothness in relative geometry.* e-print math.190812.1152.

[RV] F. Rodriguez–Villegas. *On the zeroes of certain polynomials.* Proc. AMS, 130:8 (2002), 2251–2254.

[Sm1] A. L. Smirnov. *Hurwitz inequalities for number fields. (Russian).* Algebra i Analiz 4 (1992), no. 2, 186–209; translation in St. Petersburg Math. J. 4 (1993), no. 2, 357–375.

[Sm2] A. L. Smirnov. *Absolute determinants and Hilbert symbols.* Preprint MPI 94/72, Bonn, 1994.

[So] C. Soulé. *Les variétés sur le corps à un élément.* Mosc. Math. J. 4:1 (2004), 217–244.

[Ti] J. Tits. *Sur les analogues algébriques des groupes semi–simples complexes.* Colloque d'algèbre supérieure, Centre Belge de Recherches Mathématiques, Établissement Ceuterick, Louvain, 1957, 261–289.

[TV] B. Toën, M. Vaquié. *Au–dessous de Spec* $\mathbf{Z}$. J. K.-Theory, 3 (2009), 437–500. e-print math.AG/0509684

[van D] D. van Danzig. *Nombres universels ou $\nu!$–adiques avec une introduction sur l'algèbre topologique.* Ann. Sci. de l' ÉNS, 3^e série, tome 53 (1936), 275–307.

[Za] D. Zagier. *Vassiliev invariants and a strange identity related to the Dedekind eta–function.* Topology, 40 (2001), 945–960.

Max–Planck–Institut für Mathematik, Bonn, Germany, and Northwestern University, Evanston, USA

Clay Mathematics Proceedings
Volume 11, 2010

Motivic Renormalization and Singularities

Matilde Marcolli

To Alain Connes, on his 60th birthday and many other occasions

Così tra questa
infinità s'annega il pensier mio:
e 'l naufragar m'è dolce in questo mare.
(Giacomo Leopardi, *L'infinito*, second handwritten version)

Abstract. We consider parametric Feynman integrals and their dimensional regularization from the point of view of differential forms on hypersurface complements and the approach to mixed Hodge structures via oscillatory integrals. We consider restrictions to linear subspaces that slice the singular locus, to handle the presence of non-isolated singularities. In order to account for all possible choices of slicing, we encode this extra datum as an enrichment of the Hopf algebra of Feynman graphs. We introduce a new regularization method for parametric Feynman integrals, which is based on Leray coboundaries and, like dimensional regularization, replaces a divergent integral with a Laurent series in a complex parameter. The Connes–Kreimer formulation of renormalization can be applied to this regularization method. We relate the dimensional regularization of the Feynman integral to the Mellin transforms of certain Gelfand–Leray forms and we show that, upon varying the external momenta, the Feynman integrals for a given graph span a family of subspaces in the cohomological Milnor fibration. We show how to pass from regular singular Picard–Fuchs equations to irregular singular flat equisingular connections. In the last section, which is more speculative in nature, we propose a geometric model for dimensional regularization in terms of logarithmic motives and motivic sheaves.

Contents

2010 *Mathematics Subject Classification.* Primary 81Q30; Secondary 14C15, 14C30, 16T05, 32S35, 32S40, 34M35.

1. Introduction

We consider here perturbative quantum field theories governed by a Lagrangian, which in a Lorentzian metric of signature $(+1,-1,\dots,-1)$ on the flat D-dimensional spacetime $\mathbb{R}^D$, is given in the form

$$\mathcal{L}(\phi) = \frac{1}{2}(\partial\phi)^2 - \frac{m^2}{2}\phi^2 - \mathcal{L}_{int}(\phi), \tag{1.1}$$

where the interaction term $\mathcal{L}_{int}(\phi)$ is polynomial in ϕ of degree at least three. The corresponding action functional $S(\phi) = \int \mathcal{L}(\phi) d^D x$ involves a single scalar field ϕ. This is the simplest case, considered in the work of Connes–Kreimer. Generalizations of the Connes–Kreimer formalism for other theories have been developed more recently (see for instance [**40**] for the case of gauge theories), but for the purposes of the present paper we restrict our attention to scalar theories.

The purpose of this paper is to relate the approach to renormalization of Connes–Kreimer [**21**], via Birkhoff factorization of loops in Lie groups of characters of the Hopf algebra of Feynman graphs, and its successive reformulation of Connes–Marcolli [**22**] in terms of Galois theory of a category of flat equisingular connections with irregular singularities, to the approach via parametric Feynman integrals, periods of complements of graph hypersurfaces, and motives, developed by Bloch–Esnault–Kreimer in [**14**], [**13**].

The main approach we follow in this paper, in order to bridge between these two different approaches, is a formulation of the dimensionally regularized Feynman integrals in terms of Mellin transforms of certain Gelfand–Leray forms, as in the approach of Varchenko [**42**], [**43**] to the theory of singularities and asymptotic mixed Hodge structures on the cohomological Milnor fibration, in terms of asymptotic properties of oscillatory integrals.

We deal with the fact that the graph hypersurfaces tend to have non-isolated singularities by slicing the Feynman integral along generic linear spaces of dimension at most equal to the codimension of the singular locus, using the same kind of techniques used in the integral geometry of Radon transforms in projective spaces developed by Gelfand–Gindikin–Graev [**28**]. Since typically the singular locus is rather large in dimension, the slices obtained in this way will often be singular curves in $\mathbb{P}^2$ or singular surfaces in $\mathbb{P}^3$. Instead of considering a single choice of a slicing, which would mean losing too much information on the graph hypersurface, one considers all possible choices and implements the data of the cutting linear space as part of the Hopf algebra of graphs, much like what one does with the choice of the external momenta, so that all possible choices are considered as part of the structure.

The formulation one obtains in this way, in terms of Gelfand–Leray forms, suggests a new method of regularization of parametric Feynman integrals, which, as in the case of dimensional regularization, replaces a divergent integral with a Laurent series in a complex variable ϵ, but which is defined using Leray coboundaries to avoid the singular locus, by integrating around it along the fibers of a circle bundle. We check that the formulation of renormalization in terms of Hopf algebras and

Birkhoff factorization developed in Connes–Kreimer [**21**] applies without changes if one uses this new regularization method instead of the customary dimensional regularization.

The interpretation of the dimensionally regularized Feynman integrals as Mellin transforms of Gelfand–Leray forms provides a direct link between Feynman integrals and the cohomological Milnor fibration. In particular, we prove that, upon varying the external momenta and the spacetime dimension $D \in \mathbb{N}$ in which the scalar theory is considered, the corresponding Feynman integrals determine a family of subspaces of the cohomological Milnor fibration, which inherit a Hodge and a weight filtration from the asymptotic mixed Hodge structure of Varchenko. It remains to be seen when this subspace recovers the full Milnor fiber cohomology and/or when these induced filtrations still define a mixed Hodge structure.

Another important question, in trying to compare the approaches of [**22**] and [**14**], is the use of irregular, as opposed to regular singular, connections. In fact, from the point of view of motives or mixed Hodge structures, what one expects to find is regular singular connections. These appear naturally in the form of Picard–Fuchs equations and Gauss–Manin connections. However, the Galois theory approach to the classification of divergences in perturbative quantum field theory developed in [**22**] relies on the use of irregular singular connections and a form of the Riemann–Hilbert correspondence based on Ramis' wild fundamental group. We reconcile these two approaches by showing that, upon passing to Mellin transforms of solutions of a regular singular Picard–Fuchs equation, one obtains solutions of differential equations with irregular singularities. More precisely, we first recall the construction and properties of the irregular singular connections considered in [**22**] and the equisingularity condition that characterizes them. We then prove that solutions of the regular singular Picard–Fuchs equations at the singular points of a graph hypersurface (sliced with a linear space of a suitable dimension so that singularities are isolated) can be assembled to give rise to a solution of a differential system of the type considered in [**22**], with irregular singularities and with coefficients in the Lie algebra of the affine group scheme of the Hopf algebra of Feynman graphs of the theory, suitably enriched to account for the choice of the slicing of the Feynman integrals by linear spaces of the appropriate dimension.

Finally, we propose a motivic interpretation for dimensional regularization, in terms of the logarithmic extensions of Tate motives (the Kummer motives), and their pullbacks via the polynomial function defining the graph hypersurface. This amounts to associating to the Feynman graphs of a given scalar theory a subcategory of the Arapura category of motivic sheaves of [**3**]. We expect that this may provide a way of interpreting the relation between dimensionally regularized Feynman integrals and cohomological Milnor fibrations in terms of a motivic version of the Milnor fiber. We hope to relate, in this way, a motivic zeta function associated to the resulting mixed motive with the dimensionally regularized Feynman integral.

An in-depth study of parametric Feynman integrals in perturbative renormalization and their relation to mixed Hodge structures was carried out in very recent work of Bloch and Kreimer [**15**]. The topics covered in this paper, as well as other recent developments arising from the collaboration of Aluffi and the author, are reviewed in the recent monograph [**36**].

Acknowledgment. Part of this work was carried out during a stay of the author at Florida State University, where the MAS6396 class provided a good sounding board for various related topics. The author is partially supported by NSF grants DMS-0651925, DMS-0901221, and DMS-1007207.

2. Parametric Feynman integrals

In this section we recall the Feynman parametric formulation of the momentum integrals associated to the Feynman graphs in the perturbative expansion of a scalar field theory. We also recall the Dimensional Regularization method and the form of the regularized integrals. These are all well known techniques, but we review them briefly for completeness. We also recall the explicit form of the graph polynomials $\Psi_\Gamma(t)$ and $P_\Gamma(t,p)$ and their properties, as well as the explicit mass scale dependence of the dimensionally regularized Feynman integrals. Moreover, in §2.4 we give a reformulation of the Feynman integrals in terms of differential forms on hypersurface complements in projective spaces.

2.1. Feynman parameters and algebraic varieties. We recall briefly the method for the computation of Feynman integrals based on the *parametric representation.* This is well known material in the physics literature, see *e.g.* §6-2-3 of [**31**], §18 of [**11**], and §6 of [**37**]. However, since it is not part of the standard mathematician's toolkit, we prefer to spend a few words here recalling the basic ideas.

The terms in the formal asymptotic expansion of functional integrals

$$\int \mathcal{O}(\phi) e^{\frac{i}{\hbar} S(\phi)} \mathcal{D}[\phi],$$

obtained by treating the interaction terms $S_{int}(\phi) = \int \mathcal{L}_{int}(\phi) d^D x$ as a perturbation, are labeled by Feynman graphs of the theory. The topology of these graphs is constrained by the requirement that the valence of each vertex is equal to the degree of one of the monomials in the Lagrangian. The edges of the graph are divided into internal lines, each connecting two vertices, and external lines, which are half-lines with one end attached to a vertex of the graph and one open end. The order in the expansion is given by the loop number of the graph, or by the number of internal lines. Each external line carries a datum of an external momentum $p \in \mathbb{R}^D$ with a conservation law

$$\sum_{e \in E_{ext}(\Gamma)} p_e = 0, \tag{2.1}$$

where $E_{ext}(\Gamma)$ is the set of external edges of Γ.

We assume that all our graphs are one-particle-irreducible (1PI), *i.e.* that they cannot be disconnected by cutting a single internal edge.

The Feynman rules assign to a Feynman graph a function

$$U(\Gamma) = U(\Gamma, p_1 \ldots, p_N)$$

of the external momenta obtained by integrating, over momentum variables k_e assigned to each internal edge of Γ, an expression involving propagators for each internal line and momentum conservations at each vertex, in the form

$$U(\Gamma) = \int \frac{\delta(\sum_{i=1}^n \epsilon_{v,i} k_i + \sum_{j=1}^N \epsilon_{v,j} p_j)}{q_1 \cdots q_n} \, d^D k_1 \cdots d^D k_n, \tag{2.2}$$

where $n = \#E_{int}(\Gamma)$, the number of internal edges in the graph, $N = \#E_{ext}(\Gamma)$ and $\epsilon_{v,e}$ is the incidence matrix

$$\epsilon_{v,e} = \begin{cases} +1 & t(e) = v \\ -1 & s(e) = v \\ 0 & \text{otherwise,} \end{cases} \tag{2.3}$$

with $s(e)$ and $t(v)$ the source and target vertices of the oriented edge e.

The q_i, for $i = 1, \ldots, n$ are the quadratic forms defining the free field propagator associated to the corresponding line in the graph, namely

$$q_i(k) = k_i^2 - m^2 + i\epsilon \quad \text{or} \quad q_i(k) = k_i^2 + m^2, \tag{2.4}$$

respectively in the Lorentzian and in the Eucidean signature case. In the following, we work primarily in the Euclidean setting.

We refer to $U(\Gamma)$ as the *unrenormalized Feynman integral*. The parametric form of $U(\Gamma)$ is obtained by first introducing the *Schwinger parameters*, using the identity

$$\frac{1}{q} = \int_0^\infty e^{-sq} ds.$$

This gives the expression

$$\frac{1}{q_1 \cdots q_n} = \int_0^\infty \cdots \int_0^\infty e^{-(s_1 q_1 + \cdots + s_n q_n)}\; ds_1 \cdots ds_n, \tag{2.5}$$

which is a special case of the more general identity

(2.6)

$$\frac{1}{q_1^{k_1} \cdots q_n^{k_n}} = \frac{1}{\Gamma(k_1) \cdots \Gamma(k_n)} \int_0^\infty \cdots \int_0^\infty e^{-(s_1 q_1 + \cdots + s_n q_n)}\, s_1^{k_1 - 1} \cdots s_n^{k_n - 1}\; ds_1 \cdots ds_n.$$

The Feynman parametric form is obtained from this expression by a change of variables that replaces the Schwinger parameters $s_i \in \mathbb{R}_+$ with new variables $t_i \in [0, 1]$, by setting $s_i = St_i$ with $S = s_1 + \cdots + s_n$. This gives

(2.7)

$$\frac{1}{q_1^{k_1} \cdots q_n^{k_n}} = \frac{\Gamma(k_1 + \cdots + k_n)}{\Gamma(k_1) \cdots \Gamma(k_n)} \int_0^1 \cdots \int_0^1 \frac{t_1^{k_1 - 1} \cdots t_n^{k_n - 1}\, \delta(1 - \sum_{i=1}^n t_i)}{(t_1 q_1 + \cdots + t_n q_n)^{k_1 + \cdots + k_n}}\; dt_1 \cdots dt_n,$$

hence in particular one obtains

$$\frac{1}{q_1 \cdots q_n} = (n-1)! \int \frac{\delta(1 - \sum_{i=1}^n t_i)}{(t_1 q_1 + \cdots + t_n q_n)^n}\; dt_1 \cdots dt_n, \tag{2.8}$$

as an integration in the Feynman parameters $t = (t_i)$ over the simplex

$$\Sigma = \{t = (t_i) \in \mathbb{R}_+^n \mid \sum_i t_i = 1\}. \tag{2.9}$$

Next one introduces a further change of variables involving another matrix naturally associated to the graph, the circuit matrix η_{ik}, defined in terms of an orientation of the edges $e_i \in E(\Gamma)$ and a choice of a basis for the first homology group, $l_k \in H_1(\Gamma, \mathbb{Z})$, with $k = 1, \ldots, \ell = b_1(\Gamma)$, by setting

$$\eta_{ik} = \begin{cases} +1 & \text{edge } e_i \in \text{loop } l_k, \text{ same orientation} \\ -1 & \text{edge } e_i \in \text{loop } l_k, \text{ reverse orientation} \\ 0 & \text{otherwise.} \end{cases} \tag{2.10}$$

We also define $M_\Gamma(t)$ to be the matrix

$$(M_\Gamma)_{kr}(t) = \sum_{i=0}^{n} t_i \eta_{ik} \eta_{ir}. \tag{2.11}$$

Notice that, while the matrix $M_\Gamma(t)$ depends on the choice of the orientation of the edges and of the choice of a basis for the first homology of Γ, the determinant $\det(M_\Gamma(t))$ is independent of both choices.

One then makes a change of variables in the quadratic forms q_i of (2.4), by setting

$$k_i = u_i + \sum_{k=1}^{\ell} \eta_{ik} x_k, \tag{2.12}$$

with the constraint

$$\sum_{i=0}^{n} t_i u_i \eta_{ik} = 0, \tag{2.13}$$

for all $k = 1, \ldots, \ell$. The momentum conservation relations

$$\sum_{i=1}^{n} \epsilon_{v,i} k_i + \sum_{j=1}^{N} \epsilon_{v,j} p_j = 0$$

of (2.2) shows that the u_i in (2.12) also satisfy

$$\sum_{i=1}^{n} \epsilon_{v,i} u_i + \sum_{j=1}^{N} \epsilon_{v,j} p_j = 0. \tag{2.14}$$

This uses the fact that the incidence matrix $\epsilon = (\epsilon_{v,e})$ and the circuit matrix $\eta = (\eta_{e,k})$ satisfy $\epsilon\eta = \sum_{e\in E(\Gamma)} \epsilon_{v,e}\eta_{e,k} = 0$, *cf.* [**37**], §3. The two equations (2.13) and (2.14) are the analogs for momenta in Feynman graphs of the Kirchhoff laws of circuits, respectively giving the conservation laws for the sum of voltage drops along a loop in a circuit and of incoming currents at a vertex, with momenta replacing currents and the Feynman parameters in the role of resistances ([**11**], §18).

The u_i are determined by (2.13) and (2.14), and one can write the term $\sum_i t_i(u_i^2 + m^2)$ in the form of a function of the Feynman parameters t and the external momenta p of the form

$$V_\Gamma(t,p) := p^\tau R_\Gamma(t) p + m^2, \tag{2.15}$$

where we use the fact that $\sum_i t_i = 1$. The $N \times N$ matrix $R(t)$, with $N = \#E_{ext}(\Gamma)$ is constructed out of another matrix associated to the graph. This is obtained as follows (*cf.* [**31**], §6-2-3). Let $D_\Gamma(t)$ denote the matrix

$$(D_\Gamma(t))_{v,v'} = \sum_{i=1}^{n} \epsilon_{v,i}\, \epsilon_{v',i}\, t_i^{-1}, \tag{2.16}$$

with $n = \#E_{int}(\Gamma)$ and with $\epsilon_{v,i}$ the incidence matrix as in (2.3). Then the quadratic form $p^\tau R(t) p$ of (2.15) has the form

$$p^\tau R_\Gamma p = \sum_{v,v'} P_v (D_\Gamma(t)^{-1})_{v,v'} P_{v'}, \tag{2.17}$$

where

$$P_v = \sum_{e \in E_{ext}(\Gamma), t(e)=v} p_e \tag{2.18}$$

is the sum of the incoming external momenta at the vertex v.

Summarizing the previous discussion, the result of the change of variables (2.12) is that we can rewrite the original Feynman integral (2.2) in the following form.

Lemma 2.1. *For $n - D\ell/2 > 0$, the Feynman integral* (2.2) *can be written, after the change of variables* (2.12), *in the form*

$$\int_\Sigma \frac{\delta(1-\sum_i t_i)}{\det(M_\Gamma(t))^{D/2} V_\Gamma(t,p)^{n-D\ell/2}}\, dt_1 \cdots dt_n, \tag{2.19}$$

up to a multiplicative constant.

Proof. This follows [**11**], §18 and [**31**] p.376. First recall the well known identity for the Gaussian integral

$$\int e^{-\frac{1}{2}x^\tau A x} d^D x_1 \cdots d^D x_\ell = \frac{(2\pi)^{D\ell/2}}{\det(A)^{D/2}}, \tag{2.20}$$

for A an $\ell \times \ell$ real symmetric matrix. We then have

$$\frac{1}{(4\pi)^{D\ell/2}} \int e^{-x^\tau A x} d^D x_1 \cdots d^D x_\ell = \det(A)^{-D/2}.$$

With the change of variable (2.12) and the conditions (2.13) and (2.14), one can rewrite the integral $U(\Gamma)$ of (2.2) in the form

$$U(\Gamma) = \int_{\mathbb{R}_+^n} e^{-V_\Gamma(t,p)} \left(\int e^{-x^\tau M_\Gamma(t) x} d^D x_i \cdots d^D x_\ell \right) dt_1 \cdots dt_n, \tag{2.21}$$

where $\ell = b_1(\Gamma)$ is the number of loops in the graph. After performing the Gaussian integration and rewriting the expression in the external momenta as described above in the form (2.15) and (2.17), this becomes of the form

$$U(\Gamma) = (4\pi)^{-\ell D/2} \int_{\mathbb{R}_+^n} \frac{e^{-V_\Gamma(t,p)}}{\Psi_\Gamma(t)^{D/2}}\, dt_1 \cdots dt_n, \tag{2.22}$$

with

$$\Psi_\Gamma(t) = \det M_\Gamma(t). \tag{2.23}$$

Then using the identity $1 = \int_0^\infty d\lambda\, \delta(\lambda - \sum_{i=1}^n t_i)$ and scaling $t_i \mapsto t_i \lambda$, one rewrites (2.22) in the form

(2.24)

$$U(\Gamma) = (4\pi)^{-\ell D/2} \int_0^\infty \left(\int_{[0,1]^n} \delta(1 - \sum_i t_i) \frac{e^{-\lambda V_\Gamma(t,p)}}{\Psi_\Gamma(t)^{D/2}}\, dt_1 \cdots dt_n \right) \lambda^{n-D\ell/2} \frac{d\lambda}{\lambda}.$$

Using again the special form

$$V_\Gamma^{-n+D\ell/2} = \frac{1}{\Gamma(n-D\ell/2)} \int_0^\infty e^{-\lambda V_\Gamma} \lambda^{n-D\ell/2-1}\, d\lambda \tag{2.25}$$

of the general identity (2.6), one then obtains the parametric form

$$U(\Gamma) = \frac{\Gamma(n-D\ell/2)}{(4\pi)^{\ell D/2}} \int_{[0,1]^n} \frac{\delta(1-\sum_i t_i)}{\Psi_\Gamma(t)^{D/2} V_\Gamma(t,p)^{n-D\ell/2}}\, dt_1 \cdots dt_n. \tag{2.26}$$

The condition $n - D\ell/2 = \ell(1 - D/2) + \#V(\Gamma) - 1 > 0$ ensures the convergence at $\lambda = 0$ of the integral (2.25). $\square$

A graph is said to be *log divergent* if $n = D\ell/2$, in which case the Feynman integral reduces to the simpler form

$$\int_\Sigma \frac{\omega}{\det(M_\Gamma(t))^{D/2}}, \tag{2.27}$$

with $\omega = \delta(1 - \sum_i t_i) dt_1 \cdots dt_n$ the volume form on the simplex Σ defined by the integration (2.26).

Remark 2.2. For the purpose of establishing relations between values of Feynman integrals and periods of motives, it is important to check that the multiplicative constant one is neglecting in passing from (2.2) to the parametric form (2.19) in fact belongs to $\mathbb{Q}(\pi)$, *cf.* [**14**]. In (2.26) one sees in fact that the multiplicative constant is of the form $\Gamma(n - D\ell/2)(4\pi)^{-\ell D/2}$. This either gives a divergent factor, at the poles of the Gamma function, in which case one considers the residue, or else, when convergent, it gives a multiplicative factor in $\mathbb{Q}(\pi)$.

The function $\Psi_\Gamma(t) = \det(M_\Gamma(t))$ has an equivalent expression in terms of the connectivity of the graph Γ as the polynomial (see [**31**], §6-2-3 and [**37**] §1.3-2)

$$\Psi_\Gamma(t) = \sum_S \prod_{e \in S} t_e, \tag{2.28}$$

where S ranges over all the sets $S \subset E_{int}(\Gamma)$ of $\ell = b_1(\Gamma)$ internal edges of Γ, such that the removal of all the edges in S leaves a connected graph. This can be equivalently formulated in terms of spanning trees of the graph Γ ([**37**] §1.3), *i.e.* $\Psi_\Gamma(t)$ is given by the Kirchhoff polynomial

$$\Psi_\Gamma(t) = \sum_T \prod_{e \notin T} t_e, \tag{2.29}$$

with the sum over spanning trees T of the graph. Each spanning tree, in fact, has $\#V - 1$ edges and is the complement of a cut-set S.

Lemma 2.3. *The graph polynomial Ψ_Γ is a homogeneous polynomial of degree*

$$\deg \Psi_\Gamma = b_1(\Gamma). \tag{2.30}$$

In the massless case *with $m = 0$, the function $V_\Gamma(t,p)$, for fixed p, is homogeneous of degree one and given by the ratio of a homogeneous polynomial $P_\Gamma(t,p)$ by $\Psi_\Gamma(t)$.*

Proof. We have $\deg \Psi_\Gamma = \#E(\Gamma) - \#E(T)$, where $\#E(T) = \#V(\Gamma) - 1$ is the number of edges in a (hence any) spanning tree, hence from the Euler characteristic formula $\#V(\Gamma) - \#E(\Gamma) = 1 - b_1(\Gamma)$ we get (2.30). We write the polynomial $V_\Gamma(t,p) = p^\tau R_\Gamma(t) p + \sum_i t_i m^2$. In the massless case, using the reformulation given in (6-87) and (6-88) of [**31**], p.297, we rewrite the function $V_\Gamma(t,p)$ in the form of the ratio

$$V_\Gamma(t,p) = \frac{P_\Gamma(t,p)}{\Psi_\Gamma(t)} \tag{2.31}$$

of a homogeneous polynomial P_Γ of degree $\ell + 1 = b_1(\Gamma) + 1$, divided by the polynomial Ψ_Γ, which is homogeneous of degree $b_1(\Gamma)$. In fact, we have ([**31**], §6-2-3)

$$P_\Gamma(p,t) = \sum_{C \subset \Gamma} s_C \prod_{e \in C} t_e, \tag{2.32}$$

where the sum is over the cut-sets $C \subset \Gamma$, *i.e.* the collections of $b_1(\Gamma)+1$ edges that divide the graph Γ into exactly two connected components $\Gamma_1 \cup \Gamma_2$. The coefficient s_C is a function of the external momenta attached to the vertices in either one of the two components:

$$s_C = \left(\sum_{v \in V(\Gamma_1)} P_v \right)^2 = \left(\sum_{v \in V(\Gamma_2)} P_v \right)^2, \tag{2.33}$$

where the P_v are defined as in (2.18), as the sum of the incoming external momenta (see [**31**], (6-87) and (6-88)). □

In the following, we work under the following assumption on the graph Γ.

Definition 2.4. *A 1PI graph Γ satisfies the* generic condition *on the external momenta if, for p in a dense open set in the space of external momenta, the polynomials $P_\Gamma(t,p)$ and $\Psi_\Gamma(t)$ have no common factor.*

To understand better the nature of this condition, it is useful to reformulate the polynomial $P_\Gamma(t,p)$ of (2.32) in terms of spanning trees of the graph. One has, in the case where $m = 0$,

$$P_\Gamma(p,t) = \sum_T \sum_{e' \in T} s_{T,e'}\, t_{e'} \prod_{e \in T^c} t_e, \tag{2.34}$$

where $s_{T,e'} = s_C$ for the cut-set $C = T^c \cup \{e'\}$.

The parameter space of the external momenta is the hyperplane in the affine space $\mathbb{A}^{D \cdot \# E_{ext}(\Gamma)}$ obtained by imposing the conservation law

$$\sum_{e \in E_{ext}(\Gamma)} p_e = 0. \tag{2.35}$$

Thus, the simplest possible configuration of external momenta is the one where one puts all the external momenta to zero, except for a pair $p_{e_1} = p = -p_{e_2}$ associated to a choice of a pair of external edges $\{e_1, e_2\} \subset E_{ext}(\Gamma)$. Let v_i be the unique vertex attached to the external edge e_i of the chosen pair. We then have, in this case, $P_{v_1} = p = -P_{v_2}$. Upon writing the polynomial $P_\Gamma(t,p)$ in the form (2.34), we obtain in this case

$$P_\Gamma(p,t) = p^2 \sum_T \Big(\sum_{e' \in T_{v_1,v_2}} t_{e'} \Big) \prod_{e \notin T} t_e, \tag{2.36}$$

where $T_{v_1,v_2} \subset T$ is the unique path in T without backtrackings connecting the vertices v_1 and v_2. We use (2.33) to get $s_C = p^2$ for all the nonzero terms in this (2.36). These are all the terms that correspond to cut-sets C such that the vertices v_1 and v_2 belong to different components. These cut-sets consist of the complement of a spanning tree T and an edge of T_{v_1,v_2}.

In the following we will make use of the notation

$$L_T(t) = p^2 \sum_{e \in T_{v_1,v_2}} t_e \tag{2.37}$$

for the linear functions in (2.36).

If the polynomial $\Psi_\Gamma(t)$ of (2.29) divides (2.36), one has

$$P_\Gamma(p,t) = \Psi_\Gamma(t) \cdot L(t),$$

for a degree one polynomial $L(t)$, which gives

$$\sum_T (L_T(t) - L(t)) \prod_{e \notin T} t_e \equiv 0,$$

for all t. One then sees, for example, that the 1PI condition on the graph Γ is necessary in order to have the condition of Definition 2.4. In fact, for a graph that is not 1PI, one may be able to find vertices and momenta as above such that the degree one polynomials $L_T(t)$ are all equal to the same $L(t)$. Generally, the validity of the condition of Definition 2.4 can be checked algorithmically for a given graph.

One does not need to assume the condition of Definition 2.4. However, several of our formulae become more complicated if we allow the case where the polynomials Ψ_Γ and $P_\Gamma(t,p)$ have common factors. Thus, for our purposes we assume that we are working under the hypothesis that the "generic condition on the external momenta" holds.

Definition 2.5. *The affine graph hypersurface $\hat{X}_\Gamma$ is the zero locus of the Kirchhoff polynomial*

$$\hat{X}_\Gamma = \{t \in \mathbb{A}^n \,:\, \Psi_\Gamma(t) = 0\}, \tag{2.38}$$

with $n = \#E_{int}(\Gamma)$. The locus of zeros of the polynomial $P_\Gamma(t,p)$, for fixed external momenta p, also defines a hypersurface

$$\hat{Y}_\Gamma = \hat{Y}_\Gamma(p) := \{t \in \mathbb{A}^n \mid P_\Gamma(t,p) = 0\}. \tag{2.39}$$

Since both $\Psi_\Gamma(t)$ and $P_\Gamma(t,p)$ are homogeneous polynomials in t, we can consider the corresponding projective hypersurfaces

$$X_\Gamma = \{t = (t_1 : \cdots : t_n) \in \mathbb{P}^{n-1} \,:\, \Psi_\Gamma(t) = 0\} \tag{2.40}$$

of degree $b_1(\Gamma)$ and

$$Y_\Gamma = Y_\Gamma(p) := \{t = (t_1 : \cdots : t_n) \in \mathbb{P}^{n-1} \mid P_\Gamma(t,p) = 0\}. \tag{2.41}$$

of degree $b_1(\Gamma) + 1$.

In the case of log divergent graphs, or of arbitrary graphs in the range with sufficiently large spacetime dimension D (*i.e.* for D satisfying $-n + D\ell/2 \geq 0$, with $n = \#E_{int}(\Gamma)$ and $\ell = b_1(\Gamma)$), the possible divergences of the Feynman integral $U(\Gamma)$ depend on the intersection of the domain of integration Σ with the graph hypersurface $\hat{X}_\Gamma$ in $\mathbb{P}^{n-1}$. Notice that the intersections $\Sigma \cap \hat{X}_\Gamma$ can only happen on the boundary $\partial\Sigma$, as in the interior of Σ the polynomial Ψ_Γ takes strictly positive real values. See [**14**] and [**13**] for a detailed analysis of this case and for its motivic interpretation. More generally, for non-log-divergent integrals of the form (2.19), outside of the range $-n + D\ell/2 \geq 0$, the singularities of the integral also involve the intersections of the hypersurfaces $\hat{Y}_\Gamma(t,p)$ with the domain of integration Σ.

This case requires in general a more detailed analysis, as in this case some of the intersections may also appear away from the boundary of Σ, depending on the values of the external momenta p, see *e.g.* [**11**], §18.

2.2. Dimensional Regularization. One of the main problems that emerged in the historic development of perturbative quantum field theory is how to "cure" the divergences that occur systematically in the Feynman integrals (2.2), *i.e.* the problem of renormalization. Usually this is treated by choosing a *regularization* method, combined with a *renormalization* procedure. Regularization replaces a divergent integral (2.2) with a function of additional parameters that happens to have a pole or singularity at the special value of the parameter that corresponds to the original integral, but which is otherwise well defined and finite at nearby values of the parameter. Renormalization, on the other hand, gives a method for extracting finite values from the regularized expressions in a way that is consistent with the combinatorics of nested subdivergences, *i.e.* subgraphs of graphs with divergent Feynman integrals, which themselves contribute divergences.

The Connes–Kreimer theory [**21**] uses the regularization method known as *dimensional regularization and minimal subtraction*, combined with the renormalization procedure of Bogoliubov–Parasiuk–Hepp–Zimmermann (BPHZ). It was later shown (see *e.g.* [**27**]) that the main results of Connes–Kreimer may be applied to other regularization procedures, as long as the "subtraction of infinities" can be formulated in terms of a Rota–Baxter operator. The projection of a Laurent series onto its polar part is an example of such an operator, which corresponds to the "minimal subtraction" case. Using this more general formulation, it is possible to extend the Connes–Kreimer theory to other regularization methods, which makes it possible, for instance, to extend it to the case of curved backgrounds as in [**1**]. We concentrate here on the Dimensional Regularization and Minimal Subtraction procedure. In fact, our purpose is to compare the approach to motives and renormalization of [**22**] with the one of [**14**], and we prefer to remain close to the formulation given in [**22**] using DimReg.

Dimensional Regularization consists of formally extending the usual Gaussian integration (2.20) from the case of integer dimension $D \in \mathbb{N}$ to the case of a "complexified dimension" $z \in \mathbb{C}$, in a small neighborhood of $z = 0$, by setting

$$\int e^{-\frac{1}{2}x^\tau Ax} d^{D+z}x_1 \cdots d^{D+z}x_\ell := \frac{(2\pi)^{(D+z)\ell/2}}{\det(A)^{(D+z)/2}}, \tag{2.42}$$

This results in the analytic continuation of the parametric Feynman integral formulae (2.22), (2.24), (2.26) to complex values of the dimension D.

LEMMA 2.6. *Upon replacing the integer dimension D by a complexified dimension $D \mapsto D + z$, with $z \in \Delta^*$ a small punctured disk around $z = 0$, the integral* (2.21) *becomes of the form*

(2.43)
$$U(\Gamma)(z) = \frac{\Gamma(n - \frac{(D+z)\ell}{2})}{(4\pi)^{\frac{\ell(D+z)}{2}}} \int_{[0,1]^n} \frac{\delta(1 - \sum_i t_i)}{\Psi_\Gamma(t)^{(D+z)/2} V_\Gamma(t,p)^{n-(D+z)\ell/2}}\, dt_1 \cdots dt_n.$$

PROOF. One uses the same argument as in Lemma 2.1, but using (2.42) instead of (2.20) in (2.21). This gives

$$U(\Gamma) = (4\pi)^{-\ell(D+z)/2} \int_{\mathbb{R}_+^n} \frac{e^{-V_\Gamma(t,p)}}{\Psi_\Gamma(t)^{(D+z)/2}} \, dt_1 \cdots dt_n, \tag{2.44}$$

We then use the same argument as in Lemma 2.1 to write this in the form

$$U(\Gamma) = (4\pi)^{-\ell(D+z)/2} \int_0^\infty \left(\int_{[0,1]^n} \delta(1 - \sum_i t_i) \frac{e^{-\lambda V_\Gamma(t,p)}}{\Psi_\Gamma(t)^{(D+z)/2}} \, dt_1 \cdots dt_n \right) \lambda^{n-(D+z)\ell/2} \frac{d\lambda}{\lambda} \tag{2.45}$$

and we use

$$V_\Gamma^{-n+(D+z)\ell/2} = \frac{1}{\Gamma(n-(D+z)\ell/2)} \int_0^\infty e^{-\lambda V_\Gamma} \lambda^{n-(D+z)\ell/2-1} \, d\lambda \tag{2.46}$$

to obtain (2.43). One recovers the parametric form (2.19) from (2.42). □

2.3. Mass scale dependence. It is well known that, when one regularizes the integrals $U(\Gamma)$ using dimensional regularization, as recalled above, one introduces an explicit dependence on the mass scale, which plays a very important role in the renormalization process and is the source of the nontrivial action of the renormalization group (see [**20**], [**21**], [**22**], [**23**]).

The source of the mass scale dependence is the fact that, in order to maintain the physical units, the integral (2.42) should in fact be written in the form

$$\int e^{-\frac{1}{2} x^\tau A x} \mu^{-z} d^{D+z} x_1 \cdots \mu^{-z} d^{D+z} x_\ell := \mu^{-z\ell} \frac{(2\pi)^{(D+z)\ell/2}}{\det(A)^{(D+z)/2}}, \tag{2.47}$$

where μ has the physical units of a mass (energy), so that the $\mu^{-z} d^{D+z} x_i$ still have the same physical units as the original $d^D x_i$ (see [**20**]).

LEMMA 2.7. *The dimensional regularization* $U(\Gamma)(z)$ *of* (2.43) *depends on the mass scale* μ *in the form*

$$U_\mu(\Gamma)(z) = \mu^{-z\ell} \frac{\Gamma(n - \frac{(D+z)\ell}{2})}{(4\pi)^{\frac{\ell(D+z)}{2}}} \int_{[0,1]^n} \frac{\delta(1 - \sum_i t_i) dt_1 \cdots dt_n}{\Psi_\Gamma(t)^{\frac{(D+z)}{2}} V_\Gamma(t,p)^{n - \frac{(D+z)\ell}{2}}} . \tag{2.48}$$

PROOF. In the derivation of the parametric form of the Feynman integral with dimensional regularization, we see that we have in (2.44) a mass scale dependence

$$U_\mu(\Gamma)(z) = (4\pi)^{-\ell(D+z)/2} \mu^{-z\ell} \int_{\mathbb{R}_+^n} \frac{e^{-V_\Gamma(t,p)}}{\Psi_\Gamma(t)^{(D+z)/2}} \, dt_1 \cdots dt_n. \tag{2.49}$$

The rest of the argument of Lemma 2.6 is unchanged. In particular, no further μ dependence is introduced by the term in $V_\Gamma(t,p)$, so that we obtain (2.48). □

2.4. Integrals on projective spaces. As remarked above, due to the homogeneity of the polynomials Ψ_Γ and P_Γ, it is natural to regard the graph hypersurfaces as projective hypersurfaces X_Γ and Y_Γ in $\mathbb{P}^{n-1}$, with n the number of internal lines of the graph Γ. Thus, we want to think of the parametric Feynman integrals as being computed in projective space.

In order to reformulate in projective space $\mathbb{P}^{n-1}$ integrals originally defined in affine space $\mathbb{A}^n$, one needs to work with the projective analog (*cf.* [**28**], §II) of the volume form

$$\omega_n = dt_1 \wedge \cdots \wedge dt_n.$$

This is given by the form

$$\Omega = \sum_{i=1}^{n} (-1)^{i+1} t_i \, dt_1 \wedge \cdots \wedge \widehat{dt_i} \wedge \cdots \wedge dt_n. \tag{2.50}$$

The relation between the volume form $dt_1 \wedge \cdots \wedge dt_n$ and the homogeneous form Ω of degree n of (2.50) is given by (*cf.* [**25**], p.180)

$$\Omega = \Delta(\omega_n), \tag{2.51}$$

where $\Delta : \Omega^k \to \Omega^{k-1}$ is the operator of contraction with the Euler vector field

$$E = \sum_i t_i \frac{\partial}{\partial t_i}, \tag{2.52}$$

$$\Delta(\omega)(v_1, \cdots, v_{k-1}) = \omega(E, v_1, \cdots, v_{k-1}). \tag{2.53}$$

In the parametric Feynman integrals, we consider as region of definition of the integrand (in the log divergent case, or in the case of integrals in the range $-n + D\ell/2 \geq 0$) the hypersurface complement

$$\mathcal{D}(\Psi_\Gamma) = \{ t \in \mathbb{A}^n \mid \Psi_\Gamma(t) \neq 0 \} = \mathbb{A}^n \smallsetminus \hat{X}_\Gamma, \tag{2.54}$$

while in the formulation (2.26) outside of the range $-n+D\ell/2 \geq 0$, we also need to avoid the second hypersurface $\hat{Y}_\Gamma$ defined by the vanishing of P_Γ (for assigned external momenta), as in (2.39). In this case the domain of definition of the integrand is

$$\begin{aligned} \mathcal{D}(\Psi_\Gamma, P_\Gamma) = &\quad \{ t \in \mathbb{A}^n \mid \Psi_\Gamma(t) \neq 0 \text{ and } P_\Gamma(t,p) \neq 0 \} \\ = &\quad \mathcal{D}(\Psi_\Gamma) \cap \mathcal{D}(P_\Gamma) = \mathbb{A}^n \smallsetminus (\hat{X}_\Gamma \cup \hat{Y}_\Gamma). \end{aligned} \tag{2.55}$$

Let $\mathcal{U}(\Psi_\Gamma)$ and $\mathcal{U}(\Psi_\Gamma, P_\Gamma)$ denote the corresponding hypersurface complements in projective space, namely

$$\begin{aligned} \mathcal{U}(\Psi_\Gamma) = &\quad \{ t \in \mathbb{P}^{n-1} \mid \Psi_\Gamma(t) \neq 0 \} = \mathbb{P}^{n-1} \smallsetminus X_\Gamma \\ \mathcal{U}(\Psi_\Gamma, P_\Gamma) = &\quad \{ t \in \mathbb{P}^{n-1} \mid \Psi_\Gamma(t) \neq 0 \text{ and } P_\Gamma(t,p) \neq 0 \} \\ = &\quad \mathcal{U}(\Psi_\Gamma) \cap \mathcal{U}(P_\Gamma) = \mathbb{P}^{n-1} \smallsetminus (X_\Gamma \cup Y_\Gamma). \end{aligned} \tag{2.56}$$

As we see in more detail in (2.70) and Proposition 2.9 below, in both the affine and the projective case, we can describe $\mathcal{D}(\Psi_\Gamma, P_\Gamma)$ and $\mathcal{U}(\Psi_\Gamma, P_\Gamma)$ as hypersurface complements, by identifying $X_\Gamma \cup Y_\Gamma$ with the hypersurface defined by the vanishing of a homogeneous polynomial given by a product $\Psi_\Gamma^{n_1} \cdot P_\Gamma^{n_2}$, a homogeneous polynomial of degree $n_1 b_1(\Gamma) + n_2(b_1(\Gamma)+1)$, where the component hypersurfaces X_Γ and Y_Γ are counted with multiplicities n_1 and n_2. These multiplicities depend on the number of edges and loops of the graph and on the spacetime dimension, and are defined more precisely in (2.70) below. Thus, in the following, wherever needed, we write $\mathcal{D}(\Psi_\Gamma, P_\Gamma) = \mathcal{D}(f)$ and $\mathcal{U}(\Psi_\Gamma, P_\Gamma) = \mathcal{U}(f)$, with $f = \Psi_\Gamma^{n_1} \cdot P_\Gamma^{n_2}$, as in the various cases of (2.70) below.

We introduce here some notation that will be useful in the following (*cf.* [**25**], p.177). Let $\mathcal{R} = \mathbb{C}[t_1, \dots, t_n]$ be the ring of polynomials of $\mathbb{A}^n$. Let $\mathcal{R}_m$ denote the subset of homogeneous polynomials of degree m. Similarly, let Ω^k denote the $\mathcal{R}$-module of k-forms on $\mathbb{A}^n$ and let Ω^k_m denote the subset of k-forms that are homogeneous of degree m.

We recall the following general fact (*cf.* [**25**], p.178) about hypersurface complements. Let $\pi : \mathbb{A}^n \smallsetminus \{0\} \to \mathbb{P}^{n-1}$ be the standard projection $t = (t_1, \dots, t_n) \mapsto t = (t_1 : \cdots : t_n)$. Suppose given a homogeneous polynomial function f on $\mathbb{A}^n$ of degree $\deg(f)$. Let $\mathcal{D}(f) \subset \mathbb{A}^n$ and $\mathcal{U}(f) \subset \mathbb{P}^{n-1}$ be the hypersurface complements, *i.e.* the complements, in $\mathbb{A}^n$ and $\mathbb{P}^{n-1}$ respectively, of the locus of zeros $X_f = \{t \,|\, f(t) = 0\}$. With the notation introduced here above, we can always write a form $\omega \in \Omega^k(\mathcal{D}(f))$ as

$$\omega = \frac{\eta}{f^m}, \quad \text{with} \quad \eta \in \Omega^k_{m\deg(f)}. \tag{2.57}$$

We then have the following characterization of the pullback along $\pi : \mathcal{D}(f) \to \mathcal{U}(f)$ of forms on $\mathcal{U}(f)$ (see [**25**], p.180 and [**26**]). Given $\omega \in \Omega^k(\mathcal{U}(f))$, the pullback $\pi^*(\omega) \in \Omega^k(\mathcal{D}(f))$ is characterized by the properties of being invariant under the $\mathbb{G}_m$ action on $\mathbb{A}^n \smallsetminus \{0\}$ and of satisfying $\Delta(\pi^*(\omega)) = 0$, where Δ is the contraction (2.53) with the Euler vector field E of (2.52). Thus, since the sequence

$$0 \to \Omega^n \overset{\Delta}{\to} \Omega^{n-1} \overset{\Delta}{\to} \cdots \overset{\Delta}{\to} \Omega^1 \overset{\Delta}{\to} \Omega^0 \to 0$$

is exact at all but the last term, one can write

$$\pi^*(\omega) = \frac{\Delta(\eta)}{f^m}, \quad \text{with} \quad \eta \in \Omega^k_{m\deg(f)}. \tag{2.58}$$

Thus, in particular, any $(n-1)$-form on $\mathcal{U}(f) \subset \mathbb{P}^{n-1}$ can be written as

$$\frac{h\Omega}{f^m}, \quad \text{with} \quad h \in \mathcal{R}_{m\deg(f)-n} \tag{2.59}$$

and with $\Omega = \Delta(dt_1 \wedge \cdots \wedge dt_n)$ the $(n-1)$-form (2.50), homogeneous of degree n.

PROPOSITION 2.8. *Let $\omega \in \Omega^k_{m\deg(f)}$ be a closed k-form which is homogeneous of degree $m\deg(f)$, and consider the form ω/f^m on $\mathbb{A}^n$. Let $\Sigma \subset \mathbb{A}^n \smallsetminus \{0\}$ be a k-dimensional domain with boundary $\partial\Sigma \neq \emptyset$. Then the integration of ω/f^m on Σ satisfies*

$$m\deg(f) \int_\Sigma \frac{\omega}{f^m} = \int_{\partial\Sigma} \frac{\Delta(\omega)}{f^m} + \int_\Sigma df \wedge \frac{\Delta(\omega)}{f^{m+1}}. \tag{2.60}$$

PROOF. Recall that we have ([**25**], [**26**])

$$d\left(\frac{\Delta(\omega)}{f^m}\right) = -\frac{\Delta(d_f\omega)}{f^{m+1}}, \tag{2.61}$$

where, for a form ω that is homogeneous of degree $m\deg(f)$,

$$d_f\omega = f\, d\omega - m\, df \wedge \omega. \tag{2.62}$$

Thus, we have

$$d\left(\frac{\Delta(\omega)}{f^m}\right) = -\frac{\Delta(d\omega)}{f^m} + m\frac{\Delta(df \wedge \omega)}{f^{m+1}}. \tag{2.63}$$

Since the form ω is closed, $d\omega = 0$, and we have

$$\Delta(df \wedge \omega) = \deg(f)\, f\,\omega - df \wedge \Delta(\omega), \tag{2.64}$$

we obtain from the above

$$d\left(\frac{\Delta(\omega)}{f^m}\right) = m \deg(f)\frac{\omega}{f^m} - \frac{df \wedge \Delta(\omega)}{f^{m+1}}. \tag{2.65}$$

By Stokes' theorem we have

$$\int_{\partial\Sigma} \frac{\Delta(\omega)}{f^m} = \int_{\Sigma} d\left(\frac{\Delta(\omega)}{f^m}\right).$$

Using (2.65) this gives

$$\int_{\partial\Sigma} \frac{\Delta(\omega)}{f^m} = m \deg(f) \int_{\Sigma} \frac{\omega}{f^m} - \int_{\Sigma} \frac{df \wedge \Delta(\omega)}{f^{m+1}}. \tag{2.66}$$

□

We can use this result to reformulate the parametric Feynman integrals in terms of integrals of forms that are pullbacks to $\mathbb{A}^n \smallsetminus \{0\}$ of forms on a hypersurface complement in $\mathbb{P}^{n-1}$. For simplicity, we remove here the divergent Γ-factor from the parametric Feynman integral and we concentrate on the residue given by the integration on the simplex Σ as in (2.67) below.

PROPOSITION 2.9. *Under the generic condition on the external momenta, the parametric Feynman integral*

$$\mathbb{U}(\Gamma) = \int_{\Sigma} \frac{\omega_n}{\Psi_\Gamma^{D/2} V_\Gamma^{n-D\ell/2}} \tag{2.67}$$

can be computed as

$$\mathbb{U}(\Gamma) = \frac{1}{C(n,D,\ell)} \left(\int_{\partial\Sigma} \pi^*(\eta) + \int_{\Sigma} df \wedge \frac{\pi^*(\eta)}{f} \right), \tag{2.68}$$

where $\pi : \mathbb{A}^n \smallsetminus \{0\} \to \mathbb{P}^{n-1}$ is the projection and η is the form on the hypersurface complement $\mathcal{U}(f)$ in $\mathbb{P}^{n-1}$ with

$$\pi^*(\eta) = \frac{\Delta(\omega)}{f^m}, \tag{2.69}$$

on $\mathbb{A}^n$, where

$$f = \begin{cases} P_\Gamma & n - \frac{D(\ell+1)}{2} \geq 0 \\ P_\Gamma^{\frac{2n-D\ell}{2m}} \Psi_\Gamma^{\frac{D}{2m}} & n - \frac{D(\ell+1)}{2} < 0 < n - \frac{D\ell}{2} \\ & m = \gcd\{n - D\ell/2, D/2\} \\ \Psi_\Gamma & n - \frac{D\ell}{2} \leq 0, \end{cases} \tag{2.70}$$

with

$$m = \begin{cases} n - D\ell/2 & n - \frac{D(\ell+1)}{2} \geq 0 \\ \gcd\{n - D\ell/2, D/2\} & n - \frac{D(\ell+1)}{2} < 0 < n - \frac{D\ell}{2} \\ -n + D(\ell+1)/2 & n - \frac{D\ell}{2} \leq 0, \end{cases} \tag{2.71}$$

and with

$$\omega = \begin{cases} \Psi_\Gamma^{n-D(\ell+1)/2}\omega_n & n - \frac{D(\ell+1)}{2} \geq 0 \\ \Psi_\Gamma^{n-D\ell/2}\omega_n & n - \frac{D(\ell+1)}{2} < 0 < n - \frac{D\ell}{2} \\ P_\Gamma^{-n+D\ell/2}\omega_n & n - \frac{D\ell}{2} \leq 0, \end{cases} \tag{2.72}$$

where $\omega_n = dt_1 \wedge \cdots \wedge dt_n$ *on* $\mathbb{A}^n$*, with* $\Omega = \Delta(\omega_n)$ *as in* (2.50)*. The coefficient* $C(n, D, \ell)$ *in* (2.68) *is given by*

$$C(n, D, \ell) = \begin{cases} (n - D\ell/2)(\ell+1) & n - \frac{D(\ell+1)}{2} \geq 0 \\ (n - D\ell/2)\ell + n & n - \frac{D(\ell+1)}{2} < 0 < n - \frac{D\ell}{2} \\ -(n - D(\ell+1)/2)\ell & n - \frac{D\ell}{2} \leq 0. \end{cases} \tag{2.73}$$

Proof. Consider on $\mathbb{A}^n$ the form given by $\Delta(\omega)/f^m$, with f, m, and ω, respectively as in (2.70), (2.71) and (2.72). We assume the condition of Definition 2.4, *i.e.* for a generic choice of the external momenta the polynomials P_Γ and Ψ_Γ have no common factor. First notice that, since the polynomial Ψ_Γ is homogeneous of degree ℓ and P_Γ is homogeneous of degree $\ell+1$, the form $\Delta(\omega)/f^m$ is $\mathbb{G}_m$-invariant on $\mathbb{A}^n \smallsetminus \{0\}$. Moreover, since it is of the form $\alpha = \Delta(\omega)/f^m$, it also satisfies $\Delta(\alpha) = 0$, hence it is the pullback of a form η on $\mathcal{U}(f) \subset \mathbb{P}^{n-1}$. Also notice that the domain of integration $\Sigma \subset \mathbb{A}^n$, given by the simplex $\Sigma = \{\sum_i t_i = 1, \, t_i \geq 0\}$, is contained in a fundamental domain of the action of the multiplicative group $\mathbb{C}^*$ on $\mathbb{C}^n \smallsetminus \{0\}$.

Applying the result of Proposition 2.8 above, we obtain

$$\begin{aligned} \int_\Sigma \frac{dt_1 \wedge \cdots \wedge dt_n}{\Psi_\Gamma^{D/2} V_\Gamma^{n-D\ell/2}} &= \int_\Sigma \frac{\omega}{f^m} \\ = \frac{1}{m \deg(f)} &\left(\int_{\partial\Sigma} \frac{\Delta(\omega)}{f^m} + \int_\Sigma df \wedge \frac{\Delta(\omega)}{f^{m+1}} \right) \\ = C(n, D, \ell)^{-1} &\left(\int_{\partial\Sigma} \frac{\Delta(\omega_n)}{\Psi_\Gamma^{D/2} V_\Gamma^{n-D\ell/2}} + \int_\Sigma df \wedge \frac{\Delta(\omega_n)}{\Psi_\Gamma^a P_\Gamma^b} \right), \end{aligned}$$

where f is as in (2.70) and

$$a = \begin{cases} D(\ell+1)/2 - n & n - \frac{D(\ell+1)}{2} \geq 0 \\ \frac{D}{2}(1 + \frac{1}{m}) & n - \frac{D(\ell+1)}{2} < 0 < n - \frac{D\ell}{2} \\ -n + \frac{D(\ell+1)}{2} + 1 & n - \frac{D\ell}{2} \leq 0, \end{cases} \tag{2.74}$$

$$b = \begin{cases} n - \frac{D\ell}{2} + 1 & n - \frac{D(\ell+1)}{2} \geq 0 \\ (n - \frac{D\ell}{2})(1 + \frac{1}{m}) & n - \frac{D(\ell+1)}{2} < 0 < n - \frac{D\ell}{2} \\ n - \frac{D\ell}{2} & n - \frac{D\ell}{2} \leq 0. \end{cases} \tag{2.75}$$

In fact, the cases of $n - \frac{D(\ell+1)}{2} \geq 0$ and $n - \frac{D\ell}{2} \leq 0$ are clear, while in the range with $n - \frac{D(\ell+1)}{2} < 0$ and $n - \frac{D\ell}{2} > 0$ we have

$$f^{m+1} = P_\Gamma^{(n-D\ell/2)(1+\frac{1}{m})} \Psi_\Gamma^{\frac{D}{2}(1+\frac{1}{m})}.$$

The coefficient $C(n, D, \ell)$ is given by $C(n, D, \ell) = m \deg(f)$, with m and f as in (2.71) and (2.70). Thus, it is given by (2.73), where in the second case we use

$$m((\ell+1)(n - D\ell/2)/m + D\ell/2m) = (n - D\ell/2)\ell + n,$$

for $m = \gcd\{n - D\ell/2, D/2\}$. $\square$

3. Singularities, slicing, and Milnor fiber

3.1. Non-isolated singularities. One problem in trying to use in our setting the techniques developed in singularity theory (*cf.* [**5**]) to study mixed Hodge structures in terms of oscillatory integrals is that the graph hypersurfaces $X_\Gamma \subset \mathbb{P}^{n-1}$ defined by the vanishing of the polynomial $\Psi_\Gamma(t) = \det(M_\Gamma(t))$ usually have non-isolated singularities. This can easily be seen by the following observation.

LEMMA 3.1. *Let Γ be a graph with $\deg \Psi_\Gamma > 2$. The singular locus of X_Γ is given by the intersection of cones over the hypersurfaces X_{Γ_e}, for $e \in E(\Gamma)$, where Γ_e is the graph obtained by removing the edge e of Γ. The cones $C(X_{\Gamma_e})$ do not intersect transversely.*

PROOF. First observe that, since X_Γ is defined by a homogeneous equation $\Psi_\Gamma(t) = 0$, with Ψ_Γ a polynomial of degree m, the Euler formula $m\Psi_\Gamma(t) = \sum_e t_e \frac{\partial}{\partial t_e} \Psi_\Gamma(t)$ implies that $\cap_e Z(\partial_e \Psi_\Gamma) \subset X_\Gamma$, where $Z(\partial_e \Psi_\Gamma)$ is the zero locus of the t_e-derivative. Thus, the singular locus of X_Γ is just given by the equations $\partial_e \Psi_\Gamma = 0$. The variables t_e appear in the polynomial $\Psi_\Gamma(t)$ only with degree zero or one, hence the polynomial $\partial_e \Psi_\Gamma$ consists of only those monomials of Ψ_Γ that contain the variable t_e, where one sets $t_e = 1$. The resulting polynomial is therefore of the form Ψ_{Γ_e}, where Γ_e is the graph obtained from Γ by removing the edge e. In fact, one can see in terms of spanning trees that, if T is a spanning tree containing the edge e then $T \smallsetminus e$ is no longer a spanning tree of Γ_e, so the corresponding terms disappear in passing from Ψ_Γ to Ψ_{Γ_e}, while if T is a spanning tree of Γ which does not contain e, then T is still a spanning tree of Γ_e and the corresponding monomial m_T of Ψ_{Γ_e} is the same as the monomial m_T in Ψ_Γ without the variable t_e. Thus, the zero locus $Z(\Psi_{\Gamma_e}) \subset \mathbb{P}^{n-1}$ is a cone $C(X_{\Gamma_e})$ over the graph hypersurface $X_{\Gamma_e} \subset \mathbb{P}^{n-2}$ with vertex at the coordinate point $v_e = (0, \ldots, 0, 1, 0, \ldots 0)$ with $t_e = 1$. To see that these cones do not intersect transversely, notice that, in the case where $\deg \Psi_\Gamma > 2$, given any two $C(X_{\Gamma_e})$ and $C(X_{\Gamma_{e'}})$ the vertex of one cone is contained in the graph hypersurface spanning the other cone. $\square$

The work of Bergbauer–Rej [**10**] gives a more detailed analysis of the singular locus of the graph hypersurfaces, using a formula for the Kirchhoff polynomials under insertion of subgraphs at vertices.

3.2. Projective Radon transform. Among various techniques introduced for the study of non-isolated singularities, a common procedure consists of cutting the ambient space with linear spaces of dimension complementary to that of the singular locus of the hypersurface (*cf. e.g.* [**41**]). In this case, the restriction of the function defining the hypersurface to these linear spaces defines hypersurfaces with isolated singularities, to which the usual invariants and constructions for isolated singularities can be applied.

One finds that, in typical cases, the graph hypersurfaces have singular locus of codimension one, which means that the slicing is given by planes $\mathbb{P}^2$ intersecting the hypersurface along a curve with isolated singular points. When the singular locus is of codimension two in the hypersurface, the slicing is given by 3-dimensional spaces cutting the hypersurface into a family of surfaces in $\mathbb{P}^3$ with isolated singularities.

In our setting, we are interested in computing integrals of the form (2.67). From this point of view, the procedure of restricting the function defining the hypersurface to linear spaces of a fixed dimension corresponds to an integral transform analogous to a Radon transform in projective space (*cf.* [**28**]).

We recall the basic setting for integral transforms on projective spaces (*cf.* §II of [**28**]). On any k-dimensional subspace $\mathbb{A}^k \subset \mathbb{A}^n$ there is a unique (up to a multiplicative constant) $(k-1)$-form that is invariant under the action of SL_k. It is given as in (2.50) by the expression

$$\Omega_k = \sum_{i=1}^{k} (-1)^{i+1} t_i \, dt_1 \wedge \cdots \wedge \widehat{dt_i} \wedge \cdots \wedge dt_k. \tag{3.1}$$

The form (3.1) is homogeneous of degree k. Suppose given a function f on $\mathbb{A}^n$ which satisfies the homogeneity condition

$$f(\lambda t) = \lambda^{-k} f(t), \quad \forall t \in \mathbb{A}^n, \, \lambda \in \mathbb{G}_m. \tag{3.2}$$

Then the integrand $f\Omega_k$ is well defined on the corresponding projective space $\mathbb{P}^{k-1} \subset \mathbb{P}^{n-1}$ and one defines the integral by integrating on a fundamental domain in $\mathbb{A}^k \smallsetminus \{0\}$, *i.e.* on a surface that intersect each line from the origin once.

Suppose given dual vectors $\xi_i \in (\mathbb{A}^n)'$, for $i = 1, \ldots, n-k$. These define a k-dimensional linear subspace $\Pi = \Pi_\xi \subset \mathbb{A}^n$ by the vanishing condition

$$\Pi_\xi = \{ t \in \mathbb{A}^n \mid \langle \xi_i, t \rangle = 0, \, i = 1, \ldots, n-k \}. \tag{3.3}$$

Given a choice of a subspace Π_ξ, there exists a $(k-1)$-form Ω_ξ on $\mathbb{A}^n$ satisfying

$$\langle \xi_1, dt \rangle \wedge \cdots \wedge \langle \xi_{n-k}, dt \rangle \wedge \Omega_\xi = \Omega_n, \tag{3.4}$$

with Ω_n the $(n-1)$-form of (2.50), *cf.* (3.1). The form Ω_ξ is not uniquely defined on $\mathbb{A}^n$, but its restriction to Π_ξ is uniquely defined by (3.4). Then, given a function f on $\mathbb{A}^n$ with the homogeneity condition (3.2), one can consider the integrand $f\Omega_\xi$ and define its integral on the projective space $\pi(\Pi_\xi) \subset \mathbb{P}^{n-1}$ as above. This defines the integral transform, that is, the $(k-1)$-dimensional projective Radon transform (§II of [**28**]), as

$$\mathcal{F}_k(f)(\xi) = \int_{\pi(\Pi_\xi)} f(t) \, \Omega_\xi(t) = \int_{\mathbb{P}^{n-1}} f(t) \prod_{i=1}^{n-k} \delta(\langle \xi_i, t \rangle) \, \Omega_\xi(t). \tag{3.5}$$

For our purposes, it is convenient to consider also the following variant of the Radon transform (3.5).

DEFINITION 3.2. *Let $\Sigma \subset \mathbb{A}^n$ be a compact region that is contained in a fundamental domain of the action of $\mathbb{G}_m$ on $\mathbb{A}^n \smallsetminus \{0\}$. The partial $(k-1)$-dimensional* *projective Radon transform is given by the expression*

$$\mathcal{F}_{\Sigma,k}(f)(\xi) = \int_{\Sigma\cap\pi(\Pi_\xi)} f(t)\,\Omega_\xi(t) = \int_{\Sigma\cap\pi(\Pi_\xi)} f(t) \prod_{i=1}^{n-k} \delta(\langle \xi_i, t\rangle)\,\Omega_\xi(t), \tag{3.6}$$

where one identifies Σ *with its image* $\pi(\Sigma) \subset \mathbb{P}^{n-1}$.

Let us now return to the parametric Feynman integrals we are considering.

PROPOSITION 3.3. *The Feynman integral* (2.22) *can be reformulated as*

$$U(\Gamma) = \frac{\Gamma(k - \frac{D\ell}{2})}{(4\pi)^{\ell D/2}} \int \mathcal{F}_{\Sigma,k}(f_\Gamma)(\xi)\,\langle \xi, dt\rangle, \tag{3.7}$$

where ξ *is an* $(n-k)$*-frame in* $\mathbb{A}^n$ *and* $\mathcal{F}_{\Sigma,k}(f)$ *is the Radon transform, with* Σ *the simplex* $\sum_i t_i = 1$, $t_i \geq 0$, *and with*

$$f_\Gamma(t) = \frac{V_\Gamma(t,p)^{-k+D\ell/2}}{\Psi_\Gamma(t)^{D/2}}. \tag{3.8}$$

PROOF. Consider first the form (2.22) of the Feynman integral, which we write equivalently as

$$U(\Gamma) = (4\pi)^{-\ell D/2} \int_{\mathbb{A}^n} \chi_+(t) \frac{e^{-V_\Gamma(t,p)}}{\Psi_\Gamma(t)^{D/2}}\, dt_1 \cdots dt_n, \tag{3.9}$$

where $\chi_+(t)$ is the characteristic function of the domain $\mathbb{R}^n_+$.

Given a choice of an $(n-k)$-frame ξ, we can then write the Feynman integrals in the form

$$U(\Gamma) = (4\pi)^{-\ell D/2} \int \left(\int_{\Pi_\xi} \chi_+(t) \frac{e^{-V_\Gamma(t,p)}}{\Psi_\Gamma(t)^{D/2}} \omega_\xi \right) \langle \xi, dt\rangle, \tag{3.10}$$

where $\langle \xi, dt\rangle$ is a shorthand notation for

$$\langle \xi, dt\rangle = \langle \xi_1, dt\rangle \wedge \cdots \wedge \langle \xi_{n-k}, dt\rangle$$

and ω_ξ satisfies

$$\langle \xi, dt\rangle \wedge \omega_\xi = \omega_n = dt_1 \wedge \cdots \wedge dt_n. \tag{3.11}$$

We then apply the same procedure as in (2.24) and (2.25) to the integral on Π_ξ and write it in the form

$$\int_{\Pi_\xi} \chi_+(t) \frac{e^{-V_\Gamma(t,p)}}{\Psi_\Gamma(t)^{D/2}} \omega_\xi(t) = \Gamma(k - \frac{D\ell}{2}) \int_{\Pi_\xi} \delta(1 - \sum_i t_i)\, \frac{\omega_\xi(t)}{\Psi_\Gamma(t)^{D/2} V_\Gamma(t,p)^{k-D\ell/2}}. \tag{3.12}$$

The function $f_\Gamma(t)$ of (3.8) satisfies the scaling property (3.2) and the integrand

$$\frac{\omega_\xi(t)}{\Psi_\Gamma(t)^{D/2} V_\Gamma(t,p)^{k-D\ell/2}}$$

is therefore $\mathbb{G}_m$-invariant, since the form ω_ξ is homogeneous of degree k. Moreover, the domain Σ of integration is contained in a fundamental domain for the action of $\mathbb{G}_m$. Thus, we can reformulate the integral (3.12) in projective space in terms of the Radon transform as

$$\Gamma(k - \frac{D\ell}{2})\,(4\pi)^{-\ell D/2} \int \mathcal{F}_{\Sigma,k}(f_\Gamma)(\xi)\,\langle \xi, dt\rangle, \tag{3.13}$$

where $\mathcal{F}_{\Sigma,k}(f_\Gamma)$ is the Radon transform over the simplex Σ, as in Definition 3.2. $\square$

In the following, we will then consider integrals of the form

$$\mathbb{U}(\Gamma)_\xi = \mathcal{F}_{\Sigma,k}(f_\Gamma)(\xi) = \int_{\Pi_\xi} \delta(1-\sum_i t_i)\, \frac{\omega_\xi(t)}{\Psi_\Gamma(t)^{D/2} V_\Gamma(t,p)^{k-D\ell/2}} \tag{3.14}$$

$$= \int_{\Sigma_\xi} \frac{\omega_\xi(t)}{\Psi_\Gamma(t)^{D/2} V_\Gamma(t,p)^{k-D\ell/2}}$$

as well as their dimensional regularizations

$$\mathbb{U}(\Gamma)_\xi(z) = \int_{\Sigma_\xi} \frac{\omega_\xi(t)}{\Psi_\Gamma(t)^{(D+z)/2} V_\Gamma(t,p)^{k-(D+z)\ell/2}}, \tag{3.15}$$

where Π_ξ is a generic linear subspace of dimension equal to the codimension of the singular locus of the hypersurface $X_\Gamma \cup Y_\Gamma$.

3.3. The polar filtration. As we recalled already in §2.4 above (*cf.* [**25**]) algebraic differential forms $\omega \in \Omega^k(\mathcal{D}(f))$ on a hypersurface complement can always be written in the form $\omega = \eta/f^m$ as in (2.57), for some $m \in \mathbb{N}$ and some $\eta \in \Omega^k_{m\deg(f)}$. The minimal m such that ω can be written in the form $\omega = \eta/f^m$ is called the order of pole of ω along the hypersurface X and is denoted by $\mathrm{ord}_X(\omega)$. The order of pole induces a filtration, called the *polar filtration*, on the de Rham complex of differential forms on the hypersurface complement. One denotes by $P^r\Omega^k_{\mathbb{P}^n} \subset \Omega^k_{\mathbb{P}^n}$ the subspace of forms of order $\mathrm{ord}_X(\omega) \leq k-r+1$, if $k-r+1 \geq 0$, or $P^r\Omega^k = 0$ for $k-r+1 < 0$. The polar filtration $P^\bullet$ is related to the Hodge filtration $F^\bullet$ by $P^r\Omega^m \supset F^r\Omega^m$, by a result of [**24**].

PROPOSITION 3.4. *Under the generic condition on the external momenta, the forms*

$$\frac{\Omega_\xi}{\Psi_\Gamma^{D/2} V_\Gamma^{k-D\ell/2}} \tag{3.16}$$

span subspaces $P^{r,k}_\xi$ of the polar filtration $P^r\Omega^{k-1}_{\mathbb{P}^{n-1}}$ of a hypersurface complement $\mathcal{U}(f) \subset \mathbb{P}^{n-1}$, where

$$f = \begin{cases} P_\Gamma & k - D(\ell+1)/2 \geq 0 \\ P_\Gamma^{(k-D\ell/2)/m}\Psi_\Gamma^{D/(2m)} & k-D(\ell+1)/2 < 0 < k - D\ell/2, \\ & m = \gcd\{k-D\ell/2, D/2\} \\ \Psi_\Gamma & k - D\ell/2 \leq 0, \end{cases} \tag{3.17}$$

and for the index r of the filtration in the range

$$\begin{cases} r \leq D\ell/2 & k - D(\ell+1)/2 \geq 0 \\ r \leq k - \gcd\{k-D\ell/2, D/2\} & k - D(\ell+1)/2 < 0 < k - D\ell/2 \\ r \leq 2k - D(\ell+1)/2 & k - D\ell/2 \leq 0. \end{cases} \tag{3.18}$$

PROOF. We are assuming that P_Γ and Ψ_Γ have no common factor, for generic external momenta. Consider first the case where $k - D\ell/2 \geq 0$. This is further divided into two cases: the case where also $k - D(\ell+1)/2 \geq 0$ and the case where $k - D(\ell+1)/2 < 0$. In the first case, the form (3.16) can be written, using (2.31), as

$$\frac{\Delta(\alpha)}{f^m} = \frac{\Psi_\Gamma^{k-D(\ell+1)/2}\Omega_\xi}{P_\Gamma^{k-D\ell/2}}, \tag{3.19}$$

where

$$\alpha = \Psi_\Gamma^{k-D(\ell+1)/2}\omega_\xi \quad \text{and} \quad f = P_\Gamma, \quad \text{with} \quad m = k - D\ell/2. \tag{3.20}$$

Thus, in this case we consider the polar filtration for differential forms on the complement of the projective hypersurface Y_Γ of degree $\ell+1$ defined by $P_\Gamma = 0$. The forms (3.19), for a generic choice of the $(n-k)$-frame ξ, and for varying external momenta p, span a subspace $P_\xi^{r,k}$ of the polar filtration $P^r\Omega^{k-1}_{\mathbb{P}^{n-1}}$, for all $r \leq D\ell/2$. Notice that $r \leq D\ell/2$ also implies $r \leq k$ so that one remains within the nontrivial range $k - r \geq 0$ of the filtration.

In the case where we still have $k - D\ell/2 \geq 0$ but $k - D(\ell+1)/2 < 0$, we let

$$m = \gcd\{k - D\ell/2, D/2\},$$

so that $k - D\ell/2 = n_1 m$ and $D/2 = n_2 m$. We then write (3.16) in the form

$$\frac{\Delta(\alpha)}{f^m} = \frac{\Psi_\Gamma^{k-D\ell/2}\Omega_\xi}{P_\Gamma^{k-D\ell/2}\Psi_\Gamma^{D/2}}, \tag{3.21}$$

with

$$\alpha = \Psi_\Gamma^{k-D\ell/2}\omega_\xi, \quad \text{and} \quad f = P_\Gamma^{n_1}\Psi_\Gamma^{n_2} \quad \text{and} \quad m = \gcd\{k - D\ell/2, D/2\}. \tag{3.22}$$

In this case, we consider the polar filtration associated to the complement of the projective hypersurface defined by the equation $P_\Gamma^{n_1}\Psi_\Gamma^{n_2} = 0$. For a generic choice of the $(n-k)$-frame ξ, and for varying external momenta p, we obtain in this case a subspace $P_\xi^{r,k}$ of the polar filtration $P^r\Omega^{k-1}_{\mathbb{P}^{n-1}}$, for all $r \leq k - \gcd\{k - D\ell/2, D/2\}$.

The remaining case is when $k - D\ell/2 < 0$, so that also $k - D(\ell+1)/2 < 0$. In this case, we write (3.16) in the form

$$\frac{\Delta(\alpha)}{f^m} = \frac{P_\Gamma^{-k+D\ell/2}\Omega_\xi}{\Psi_\Gamma^{-k+D(\ell+1)/2}}, \tag{3.23}$$

where

$$\alpha = P_\Gamma^{-k+D\ell/2}\omega_\xi, \quad \text{and} \quad f = \Psi_\Gamma \quad \text{and} \quad m = -k + D(\ell+1)/2. \tag{3.24}$$

We are considering here the polar filtration on forms on the complement of the hypersurface X_Γ defined by $\Psi_\Gamma = 0$. We then obtain, for generic ξ and varying p, a subspace of $P_\xi^{r,k}$ of the filtration $P^r\Omega^{k-1}_{\mathbb{P}^{n-1}}$, for all $r \leq 2k - D(\ell+1)/2$. $\square$

3.4. Milnor fiber. Suppose then that $k = \operatorname{codim} \operatorname{Sing}(X)$, where $\operatorname{Sing}(X)$ is the singular locus of the hypersurface $X = \{f = 0\}$, with f as in Proposition 3.4 above. In this case, for generic ξ, the linear space Π_ξ cuts the singular locus $\operatorname{Sing}(X)$ transversely and the restriction $X_\xi = X \cap \Pi_\xi$ has isolated singularities.

Recall that, in the case of isolated singularities, there is an isomorphism between the cohomology of the Milnor fiber F_ξ of X_ξ and the total cohomology of the Koszul–de Rham complex of forms (2.57) with the total differential $d_f \omega = f\, d\omega - m\, df \wedge \omega$ as above. The explict isomorphism is given by the Poincaré residue map and can be written in the form

$$[\omega] \mapsto [j^* \Delta(\omega_\xi)], \tag{3.25}$$

where $j : F_\xi \hookrightarrow \Pi_\xi$ is the inclusion of the Milnor fiber in the ambient space (see [**25**], §6).

Let $M(f)$ be the Milnor algebra of f, *i.e.* the quotient of the polynomial ring in the coordinates of the ambient projective space by the ideal of the derivatives of f. When f has isolated singularities, the Milnor algebra is finite dimensional. One denotes by $M(f)_m$ the homogeneous component of degree m of $M(f)$.

It then follows from the identification (3.25) above ([**25**],§6.2) that, in the case of isolated singularities, a basis for the cohomology $H^r(F_\xi)$ of the Milnor fiber, with $r = \dim \Pi_\xi - 1$ is given by elements of the form

$$\omega_\alpha = \frac{t^\alpha \Delta(\omega_\xi)}{f^m}, \qquad \text{with} \qquad t^\alpha \in M(f)_{m \deg(f) - k}, \tag{3.26}$$

where f is the restriction to Π_ξ of the function of (3.17). We then have the following consequence of Proposition 3.4.

Corollary 3.5. *For a generic $(n-k)$-frame ξ with $n-k = \dim \operatorname{Sing}(X)$, with X the hypersurface of Proposition 3.4, and for a fixed generic choice of the external momenta p under the assumption of Definition 2.4, the Feynman integrand* (3.16) *of* (3.14) *defines a cohomology class in $H^r(F_\xi)$, with $r = \dim \Pi_\xi - 1$ and $F_\xi \subset \Pi_\xi$ the Milnor fiber of the hypersurface with isolated singularities $X_\xi = X \cap \Pi_\xi \subset \Pi_\xi$.*

Proof. By Proposition 3.4, the form (3.16) can be written as

$$\frac{h \Delta(\omega_\xi)}{f^m}, \tag{3.27}$$

where f is as in (3.17), and h is a polynomial of the form

$$h = \begin{cases} \Psi_\Gamma^{k - D(\ell+1)/2} & k - D(\ell+1)/2 \geq 0 \\ \Psi_\Gamma^{k - D\ell/2} & k - D(\ell+1)/2 < 0 < k - D\ell/2 \\ P_\Gamma^{-k + D\ell/2} & k - D\ell/2 \leq 0. \end{cases} \tag{3.28}$$

Let $\mathcal{I}_\xi$ denote the ideal of derivatives of the restriction $f|_{\Pi_\xi}$ of f to Π_ξ. Then let

$$h_\xi = h \mod \mathcal{I}_\xi. \tag{3.29}$$

For a fixed generic choice of the external momenta, this defines an element in the Milnor algebra $M(f|_{\Pi_\xi})$, which lies in the homogeneous component

$$M(f|_{\Pi_\xi})_{m \deg(f) - k},$$

for

$$m = \begin{cases} k - D\ell/2 & k - D(\ell+1)/2 \geq 0 \\ \gcd\{k - D\ell/2, D/2\} & k - D(\ell+1)/2 < 0 < k - D\ell/2 \\ -k + D(\ell+1)/2 & k - D\ell/2 \leq 0. \end{cases} \tag{3.30}$$

Thus, the form (3.27) defines a class in the cohomology $H^r(F_\xi)$ with

$$r = \dim \Pi_\xi - 1.$$

□

3.5. The Feynman integral: slicing. As in Proposition 2.9, we can reformulate the integral (3.14) in terms of integrals of pullbacks of forms on a hypersurface complement in projective space, using the explicit description of Proposition 3.4 above.

PROPOSITION 3.6. *The integral* (3.14) *can be computed in the form*

$$\mathbb{U}(\Gamma)_\xi = \frac{1}{C(k,D,\ell)} \left(\int_{\partial\Sigma\cap\Pi_\xi} \pi^*(\eta_\xi) + \int_{\Sigma\cap\Pi_\xi} df|_{\Pi_\xi} \wedge \frac{\pi^*(\eta_\xi)}{f|_{\Pi_\xi}} \right), \tag{3.31}$$

where $\pi : \mathbb{A}^n \smallsetminus \{0\} \to \mathbb{P}^{n-1}$ *is the projection and* η_ξ *satisfies*

$$\pi^*(\eta_\xi) = \frac{h|_{\Pi_\xi}\Omega_\xi}{(f|_{\Pi_\xi})^m} \tag{3.32}$$

on $\mathbb{A}^n$*, where* Ω_ξ *is given by* (3.4) *and* f*,* m *and* h *are as in Proposition 3.4 and Corollary 3.5. The coefficient* $C(k,D,\ell)$ *is given as in* (2.73).

PROOF. As in the case of Proposition 2.9, the result follows by applying Proposition 2.8 and Proposition 3.4, together with the fact that $\Omega_\xi = \Delta(\omega_\xi)$, which can be seen by writing

$$\Omega_\xi = \prod_{i=1}^{n-k} \delta(\langle \xi_i, t \rangle)\, \Omega_n.$$

The coefficient $C(k,D,\ell)$ is given by $C(k,D,\ell) = m \deg(f)$, with m and f as in (3.30) and (3.17). □

4. Oscillatory integrals: Leray and Dimensional Regularizations

A well known method for studying integrals of holomorphic forms on vanishing cycles of a singularity and to relate these to mixed Hodge structures is via oscillatory integrals and their asymptotic expansion (see [**4**] and Vol.II of [**5**]). Our main result in this section will be to show that the dimensionally regularized parametric Feynman integrals can be related to the Mellin transform of a Gelfand–Leray form, whose Fourier transform is the oscillatory integral usually considered in the context of singularity theory.

4.1. Oscillatory integrals and the Gelfand–Leray forms. We recall here briefly some results on oscillatory integrals and their asymptotic expansion. We refer the reader to §2, Vol.II of [**5**] for more details. In general, an *oscillatory integral* is an expression of the form

$$I(\alpha) = \int_{\mathbb{R}^n} e^{i\alpha f(x)} \phi(x)\, dx_1 \cdots dx_n, \tag{4.1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ and $\phi : \mathbb{R}^n \to \mathbb{R}$ are smooth functions and $\alpha \in \mathbb{R}_+^*$ is a real parameter. It is well known that, if the phase $f(x)$ is an analytic function in a neighborhood of a critical point x_0, then (4.1) has an asymptotic development for $\alpha \to \infty$ given by a series

$$I(\alpha) \sim e^{i\alpha f(x_0)} \sum_u \sum_{k=0}^{n-1} a_{k,u}(\phi)\, \alpha^u (\log \alpha)^k, \tag{4.2}$$

where u runs over a finite set of arithmetic progressions of negative rational numbers depending only on the phase $f(x)$, and the $a_{k,u}$ are distributions supported on the critical points of the phase, *cf.* §2.6.1, Vol.II of [**5**].

It is also well known that the integral (4.1) can be reformulated in terms of one-dimensional integrals using the Gelfand–Leray form

$$I(\alpha) = \int_{\mathbb{R}} e^{i\alpha t} \left(\int_{X_t(\mathbb{R})} \phi(x) \omega_f(x,t) \right) dt \tag{4.3}$$

where $X_t(\mathbb{R}) \subset \mathbb{R}^n$ is the level set $X_t(\mathbb{R}) = \{x \in \mathbb{R}^n \,:\, f(x) = t\}$ and $\omega_f(x,t)$ is the Gelfand–Leray form, that is, the unique $(n-1)$-form on the level hypersurface X_t with the property that

$$df \wedge \omega_f(x,t) = dx_1 \wedge \cdots \wedge dx_n. \tag{4.4}$$

Notice that, as in the case of the forms (3.4), there is an ambiguity in the choice of an $(n-1)$-form satisfying (4.4), but the restriction to X_t is unique so that the Gelfand–Leray form on X_t is well defined. Notice also that, up to throwing away a set of measure zero, we can assume here that the integration is over the values $t \in \mathbb{R}$ such that the level set X_t is a smooth hypersurface.

The Gelfand–Leray form $\omega_f(x,t)$ is often written in the notation

$$\omega_f(x,t) = \frac{dx_1 \wedge \cdots \wedge dx_n}{df}. \tag{4.5}$$

It is given by the Poincaré residue

$$\frac{\omega}{df} = \mathrm{Res}_{\epsilon=0} \frac{\omega}{f - \epsilon}. \tag{4.6}$$

The Gelfand–Leray function is the associated function

$$J(t) := \int_{L_t} \phi(x) \omega_f(x,t). \tag{4.7}$$

For more details, see §2.6 and §2.7, Vol.II [**5**].

We recall here a property of the Gelfand–Leray forms that will be useful in the following, where we consider complex hypersurfaces $X \subset \mathbb{A}^n = \mathbb{C}^n$, with defining polynomial equation $f = 0$ and the hypersurface complement $\mathcal{D}(f) \subset \mathbb{A}^n$, such that the restriction of f to the interior of the domain of integration $\Sigma \subset \mathbb{A}^n$ takes values in $\mathbb{R}_+^*$.

Recall that the Leray coboundary of a k-chain σ in X is a $(k+1)$-chain in $\mathcal{D}(f)$ obtained by considering a tubular neighborhood of X in $\mathbb{A}^n$, in the following way. Since X is a hypersurface, the boundary of its tubular neighborhood is a circle bundle over X. One considers the preimage of σ under the projection map as a chain in $\mathcal{D}(f)$. We denote the resulting chain by $\mathcal{L}(\sigma)$. It is called the Leray coboundary of σ (see [**5**] p.282). The Leray coboundary $\mathcal{L}(\sigma)$ is a cycle if σ is a cycle, and if one changes σ by a boundary then $\mathcal{L}(\sigma)$ also changes by a boundary.

Lemma 4.1. *Let σ_ϵ be a k-chain in $X_\epsilon = \{t \in \mathbb{A}^n | f(t) = \epsilon\}$ and let $\mathcal{L}(\sigma_\epsilon)$ be its Leray coboundary in $\mathcal{D}(f-\epsilon)$. Then, for a form $\alpha \in \Omega^k$ that admits a Gelfand–Leray form, one has*

$$\frac{1}{2\pi i}\int_{\mathcal{L}(\sigma(\epsilon))} df \wedge \frac{\alpha}{f-\epsilon} = \int_{\sigma(\epsilon)} \alpha, \tag{4.8}$$

where

$$\frac{d}{d\epsilon}\int_{\sigma(\epsilon)} \alpha = \int_{\sigma(\epsilon)} \frac{d\alpha}{df} - \int_{\partial\sigma(\epsilon)} \frac{\alpha}{df}. \tag{4.9}$$

Proof. First let us show that if α has a Gelfand–Leray form then $d\alpha$ also does. We have a form α/df such that

$$df \wedge \frac{\alpha}{df} = \alpha.$$

Its differential gives

$$d\alpha = d\left(df \wedge \frac{\alpha}{df}\right) = -df \wedge d\left(\frac{\alpha}{df}\right).$$

Thus, the form

$$\frac{d\alpha}{df} = -d\left(\frac{\alpha}{df}\right)$$

is a Gelfand–Leray form for $d\alpha$.

Then we proceed to prove the first statement. One can write

$$\frac{1}{2\pi i}\int_{\mathcal{L}(\sigma(\epsilon))} df \wedge \frac{\alpha}{f-\epsilon} = \frac{1}{2\pi i}\int_{\gamma}\left(\int_{\sigma(s)} \alpha\right)\frac{ds}{s-\epsilon},$$

where $\gamma \cong S^1$ is the boundary of a small disk centered at $\epsilon \in \mathbb{C}$. This can then be written as

$$= \frac{1}{2\pi i}\int_{\gamma}\int_{\sigma(\epsilon)} \alpha\,\frac{ds}{s-\epsilon} + \left(\frac{1}{2\pi i}\int_{\gamma}\int_{\sigma(s)} \alpha\,\frac{ds}{s-\epsilon} - \frac{1}{2\pi i}\int_{\gamma}\int_{\sigma(\epsilon)} \alpha\,\frac{ds}{s-\epsilon}\right).$$

The last term can be made arbitrarily small, so one gets (4.8). To obtain (4.9) notice that

$$\frac{1}{2\pi i}\frac{d}{d\epsilon}\int_{\mathcal{L}(\sigma(\epsilon))} df \wedge \frac{\alpha}{f-\epsilon} = \frac{1}{2\pi i}\int_{\mathcal{L}(\sigma(\epsilon))} df \wedge \frac{\alpha}{(f-\epsilon)^2}.$$

One then uses

$$d\left(\frac{\alpha}{f-\epsilon}\right) = \frac{d\alpha}{f-\epsilon} - \frac{\alpha}{(f-\epsilon)^2}$$

to rewrite the above as

$$\frac{1}{2\pi i}\left(\int_{\mathcal{L}(\sigma(\epsilon))} \frac{d\alpha}{f-\epsilon} - \int_{\mathcal{L}(\sigma(\epsilon))} d\left(\frac{\alpha}{f-\epsilon}\right)\right)$$

$$= \frac{1}{2\pi i} \int_{\mathcal{L}(\sigma(\epsilon))} \frac{df \wedge \frac{d\alpha}{df}}{f-\epsilon} - \frac{1}{2\pi i} \int_{\mathcal{L}(\partial\sigma(\epsilon))} \frac{\alpha}{f-\epsilon}$$

$$= \frac{1}{2\pi i} \int_{\mathcal{L}(\sigma(\epsilon))} \frac{df \wedge \frac{d\alpha}{df}}{f-\epsilon} - \frac{1}{2\pi i} \int_{\mathcal{L}(\partial\sigma(\epsilon))} \frac{df \wedge \frac{\alpha}{df}}{f-\epsilon},$$

where $d\alpha/df$ is a Gelfand–Leray form such that

$$df \wedge \frac{d\alpha}{df} = d\alpha,$$

and α/df is a Gelfand–Leray form with the property that

$$df \wedge \frac{\alpha}{df} = \alpha.$$

This then gives by (4.8)

$$\frac{d}{d\epsilon} \int_{\sigma(\epsilon)} \alpha = \int_{\sigma(\epsilon)} \frac{d\alpha}{df} - \int_{\partial\sigma(\epsilon)} \frac{\alpha}{df}.$$

This completes the proof. □

4.2. Leray coboundary regularization and subtraction. The formulation (2.68) of the parametric Feynman integrals, in the form of Proposition 2.9, suggests a regularization procedure different from Dimensional Regularization, but with the similar effect of replacing a divergent integral with a meromorphic function to which the "minimal subtraction" procedure can be applied to remove the polar part and extract a finite value.

Since the singularities arise where the domain of integration Σ meets the hypersurface $X = \{f = 0\}$, with f as in (2.70), we can concentrate on only the part of the integral that is supported near this intersection.

Let $D_\epsilon(X)$ denote a neighborhood of the hypersurface X in $\mathbb{P}^{n-1}$, given by level sets

$$D_\epsilon(X) = \cup_{s \in \Delta^*_\epsilon} X_s, \tag{4.10}$$

where $X_s = \{t | f(t) = s\}$ and $\Delta^*_\epsilon \subset \mathbb{C}^*$ is a small punctured disk of radius $\epsilon > 0$. The boundary $\partial D_\epsilon(X)$ is given by

$$\partial D_\epsilon(X) = \cup_{s \in \partial\Delta^*_\epsilon} X_s. \tag{4.11}$$

It is a circle bundle over the generic fiber X_ϵ, with projection $\pi_\epsilon : \partial D_\epsilon(X) \to X_\epsilon$. Given a domain of integration Σ, we consider the intersection $\Sigma \cap D_\epsilon(X)$. This is the region that contains the locus $\Sigma \cap X$ where the divergence in the Feynman integral can occur. We let $\mathcal{L}_\epsilon(\Sigma)$ denote the set

$$\mathcal{L}_\epsilon(\Sigma) = \pi_\epsilon^{-1}(\Sigma \cap X_\epsilon). \tag{4.12}$$

This enjoys the same properties of the Leray coboundary discussed above. In particular, notice that $\mathcal{L}_\epsilon(\partial\Sigma) = \partial\mathcal{L}_\epsilon(\Sigma)$.

We consider forms $\pi^*(\eta)$ as in (2.69). To keep track explicitly of the order of pole of such forms along the hypersurface X, we modify the notation and write

$$\pi^*(\eta_m) = \frac{\Delta(\omega)}{f^m}, \tag{4.13}$$

with ω and f as in (2.69).

We then make the following proposal for a regularization method for the Feynman integrals (2.68). We call it *Leray regularization*, because it is based on the use of Leray coboundaries. (Notice that this procedure of regularization and subtraction happens after having already removed the divergent Γ-factor from the parametric Feynman integrals and passing to residues. It is meant in fact to take care of the remaining singularities that arise from the intersections of the hypersurface with the domain of integration.)

Definition 4.2. *The Leray regularized Feynman integral is obtained from* (2.68) *by replacing the part*

$$\int_{\partial\Sigma\cap D_\epsilon(X)} \pi^*(\eta_m) + \int_{\Sigma\cap D_\epsilon(X)} df \wedge \frac{\pi^*(\eta_m)}{f} \tag{4.14}$$

of (2.68) *with the integral*

$$\int_{\mathcal{L}_\epsilon(\partial\Sigma)} \frac{\pi^*(\eta_{m-1})}{f-\epsilon} + \int_{\mathcal{L}_\epsilon(\Sigma)} df \wedge \frac{\pi^*(\eta_m)}{f-\epsilon}. \tag{4.15}$$

Thus, the Leray regularization introduced here consists of replacing the integral over $\Sigma\cap D_\epsilon(X)$ with an integral over $\mathcal{L}_\epsilon(\Sigma) \simeq (\Sigma\cap X_\epsilon)\times S^1$, which avoids the locus $\Sigma\cap X$ where the divergence can occur by going around it along a circle of small radius $\epsilon > 0$.

Using the result of Lemma 4.1, we can formulate (4.15) equivalently in the following form.

Lemma 4.3. *The Leray regularization of the Feynman integral* (2.68) *can be equivalently written in the form*

$$\begin{aligned} \mathbb{U}(\Gamma)_\epsilon = {} & \frac{1}{C(n,D,\ell)}\left(\int_{\partial\Sigma\cap D_\epsilon(X)^c} \pi^*(\eta_m) + \int_{\Sigma\cap D_\epsilon(X)^c} df \wedge \frac{\pi^*(\eta_m)}{f}\right) \\ + {} & \frac{2\pi i}{C(n,D,\ell)}\left(\int_{\partial\Sigma\cap X_\epsilon} \frac{\pi^*(\eta_{m-1})}{df} + \int_{\Sigma\cap X_\epsilon} \pi^*(\eta_m), \right) \end{aligned} \tag{4.16}$$

with $\pi^*(\eta_m) = \Delta(\omega)/f^m$ *as in* (4.13) *and Proposition 2.9.*

Proof. The result follows directly from Proposition 2.9 and Lemma 4.1 applied to (4.15). ☐

In (4.16) we use the notation $D_\epsilon(X)^c$ to denote the complement of $D_\epsilon(X)$. Notice how only the part of the integral (2.68) that is computed inside $D_\epsilon(X)$ is replaced by (4.15) in the Leray regularization, while the part of the integral (2.68) computed outside of $D_\epsilon(X)$ remains unchanged.

We now study the dependence on the parameter $\epsilon > 0$ of the Leray regularized Feynman integral (4.15), that is, of the integral

$$I_\epsilon := \int_{\partial\Sigma\cap X_\epsilon} \frac{\pi^*(\eta_{m-1})}{df} + \int_{\Sigma\cap X_\epsilon} \pi^*(\eta_m). \tag{4.17}$$

Theorem 4.4. *The function* I_ϵ *of* (4.17) *is infinitely differentiable in* ϵ. *Moreover, it extends to a holomorphic function for* $\epsilon\in\Delta^*\subset\mathbb{C}$, *a small punctured disk, with a pole of order at most* m *at* $\epsilon = 0$, *with* m *as in* (2.71).

PROOF. To prove the differentiability of I_ϵ, let us write

$$A_\epsilon(\eta) = \int_{\Sigma\cap X_\epsilon} \pi^*(\eta), \tag{4.18}$$

with $\pi^*(\eta)$ as in (2.69). By Lemma 4.1 above, and the fact that $d\pi^*(\eta) = 0$, we obtain

$$\frac{d}{d\epsilon} A_\epsilon(\eta) = -\int_{\partial\Sigma\cap X_\epsilon} \frac{\pi^*(\eta)}{df}, \tag{4.19}$$

where f is as in (2.70) and $\pi^*(\eta)/df$ is the Gelfand–Leray form of $\pi^*(\eta)$. Thus, we can write

$$I_\epsilon = A_\epsilon(\eta_m) - \frac{d}{d\epsilon} A_\epsilon(\eta_{m-1}).$$

Thus, to check the differentiability in the variable ϵ to all orders of I_ϵ is equivalent to checking that of A_ϵ. We define then $\Upsilon : \Omega^n \to \Omega^n$ by setting

$$\Upsilon(\alpha) = d\left(\frac{\alpha}{df}\right), \tag{4.20}$$

where α/df is a Gelfand–Leray form for α. In turn, the n-form $\Upsilon(\alpha)$ also has a Gelfand–Leray form, which we denote by

$$\delta(\alpha) = \frac{\Upsilon(\alpha)}{df} = \frac{d\left(\frac{\alpha}{df}\right)}{df}. \tag{4.21}$$

We then prove that, for $k \geq 2$,

$$\frac{d^k}{d\epsilon^k} A_\epsilon = -\int_{\partial\Sigma\cap X_\epsilon} \delta^{k-1}\left(\frac{\pi^*(\eta)}{df}\right). \tag{4.22}$$

This follows by induction. In fact, we first see that

$$\frac{d^2}{d\epsilon^2} A_\epsilon = -\frac{d}{d\epsilon}\int_{\partial\Sigma\cap X_\epsilon} \frac{\pi^*(\eta)}{df}$$

which, applying Lemma 4.1 gives

$$= -\int_{\partial\Sigma\cap X_\epsilon} \frac{d\left(\frac{\pi^*(\eta)}{df}\right)}{df}.$$

Assuming then that

$$\frac{d^k}{d\epsilon^k} A_\epsilon = -\int_{\partial\Sigma\cap X_\epsilon} \delta^{k-1}\left(\frac{\pi^*(\eta)}{df}\right)$$

we obtain again by a direct application of Lemma 4.1

$$\begin{aligned}
\frac{d^{k+1}}{d\epsilon^{k+1}} A_\epsilon &= -\int_{\partial\Sigma\cap X_\epsilon} \frac{d\left(\frac{\delta^{k-1}\left(\frac{\pi^*(\eta)}{df}\right)}{df}\right)}{df} \\
&= -\int_{\partial\Sigma\cap X_\epsilon} \delta^k\left(\frac{\pi^*(\eta)}{df}\right).
\end{aligned}$$

This proves differentiability to all orders.

Notice then that, while the expression (4.15) used in Definition 4.2 is, a priori, only defined for $\epsilon > 0$, the equivalent expression given in the second line of (4.16) and in (4.17) is clearly defined for any complex $\epsilon \in \Delta^*$ in a punctured disk around

$\epsilon = 0$ of sufficiently small radius. It can then be seen that the expression (4.17) depends holomorphically on the parameter ϵ by the general argument on holomorphic dependence on parameters given in Part III, §10.2 of Vol.II of [**5**].

Finally, to see that I_ϵ has a pole of order at most m at $\epsilon = 0$, notice that the form $\pi^*(\eta_m)$ of (4.13) is given by $\Delta(\omega)/f^m$ and has a pole of order at most m at X. This is evident in the two cases with $n - \frac{D(\ell+1)}{2} \geq 0$ or $n - \frac{D\ell}{2} \leq 0$. It also holds in the intermediate case with $n - \frac{D(\ell+1)}{2} < 0 < n - \frac{D\ell}{2}$, since we are using the convention that, in the case of a hypersurface X defined by a polynomial $f = f_1^{n_1} f_2^{n_2}$, a form $\Delta(\omega)/f^m$ has pole order m along X, even though on the individual components it has order mn_1 and mn_2, respectively. $\square$

In particular, the result of Proposition 4.4 shows that we can use the Leray regularization as an alternative to dimensional regularization to replace a divergent Feynman integral by a meromorphic function of a complex variable ϵ with a pole at $\epsilon = 0$. It is then possible to proceed as in dimensional regularization and apply "minimal subtraction", namely subtract the polar part of the resulting Laurent series in ϵ and evaluate the remaining part at $\epsilon = 0$.

It is clear that this regularization method is subject to the same problems as dimensional regularization when it comes to considering Feynman integrals associated to graphs that contain subdivergences. One can organize the hierarchy of subdivergences using the Bogoliubov-Parasiuk preparation, as in the case of dimensional regularization.

4.3. Birkhoff factorization and renormalization. Connes and Kreimer [**21**] showed that the BPHZ renormalization procedure, in the DimReg+MS regularization scheme, can be understood conceptually as the Birkhoff factorization of loops in the Lie group of complex points of the affine group scheme G dual to a commutative Hopf algebra $\mathcal{H}$ generated by the Feynman diagrams of the given physical theory. The Hopf algebra $\mathcal{H}$, at the discrete combinatorial level, is the commutative algebra generated by the one-particle-irreducible (1PI) graphs of the theory, with the coproduct

$$\Delta(\Gamma) = \Gamma \otimes 1 + 1 \otimes \Gamma + \sum \gamma \otimes \Gamma/\gamma, \tag{4.23}$$

where the sum is over proper subgraphs $\gamma \subset \Gamma$ satisfying a set of properties such as being Feynman diagrams of the same theory (see for instance [**23**] for a detailed discussion of the assumptions on the family of subgraphs involved in the coproduct). The quotient Γ/γ denotes the graph obtained by contracting each component of γ to a single vertex. It is sometimes denoted in the literature with the notation $\Gamma//\gamma$. The Hopf algebra is graded by the number of internal lines of graphs.

After identifying loops $\gamma : \Delta^* \to G(\mathbb{C})$, defined on an infinitesimal punctured disk Δ^* around $z = 0$, with elements $\phi \in G(K) = \mathrm{Hom}(\mathcal{H}, K)$, where K is the field of germs of meromorphic functions at $z = 0$, Connes and Kreimer showed that the BPHZ formula for renormalization is the recursive formula

$$\begin{aligned} \phi_-(x) &= -T(\phi(x) + \textstyle\sum \phi_-(x')\phi(x'')) \\ \phi_+(x) &= \phi(x) + \phi_-(x) + \textstyle\sum \phi_-(x')\phi(x''), \end{aligned} \tag{4.24}$$

with $\Delta(x) = x \otimes 1 + 1 \otimes x + \sum x' \otimes x''$, and x', x'' of lower degree, and with T the projection of a Laurent series onto its polar part. The original BPHZ formula is obtained by applying (4.24) to the element $\phi \in \mathrm{Hom}(\mathcal{H}, K)$ that assigns to a generator Γ of $\mathcal{H}$ its unrenormalized Feynman integral $U(\Gamma)$. As shown in [**21**], the formula (4.24) is the recursive formula that gives the Birkhoff factorization

$$\gamma(z) = \gamma_-(z)^{-1}\gamma_+(z) \tag{4.25}$$

of the loop γ into a part γ_+ that is holomorphic on Δ and a part γ_- that is holomorphic at $\infty \in \mathbb{P}^1(\mathbb{C})$, where one identifies γ_+ with $\phi_+ \in \mathrm{Hom}(\mathcal{H}, \mathcal{O})$, with $\mathcal{O}$ the algebra of germs of holomorphic functions at $z = 0$ and γ_- with $\phi_- \in \mathrm{Hom}(\mathcal{H}, \mathcal{Q})$ wth $\mathcal{Q} = \mathbb{C}[z^{-1}]$ so that

$$\phi = (\phi_- \circ S) * \phi_+, \tag{4.26}$$

with S the antipode of $\mathcal{H}$ and $*$ the product in the affine group scheme G, dual to the coproduct of $\mathcal{H}$.

The formulation in terms of Birkhoff factorization of loops with values in the Lie group of complex points of the affine group scheme of diffeographisms is applied in [**21**] to the Dimensional Regularization of Feynman integrals. Namely, the dimensionally regularized Feynman integrals $U(\Gamma)(z)$ of (2.43) define an element $\phi \in \mathrm{Hom}(\mathcal{H}, K)$, with $\mathcal{H}$ the Connes–Kreimer Hopf algebra of Feynman graphs of the theory and K the field of germs of meromorphic functions at $z = 0$, given by assigning as values on the generators of the Hopf algebra

$$\phi(\Gamma) = U(\Gamma) \in K. \tag{4.27}$$

In the case of dimensional regularization of Feynman integrals, the fact that the $U(\Gamma)(z)$ define meromorphic functions is very delicate, see the discussion in §1.4 of [**23**], especially Lemma 1.7, Lemma 1.8, and Theorem 1.9. On the contrary, we have seen that, using the Leray coboundary regularization introduced above, one easily obtains meromorphic functions of the parameter ϵ. We return to discuss the analytic continuation to meromorphic functions of the dimensionally regularized integrals via a different approach in §4.4 below.

By the results of §4.2 above, we can apply the same BPHZ renormalization procedure to the Leray coboundary regularization introduced in Definition 4.3. We thus consider the element $\phi \in \mathrm{Hom}(\mathcal{H}, K)$ defined by assigning on generators

$$\phi(\Gamma)(\epsilon) = U(\Gamma)_\epsilon \tag{4.28}$$

defined as in (4.16). By Proposition 4.4, we know that $U(\Gamma)_\epsilon$ defines a germ of a meromorphic function for $\epsilon \in \Delta^*$, an infinitesimal punctured disk around $\epsilon = 0$, hence it defines an element in K. We can then apply the Birkhoff factorization of ϕ, as in (4.26). This provides the counterterms, in the form

$$C(\Gamma)_\epsilon = \phi_-(\Gamma)(\epsilon), \tag{4.29}$$

which, as a function of ϵ, is an element in $\mathcal{Q}$, and the renormalized value of the Feynman integral, given by the finite value at zero

$$R(\Gamma) = \phi_+(\Gamma)(0), \tag{4.30}$$

where $\phi_+(\Gamma)(\epsilon)$ defines an element in the ring of convergent power series $\mathcal{O} \subset K$.

4.4. Mellin transform and the DimReg integral. We now return to consider the method of Dimensional Regularization and reinterpret it in terms of oscillatory integrals and mixed Hodge structures. As we recalled briefly in §4.1 above, the oscillatory integrals used in the theory of singularities and mixed Hodge structures can be seen as Fourier transforms (4.3) of a Gelfand–Leray function (4.7). One can also consider, instead of a Fourier transform, a Mellin transform of the same Gelfand–Leray function. Since the Mellin and Fourier transforms determine each other by well known formulae, the information obtained in this way is equivalent. In the context of singularity theory, the Mellin transforms of Gelfand–Leray functions and their relation to the oscillatory integral is discussed, for instance, in Part II, §7.2.1, of [**5**], Vol.II.

It was already proved by Belkale and Brosnan in [**9**] that, in the case of log-divergent graphs, the dimensionally regularized parametric Feynman integral can be written as a local Igusa L-function. This was later generalized to the non-log-divergent case in the work of Bogner and Weinzierl [**16**], [**17**], [**18**]. Our approach here is closely related to these results, though we do not discuss in detail the explicit relation. Moreover, we simplify the form of the integrals with respect to the case considered by Bogner and Weinzierl, so that we do not have to perform the cutting into sectors and blowups. We rely, in fact, on the formulation in terms of the exponential of the rational function $V_\Gamma(t,p)$ and its expansion, and we analyze the resulting terms individually. A more detailed analysis using the formulation of Bogner–Weinzierl and Belkale–Brosnan is possible, but we do not consider it here.

In order to relate the dimensionally regularized parametric Feynman integral to the oscillatory integrals and the Mellin transforms of Gelfand–Leray functions, consider again the integrals of the form (2.43), or better, the similar integrals computed after slicing with a k-plane Π_ξ as in §3.5, so that the intersection $X_\Gamma \cap \Pi_\xi$ has isolated singularities.

As shown in Lemma 2.6, we can equivalently compute the dimensionally regularized Feynman integral (2.43) using the form (2.44). Thus, we first consider an integral of the form

$$\int_{\Pi_\xi^+} \frac{e^{-V_\Gamma(t,p)}}{\Psi_\Gamma(t)^{(D+z)/2}}\,\omega_\xi \tag{4.31}$$

where $\Pi_\xi^+ = \Pi_\xi \cap \mathbb{R}_+^n$ and ω_ξ is as in (3.11). After expanding the exponential term and using (2.31), we are reduced to considering integrals of the form

$$\int_{\Pi_\xi^+} \frac{P_\Gamma(t,p)^\ell}{\Psi_\Gamma(t)^{\ell+(D+z)/2}}\,\omega_\xi. \tag{4.32}$$

Thus, we concentrate here on integrals of the form

$$F_{\Gamma,\xi}(z) = \int \Psi_\Gamma^z\,\chi_\xi\,P_\Gamma^\ell \Omega_\xi, \tag{4.33}$$

with Ω_ξ is as in (3.4), and for some integer $\ell \geq 0$. We have made here a simple change of coordinates on the complex variable z, whose meaning will become apparent in a moment.

The function χ_ξ in (4.33) is the characteristic function of the domain of integration. In order to show that one can extract from these dimensionally regularized Feynman integrals information on the singularities of the graph hypersurface X_Γ

(through its slices $X_\Gamma \cap \Pi_\xi$), it suffices to concentrate on the part of the domain of integration that is close to the hypersurface X_Γ. Thus, we can include in the function χ_ξ an additional cutoff of the integral that is supported in a neighborhood of the intersection $\Sigma_\xi \cap X_\Gamma$ of the original domain of integration in Π_ξ with the graph hypersurface.

In the following, for simplicity of notation, we just write (4.33) as

$$F_{\Gamma,\xi}(z) = \int \Psi_\Gamma^z \, \alpha_\xi, \tag{4.34}$$

where

$$\alpha_\xi = \chi_\xi \, P_\Gamma^\ell \, \Omega_\xi. \tag{4.35}$$

LEMMA 4.5. *The function* (4.34) *is the Mellin transform of the Gelfand-Leray function*

$$J_{\Gamma,\xi}(\epsilon) = \int_{X_\epsilon} \frac{\alpha_\xi}{df}, \tag{4.36}$$

with $f = \Psi_\Gamma|_{\Pi_\xi}$.

PROOF. First observe that both functions $\Psi_\Gamma(t)$ and $P_\Gamma(t,p)$ are real when restricted to the domain $\Sigma \subset \mathbb{R}^n_+$, with $\Psi_\Gamma(t) > 0$ on the interior of this domain. Thus, we can write the function $F_{\Gamma,\xi}(z)$ of (4.34) in the form

$$F_{\Gamma,\xi}(z) = \int_0^\infty s^z \left(\int_{X_s} \frac{\alpha_\xi}{df} \right) ds. \tag{4.37}$$

One can then recognize that (4.37) is in fact the Mellin transform

$$F_{\Gamma,\xi}(z) = \int_0^\infty s^z \, J_{\Gamma,\xi}(s) \, ds, \tag{4.38}$$

for $J_{\Gamma,\xi}$ as in (4.36), the corresponding Gelfand–Leray function. $\square$

The identification of $F_{\Gamma,\xi}(z)$ with the Mellin transform (4.38) also provides an answer to the problem of the analytic continuation to meromorphic functions in the complex plane for functions of the form (4.33). This analytic continuation is needed in order to justify our change of variables in z in passing from (4.32) to (4.33), as well as the use of integrals of the form (4.33) to derive conclusions about the original dimensionally regularized integrals (4.32). In fact, the existence of an analytic continuation to meromorphic functions for the functions $F_{\Gamma,\xi}(z)$ follows from the existence of an asymptotic expansion for Gelfand–Leray functions of the form

$$J(s) = \int_{X_s} \frac{\alpha}{df}, \quad \alpha = h\chi\omega_n, \tag{4.39}$$

with h a polynomial term and χ a compactly supported smooth function, supported near an isolated singularity of the hypersurface $f = 0$. The asymptotic expansion is given by

$$J(s) \sim \sum_{\lambda \in \Xi} \sum_{r=0}^{n-1} a_{r,\lambda} \, s^\lambda \log(s)^r, \quad s \to 0^+ \tag{4.40}$$

with Ξ a discrete subset of $\mathbb{R}$. The points $\lambda \in \Xi$ depend on the set of multiplicities of an embedded resolution of the singularity, see Part II, §7 of [**5**] and [**42**]. This implies the following result (*cf.* [**5**]), for generic choice of the slicing Π_ξ and of the external momenta.

COROLLARY 4.6. *Suppose that the cutoff function* χ_ξ *in* (4.35) *is supported in a small neighborhood of an isolated singularity of* $X_\Gamma \cap \Pi_\xi$*. Then the function* $F_{\Gamma,\xi}(z)$*, defined as in* (4.34) *for* $\Re(z) > 0$ *sufficiently large, admits an analytic continuation to a meromorphic function over the whole complex plane, with poles at the discrete set of points* $z = -(\lambda+1)$*, with* $\lambda \in \Xi$ *as in* (4.40)*, with the coefficient of* $(z+\lambda+1)^{-(r+1)}$ *in the Laurent series expansion given by* $(-1)^r r! a_{r,\lambda}$*, with* $a_{r,\lambda}$ *as in* (4.40)*.*

4.5. Dimensional regularization and mixed Hodge structures. We use the results of the previous section relating the dimensional regularization of the Feynman integrals to the Mellin transform of Gelfand–Leray functions, and the results of §3.4 on the interpretation in terms of cohomology of the Milnor fiber, to relate the dimensionally regularized Feynman integrals to limiting mixed Hodge structures.

We assume here to be in the case of isolated singularities, possibly after replacing the original Feynman integrals with their slices along planes Π_ξ of dimension complementary to that of the singular locus of the hypersurface, as discussed in §§3.2 and 3.5 above.

The cohomological Milnor fibration has fiber over ϵ given by the complex vector space $H^{k-1}(F_\epsilon, \mathbb{C})$, where the Milnor fiber F_ϵ of X_ξ is homotopically a bouquet of μ spheres S^{k-1}, with $k = \dim \Pi_\xi - 1$ and with μ the Milnor number of the isolated singularity. A holomorphic k-form $\alpha = h\omega_\xi / f^m$ determines a section of the cohomological Milnor fibration by taking the classes

$$\left[\frac{\alpha}{df} |_{F_\epsilon} \right] \in H^{k-1}(F_\epsilon, \mathbb{C}). \tag{4.41}$$

We then have the following results ([**5**], Vol.II §13). The asymptotic formula (4.40) for the Gelfand–Leray functions implies that the function of ϵ obtained by pairing the section (4.41) with a locally constant section of the homological Milnor fibration has an asymptotic expansion

$$\left\langle \left[\frac{\alpha}{df} \right], \delta \right\rangle \sim \sum_{\lambda,r} \frac{a_{r,\lambda}}{r!} \epsilon^\lambda \log(\epsilon)^r, \tag{4.42}$$

for $\epsilon \to 0$, where $\delta(\epsilon) \in H_{k-1}(F_\epsilon, \mathbb{Z})$. Moreover, there exist classes

$$\eta^\alpha_{r,\lambda}(\epsilon) \in H^{k-1}(F_\epsilon, \mathbb{C}) \tag{4.43}$$

such that the coefficients $a_{r,\lambda}$ of (4.42) are given by

$$\langle \eta^\alpha_{r,\lambda}(\epsilon), \delta(\epsilon) \rangle = a_{r,\lambda}. \tag{4.44}$$

Thus, one defines the "geometric section" associated to the holomorphic k-form α as

$$\sigma(\alpha) := \sum_{r,\lambda} \eta^\alpha_{r,\lambda}(\epsilon) \frac{\epsilon^\lambda \log(\epsilon)^r}{r!}. \tag{4.45}$$

The order of the geometric section $\sigma(\alpha)$ is defined as being the smallest λ in the discrete set $\Xi \subset \mathbb{R}$ such that $\eta^{\alpha}_{0,\lambda} \neq 0$. One denotes it by λ_α. The *principal part* of $\sigma(\alpha)$ is then defined as

$$\sigma_{\max}(\alpha)(\epsilon) := \epsilon^{\lambda_\alpha} \left(\eta^{\alpha}_{0,\lambda_\alpha} + \cdots + \frac{\log(\epsilon)^{k-1}}{(k-1)!} \eta^{\alpha}_{k-1,\lambda_\alpha} \right), \tag{4.46}$$

where one knows that

$$\eta^{\alpha}_{r,\lambda} = \mathcal{N}^r \eta^{\alpha}_{0,\lambda}, \tag{4.47}$$

where $\mathcal{N}$ is the nilpotent operator given by the logarithm of the unipotent monodromy, given by

$$\mathcal{N} = -\frac{1}{2\pi i} \log \mathcal{T}$$

with $\log \mathcal{T} = \sum_{r\geq 1} (-1)^{r+1} (\mathcal{T} - id)^r / r$.

The *asymptotic mixed Hodge structure* on the fibers of the cohomological Milnor fibration constructed by Varchenko ([**43**], [**44**]) has as the Hodge filtration the subspaces $F^r \subset H^{k-1}(F_\epsilon, \mathbb{C})$ defined by

$$F^r = \{[\alpha/df] \,|\, \lambda_\alpha \leq k - r - 1\} \tag{4.48}$$

and as weight filtration $W_\ell \subset H^{k-1}(F_\epsilon, \mathbb{C})$ the filtration associated to the nilpotent monodromy operator $\mathcal{N}$. This mixed Hodge structure has the same weight filtration as the *limiting mixed Hodge structure* constructed by Steenbrink ([**38**], [**39**]), but the Hodge filtration is different, though the two agree on the graded pieces of the weight filtration.

We now use a refined version of the results of §3.4, and in particular Corollary 3.5 for Feynman integrands as in (3.16). We show that, upon varying the choice of the external momenta p and of the spacetime dimension D, the corresponding Feynman integrands, in a neighborhood of an isolated singular point of $X_\Gamma \cap \Pi_\xi$, determine a subspace of the cohomology $H^{k-1}(F_\xi, \mathbb{C})$ of the Milnor fiber of $X_\Gamma \cap \Pi_\xi$. This inherits a Hodge and a weight filtration from the Milnor fiber cohomology with its asymptotic mixed Hodge structure. We concentrate on the case where $k - D\ell/2 \leq 0$, so that we can consider, for fixed k, arbitrarily large values of $D \in \mathbb{N}$.

PROPOSITION 4.7. *Consider Feynman integrals, sliced along a linear space* Π_ξ *as in* (3.14). *We write the integrand in the form*

$$\alpha_\xi = \frac{h \Omega_\xi}{f^m}, \tag{4.49}$$

with

$$\begin{cases} h = P_\Gamma^{-k+D\ell/2} \\ f = \Psi_\Gamma \\ m = -k + D(\ell+1)/2, \end{cases} \tag{4.50}$$

as in (3.24), *with* $k - D\ell/2 \leq 0$. *Upon varying the external momenta* p *in* $P_\Gamma(p,t)$ *and the spacetime dimension* $D \in \mathbb{N}$, *with* $k - D\ell/2 \leq 0$, *the forms* α_ξ *as above determine a subspace*

$$H^{k-1}_{\text{Feynman}}(F_\epsilon, \mathbb{C}) \subset H^{k-1}(F_\epsilon, \mathbb{C}),$$

of the fibers of the cohomological Milnor fibration, spanned by elements of the form (4.49)*, where the polynomials* $h = h_{T,v,w,p}$ *are of the form*

$$h(t) = \prod_{i=1}^{-k+D\ell/2} L_{T_i}(t) \prod_{e \notin T_i} t_e, \tag{4.51}$$

where the T_i *are spanning trees and the* $L_{T_i}(t)$ *are the linear functions of* (2.37)*.*

PROOF. Consider the explicit expression (2.32) of the polynomial $P_\Gamma(t,p)$ as a function of the external momenta, through the coefficients s_C of (2.33). One can see that, by varying arbitrarily the external momenta, subject to the global conservation law (2.35), one can reduce to the simplest possible case, where all external momenta are zero except for a pair of opposite momenta $P_{v_1} = p = -P_{v_2}$ associated to a pair of external edges attached to a pair of vertices v_1, v_2. In such a case, the polynomial $P_\Gamma(t,p)$ becomes of the form (2.36). Thus, when considering powers $P_\Gamma(t,p)^{-k+D\ell/2}$ for varying D, we obtain all polynomials of the form (4.51). □

We denote by $H^{k-1}_{\text{Feynman}}(F_\epsilon, \mathbb{C})$ the subspace of the cohomology $H^{k-1}(F_\epsilon, \mathbb{C})$ of the Milnor fiber spanned by the classes $[\alpha_\xi/df]$ with α_ξ of the form (4.49), with h of the form (4.51), considered modulo the ideal generated by the derivatives of $f = \Psi_\Gamma$ and localized at an isolated singular point, *i.e.* viewed as elements in the Milnor algebra $M(f)$. The subspace $H^{k-1}_{\text{Feynman}}(F_\epsilon, \mathbb{C})$ inherits a Hodge and a weight filtration $F^\bullet \cap H^{k-1}_{\text{Feynman}}$ and $W_\bullet \cap H^{k-1}_{\text{Feynman}}$ from the asymptotic mixed Hodge structure of Varchenko on $H^{k-1}(F_\epsilon, \mathbb{C})$. It is an interesting problem to see whether the subspace H^{k-1}_{Feynman} recovers the full $H^{k-1}(F_\epsilon, \mathbb{C})$ and if

$$(F^\bullet \cap H^{k-1}_{\text{Feynman}}, W_\bullet \cap H^{k-1}_{\text{Feynman}})$$

still give a mixed Hodge structure, at least for some classes of graphs Γ.

5. Regular and irregular singular connections

An important and still mysterious aspect of the motivic approach to Feynman integrals and renormalization is the problem of reconciling the Riemann–Hilbert correspondence of perturbative renormalization formulated by Connes–Marcolli in [**22**] (see also [**23**]), which is based on equivalence classes of certain *irregular singular* connections, with the setting of motives (especially mixed Tate motives) and mixed Hodge structures, which are naturally related to *regular singular* connections. The irregular singular connections of [**22**] have values in the Lie algebra of the Connes–Kreimer group of diffeographisms and are defined on a fibration over a punctured disk with fiber the multiplicative group, respectively representing the complex variable z of dimensional regularization and the energy scale μ (or rather μ^z) upon which the dimensionally regularized Feynman integrals depend. On the other hand, in the case of hypersurfaces in projective spaces, the natural associated regular singular connection is the Gauss–Manin connection on the cohomology of the Milnor fiber and the Picard–Fuchs equation for the vanishing cycles. We sketch here a relation between this regular singular connection and the irregular equisingular connections of [**22**]. (To avoid any possible confusion, the reader should keep in mind that the use of the term "equisingular" in [**22**] is not the same as the well established use in singularity theory, as in [**41**] for instance.)

5.1. Picard–Fuchs equation and Gauss–Manin connection. In the following we let

$$\left[\frac{\omega_i}{df}\right] \qquad i = 1, \ldots, \mu \tag{5.1}$$

be a basis for the vanishing cohomology bundle, written with the same notation we used above for the Gelfand–Leray form. Then the Gauss–Manin connection on the vanishing cohomology bundle, which is defined by the integer cohomology lattice in each real cohomology fiber, acts on the basis (5.1) by

$$\nabla_s^{GM} \left[\frac{\omega_i}{df}\right]_s = \sum_j p_{ij}(s) \left[\frac{\omega_j}{df}\right]_s, \tag{5.2}$$

where the $p_{ij}(s)$ are holomorphic away from $s = 0$ and have a pole at $s = 0$. The Gauss–Manin connection is regular singular and its monodromy agrees with the monodromy of the singularity (see [**4**], §2.3). Given a covariantly constant section $\delta(s)$ of the vanishing homology bundle, the function

$$I(s) = \left(\int_{\delta(s)} \frac{\omega_1}{df}, \ldots, \int_{\delta(s)} \frac{\omega_\mu}{df} \right) \tag{5.3}$$

is a solution of the regular-singular Picard–Fuchs equation

$$\frac{d}{ds} I(s) = P(s) I(s), \qquad \text{with} \quad P(s)_{ij} = p_{ij}(s). \tag{5.4}$$

Similarly, suppose given a holomorphic n-form ω and let ω/df be the corresponding Gelfand–Leray form, defining a section $[\omega/df]$ of the vanishing cohomology bundle. Let $\delta_1, \ldots, \delta_\mu$ be a basis of the vanishing homology, $\delta_i(s) \in H_{n-1}(F_s, \mathbb{Z})$. Then the function

$$I(s) = \left(\int_{\delta_1(s)} \frac{\omega}{df}, \ldots, \int_{\delta_\mu(s)} \frac{\omega}{df} \right) \tag{5.5}$$

satisfies a regular singular order ℓ differential equation

$$I^{(\ell)}(s) + p_1(s) I^{(\ell-1)}(s) + \cdots + p_\ell(s) I(s) = 0, \tag{5.6}$$

where the order is bounded above by the multiplicity of the critical point (see [**5**], §12.2.1). One refers to (5.6), or to the equivalent system of regular singular homogeneous first order equations

$$\frac{d}{ds} \mathcal{I}(s) = \mathcal{P}(s) \mathcal{I}(s), \tag{5.7}$$

with

$$\mathcal{I}_r(s) = s^{r-1} I^{(r-1)}(s), \tag{5.8}$$

as the Picard–Fuchs equation of ω. For the relation between Picard–Fuchs equations and mixed Hodge structures see §12 of [**5**] and [**32**].

5.2. Flat equisingular connections. We first recall some properties of the flat equisingular connections introduced in [**22**] (see also §1 of [**23**]). We denote by G the affine group scheme dual to the commutative Hopf algebra of Feynman diagrams, graded by loop number. We let $\mathfrak{g}$ denote the Lie algebra $\mathfrak{g} = \mathrm{Lie}(G)$. Let K denote the field of germs of meromorphic functions at $z = 0$. We also let B denote a fibration over an infinitesimal disk Δ^* with fiber the multiplicative group $\mathbb{G}_m$ and we denote by P the principal G-bundle $P = B \times G$. We consider $\mathrm{Lie}(G)$-valued flat connections ω that are *equisingular*, *i.e.* they satisfy

- The connection satisfies $\omega(z, \lambda u) = \lambda^Y \omega(z, u)$, for $\lambda \in \mathbb{G}_m$, with Y the grading operator.
- Solutions of $D\gamma = \omega$ have the property that their pullbacks $\sigma^*(\gamma) \in G(K)$ along any section $\sigma : \Delta \to B$ with fixed value $\sigma(0)$ have the same negative piece of the Birkhoff factorization $\sigma^*(\gamma)_-$.

The first condition and the flatness condition imply that the connection $\omega(z, u)$ can be written in the form

$$\omega(z, u) = u^Y(a(z))\, dz + u^Y(b(z))\, \frac{du}{u}, \tag{5.9}$$

where $a(z)$ and $b(z)$ are elements of $\mathfrak{g}(K)$ satisfying the flatness condition

$$\frac{db}{dz} - Y(a) + [a, b] = 0. \tag{5.10}$$

Recall that the Lie bracket in the Lie algebra $\mathrm{Lie}(G)$ is obtained by assigning

$$[\Gamma, \Gamma'] = \sum_{v \in V(\Gamma)} \Gamma \circ_v \Gamma' - \sum_{v' \in V(\Gamma')} \Gamma' \circ_{v'} \Gamma, \tag{5.11}$$

where $\Gamma \circ_v \Gamma'$ denotes the graph obtained by inserting Γ' into Γ at the vertex $v \in V(\Gamma)$ and the sum is over all vertices where an insertion is possible.

The equisingularity condition, which determines the behavior of pullbacks of solutions along sections of the fibration $\mathbb{G}_m \to B \to \Delta$, can be checked by writing the equation $Df = \omega$ in the more explicit form

$$\gamma^{-1} \frac{d\gamma}{dz} = a(z), \quad \text{and} \quad \gamma^{-1} Y(\gamma) = b(z). \tag{5.12}$$

When one interprets elements $\gamma \in G(K)$ as algebra homomorphisms $\phi \in \mathrm{Hom}(\mathcal{H}, K)$, one can write the above equivalently in the form

$$(\phi \circ S) * \frac{d\phi}{dz} = a, \quad \text{and} \quad (\phi \circ S) * Y(\phi) = b, \tag{5.13}$$

where S is the antipode in $\mathcal{H}$ and $*$ is the product dual to the coproduct in the Hopf algebra. This means, on generators Γ of $\mathcal{H}$,

$$\langle (\phi \circ S) \otimes \frac{d\phi}{dz}, \Delta(\Gamma) \rangle = a_\Gamma, \quad \text{and} \quad \langle (\phi \circ S) \otimes Y(\phi), \Delta(\Gamma) \rangle = b_\Gamma, \tag{5.14}$$

where

$$\Delta(\Gamma) = \Gamma \otimes 1 + 1 \otimes \Gamma + \sum_\gamma \gamma \otimes \Gamma/\gamma$$

as in (4.23), with the sum over subdivergences, and the antipode is given inductively by

$$S(X) = -X - \sum S(X') X'', \tag{5.15}$$

for $\Delta(X) = X \otimes 1 + 1 \otimes X + \sum X' \otimes X''$, with X' and X'' of lower degree.

5.3. From regular to irregular singularities. We now show how to produce a flat connection of the desired form (5.9), with irregular singularities, starting from the graph hypersuraces X_Γ, a consistent choice of slicing Π_ξ, and the regular singular Picard–Fuchs equation associated to the resulting isolated singularities of $X_\Gamma \cap \Pi_\xi$.

We begin by introducing a small modification of the Hopf algebra and coproduct, which accounts for the fact of having to choose a slicing Π_ξ. This is similar to what happens when one enriches the discrete Hopf algebra by adding the data of the external momenta.

Let $\mathcal{S}_\Gamma$ denote the manifold of planes Π_ξ in $\mathbb{A}^{\#E(\Gamma)}$ with

$$\dim \Pi_\xi \leq \operatorname{codim} \operatorname{Sing}(X_\Gamma).$$

We can write $\mathcal{S}_\Gamma$ as a disjoint union

$$\mathcal{S}_\Gamma = \bigcup_{m=1}^{\operatorname{codim} \operatorname{Sing}(X_\Gamma)} \mathcal{S}_{\Gamma,m}, \tag{5.16}$$

where $\mathcal{S}_{\Gamma,m}$ is the manifold of m-dimensional planes in $\mathbb{A}^{\#E(\Gamma)}$. We denote by $\mathcal{C}^\infty(\mathcal{S}_\Gamma)$ the space of test functions on $\mathcal{S}_\Gamma$ and by $\mathcal{C}_c^{-\infty}(\mathcal{S}_\Gamma)$ its dual space of distributions.

LEMMA 5.1. *Suppose given a subgraph $\gamma \subset \Gamma$. Then the choice of a distribution $\sigma \in \mathcal{C}_c^{-\infty}(\mathcal{S}_\Gamma)$ induces distributions $\sigma_\gamma \in \mathcal{C}_c^{-\infty}(\mathcal{S}_\gamma)$ and $\sigma_{\Gamma/\gamma} \in \mathcal{C}_c^{-\infty}(\mathcal{S}_{\Gamma/\gamma})$.*

PROOF. Given $\gamma \subset \Gamma$, neglecting external edges, we can realize the affine X_γ as a hypersurface inside a linear subspace $\mathbb{A}^{\#E(\gamma)} \subset \mathbb{A}^{\#E(\Gamma)}$ and similarly for the affine $X_{\Gamma/\gamma}$, seen as a hypersurface inside a linear subspace $\mathbb{A}^{\#E(\Gamma/\gamma)} \subset \mathbb{A}^{\#E(\Gamma)}$, where we simply identify the edges of γ or Γ/γ with a subset of the edges of the original graph Γ.

One then has a restriction map $T_\gamma : \mathcal{S}_{\Gamma,\gamma} \to \mathcal{S}_\gamma$, where $\mathcal{S}_{\Gamma,\gamma} \subset \mathcal{S}_\Gamma$ is the union of the components $\mathcal{S}_{\Gamma,m}$ of $\mathcal{S}_\Gamma$ with $m \leq \operatorname{codim} \operatorname{Sing}(X_\gamma)$,

$$\mathcal{S}_{\Gamma,\gamma} = \bigcup_{m=1}^{\operatorname{codim} \operatorname{Sing}(X_\gamma)} \mathcal{S}_{\Gamma,m}, \tag{5.17}$$

which is given by

$$T_\gamma(\Pi_\xi) = \Pi_\xi \cap \mathbb{A}^{\#E(\gamma)}. \tag{5.18}$$

This induces a map $T_\gamma : \mathcal{C}^\infty(\mathcal{S}_\gamma) \to \mathcal{C}^\infty(\mathcal{S}_\Gamma)$ given by

$$T_\gamma(f)(\Pi_\xi) = \begin{cases} f(T_\gamma(\Pi_\xi)) & \Pi_\xi \in \mathcal{S}_{\Gamma,\gamma} \\ 0 & \text{otherwise.} \end{cases} \tag{5.19}$$

In turn, this defines a map $T_\gamma : \mathcal{C}_c^{-\infty}(\mathcal{S}_\Gamma) \to \mathcal{C}_c^{-\infty}(\mathcal{S}_\gamma)$, at the level of distributions, by

$$T_\gamma(\sigma)(f) = \sigma(T_\gamma(f)). \tag{5.20}$$

The argument for Γ/γ is analogous. One sets $\sigma_\gamma = T_\gamma(\sigma_\Gamma)$ and $\sigma_{\Gamma/\gamma} = T_{\Gamma/\gamma}(\sigma_\Gamma)$. $\square$

We then enrich the original Hopf algebra $\mathcal{H}$ by adding the datum of the slicing Π_ξ. We consider the commutative algebra

$$\tilde{\mathcal{H}} = \mathrm{Sym}(\mathcal{C}_c^{-\infty}(\mathcal{S})), \tag{5.21}$$

where $\mathcal{S} = \cup_\Gamma \mathcal{S}_\Gamma$, endowed with the coproduct

$$\Delta(\Gamma,\sigma) = (\Gamma,\sigma)\otimes 1 + 1\otimes(\Gamma,\sigma) + \sum_\gamma (\gamma,\sigma_\gamma)\otimes(\Gamma/\gamma,\sigma_{\Gamma/\gamma}). \tag{5.22}$$

LEMMA 5.2. *The coproduct* (5.22) *is coassociative and* $\tilde{\mathcal{H}}$ *is a Hopf algebra.*

PROOF. The proof is analogous to the one given in [**23**], Theorem 1.27. □

We then proceed as follows. We pass to the projective instead of affine formulation and we fix a small neighborhood of an isolated singular point of $X_\Gamma \cap \Pi_\xi$, for Π_ξ a linear space of dimension at most equal to the codimension of $\mathrm{Sing}(X_\Gamma)$. Suppose given a holomorphic k-form α_ξ on Π_ξ. Then there exists an associated regular singular Picard–Fuchs equation

$$J^{(\ell)}_{\Gamma,\xi}(s) + p_1(s) J^{(\ell-1)}_{\Gamma,\xi}(s) + \cdots + p_\ell(s) J_{\Gamma,\xi}(s) = 0, \tag{5.23}$$

with the property that any solution $J_{\Gamma,\xi}(s)$ is a linear combination of the functions

$$J_{\Gamma,\xi,i}(s) = \int_{\delta_i(s)} \frac{\alpha_\xi}{df}, \tag{5.24}$$

where $\delta_1,\dots,\delta_\mu$ be a basis of locally constant sections of the homological Milnor fibration, $\delta_i(s) \in H_{k-1}(F_s,\mathbb{Z})$, and α_ξ/df is the Gelfand–Leray form associated to the holomorphic k-form α_ξ.

This depends on the choice of a singular point and can be localized in a small neighborhood of the singular point in $X_\Gamma \cap \Pi_\xi$. In fact, introducing a cutoff χ_ξ as in (4.33) that is supported near the singularities of $X_\Gamma \cap \Pi_\xi$ amounts to adding the expressions (5.24) for the different singular points. Thus, to simplify notations, we can just assume to have a single expression $J_{\Gamma,\xi}(s)$ at a unique isolated critical point.

We then have the following result, which constructs irregular singular connections as in §5.2 from solutions of the regular singular Picard–Fuchs equation.

THEOREM 5.3. *Any solution* $J_{\Gamma,\xi}$ *of the regular singular Picard–Fuchs equation* (5.23) *determines a flat* $\mathfrak{g}(K)$*-valued connection* $\omega(z,u)$ *of the form* (5.9). *Moreover, if the* k*-form* α_ξ *is given by* $P_\Gamma^\ell \Omega_\xi$ *as in* (4.35)*, then the connection is equisingular.*

PROOF. We consider the Mellin transform, as in (4.38)

$$\mathcal{F}_{\Gamma,\xi}(z) = \int_0^\infty s^z\, J_{\Gamma,\xi}(s)\, ds. \tag{5.25}$$

As in Corollary 4.6 (see §7 of [**5**]), the function $\mathcal{F}_{\Gamma,\xi}(z)$ admits an analytic continuation to meromorphic functions with poles at points $z = -(\lambda+1)$ with $\lambda \in \Xi_{\Gamma,\xi}$ a discrete set in $\mathbb{R}$ of points related to the multiplicities of an embedded resolution of the singular point of $X_\Gamma \cap \Pi_\xi$. We look at the function $\mathcal{F}_{\Gamma,\xi}(z)$ in a small neighborhood of a chosen point $z = -D$. It has an expansion as a Laurent series, with a pole at $z = -D$ if $-D \in \Xi_{\Gamma,\xi}$.

After a change of variables on the complex coordinate z, so that we have $z \in \Delta^*$, a small neighborhood of $z = 0$, we define

$$\phi_\mu(\Gamma, \sigma)(z) := \mu^{-z\, b_1(\Gamma)} \sigma \left(\mathcal{F}_{\Gamma,\xi}(-\frac{D+z}{2}) \right), \tag{5.26}$$

where we consider $\mathcal{F}_{\Gamma,\xi}$ as a function of ξ to which we apply the distribution σ. More precisely, after identifying $\mathcal{F}_{\Gamma,\xi}$ with its Laurent series expansion, we apply σ to the coefficients seen as functions of ξ. This defines an algebra homomorphism $\phi_\mu \in \mathrm{Hom}(\tilde{\mathcal{H}}, K)$, by assigning the values (5.26) on generators. Here μ is the mass scale as in §2.3 above. The homomorphism ϕ defined by (5.26) can be equivalently described as a family of $\tilde{G}(\mathbb{C})$-valued loops $\gamma_\mu : \Delta^* \to \tilde{G}(\mathbb{C})$, depending on the mass scale μ. Here $\tilde{G}$ denotes the affine group scheme dual to the commutative Hopf algebra $\tilde{\mathcal{H}}$. The dependence on μ of (5.26) implies that γ_μ satisfies the scaling property

$$\gamma_{e^t \mu}(z) = \theta_{tz}(\gamma_\mu(z)), \tag{5.27}$$

where θ_t is the one-parameter family of automorphisms of $\tilde{\mathcal{H}}$ generated by the grading, $\frac{d}{dt}\theta_t|_{t=0} = Y$. Then one sets

$$a_\mu(z) := (\phi_\mu \circ S) * \frac{d}{dz}\phi_\mu, \quad \text{and} \quad b_\mu(z) := (\phi_\mu \circ S) * Y(\phi_\mu), \tag{5.28}$$

where S and $*$ are the antipode of $\tilde{\mathcal{H}}$ and the product dual to the coproduct Δ of (5.22). These define elements a_μ, b_μ $\Omega^1(\mathfrak{g}(K))$, which one can use to define a connection $\omega(z, u)$ of the form (5.9). More precisely, for $\mu = e^t$, one has

$$\gamma_\mu^{-1} \frac{d}{dz}\gamma_\mu = \theta_t(\gamma^{-1}\frac{d}{dz}\gamma) = u^Y(a(z)),$$

where we set $u^Y = e^{tY}$ and then extend the resulting expression to $u \in \mathbb{G}_m(\mathbb{C}) = \mathbb{C}^*$. Similarly, we get $\gamma_\mu^{-1} Y(\gamma_\mu) = u^Y(b(z))$. Thus, the connection $\omega(z, u)$ defined in this way satisfies by construction the first condition of the equisingularity property, namely $\omega(z, \lambda u) = \lambda^Y \omega(z, u)$, for all $\lambda \in \mathbb{G}_m$. One can see that the connection is flat since we have

$$\begin{aligned} \frac{d}{dz} b_\mu(z) - Y(a_\mu(z)) &= \frac{d\gamma_\mu^{-1}(z)}{dz} Y(\gamma_\mu(z)) + \gamma_\mu^{-1}(z)\frac{d}{dz}(Y(\gamma_\mu(z))) \\ &\quad - Y(\gamma_\mu^{-1}(z))\frac{d}{dz}\gamma_\mu(z) - \gamma_\mu^{-1}(z)\frac{d}{dz}(Y(\gamma_\mu(z))) \\ = -\gamma_\mu^{-1}(z)\frac{d}{dz}(\gamma_\mu(z))\gamma_\mu^{-1}(z) Y(\gamma_\mu(z)) &- \gamma_\mu^{-1}(z) Y(\gamma_\mu(z))\gamma_\mu^{-1}(z) = -[a(z), b(z)]. \end{aligned}$$

The second condition of equisingularity is the property that, in the Birkhoff factorization

$$\gamma_\mu(z) = \gamma_{\mu,-}(z)^{-1}\gamma_{\mu,+}(z),$$

the negative part satisfies

$$\frac{d}{d\mu}\gamma_{\mu,-}(z) = 0.$$

By dimensional analysis on the counterterms, in the case of Dimensional Regularization and Minimal Subtraction, it is possible to show (see [**20**] §5.8.1) that the counterterms obtained by the BPHZ procedure applied to the Feynman integral $U_\mu(\Gamma)(z)$ of (2.49) and (2.48) do not depend on the mass parameter μ.

This means, as shown in [**21**] (see also Proposition 1.44 of [**23**]), that the Feynman integrals $U_\mu(\Gamma)(z)$ define a $G(\mathbb{C})$-valued loop $\gamma_\mu(z)$ with the property that $\partial_\mu \gamma_{\mu,-}(z) = 0$. The integrals (5.25) considered here, in the case where α_ξ is of the form (4.35), correspond to slices along a linear space Π_ξ of the Feynman integrals (2.49), localized by a cutoff χ_ξ near the singular points. The explicit dependence on μ in the integrals (3.31) is as in (5.26), which is unchanged with respect to that of the original dimensionally regularized Feynman integrals (2.49). Thus, the same argument of [**20**] §5.8.1 and Proposition 1.44 of [**23**] applies to this case to show that $\partial_\mu \gamma_{\mu,-}(z) = 0$. □

6. Logarithmic motives, Dimensional Regularization, and motivic sheaves

In this section we propose a candidate for a motivic formulation of dimensional regularization. As we discussed already in §2.2 above, in physics dimensional regularization is intended as a purely formal recipe that assigns a meaning to Gaussian integrals in "complexified dimension" $z \in \mathbb{C}$ by continuation to non-integer values of the usual formula for integer dimensions

$$\int e^{-\lambda t^2} d^z t := \pi^{z/2} \lambda^{-z/2}. \tag{6.1}$$

Usually, in so doing, one does not attempt to give a geometric meaning to the space of integration as a "space in complexified dimension $z \in \mathbb{C}$". The question of whether one can actually make sense of a geometry in complexified dimension was considered in [**23**], from the point of view of noncommutative geometry, where the usual notion of dimension of a space is replaced by the dimension spectrum, which is a set of complex numbers. A geometric model for a space whose dimension spectrum consists of a single point $z \in \mathbb{C}^*$ is given in §I.19.2 of [**23**], and it is shown that the formula (6.1) can be recovered from the properties of the Dirac operator on this space.

Here we also consider the question of giving geometric meaning to the complexified dimension, but we try to construct a geometric model underlying the operation of dimensional regularization using motives. We propose a candidate for a motive describing the dimensional regularization of a given Feynman graph. This is defined as an extension (in fact as a pro-motive) in the category of mixed motives, which is obtained from the logarithmic extension of Tate motives and the motive of the graph hypersurface. Just as in the case of noncommutative geometry, where the operation of dimensional regularization is understood as a product of the ordinary space in integer dimension by the "space of dimension z", here we also find that the dimensionally regularized Feynman integral is recovered by taking the product, in a category of motivic sheaves, of the motive associated to the graph hypersurface of a given Feynman graph by this pro-motive representing the "space of dimension z". It would be interesting to find a more explicit relation between this motivic description of dimensional regularization and the one based on noncommutative geometry.

6.1. Mixed Tate motives and the logarithmic extensions. We recall briefly the definition of the logarithmic motives, as given in [**6**]. Let $\mathcal{DM}(\mathbb{G}_m)$

be the Voevodsky category of mixed motives (motivic sheaves) over the multiplicative group $\mathbb{G}_m$. We will assume that the base field $\mathbb{K}$ is a number field (in fact, we can work over $\mathbb{Q}$) so that the extensions considered here take place in an abelian category of mixed Tate motives (*cf.* [**2**], [**34**]). Recall that the extensions $\mathrm{Ext}^1_{\mathcal{DM}(\mathbb{K})}(\mathbb{Q}(0),\mathbb{Q}(1))$ of Tate motives are given by the Kummer motives $M=[\mathbb{Z}\stackrel{u}{\to}\mathbb{G}_m]$ with $u(1)=q\in\mathbb{K}^*$. This extension has period matrix of the form

$$\begin{pmatrix} 1 & 0 \\ \log q & 2\pi i \end{pmatrix}. \tag{6.2}$$

When, instead of working with motives over the base field $\mathbb{K}$, one works with the relative setting of motivic sheaves over a base scheme S, instead of the Tate motives $\mathbb{Q}(n)$ one considers the Tate sheaves $\mathbb{Q}_S(n)$. These correspond to the constant sheaf with the motive $\mathbb{Q}(n)$ over each point $s\in S$. In the case where $S=\mathbb{G}_m$, there is a natural way to assemble the Kummer motives into a unique extension in $\mathrm{Ext}^1_{\mathcal{DM}(\mathbb{G}_m)}(\mathbb{Q}_{\mathbb{G}_m}(0),\mathbb{Q}_{\mathbb{G}_m}(1))$. This is the Kummer extension

$$\mathbb{Q}_{\mathbb{G}_m}(1)\to\mathcal{K}\to\mathbb{Q}_{\mathbb{G}_m}(0)\to\mathbb{Q}_{\mathbb{G}_m}(1)[1], \tag{6.3}$$

where over the point $s\in\mathbb{G}_m$ one is taking the Kummer extension $M_s=[\mathbb{Z}\stackrel{u}{\to}\mathbb{G}_m]$ with $u(1)=s$. Because of the logarithm function $\log(s)$ that appears in the period matrix for this extension, the Kummer extension (6.3) is also referred to as the *logarithmic motive*. We use the notation $\mathcal{K}=\mathrm{Log}$ as in [**6**] to refer to this extension, *cf.* [**7**].

When working with $\mathbb{Q}$-coefficients, so that one can include denominators in the definition of projectors, one can then consider the logarithmic motives Log^n, defined as in [**6**] by setting

$$\mathrm{Log}^n=\mathrm{Sym}^n(\mathcal{K}), \tag{6.4}$$

where the symmetric powers of an object in $\mathcal{DM}_{\mathbb{Q}}(\mathbb{G}_m)$ are defined as

$$\mathrm{Sym}^n(X)=\frac{1}{\#\Sigma_n}\sum_{\sigma\in\Sigma_n}\sigma(X^n). \tag{6.5}$$

Recall that the polylogarithms appear naturally as period matrices for extensions involving the symmetric powers $\mathrm{Log}^n=\mathrm{Sym}^n(\mathcal{K})$, in the form [**12**]

$$0\to\mathrm{Log}^{n-1}(1)\to\mathcal{L}^n\to\mathbb{Q}(0)\to 0, \tag{6.6}$$

where $M(1)=M\otimes\mathbb{Q}(1)$ and $\mathcal{L}^1=\mathrm{Log}$. The mixed motive $\mathcal{L}^n$ has period matrix

$$\begin{pmatrix} 1 & 0 \\ M^{(n)}_{\mathrm{Li}} & M_{\mathrm{Log}^{n-1}(1)} \end{pmatrix} \tag{6.7}$$

with

$$M^{(n)}_{\mathrm{Li}}=(-\mathrm{Li}_1(s),-\mathrm{Li}_2(s),\ldots,-\mathrm{Li}_n(s))^{\tau}, \tag{6.8}$$

where τ means transpose and where

$$\mathrm{Li}_1(s)=-\log(1-s),\qquad\text{and}\qquad \mathrm{Li}_n(s)=\int_0^s\mathrm{Li}_{n-1}(u)\,\frac{du}{u},$$

equivalently defined (on the principal branch) using the power series

$$\mathrm{Li}_n(s)=\sum_k\frac{s^k}{k^n},$$

and with
(6.9)

$$M_{\mathrm{Log}^n(1)} = \begin{pmatrix} 2\pi i & 0 & 0 & \cdots & 0 \\ 2\pi i \log(s) & (2\pi i)^2 & 0 & \cdots & 0 \\ 2\pi i \frac{\log^2(s)}{2!} & (2\pi i)^2 \log(s) & (2\pi i)^3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 2\pi i \frac{\log^n(s)}{n!} & (2\pi i)^2 \frac{\log^{n-1}(s)}{(n-1)!} & (2\pi i)^3 \frac{\log^{n-2}(s)}{(n-2)!} & \cdots & (2\pi i)^n \end{pmatrix}.$$

The period matrices for the motives Log^n correspond to the description of Log^n as extension of $\mathbb{Q}(0)$ by $\mathrm{Log}^{n-1}(1)$, *i.e.* to the distinguished triangles in $\mathcal{DM}(\mathbb{G}_m)$ of the form

$$\mathrm{Log}^{n-1}(1) \to \mathrm{Log}^n \to \mathbb{Q}(0) \to \mathrm{Log}^{n-1}(1)[1]. \tag{6.10}$$

The motives Log^n form a projective system under the canonical maps

$$\beta_n : \mathrm{Log}^{n+1} \to \mathrm{Log}^n$$

given by the composition of the morphisms $\mathrm{Sym}^{n+m}(\mathcal{K}) \to \mathrm{Sym}^n(\mathcal{K}) \otimes \mathrm{Sym}^m(\mathcal{K})$, as in [**6**], Lemma 4.35, given by the fact that $\mathrm{Sym}^{n+m}(\mathcal{K})$ is canonically a direct factor of $\mathrm{Sym}^n(\mathcal{K}) \otimes \mathrm{Sym}^m(\mathcal{K})$, and the map $\mathrm{Sym}^m(\mathcal{K}) \to \mathbb{Q}(0)$ of (6.10), in the particular case $m = 1$. Let Log^∞ denote the pro-motive obtained as the projective limit

$$\mathrm{Log}^\infty = \varprojlim_n \mathrm{Log}^n. \tag{6.11}$$

The analog of the period matrix (6.9) then becomes the infinite matrix
(6.12)

$$M_{\mathrm{Log}^\infty(1)} = \begin{pmatrix} 2\pi i & 0 & 0 & \cdots & 0 & \cdots \\ 2\pi i \log(s) & (2\pi i)^2 & 0 & \cdots & 0 & \cdots \\ 2\pi i \frac{\log^2(s)}{2!} & (2\pi i)^2 \log(s) & (2\pi i)^3 & \cdots & 0 & \cdots \\ \vdots & \vdots & \vdots & \cdots & \vdots & \cdots \\ 2\pi i \frac{\log^n(s)}{n!} & (2\pi i)^2 \frac{\log^{n-1}(s)}{(n-1)!} & (2\pi i)^3 \frac{\log^{n-2}(s)}{(n-2)!} & \cdots & (2\pi i)^n & \cdots \\ \vdots & \vdots & \vdots & \cdots & \vdots & \cdots \end{pmatrix}.$$

In other words, the mixed Hodge structure associated to the motives Log^n is the one that has as the weight filtration W_{-2k} the range of multiplication by the matrix M_{Log^n} defined as in (6.9) on vectors in $\mathbb{Q}^n$ with the first $k-1$ entries equal to zero, while the Hodge filtration F^{-k} is given by the range of multiplication of M_{Log^n} on vectors of $\mathbb{C}^n$ with the entries from $k+1$ to n equal to zero [**12**].

Thus, in this Hodge realization, the H^0 piece corresponds to the first column of the matrix M_{Log^n}, where the k-th entry corresponds to the k-th graded piece of the weight filtration. Let us consider the corresponding grading operator, that multiplies the k-th entry by T^k. One can then associate to the h^0-piece of the Log^∞ motive the following formal expression that corresponds in the period matrix (6.12) to the H^0 part in the MHS realization:

$$\mathbb{Q} \cdot \sum_k \frac{\log^k(s)}{k!} T^k =: \mathbb{Q} \cdot s^T. \tag{6.13}$$

The formal expression (6.13) has in fact an interpretation in terms of periods. This follows from a well known result (*cf. e.g.* [**30**], Lemma 2.10) expressing the powers of the logarithm in terms of iterated integrals. For iterated integrals we use the notation as in [**30**]

$$\int_a^b \frac{ds}{s}\circ\frac{ds}{s}\circ\cdots\circ\frac{ds}{s} = \int_{a\leq s_1\leq\cdots\leq s_n\leq b} \frac{ds_1}{s_1}\wedge\cdots\wedge\frac{ds_n}{s_n}. \tag{6.14}$$

We also denote by $\Lambda_{a,b}(n)$ the domain

$$\Lambda_{a,b}(n) = \{(s_1,\ldots,s_n)\,|\, a\leq s_1\leq\cdots\leq s_n\leq b\}. \tag{6.15}$$

Lemma 6.1. *The expression* (6.13) *is obtained as rational multiples of the pairing*

$$s^T = \int_{\Lambda_{1,s}(\infty)} \eta(T), \tag{6.16}$$

with $\Lambda_{1,s}(\infty) = \cup_n \Lambda_{1,s}(n)$ *and the form*

$$\eta(T) := \sum_n \frac{ds_1}{s_1}\wedge\cdots\wedge\frac{ds_n}{s_n}\, T^n. \tag{6.17}$$

Proof. The result follows from the basic identity (*cf.* [**30**], Lemma 2.10)

$$\int_{\Lambda_{a,b}(n)} \frac{ds_1}{s_1}\wedge\cdots\wedge\frac{ds_n}{s_n} = \frac{\log\left(\frac{b}{a}\right)}{n!}. \tag{6.18}$$

∎

6.2. Motivic sheaves and graph hypersurfaces. Arapura constructed in [**3**] a category of motivic sheaves over a base scheme S, modeled on Nori's approach to the construction of categories of mixed motives. We discuss briefly how a similar formalism may be applied to the Feynman motives associated to the graph hypersurfaces with the corresponding periods of the form (2.19).

The category of motivic sheaves constructed in [**3**] is based on Nori's construction of categories of motives via representations of graphs made of objects and morphisms (*cf.* [**19**]). In Arapura's case, one constructs a category of motivic sheaves over a scheme S, by taking as vertices of the corresponding graph objects of the form

$$(f: X\to S, Y, i, w)\,, \tag{6.19}$$

where $f: X\to S$ is a quasi-projective morphism, $Y\subset X$ is a closed subvariety, $i\in\mathbb{N}$, and $w\in\mathbb{Z}$. One thinks of such an object as determining a motivic version $h^i_S(X,Y)(w)$ of the local system given by the (Tate twisted) fiberwise cohomology of the pair $H^i_S(X,Y;\mathbb{Q}) = R^i f_* j_! \mathbb{Q}_{X\smallsetminus Y}$, where $j = j_{X\smallsetminus Y}: X\smallsetminus Y\hookrightarrow X$ is the open inclusion, *i.e.* the sheaf defined by

$$U\mapsto H^i(f^{-1}(U), f^{-1}(U)\cap Y;\mathbb{Q}).$$

The edges are given by the geometric morphisms, *i.e.* morphisms of varieties over S,

$$(f_1: X_1\to S, Y_1, i, w)\to (f': X_2\to S, Y_2 = F(Y), i, w), \quad \text{with} \quad f_2\circ F = f_1; \tag{6.20}$$

the connecting morphisms

$$(f : X \to S, Y, i+1, w) \to (f|_Y : Y \to S, Z, i, w), \quad \text{for} \quad Z \subset Y \subset X; \tag{6.21}$$

and the twisted projection morphisms

$$(f : X \times \mathbb{P}^1 \to S, Y \times \mathbb{P}^1 \cup X \times \{0\}, i+2, w+1) \to (f : X \to S, Y, i, w). \tag{6.22}$$

The product in the category of motivic sheaves of [**3**] is given by the fiber product

$$\begin{aligned} &(X \to S, Y, i, w) \times (X' \to S, Y', i', w') = \\ &(X \times_S X' \to S, Y \times_S X' \cup X \times_S Y', i+i', w+w'). \end{aligned} \tag{6.23}$$

This has the following effect on period computations.

LEMMA 6.2. *Suppose then given $\Sigma \subset X$ and $\Sigma' \subset X'$, defining relative homology cycles for (X, Y) and (X', Y'), respectively. One then has, for the fiber product* (6.23)*, the period pairing*

$$\int_{\Sigma \times_S \Sigma'} \pi_X^*(\omega) \wedge \pi_{X'}^*(\eta) = \int_\Sigma \omega \wedge f^* f'_*(\eta), \tag{6.24}$$

where $f : \Sigma \to S$ and $f' : \Sigma' \to S$ are the restrictions of the maps $X \to S$ and $X' \to S$.

PROOF. First recall that, when integrating a differential form on a fiber product, one has the formula

$$\int_{X \times_S X'} \pi_X^*(\omega) \wedge \pi_{X'}^*(\eta) = \int_X \omega \wedge (\pi_X)_* \pi_{X'}^*(\eta) = \int_X \omega \wedge f^* f'_*(\eta), \tag{6.25}$$

which corresponds to the diagram

$$\begin{tikzcd} & X \times_S X' \arrow[dl, "\pi_X"'] \arrow[dr, "\pi_{X'}"] & \\ X \arrow[dr, "f"'] & & X' \arrow[dl, "f'"] \\ & S & \end{tikzcd} \tag{6.26}$$

Suppose then given $\Sigma \subset X$ such that $\partial\Sigma \subset Y$ and $\Sigma' \subset X'$ with $\partial\Sigma' \subset Y'$. One has

$$\partial(\Sigma \times_S \Sigma') = \partial\Sigma \times_S \Sigma' \cup \Sigma \times_S \partial\Sigma' \subset Y \times_S X' \cup X \times_S Y',$$

so that $\Sigma \times_S \Sigma'$ defines a relative homology class in $(X \times_S X', Y \times_S X' \cup X \times_S Y')$. Given elements $[\omega] \in H^\cdot_S(X, Y)$ and $[\eta] \in H^\cdot_S(X', Y')$, we then apply the formula (6.25) to the integration on $\Sigma \times_S \Sigma'$ and obtain (6.24). $\square$

6.3. Logarithmic Feynman motives. Consider then the graph polynomial $\Psi_\Gamma(s) = \det(M_\Gamma(s))$. By removing the set of zeros of Ψ_Γ, *i.e.* the graph hypersurface X_Γ, we can consider Ψ_Γ as a morphism

$$\Psi_\Gamma : \mathbb{A}^{\#E_\Gamma} \smallsetminus \hat{X}_\Gamma \to \mathbb{G}_m. \tag{6.27}$$

We can then consider the pullback of the logarithmic motive $\mathrm{Log} \in \mathcal{DM}(\mathbb{G}_m)$ by this morphism, as in the construction of the logarithmic specialization system given in [**6**]. This gives a motive

$$\mathrm{Log}_\Gamma := \Psi_\Gamma^*(\mathrm{Log}) \in \mathcal{DM}(U_\Gamma), \tag{6.28}$$

where $U_\Gamma = \mathbb{A}^{\#E_\Gamma} \smallsetminus \hat{X}_\Gamma$.

In fact, a more sophisticated approach would involve considering here the "log complex" as in §9.2 of [**35**], *cf.* also §9.4 of [**35**], see also [**29**].

In the context of the category of motivic sheaves of Arapura recalled above, we can define the Feynman motives as follows.

Definition 6.3. *The category of Feynman motivic sheaves, for a fixed scalar quantum field theory, is the subcategory of the Arapura category of motivic sheaves over $\mathbb{G}_m$ spanned by the objects of the form*

$$(\Psi_\Gamma : \mathbb{A}^{\#E(\Gamma)} \smallsetminus \hat{X}_\Gamma \to \mathbb{G}_m, \Lambda \smallsetminus (\Lambda \cap \hat{X}_\Gamma), \#E(\Gamma) - 1, \#E(\Gamma) - 1), \tag{6.29}$$

where Γ ranges over the Feynman graphs of the given scalar field theory, and where

$$\Lambda = \{t \in \mathbb{A}^{\#E(\Gamma)} \mid \prod_i t_i = 0\} \tag{6.30}$$

is the union of the coordinate hyperplanes.

The above correspond to the local systems

$$H^{n-1}_{\mathbb{G}_m}(\mathbb{A}^n \smallsetminus \hat{X}_\Gamma, \Lambda \smallsetminus (\Lambda \cap \hat{X}_\Gamma), \mathbb{Q}(n-1)), \tag{6.31}$$

with $n = \#E_{int}(\Gamma)$.

One can also include as part of the data the slicing by all possible k-dimensional linear spaces $\Pi_\xi \subset \mathbb{A}^{\#E(\Gamma)}$, with $k \leq \operatorname{codim} \operatorname{Sing}(X_\Gamma)$, as we did in our previous discussions, and consider instead of the (6.29) objects of the form

$$(\Psi_\Gamma|_{\Pi_\xi} : \Pi_\xi \smallsetminus (\hat{X}_\Gamma \cap \Pi_\xi) \to \mathbb{G}_m, (\Lambda \cap \Pi_\xi) \smallsetminus (\Lambda \cap \hat{X}_\Gamma \cap \Pi_\xi), k-1, w). \tag{6.32}$$

Remark 6.4. The reason for taking the cohomology (6.31) relative to the *algebraic simplex* Λ, that is, the union of the coordinate hyperplanes defined by (6.30), is that, in this way, we can regard the *topological simplex* $\Sigma = \{t \in \mathbb{R}^n_+ \mid \sum_{i=1}^n t_i = 1\}$ as defining a homology cycle, since $\partial\Sigma \subset \Lambda$.

6.4. Dimensional Regularization and motives. In these terms, the procedure of dimensional regularization can then be described as follows. Consider again the logarithmic (pro-)motive, viewed itself as a motivic sheaf $X_{\mathrm{Log}^\infty} \to \mathbb{G}_m$ over $\mathbb{G}_m$. One can then take the product of a Feynman motive

$$(\Psi_\Gamma : \mathbb{A}^n \smallsetminus \hat{X}_\Gamma \to \mathbb{G}_\mathfrak{m}, \Lambda \smallsetminus (\Lambda \cap \hat{X}_\Gamma), k-1, k-1),$$

or more generally one of the form (6.32), by the (pro-)motive

$$(X^\infty_{\mathrm{Log}} \to \mathbb{G}_m, \Lambda_\infty, 0, 0), \tag{6.33}$$

where Λ_∞ is such that the domain of integration $\Lambda_{1,t}(\infty)$ of the period computation of Lemma 6.1 defines a cycle. The product is given by a fiber product as in (6.23), namely

$$\begin{tikzcd} \Psi_\Gamma^*(\mathrm{Log}^\infty) = (\mathbb{A}^n \smallsetminus \hat{X}_\Gamma) \rtimes_{\mathbb{G}_m} X_{\mathrm{Log}^\infty} \arrow[r] \arrow[d] & X_{\mathrm{Log}^\infty} \arrow[d] \\ \mathbb{A}^n \smallsetminus \hat{X}_\Gamma \arrow[r, "\Psi_\Gamma"] & \mathbb{G}_m \end{tikzcd} \tag{6.34}$$

We then have the following interpretation of the dimensionally regularized Feynman integrals.

Proposition 6.5. *The dimensionally regularized Feynman integral $F_{\Gamma,\xi}(z)$ of* (4.34) *are periods on the product, in the category of motivic sheaves enlarged to include projective limits, of the Feynman motive* (6.32) *by the logarithmic pro-motive* Log^∞ *seen as the motivic sheaf* (6.33)*.*

Proof. Consider the product (6.34), with the two projections

$$\pi_X : (\Pi_\xi \smallsetminus (\hat{X}_\Gamma \cap \Pi_\xi)) \times_{\mathbb{G}_m} X_{\mathrm{Log}^\infty} \to \Pi_\xi \smallsetminus (\hat{X}_\Gamma \cap \Pi_\xi)$$
$$\pi_L : (\Pi_\xi \smallsetminus (\hat{X}_\Gamma \cap \Pi_\xi)) \times_{\mathbb{G}_m} X_{\mathrm{Log}^\infty} \to X_{\mathrm{Log}^\infty}.$$

and the form $\pi_X^*(\alpha_\xi) \wedge \pi_L^*(\eta(T))$, where α_ξ is as in (4.35), and $\eta(T)$ is the form on X_{Log^∞} that gives the period (6.16). The period computation of Lemma 6.1 gives

$$\Psi_\Gamma^*\left(\int_{\Lambda_{1,s}(\infty)} \eta(T)\right) = \int_{\Lambda_{1,\Psi_\Gamma(t)}(\infty)} \eta(T) = \sum_n \frac{\log(\Psi_\Gamma(t))^n}{n!} T^n = \Psi_\Gamma(t)^T. \tag{6.35}$$

We then have, by (6.24),

$$\int_{(\Sigma\cap\Pi_\xi)\times_{\mathbb{G}_m}\Lambda_{1,\Psi_\Gamma(t)}(\infty)} \pi_X^*(\alpha_\xi) \wedge \pi_L^*(\eta(T)) =$$

$$\int_{\Sigma\cap\Pi_\xi} \alpha_\xi \wedge (\pi_X)_* \pi_L^*(\eta(T)) = \int_{\Sigma\cap\Pi_\xi} \Psi_\Gamma^T \alpha_\xi.$$

This is the integral (4.34), up to replacing the formal variable T of (6.13) with the complex DimReg variable z. ∎

The interpretation that emerges from this calculation is that performing the dimensional regularization of a Feynman integral can be thought of as taking the product in the category of motivic sheaves of the motive (motivic sheaf) of the graph hypersurface by the projective limit of the logarithmic motives. The variable $z \in \mathbb{G}_m$ that gives the complexified dimension of dimensional regularization corresponds to the 1-parameter group generated by the grading operator associated to the weight filtration of the logarithmic motives. The dimensionally regularized integral is then a period of this product motive.

6.5. Motivic zeta function and the DimReg integral. Kapranov introduced a notion of motivic zeta function by defining

$$Z_X(T) := \sum_{n\geq 0} \mathrm{Sym}^{\mathrm{n}}(X)\, T^n, \tag{6.36}$$

where the $\mathrm{Sym}^n(X)$ can be regarded as objects in an abelian category of motives, or as classes $[\mathrm{Sym}^n(X)]$ in the corresponding Grothendieck ring. Kapranov proved that, when X is the motive of a curve, then the zeta function is a rational function, in the sense that, given a motivic measure $\mu : K_0(\mathcal{M}) \to A$, the zeta function $Z_{X,\mu}(T) \in A[[T]]$ is a rational function of T. Later, Larsen and Lunts showed that in general this is not true in the case of algebraic surfaces [**33**].

Here we consider the motivic zeta function of the pullback of the logarithmic motive along the function Ψ_Γ as in (6.27). Namely, we consider the motivic zeta function

$$Z_{\mathrm{Log},\Gamma}(T) := \sum_{n\geq 0} \mathrm{Sym}^{\mathrm{n}}(\mathrm{Log}_\Gamma)\, T^n. \tag{6.37}$$

An interesting question, which we do not address in the present paper, is whether one can define a motivic lift of the Dimensional Regularization of the Feynman integral associated to a Feynman graph Γ using the motivic zeta function (6.37). In other words, whether one can obtain the zeta function

$$Z_\Gamma(T) := \sum_{n\geq 0} \frac{\log^n \Psi_\Gamma}{n!} T^n = \Psi_\Gamma^T \tag{6.38}$$

and the associated integrals

$$\sum_{n\geq 0} \left(\int_{\Sigma\cap\Pi_\xi} \frac{\log^n \Psi_\Gamma}{n!} \alpha_\xi \right) T^n = \int_{\Sigma\cap\Pi_\xi} \Psi_\Gamma^T \alpha_\xi \tag{6.39}$$

in a natural way from the motivic zeta function (6.37) of $\Psi_\Gamma^*(\mathrm{Log})$. We hope to return to this and related questions in following work.

References

[1] S. Agarwala, *The geometry of renormalization*, PhD Thesis, Johns Hopkins University, 2009.

[2] Y. André, *Une introduction aux motifs*, Panoramas et Synthèses, Vol.17, Société mathématique de France, 2005.

[3] D. Arapura, *A category of motivic sheaves*, arXiv:0801.0261.

[4] V.I. Arnold, V.V. Goryunov, O.V. Lyashko, V.A. Vasilev, *Singularity theory, I*, Springer Verlag, 1998.

[5] V.I. Arnold, S.M. Gusein-Zade, A.N. Varchenko, *Singularities of differentiable maps*, Vol.II, Birkhäuser, 1988.

[6] J. Ayoub, *The motivic vanishing cycles and the conservation conjecture*, in "Algebraic Cycles and Motives", London Mathematical Society Lecture Note Series, Vol.343, pp.3–54, Cambridge University Press, 2007.

[7] A. Beĭlinson, P. Deligne, *Interprétation motivique de la conjecture de Zagier reliant polylogarithmes et régulateurs*. in "Motives" (Seattle, WA, 1991), 97–121, Proc. Sympos. Pure Math., 55, Part 2, Amer. Math. Soc., Providence, RI, 1994.

[8] P. Belkale, P. Brosnan, *Matroids, motives, and a conjecture of Kontsevich*, Duke Math. Journal, Vol.116 (2003) 147–188.

[9] P. Belkale, P. Brosnan, *Periods and Igusa local zeta functions*. Int. Math. Res. Not. 2003, no. 49, 2655–2670.

[10] C. Bergbauer, A. Rej, *Insertion of graphs and singularities of graph hypersurfaces*, preprint.

[11] J. Bjorken, S. Drell, *Relativistic Quantum Mechanics*, McGraw-Hill, 1964, and *Relativistic Quantum Fields*, McGraw-Hill, 1965.
[12] S. Bloch, *Lectures on mixed motives.* Algebraic geometry—Santa Cruz 1995, 329–359, Proc. Sympos. Pure Math., 62, Part 1, Amer. Math. Soc., 1997.
[13] S. Bloch, *Motives associated to graphs*, Japan J. Math., Vol.2 (2007) 165–196.
[14] S. Bloch, E. Esnault, D. Kreimer, *On motives associated to graph polynomials*, Commun. Math. Phys., Vol.267 (2006) 181–225.
[15] S. Bloch, D. Kreimer, *Mixed Hodge structures and renormalization in physics*, Commun. Number Theory Phys. 2 (2008), no. 4, 637–718.
[16] C. Bogner, S. Weinzierl, *Periods and Feynman integrals*, J. Math. Phys. 50 (2009), no. 4, 042302, 16 pp.
[17] C. Bogner, S. Weinzierl, *Resolution of singularities for multi-loop integrals*, arXiv:0709.4092.
[18] C. Bogner, S. Weinzierl, *Blowing up Feynman integrals*, Nuclear Phys. B Proc. Suppl. 183 (2008), 256–261.
[19] A. Bruguiéres, *On a Tannakian theorem due to Nori*, preprint (2004).
[20] J. C. Collins, *Renormalization*, Cambridge University Press, 1984.
[21] A. Connes, D. Kreimer, *Renormalization in quantum field theory and the Riemann–Hilbert problem I. The Hopf algebra structure of graphs and the main theorem*, Comm. Math. Phys., Vol.210 (2000) 249–273.
[22] A. Connes, M. Marcolli, *Renormalization and motivic Galois theory.* International Math. Res. Notices (2004) N.76, 4073–4091.
[23] A. Connes, M. Marcolli, *Noncommutative Geometry, Quantum Fields, and Motives*, Colloquium Publications, Vol.55, American Mathematical Society, 2008.
[24] P. Deligne and A. Dimca, *Filtrations de Hodge et par d'ordre du pole pour les hypersurfaces singulières*, Ann. Sci. École Norm. Sup. 23 (1990) 645–656.
[25] A. Dimca, *Singularities and topology of hypersurfaces*, Universitext, Springer-Verlag, 1992.
[26] I. Dolgachev, *Weighted projective varieties*, in "Group actions and vector fields", LNM 956, pp. 34–71, Springer-Verlag, 1982.
[27] K. Ebrahimi-Fard, L. Guo, *Rota-Baxter Algebras in Renormalization of Perturbative Quantum Field Theory*, in "Universality and renormalization", 47–105, Fields Inst. Commun., 50, Amer. Math. Soc., Providence, RI, 2007.
[28] I.M. Gel'fand, S.G. Gindikin, M.I. Graev, *Integral geometry in affine and projective spaces.* Current problems in mathematics, Vol. 16, pp. 53–226, 228, Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. i Tekhn. Informatsii, Moscow, 1980.
[29] A.B. Goncharov, *Explicit regulator maps on polylogarithmic motivic complexes.* in "Motives, polylogarithms and Hodge theory, Part I" (Irvine, CA, 1998), 245–276, Int. Press Lect. Ser., 3, I, Int. Press, Somerville, MA, 2002.
[30] A.B. Goncharov, *Multiple polylogarithms and mixed Tate motives*, math.AG/0103059.
[31] C. Itzykson, J.B. Zuber, *Quantum Field Theory*, Dover Publications, 2006.
[32] V.S. Kulikov, *Mixed Hodge structures and singularities.* Cambridge University Press, 1998.
[33] M. Larsen, V. Lunts, *Motivic measures and stable birational geometry*, Moscow Math. Journal, Vol.3 (2003), N.1, 85–95.
[34] M. Levine, *Mixed motives*, Math. Surveys and Monographs, Vol. 57, AMS, 1998.
[35] M. Levine, *Motivic tubular neighborhoods.* Doc. Math. 12 (2007), 71–146.
[36] M. Marcolli, *Feynman motives*, World Scientific, 2010.
[37] N. Nakanishi, *Graph Theory and Feynman Integrals.* Gordon and Breach, 1971.
[38] J.H.M. Steenbrink, *Limits of Hodge structures.* Invent. Math. 31 (1975/76), no. 3, 229–257.
[39] J.H.M. Steenbrink, *Mixed Hodge structure on the vanishing cohomology.* in "Real and complex singularities" (Proc. Ninth Nordic Summer School/NAVF Sympos. Math., Oslo, 1976), pp. 525–563. Sijthoff and Noordhoff, Alphen aan den Rijn, 1977.
[40] W. van Suijlekom, *Renormalization of gauge fields: a Hopf algebra approach.* Comm. Math. Phys. 276 (2007), no. 3, 773–798.
[41] B. Teissier, *Introduction to equisingularity problems.* Algebraic geometry (Proc. Sympos. Pure Math., Vol. 29, Humboldt State Univ., Arcata, Calif., 1974), pp. 593–632. Amer. Math. Soc., Providence, R.I., 1975.
[42] A.N. Varchenko, *Newton polyhedra and estimates of oscillatory integrals*, Functional Analysis and Applications, Vol.10 (1976), 175–196.

[43] A.N. Varchenko, *Asymptotic behavior of holomorphic forms determines a mixed Hodge structure*. Dokl. Akad. Nauk SSSR 255 (1980), no. 5, 1035–1038.

[44] A.N. Varchenko, *Asymptotic Hodge structure on vanishing cohomology.* Izv. Akad. Nauk SSSR Ser. Mat. 45 (1981), no. 3, 540–591, 688.

Department of Mathematics, California Institute of Technology, Mail Code 253-37, 1200 E.California Blvd, Pasadena, CA 91125, USA

E-mail address: matilde@caltech.edu

Clay Mathematics Proceedings
Volume 11, 2010

Free Groups and Reduced 1-Cohomology of Unitary Representations

Florian Martin and Alain Valette

To Alain Connes, with admiration

Abstract. Guichardet [**Gui72**] showed that every unitary representation of the free group $\mathbb{F}_n$ $(2 \leq n < \infty)$ has non-zero 1-cohomology. We construct a continuum of pairwise inequivalent, irreducible, unitary representations of $\mathbb{F}_n$, with vanishing reduced 1-cohomology and such that the C^*-algebra generated by each representation is the unitized algebra of the compact operators.

1. Introduction

If G is a countable discrete group and π a unitary representation of G, we denote by $H^1(G,\pi)$ the first cohomology of G with coefficients in π, i.e. the quotient of the space $Z^1(G,\pi)$ of 1-cocycles by the space $B^1(G,\pi)$ of 1-coboundaries. Endowed with the topology of pointwise convergence, $Z^1(G,\pi)$ becomes a Fréchet space, and the *reduced* 1-cohomology $\overline{H^1}(G,\pi)$ is defined as the quotient of $Z^1(G,\pi)$ by the closure of the space of 1-coboundaries.

Reduced 1-cohomology was first considered by Guichardet [**Gui72**] and its relevance to questions of rigidity and geometric group theory was emphasized more recently in papers of Shalom (see [**Sha00**], [**Sha04**]).

This paper focuses on the free group on n generators $G = \mathbb{F}_n$ $(2 \leq n \leq \infty)$: its 1-cohomology has the following interesting property established by Guichardet (Example 1 in [**Gui72**]): $H^1(\mathbb{F}_2,\pi) \neq 0$ for every unitary representation π of $\mathbb{F}_2$. Using the dictionary between 1-cohomology and affine isometric actions on Hilbert spaces (see e.g. [**BHV08**], p.73), the geometric equivalent of this observation is: every unitary representation of $\mathbb{F}_2$ is the linear part of some affine isometric action without a globally fixed point.

We illustrate the difference between reduced and ordinary 1-cohomology by establishing:

Theorem 1.1. *Fix $n \in \mathbb{N} \cup \{\infty\}$ $(n \geq 2)$. There exists a continuum of pairwise inequivalent, unitary, irreducible representations σ of $\mathbb{F}_n$ such that*

2010 *Mathematics Subject Classification.* Primary 22D25; Secondary 22D10.

Supported by the Swiss National Fund, request No 20-65060.01.

1) $\overline{H^1}(\mathbb{F}_n, \sigma) = 0$.
2) *The C^*-algebra generated by $\sigma(\mathbb{F}_n)$ is $\tilde{\mathcal{K}}$, the unitized C^*-algebra of the algebra $\mathcal{K}$ of compact operators on an infinite-dimensional separable Hilbert space.*

There are other instances of the fact that, for a given group, the vanishing of the 1-cohomology is not equivalent to the vanishing of its reduced counterpart: for example, let λ_G be the left regular representation of a countably infinite amenable group G; then $H^1(G, \lambda_G) \neq 0$, by Théorème 1 in [**Gui72**], while $\overline{H^1}(G, \lambda_G) = 0$ by [**MV07**]. Let us mention however a remarkable result by Shalom [**Sha00**]: for a *compactly generated* locally compact group, the vanishing of reduced 1-cohomology for all unitary representations, is equivalent to the vanishing of 1-cohomology for all unitary representations (the latter being equivalent to Kazhdan's property (T), by the Delorme-Guichardet theorem, see Chapter 2 in [**BHV08**]).

2. Proof of Theorem 1.1

Let us denote by $\mathrm{Im}\,T$ the range of the linear operator T.

Lemma 2.1. *Fix an integer $n \geq 2$. Let $U_1, ..., U_n$ be unitary operators on a Hilbert space such that:*

1) *1 is not an eigenvalue of U_i, for $1 \leq i \leq n$;*
2) *for $2 \leq j \leq n$:*

$$\mathrm{Im}(U_j - 1) \cap (\sum_{i=1}^{j-1} \mathrm{Im}(U_i - 1)) = \{0\}.$$

Then the assignment $\pi(x_i) = U_i^*$ defines a unitary representation π of the free group $\mathbb{F}_n$ on n generators $x_1, ..., x_n$, such that $\overline{H^1}(\mathbb{F}_n, \pi) = 0$.

Proof of the lemma: We start the same way as Guichardet in Example 1 in [**Gui72**], in his proof of $H^1(\mathbb{F}_2, \sigma) \neq 0$ for every unitary representation σ of $\mathbb{F}_2$. For a unitary representation π of $\mathbb{F}_n$ on a Hilbert space V, the map

$$Z^1(\mathbb{F}_n, \pi) \to V^n : b \mapsto (b(x_1), ..., b(x_n))$$

is a topological isomorphism (surjectivity follows from the freeness of the group: a 1-cocycle can be defined arbitrarily on generators). In that isomorphism, $B^1(\mathbb{F}_n, \pi)$ corresponds to the image of the map

$$\psi : V \to V^n : v \mapsto ((\pi(x_1) - 1)v, ..., (\pi(x_n) - 1)v)$$

So $\overline{H^1}(\mathbb{F}_n, \pi) = 0$ if and only if ψ has dense image, if and only if $\psi^* : V^n \to V$ is injective. But

$$\psi^*(v_1, ..., v_n) = \sum_{i=1}^{n} (\pi(x_i)^* - 1)v_i.$$

With $U_i = \pi(x_i)^*$, we see that our assumptions on $U_1, ..., U_n$ exactly mean that ψ^* is injective. ☐

We now come to a problem in operator theory, namely construct families of unitary operators satisfying the conditions in Lemma 2.1. We will elaborate on Dixmier's elegant construction [**Dix49**] (see also Theorem 3.6 in [**FW71**]) to answer that question.

Proof of Theorem 1.1: On $V = L^2[0, 2\pi]$ with the trigonometric orthonormal system $(e_k)_{k\in\mathbb{Z}}$, let us construct unitary operators U_n $(n \geq 1)$ such that $U_n - 1$ is trace-class, 1 is not an eigenvalue of U_n and $\mathrm{Im}(U_n - 1) \cap (\sum_{m=1}^{n-1} \mathrm{Im}(U_m - 1)) = \{0\}$ for $n \geq 2$. Moreover U_1, U_2 will be shown to act together irreducibly on V. Taking into account the fact that every irreducible C^*-algebra intersecting $\mathcal{K}$ non-trivially, must contain it (see [**Dix77**], Corollary 4.1.10), we get the second statement in the Theorem. For $n < \infty$, the first statement (vanishing of $\overline{H^1}$) will follow straight from Lemma 2.1. For $n = \infty$, we observe that if a group Γ is the increasing union of subgroups Γ_n with $\overline{H^1}(\Gamma_n, \sigma|_{\Gamma_n}) = 0$, then clearly $\overline{H^1}(\Gamma, \sigma) = 0$.

To construct a continuum of such families of unitary operators, fix a real-valued rapidly decreasing sequence $a = (a_k)_{k\in\mathbb{Z}}$, such that $0 \neq a_k \neq a_m$ for $k \neq m \in \mathbb{Z}$, and define a diagonal, trace-class operator $A^{(a)}$ on V by

$$A^{(a)} e_k = a_k e_k \quad (k \in \mathbb{Z}).$$

Note that $A^{(a)}$ is injective with all eigenvalues of multiplicity 1, by our choice of a. Now let $(x_n)_{n>0}$ be a strictly increasing sequence in $[0, 2\pi[$, with $x_1 = 0$ and $\frac{x_2}{\pi}$ irrational. Define a function ξ_n on $[0, 2\pi[$ by

$$\xi_n(x) = \begin{cases} -1 & if \quad 0 \leq x < x_n \\ 1 & if \quad x_n \leq x < 2\pi \end{cases}$$

Let M_n be the operator of multiplication by ξ_n: this is a self-adjoint unitary operator on V; note that $M_1 = 1$. Set $A_n^{(a)} = M_n A^{(a)} M_n^*$; the following holds:

Claim: For $n \geq 2$:

$$\mathrm{Im}\, A_n^{(a)} \cap (\sum_{i=1}^{n-1} \mathrm{Im}\, A_i^{(a)}) = \{0\}.$$

To prove the claim, observe that , because a is rapidly decreasing, $\mathrm{Im}\, A^{(a)}$ is contained in the space of restrictions of real-analytic functions to $[0, 2\pi[$. Then $\sum_{i=1}^{n-1} \mathrm{Im}\, A_i^{(a)}$ is contained in the space of functions on $[0, 2\pi[$ whose restrictions to all intervals $[x_1, x_2[$, $[x_2, x_3[$, ..., $[x_{n-1}, 2\pi[$ are real analytic. On the other hand non-zero functions in $\mathrm{Im}\, A_n^{(a)}$ are not analytic in the neighborhood of $x_n \in]x_{n-1}, 2\pi[$, proving the claim.

Let then $U_n^{(a)}$ be the Cayley transform of $A_n^{(a)}$:

$$U_n^{(a)} = (1 - iA_n^{(a)})(1 + iA_n^{(a)})^{-1}.$$

Then $U_n^{(a)}$ is unitary, 1 is not an eigenvalue of $U_n^{(a)}$, and $U_n^{(a)}$ is diagonal in the basis $(M_n e_k)_{k\in\mathbb{Z}}$, with all eigenvalues of multiplicity 1. Moreover

$$U_n^{(a)} - 1 = -2iA_n^{(a)}(1 + iA_n^{(a)})^{-1}$$

so that $U_n^{(a)} - 1$ is trace-class, and $\mathrm{Im}(U_n^{(a)} - 1) \cap (\sum_{m=1}^{n-1} \mathrm{Im}(U_m^{(a)} - 1)) = \{0\}$ by the claim.

To prove that $U_1^{(a)}$, $U_2^{(a)}$ together act irreducibly on V, let S be an operator on V which commutes both with $U_1^{(a)}$ and $U_2^{(a)}$. Since $U_1^{(a)}$, $U_2^{(a)}$ have all eigenvalues of multiplicity 1, the operator S must be diagonal both in the bases $(e_k)_{k\in\mathbb{Z}}$ and $(M_2e_k)_{k\in\mathbb{Z}}$. From the Fourier series expansion of M_2e_k:

$$M_2e_k = (1-\frac{x_2}{\pi})e_k + \sum_{m\neq k}\frac{i}{\pi(m-k)}[1-e^{i(k-m)x_2}]e_m$$

and the fact that all Fourier coefficients of M_2e_k are non-zero (because $\frac{x_2}{\pi}$ is irrational), it follows that S must be scalar. Irreducibility then follows from Schur's lemma.

Finally, to get a continuum of pairwise inequivalent representations, we notice that, since $U_1^{(a)}-1$ is trace-class, the complex number

$$\mathrm{Tr}(U_1^{(a)}-1) = -2i\,\mathrm{Tr}\,A^{(a)}(1+iA^{(a)})^{-1} = -2i\sum_{k\in\mathbb{Z}}a_k(1+ia_k)^{-1}$$

is an invariant of unitary equivalence of the associated representation. So, varying a in the space of real-valued rapidly decreasing sequences satisfying $0\neq a_k\neq a_m$ for $k\neq m\in\mathbb{Z}$, we get the desired continuum. □

Theorem 1.1 motivates:

Question 1. *Is there a countable group G such that $\overline{H^1}(G,\pi)\neq 0$ for every unitary representation π of G?*

Note that such a group, if it exists, must be non-amenable: indeed, by Corollary 5.2 in [**MV07**], a countable amenable group G has $\overline{H^1}(G,\lambda_G)=0$, where λ_G is the left regular representation.

3. A remark

The irreducible representations σ constructed in Theorem 1.1 have the property that $\sigma(g)-1\in\mathcal{K}$ for every $g\in\mathbb{F}_n$. We observe that this fact alone is responsible for the non-vanishing of H^1.

Proposition 3.1. *Let G be a discrete group. Assume that there exists an infinite-dimensional unitary irreducible representation π of G with the property that $1-\pi(g)$ is a compact operator for every $g\in G$. Then $B^1(G,\pi)$ is not closed in $Z^1(G,\pi)$; in particular $H^1(G,\pi)\neq 0$.*

Proof: Observe that by irreducibility the C^*-algebra generated by $\pi(G)$ is $\tilde{\mathcal{K}}$, and consider the short exact sequence

$$0\to\mathcal{K}\to\tilde{\mathcal{K}}\xrightarrow{q}\mathbb{C}\to 0.$$

For $f\in\ell^1(G)$, we have $q(\pi(f))=\sum_{g\in G}f(g)q(\pi(g))=\sum_{g\in G}f(g)$ for every $f\in\ell^1(G)$, so that

$$|\sum_{g\in G}f(g)|\leq\|\pi(f)\|.$$

By Theorem 3.4.4 in [**Dix77**], this means exactly that π weakly contains the trivial representation of G. We conclude by applying another result by Guichardet ([**Gui72**], Théorème 1): for a unitary representation without non-zero fixed vectors, the space of 1-coboundaries is not closed in the space of 1-cocycles if and only if the representation weakly contains the trivial representation. □

References

[BHV08] B. Bekka, P. de la Harpe and A. Valette. *Kazhdan's property (T)*. Cambridge Univ. Press, 2008.

[Dix49] J. Dixmier. Etudes sur les variétés et les opérateurs de Julia, avec quelques applications. *Bull. Soc. Math. France*, 77:11–101, 1949.

[Dix77] J. Dixmier. *C*-algebras*. North Holland, 1977.

[FW71] P.A. Fillmore and J.P. Williams. On operator ranges. *Advances in Math.*, 7:254–281, 1971.

[Gui72] A. Guichardet. Sur la cohomologie des groupes topologiques II. *Bull. Sci. Math.*, 96:305–332, 1972.

[Gui80] A. Guichardet. *Cohomologie des groupes topologiques et des algèbres de Lie*. Cedic - F. Nathan, 1980.

[MV07] F. Martin and A. Valette. On the first L^p-cohomology of discrete groups. *Groups Geom. Dyn.*, 1:81-100, 2007.

[Sha00] Y. Shalom. Rigidity of commensurators and irreducible lattices. *Invent. Math.*, 141:1–54, 2000.

[Sha04] Y. Shalom Harmonic analysis, cohomology, and the large scale geometry of amenable groups. *Acta Math.*, 193:119–185, 2004.

Philip Morris International, Quai Jeanrenaud 56, CH-2000 Neuchâtel, Switzerland
E-mail address: florian.martin@pmintl.com

Institut de Mathématiques, Rue Emile Argand 11-BP158, CH-2009 Neuchâtel, Switzerland
E-mail address: alain.valette@unine.ch

Clay Mathematics Proceedings
Volume 11, 2010

Local index Formula and Twisted Spectral Triples

Henri Moscovici

Dedicated to Alain Connes, with admiration, friendship and gratitude

Abstract. We prove a local index formula for a class of twisted spectral triples of type III modeled on the transverse geometry of conformal foliations with locally constant transverse conformal factor. Compared with the earlier proofs in the untwisted case, the novel aspect resides in the fact that the twisted analogues of the JLO entire cocycle and of its retraction are no longer cocycles in their respective Connes bicomplexes. We show however that the passage to the infinite temperature limit, respectively the integration along the full temperature range against the Haar measure of the positive half-line, has the remarkable effect of curing in both cases the deviations from the cocycle and transgression identities.

Introduction

The local-global principle, as epitomized by the Atiyah-Singer index theorem but in the larger operator theoretic framework, has played a pivotal role in Alain Connes' overarching design of the foundations of Noncommutative Geometry. Although I was too overwhelmed by his brilliant intellect and fantastic mathematical insight to fully realize it at the time, this very theme was in fact the subtext of our first mathematical conversation, in the autumn of 1978, while we were both visiting the Institute for Advanced Study. As his program advanced, the theme gradually evolved into a perennial context for a substantial part of our collaborative work and, last but not least, it became a pretext for a close, lifelong friendship. As a token of my deep appreciation, I thought it would befit the occasion to try to run anew the machinery that has emerged, this time in the presence of a twist.

The basic template for a space in the Connes program is encoded in the notion of spectral triple. In our recent joint paper [**10**] we showed that this notion can be adapted to include certain type III spaces by the simple device of incorporating in it a twisting automorphism of the algebra of coordinates. The paradigmatic examples of such spaces are those describing the transverse geometry of a foliation

2010 *Mathematics Subject Classification.* Primary 58J22; Secondary 58J42.

This work was supported by the National Science Foundation awards nos. DMS-0245481 and DMS-0652167.

of codimension 1, or of a conformal foliation of arbitrary codimension. While it is always possible to associate a spectral triple of type II to any foliation by passing to the frame bundle of a complete transversal (*cf.* [**8**]), the very construction introduces a large number of additional parameters, which in turn makes the task of computing the characteristic classes quite formidable (see [**9**]). The twisting device on the other hand, whenever applicable, allows one to bypass the extra step of geometric reduction to type II.

Since the primary effect of the twisting automorphism is the replacement of the bimodule of noncommutative differential forms of a spectral triple with a bimodule of twisted differential forms, one would have expected the characteristic classes of a twisted spectral triple to be captured by a twisted version of the Connes-Chern character, landing in twisted cyclic cohomology. Somewhat counterintuitively, it turned out that no cohomological twisting is actually needed, and that Connes' original construction of the Chern character in noncommutative geometry [**3**] remains in fact operative in the twisted case as well. The natural question that arises is whether the Connes-Chern character of a twisted spectral triple can also be expressed in local terms as in [**8**], by means of a residue integral that eliminates all quantum infinitesimal perturbations of order strictly larger than 1.

In this paper we produce such a local index formula for a class of spectral triples twisted by scaling automorphisms, modeled on the geometry of a conformal foliation whose holonomy consists of germs of conformal transformations of $\mathbb{R}^n$. The proof is patterned on the strategy that evolved over a number of years of joint (and joyful) work leading up to the residue index formula [**5, 6, 7, 8**], with the important difference that the twisted counterpart of the JLO cocycle [**18**], which played a key role in our earlier proof (as well as in in Higson's [**16**]) is no longer a cocycle.

The plan of the paper is as follows. In §1 we recall from [**10**] the basic definitions concerning twisted spectral triples and their characters. Extrapolating from the expression of local Hochschild cocycle in *op.cit.* we then make a straightforward Ansatz in §2, predicting the form of the twisted local formula for the Connes-Chern character. In §3 we test the Ansatz on "real-life" examples of twisted spectral triples occurring in conformal geometry. These are the spectral triples describing the transverse geometry of the conformal foliations whose holonomy is given by germs of Möbius transformations of S^n. We conclude that the Ansatz holds true if the holonomy is restricted to the parabolic subgroup preserving a point, or equivalently to the similarity transformations of $\mathbb{R}^n$.

The main results of the paper are proved in §4, where we establish the validity of the Ansatz for an abstract class of twisted spectral triples, modeled on the conformal foliations with locally constant transverse conformal factor. While the twisted entire cochain analogous to [**18**] is no longer a cocycle, passing to the infinite temperature limit has the remarkable effect of restoring the cocycle identity, and the resulting cocycle can be expressed in terms of a residue integral as in [**8**]. To show that this residue cocycle represents the Connes-Chern character, one needs to transit through a transgressed version, as in [**7**]. In turn, the transgression process does yield a genuine cocycle because it involves integrating along the full temperature range, with respect to the Haar measure of $\mathbb{R}^+$, which miraculously again cures the deviation from the cocycle identity.

1. Twisted spectral triples and their characters

We begin by briefly reviewing the notion of twisted spectral triple of type III, introduced in [**10**], together with some of its basic properties.

1.1. Twisted spectral triple. A *twisted spectral triple* $(\mathcal{A}, \mathfrak{H}, D, \sigma)$ consists of a local Banach $*$-algebra $\mathcal{A}$ represented in the Hilbert space $\mathfrak{H}$ by bounded operators, an automorphism $\sigma \in \mathrm{Aut}\,(\mathcal{A})$, and a self-adjoint unbounded operator D such that, for any $a \in \mathcal{A}$,

$$a(D^2+1)^{-\frac{1}{2}} \in \mathcal{K}(\mathfrak{H}) \quad \text{(compact operators);} \tag{1.1}$$

$$a(\mathrm{Dom}\, D) \subset \mathrm{Dom}\, D \quad \text{and} \quad [D,a]_\sigma := D\,a - \sigma(a)\,D \quad \text{is bounded;} \tag{1.2}$$

$$\sigma(a^*) = (\sigma^{-1}(a))^*. \tag{1.3}$$

A $\mathbb{Z}_2$-*graded* (or *even*) σ-spectral triple has the additional datum of a grading operator

$$\gamma = \gamma^* \in \mathcal{L}(\mathfrak{H}), \quad \gamma^2 = I$$

which anticommutes with D, and commutes with the action of $\mathcal{A}$. In case the algebra itself is $\mathbb{Z}_2$-graded, the commutator's properties (including the twisted commutators) are understood in the graded sense.

We shall be concerned with finitely-summable twisted spectral triples, *i.e.* with those that satisfy a stronger version of (1.1), namely the (p,∞)-*summability* condition

$$a(D^2+1)^{-\frac{1}{2}} \in \mathcal{L}^{(p,\infty)}(\mathfrak{H}), \qquad \forall\, a \in \mathcal{A}, \tag{1.4}$$

for some $1 \le p < \infty$ of the same parity as the spectral triple. The notation is that of [**4**, IV, 2.α] ,

$$\mathcal{L}^{(p,\infty)}(\mathfrak{H}) = \{T \in \mathcal{K}(\mathfrak{H})\,;\ \sum_{i=0}^{N} \mu_i(T) = O(N^{1-\frac{1}{p}})\}, \qquad \text{if} \quad p > 1, \tag{1.5}$$

$$\mathcal{L}^{(1,\infty)}(\mathfrak{H}) = \{T \in \mathcal{K}(\mathfrak{H})\,;\ \sum_{i=0}^{N} \mu_i(T) = O(\log N)\}. \tag{1.6}$$

1.2. Graded double. As in the untwisted situation, there is a canonical way (*cf.* [**3**, Part I, §7]) to pass from an *ungraded* (or *odd*) twisted spectral triple $(\mathcal{A}, \mathfrak{H}, D, \sigma)$ to a $\mathbb{Z}_2$-graded twisted one over the $\mathbb{Z}_2$-graded algebra

$$\mathcal{A}_{\mathrm{gr}} = \mathcal{A} \otimes C_1.$$

Here C_1 denotes the Clifford algebra $\mathrm{Cliff}(\mathbb{R}) \otimes \mathbb{C}$; its even part is $C_1^+ = \mathbb{C}\,1$, with 1 the unit of C_1, while the odd part is $C_1^- = \mathbb{C}\,\epsilon$, with $\epsilon^2 = 1$. The automorphism remains σ, identified with $\sigma \otimes \mathrm{Id}$. One constructs an $\mathcal{A}_{\mathrm{gr}}$-module by first letting $\mathfrak{H}_1 = \mathfrak{H}_1^+ \oplus \mathfrak{H}_1^-$ be the $\mathbb{Z}_2$-graded Hilbert space with $\mathfrak{H}_1^\pm = \mathbb{C}$ on which C_1 acts via

$$\lambda 1 + \mu\epsilon \mapsto \begin{pmatrix} \lambda & \mu \\ \mu & \lambda \end{pmatrix}, \qquad \lambda, \mu \in \mathbb{C},$$

and then taking

$$\mathfrak{H}_{\mathrm{gr}} = \mathfrak{H} \otimes \mathfrak{H}_1, \quad \text{with} \quad \mathfrak{H}_{\mathrm{gr}}^\pm = \mathfrak{H} \otimes \mathfrak{H}_1^\pm,$$

on which $\mathcal{A}_{\rm gr}$ acts via the exterior tensor product representation; the corresponding grading operator is

$$\gamma \,=\, \mathrm{Id}_{\mathfrak{H}} \otimes \gamma_1, \quad \text{where} \quad \gamma_1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Finally, as operator one takes

$$D_{\rm gr} \,=\, D \otimes P_1, \quad \text{where} \quad P_1 = \begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix} : \mathfrak{H}_1 \to \mathfrak{H}_1.$$

1.3. Invertible double. When D is not invertible, we shall resort to the construction described in [**3**, Part I, §6] (akin to the passage from the Dirac operator on flat space to the Dirac Hamiltonian with mass) to canonically associate to $(\mathcal{A}, \mathfrak{H}, D, \gamma, \sigma)$ a new σ-spectral triple $(\mathcal{A}, \widetilde{\mathfrak{H}}, \widetilde{D}, \widetilde{\gamma}, \sigma)$ with invertible operator, defined as follows. With $\mathfrak{H}_1$ as above, one takes

$$\begin{aligned} \widetilde{\mathfrak{H}} \,&=\, \mathfrak{H} \hat{\otimes} \mathfrak{H}_1 \quad \text{(graded tensor product)}, \qquad \widetilde{\gamma} = \gamma \hat{\otimes} \gamma_1, \\ \widetilde{D} \,&=\, D \hat{\otimes} \mathrm{Id} + \mathrm{Id} \hat{\otimes} F_1, \quad \text{where} \quad F_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} : \mathfrak{H}_1 \to \mathfrak{H}_1. \end{aligned}$$

The algebra $\mathcal{A}$ is made to act on $\widetilde{\mathfrak{H}}$ via the representation

$$a \in \mathcal{A} \mapsto \tilde{a} := a \hat{\otimes} e_1, \quad \text{where} \quad e_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} : \mathfrak{H}_1 \to \mathfrak{H}_1.$$

1.4. Lipschitz regularity. In [**10**] a twisted spectral triple $(\mathcal{A}, \mathfrak{H}, D, \sigma)$ with invertible D was called *Lipschitz regular* if it satisfies the additional condition

$$[|D|, a]_\sigma := |D|\, a - \sigma(a)\, |D| \quad \text{is bounded}, \quad \forall\, a \in \mathcal{A}. \tag{1.7}$$

Such a twisted spectral triple can be 'untwisted' by passage to its 'phase' operator $F = D\,|D|^{-1}$. Indeed,

$$[F, a] \,=\, |D|^{-1} \big((D\, a \,-\, \sigma(a)\, D) \,-\, (|D|\, a - \sigma(a)\, |D|)\, F \big),$$

which shows that these commutators are compact operators, of the same order of magnitude as D^{-1}. Thus, $(\mathfrak{H}, F)$ is a Fredholm module over $\mathcal{A}$, defining a K_*-cycle over the norm closure C^*-algebra of $\mathcal{A}$. Moreover, if $(\mathcal{A}, \mathfrak{H}, D, \sigma)$ is (p, ∞)-summable so is $(\mathfrak{H}, F)$.

We extend the Lipschitz regularity condition to any spectral triple $(\mathcal{A}, \mathfrak{H}, D, \sigma)$ by requiring that its invertible double $(\mathcal{A}, \widetilde{\mathfrak{H}}, \widetilde{D}, \widetilde{\gamma}, \sigma)$ be Lipschitz regular. Since

$$\widetilde{F} := \widetilde{D} |\widetilde{D}|^{-1} = D(D^2+1)^{-\frac{1}{2}} \hat{\otimes} \mathrm{Id} + (D^2+1)^{-\frac{1}{2}} \hat{\otimes} F_1,$$

and therefore

$$[\widetilde{F}, a] \,=\, [D(D^2+1)^{-\frac{1}{2}}, a] \hat{\otimes} e_{11} \,+\, (D^2+1)^{-\frac{1}{2}} a \hat{\otimes} e_{21} \,-\, a(D^2+1)^{-\frac{1}{2}} \hat{\otimes} e_{12},$$

Lipschitz regularity can be alternatively phrased as the requirement

$$[D(D^2+1)^{-\frac{1}{2}}, a] \in \mathcal{K}(\mathfrak{H}), \qquad \forall\, a \in \mathcal{A}\,. \tag{1.8}$$

In the (p, ∞)-summable case these commutators belong to the ideal $\mathcal{L}^{(p,\infty)}(\mathfrak{H})$ of $\mathcal{K}(\mathfrak{H})$.

1.5. Connes-Chern character. By resorting if necessary to the doubling procedures described in 1.2 and 1.3, we may assume without essential loss of generality that the (p,∞)-summable twisted spectral triple $(\mathcal{A},\mathfrak{H},D,\sigma)$ under consideration is $\mathbb{Z}_2$-graded and D invertible. We shall often do so in the sequel without any further mention.

Let $(\mathcal{A},\mathfrak{H},D,\sigma)$ be such a twisted spectral triple which is also *Lipschitz-regular*. Then, as remarked in 1.4, it gives rise to a 'phase' Fredholm module $(\mathfrak{H},F)$ over $\mathcal{A}$. In turn, the latter has a well-defined Connes-Chern character in cyclic cohomology, *cf.* [**3**], which up to a normalizing constant is represented by the cyclic cocycle

$$\tau_F^p(a_0,a_1,\ldots,a_p) := \operatorname{Tr}(\gamma\, F\,[F,a_0]\,[F,a_1]\cdots[F,a_p])\,,\qquad a_i\in\mathcal{A}. \tag{1.9}$$

In [**10**], we have actually shown that a (p,∞)-summable twisted spectral triple as above admits a Connes-Chern character without assuming Lipschitz regularity. Indeed, we showed that the $(p+1)$-linear form on $\mathcal{A}$

$$\tau_{F_\sigma}^p(a_0,a_1,\ldots,a_p) := \operatorname{Tr}(\gamma\, D^{-1}[D,a_0]_\sigma\, D^{-1}[D,a_1]_\sigma\cdots D^{-1}[D,a_p]_\sigma), \tag{1.10}$$

is a cyclic cocycle in $Z_\lambda^n(\mathcal{A})$, by means of which one can recover the index pairing of D with $K^*(\mathcal{A})$. More precisely, up to a universal constant factor c_p, $\operatorname{Index}(\sigma(e)\,D\,e)$ is given by $\tau_{F_\sigma}^p(e,\ldots,e)$, for any class $[e]\in K^0(\mathcal{A})$.

1.6. Conformally perturbed spectral triple. An instructive class of examples of twisted spectral triples arises from conformal-type perturbations of ordinary spectral triples. Let $(\mathcal{A},\mathfrak{H},D)$ be a (p,∞)-summable spectral triple, and let $h=h^*\in\mathcal{A}$. By setting

$$D_h = e^h\,D\,e^h,\quad\text{and}\quad \sigma_h(a) = e^{2h}\,a\,e^{-2h},\quad \forall\, a\in\mathcal{A}$$

one easily sees that

$$D_h\,a - \sigma_h(a)\,D_h = e^h\,[D,\,\sigma_{h/2}(a)]\,e^h\in\mathcal{L}(\mathfrak{H})\,, \tag{1.11}$$

thus giving rise to a twisted spectral triple $(\mathcal{A},\mathfrak{H},D_h,\sigma_h)$. Noting that

$$D_h^{-1}\,[D_h,a]_\sigma = e^{-h}\,D^{-1}\,[D,\,\sigma_{h/2}(a)]\,e^h\,, \tag{1.12}$$

one obtains the identity

$$\begin{aligned}\operatorname{Tr}(\gamma\, D_h^{-1}\,[D_h,a_0]_\sigma\, D_h^{-1}\,[D_h,a_1]_\sigma\cdots D_h^{-1}\,[D_h,a_p]_\sigma)&\\ = \operatorname{Tr}(\gamma\, D^{-1}[D,\sigma_{h/2}(a_0)]\,D^{-1}[D,\sigma_{h/2}(a_1)]\cdots D^{-1}[D,\sigma_{h/2}(a_p)])\,.&\end{aligned} \tag{1.13}$$

The right hand side is a cyclic cocycle on $\mathcal{A}$ that represents, for $h=1$ and up to normalization, the Connes-Chern character

$$Ch^*(\mathcal{A},\mathfrak{H},D)\in HC^*(\mathcal{A})$$

of $(\mathcal{A},\mathfrak{H},D)$ viewed as a Fredholm module, *cf.* [**3**, Part I, §6]. It follows that the left hand side, which is the cyclic cocycle obtained by composition with the inner automorphism $\sigma_{h/2}$, determines the same class in periodic cyclic cohomology. This justifies regarding (1.13) as defining the *periodic Connes-Chern character* of the conformally perturbed spectral triple:

$$Ch^*(\mathcal{A},\mathfrak{H},D_h,\sigma_h) := Ch^*(\mathcal{A},\mathfrak{H},D)\in HP^*(\mathcal{A}). \tag{1.14}$$

1.7. Local Hochschild cocycle. With the goal of extending the local index formula of [**8**] to twisted spectral triples, we looked in [**10**] for an analogue of Connes' residue formula [**4**, IV.2.γ] for the Hochschild class of the Connes-Chern character.

We recall that if $(\mathcal{A}, \mathfrak{H}, D)$ is an (even, invertible) (p, ∞)-summable spectral triple satisfying the *smoothness condition*

$$\mathcal{A},\, [D, \mathcal{A}] \subset \bigcap_{k>0} \mathrm{Dom}\,(\delta^k), \quad \text{where} \quad \delta(T) := [|D|, T]\,, \tag{1.15}$$

the Hochschild cohomology class $I\big(Ch^*(\mathcal{A}, \mathfrak{H}, D)\big) \in HH^*(\mathcal{A})$ admits a *local* representation, given by the formula

$$\varkappa_D(a_0, a_1, \ldots, a_p) := \mathrm{Tr}_\omega\big(\gamma a_0 [D, a_1] \cdots [D, a_p]\, D^{-p}\big), \qquad a_i \in \mathcal{A}\,, \tag{1.16}$$

which defines a Hochschild cocycle. Here Tr_ω stands for a Dixmier trace (see [**4**, IV.2.β, γ]) on the ideal $\mathcal{L}^{(1,\infty)}(\mathfrak{H})$. The local nature of the above formula stems from the fact that the Dixmier trace vanishes on the subideal

$$\mathcal{L}_0^{(1,\infty)}(\mathfrak{H}) \;=\; \{T \in \mathcal{K}(\mathfrak{H})\,; \quad \sum_{i=0}^{N} \mu_i(T) = o(\log N)\},$$

which contains the trace class operators, and in particular all smoothing operators on any closed manifold. Thus, $\varkappa_D(a_0, a_1, \ldots, a_p)$ depends only on the class of the operator $a_0[D, a_1] \cdots [D, a_p]\, D^{-p} \in \mathcal{L}^{(1,\infty)}(\mathfrak{H})$ modulo $\mathcal{L}_0^{(1,\infty)}(\mathfrak{H})$, which plays the role of its symbol.

Using the identity

$$[D, a]\, D^{-k} \;=\; D^{-k+1}\left(D^k\, a\, D^{-k} \;-\; D^{k-1}\, a\, D^{-k+1}\right), \qquad \forall\, a \in \mathcal{A}\,, \tag{1.17}$$

one can successively move D^{-p} to the left and rewrite the cocycle $\varkappa_D$ in the form

$$\varkappa_D(a_0, a_1, \ldots, a_p) = \mathrm{Tr}_\omega\big(\gamma a_0 (D a_1 D^{-1} - a_1) \cdots (D^p a_p D^{-p} - D^{p-1} a_p D^{-p+1})\big).$$

In the twisted case, taking a clue from (1.12), one is led to make the formal substitution

$$D^k\, a\, D^{-k} \quad \longmapsto \quad D^k\, \sigma^{-k}(a)\, D^{-k}\,, \qquad \forall\, a \in \mathcal{A}\,, \tag{1.18}$$

and use the twisted version of (1.17), namely

$$[D, \sigma^{-k}(a)]_\sigma\, D^{-k} \;=\; D^{-k+1}\left(D^k\, \sigma^{-k}(a)\, D^{-k} \;-\; D^{k-1}\, \sigma^{-k+1}(a)\, D^{-k+1}\right), \tag{1.19}$$

to reverse the process of distributing D^{-p} among the factors. Assuming that the domain condition which permits the above operation is fulfilled, one thus arrives at the expression

$$\varkappa_{D,\sigma}(a_0, a_1, \ldots, a_p) \;:=\; \mathrm{Tr}_\omega\big(\gamma a_0 [D, \sigma^{-1}(a_1)]_\sigma \cdots [D, \sigma^{-p}(a_p)]_\sigma D^{-p}\big)\,. \tag{1.20}$$

This was shown in [**10**] to indeed be a Hochschild p-cocycle, and it will be useful to reproduce the elementary calculation that validates this statement. It relies on two basic properties of twisted spectral triples. The first is the obvious fact that the σ-bracket with D satisfies the twisted derivation rule

$$[D, ab]_\sigma \;=\; \sigma(a)\, [D, b]_\sigma \;+\; [D, a]_\sigma\, b, \qquad a, b \in \mathcal{A}. \tag{1.21}$$

The second is the observation that the positive linear functional on $\mathcal{A}$,

$$a \mapsto \operatorname{Tr}_\omega(a\,|D|^{-p}),$$

is a σ^{-p}-trace on $\mathcal{A}$, *i.e.* satisfies

$$\operatorname{Tr}_\omega(a\,b\,|D|^{-p}) \,=\, \operatorname{Tr}_\omega(b\,\sigma^{-p}(a)\,|D|^{-p}), \qquad \forall\, a, b \in \mathcal{A}\,.$$

It is in fact a σ^{-p}-hypertrace, since for any $T \in \mathcal{L}(\mathfrak{H})$ one still has

$$\operatorname{Tr}_\omega(T\,\sigma^{-p}(a)\,|D|^{-p}) \,=\, \operatorname{Tr}_\omega(a\,T\,|D|^{-p})\,. \tag{1.22}$$

Making use of the Leibniz rule (1.21), one computes the Hochschild coboundary of $\varkappa_{D,\sigma} \in Z^p(\mathcal{A}, \mathcal{A}^*)$ as follows:

$$\begin{aligned}
&b\varkappa_{D,\sigma}(a_0, a_1, ..., a_{p+1}) \,= \\
&\quad = \sum_{i=0}^{p} (-1)^i\, \varkappa_{D,\sigma}(a_0, ..., a_i a_{i+1}, ..., a_{p+1}) \,+\, (-1)^{p+1} \varkappa_{D,\sigma}(a_{p+1}a_0, a_1, ..., a_p) \\
&\quad = \operatorname{Tr}_\omega\big(\gamma\, a_0\, a_1\, [D, \sigma^{-1}(a_2)]_\sigma \cdots\, [D, \sigma^{-p}(a_{p+1})]_\sigma\, D^{-p}\big) \\
&\qquad\qquad -\operatorname{Tr}_\omega\big(\gamma\, a_0 a_1\, [D, \sigma^{-1}(a_2)]_\sigma \cdots\, [D, \sigma^{-p}(a_{p+1})]_\sigma\, D^{-p}\big) \\
&\qquad\qquad -\operatorname{Tr}_\omega\big(\gamma\, a_0\, [D, \sigma^{-1}(a_1)]_\sigma\, \sigma^{-1}(a_2) \cdots\, [D, \sigma^{-p}(a_{p+1})]_\sigma\, D^{-p}\big) + \ldots \\
&\ldots + (-1)^p \operatorname{Tr}_\omega\big(\gamma a_0 [D, \sigma^{-1}(a_1)]_\sigma \cdots \\
&\qquad\qquad \cdots [D, \sigma^{-p+1}(a_{p-1})]_\sigma \sigma^{-p+1}(a_p) [D, \sigma^{-p}(a_{p+1})]_\sigma D^{-p}\big) + \\
&+(-1)^p \operatorname{Tr}_\omega\big(\gamma a_0 [D, \sigma^{-1}(a_1)]_\sigma \cdots \\
&\qquad\qquad \cdots [D, \sigma^{-p+1}(a_{p-1})]_\sigma [D, \sigma^{-p}(a_p)]_\sigma \sigma^{-p}(a_{p+1}) D^{-p}\big) + \\
&+(-1)^{p+1} \operatorname{Tr}_\omega\big(\gamma a_{p+1} a_0\, [D, \sigma^{-1}(a_1)]_\sigma \cdots [D, \sigma^{-p}(a_p)]_\sigma\, D^{-p}\big) \qquad = \qquad 0\,.
\end{aligned}$$

The end result is 0 because the successive terms cancel in pairs, with the last two terms canceling each other thanks to the enhanced σ^{-p}-trace property Eq. (1.22).

2. Ansatz for a twisted local index formula

2.1. Local index formula for spectral triples. The local index formula that delivers in full the Connes-Chern character in cyclic cohomology was developed in [**8**], in terms of residue functionals that generalize Wodzicki's noncommutative residue, and in the framework of an abstract pseudodifferential calculus which gives a precise meaning to the notion of symbol. We briefly recall the salient notions.

Let $(\mathcal{A}, \mathfrak{H}, D)$ be a (p, ∞)-summable spectral triple (with D invertible), which satisfies the smoothness condition (1.15). We denote by $\mathcal{B}$ the algebra generated by $\bigcup_{k\geq 0} \delta^k(\mathcal{A} + [D, \mathcal{A}])$, and also set $\mathfrak{H}^\infty := \bigcap_{k\geq 0} \operatorname{Dom}(|D|^k)$. We now consider linear operators $P : \mathfrak{H}^\infty \to \mathfrak{H}^\infty$ that admit an expansion of the form

$$P \,\sim\, \sum_{k\geq 0} b_k |D|^{s-k}\,, \qquad \text{with} \qquad b_k \in \mathcal{B}\,, \quad s \in \mathbb{C}\,, \tag{2.1}$$

in the sense that

$$P - \sum_{0\le k<N} b_k |D|^{s-k} \in \mathcal{B}\cdot \mathrm{OP}^{\Re s - N}, \qquad \forall N>0,$$

where $\qquad R\in \mathrm{OP}^r \iff |D|^{-r}R \in \bigcap_{k>0} \mathrm{Dom}\,(\delta^k),$

and $\qquad \mathcal{B}\cdot\mathrm{OP}^r := \left\{ \sum_j b_j R_j\, ; \quad b_j\in\mathcal{B}, \quad R_j\in \mathrm{OP}^r\right\}.$

Thanks to the key commutation relation (see [**8**, Appendix B, Thm. B.1])

$$|D|^s\, b \sim \sum_{k\ge 0} \frac{s(s-1)\cdots(s-k+1)}{k!}\, \delta^k(b)\, |D|^{s-k}, \qquad \forall\, b\in\mathcal{B}, \quad s\in\mathbb{C}, \tag{2.2}$$

these operators form a filtered algebra.

Let now $\mathcal{D}(\mathcal{A},\mathfrak{H},D)$ be the algebra generated by $\bigcup_{k\ge 0}\nabla^k(\mathcal{A}+[D,\mathcal{A}])$, where ∇ denotes the derivation $\nabla(T)=[D^2,T]$. Its elements play the role of *differential operators*. They too have a natural order, determined by total power of ∇ involved in each monomial. Furthermore, the analogue of (2.2) holds: for any q-th order operator $T\in\mathcal{D}^q(\mathcal{A},\mathfrak{H},D)$ and $N>0$,

$$D^{2s}\, T - \sum_{0\le k<N} \frac{s(s-1)\cdots(s-k+1)}{k!}\, \nabla^k(T)\, D^{2(s-k)} \in \mathrm{OP}^{2\Re s + q - N}. \tag{2.3}$$

The intrinsic pseudodifferential calculus for the spectral triple is based on the algebra $\Psi^\bullet(\mathcal{A},\mathfrak{H},D)$, generated by the operators defined by (2.1) together with the differential operators. It is a filtered algebra, and its quotient modulo the ideal of smoothing operators $\Psi^{-\infty}(\mathcal{A},\mathfrak{H},D) = \bigcap_{N\ge 0}\mathcal{B}\cdot\mathrm{OP}^{-N}$ gives the corresponding algebra of *complete symbols* $\mathfrak{CS}^\bullet(\mathcal{A},\mathfrak{H},D) := \Psi^\bullet(\mathcal{A},\mathfrak{H},D)/\Psi^{-\infty}(\mathcal{A},\mathfrak{H},D)$.

Underlying the setup for the local index formula is the essential assumption that the spectral triple admits a *discrete dimension spectrum*, to which all singularities of zeta functions associated to elements of $\mathcal{B}$ are confined; the postulated spectrum is a *discrete subset* $\Sigma\subset\mathbb{C}$, such that the holomorphic functions

$$\zeta_b(z) = \mathrm{Tr}(b\,|D|^{-z}), \qquad \Re z > p, \quad b\in\mathcal{B}, \tag{2.4}$$

admit holomorphic extensions to $\mathbb{C}\setminus\Sigma$. This requirement is supplemented by a technical condition stipulating that *the functions* $\Gamma(z)\,\zeta_b(z)$ *decay rapidly on finite vertical strips.*

For the sake of convenience, we shall make the stronger assumption that the dimension spectrum is *simple, i.e.* Σ *consists of simple poles.* Then, for any $P\in\Psi^N(\mathcal{A},\mathfrak{H},D)$ the zeta function

$$\zeta_P(z) = \mathrm{Tr}(P\,|D|^{-z}), \qquad \Re z > p+N, \tag{2.5}$$

can be meromorphically continued to $\mathbb{C}$, with simple poles in $\Sigma + N$. Furthermore, the *residue functional*

$$⨍_D P := \mathrm{Res}_{z=0}\,\zeta_P(2z), \qquad P\in\Psi^\bullet(\mathcal{A},\mathfrak{H},D) \tag{2.6}$$

is an (algebraic) trace. By its very construction it vanishes on $\Psi^\bullet(\mathcal{A},\mathfrak{H},D)\cap\mathcal{L}^1(\mathfrak{H})$, and in particular it descends to a trace on $\mathcal{CS}^\bullet(\mathcal{A},\mathfrak{H},D)$.

The local index formula expresses the Connes-Chern character $Ch^*(\mathcal{A},\mathfrak{H},D)\in HC^*(\mathcal{A})$ in terms of a cocycle in the bicomplex $\{CC(\mathcal{A}),b,B\}$, whose components are defined by means of the symbolic trace. In the (invertible) odd case, these components are as follows: for $q=2\ell+1, \quad \ell\in\mathbb{Z}^+$,

$$\tau^q_{\rm odd}(a_0,\ldots,a_q) \;=\; \sqrt{2i}\sum_{\mathbf{k}} c_{q,\mathbf{k}} ⨍_D a_0\,[D,a_1]^{(k_1)}\cdots[D,a_q]^{(k_q)}\,|D|^{-2|\mathbf{k}|-q}\,, \tag{2.7}$$

where

$$P^{(k)} \;=\; \nabla^k(P)\,, \qquad \forall\, P\in\Psi^\bullet(\mathcal{A},\mathfrak{H},D)\,,$$

and the coefficients are given by

$$c_{q,\mathbf{k}} \;=\; \frac{(-1)^{|\mathbf{k}|}}{k_1!\cdots k_q!\,(k_1+1)\cdots(k_1+\cdots+k_q+q)}\,\Gamma\left(|\mathbf{k}|+\frac{q}{2}\right)\,, \tag{2.8}$$

$$\text{where} \qquad \mathbf{k} \;=\; (k_1,\ldots,k_q)\,, \qquad |\mathbf{k}|=k_1+\cdots+k_q\,.$$

In the even case, with $q=2\ell, \quad \ell\in\mathbb{Z}^+$,

$$\tau^0_{\rm ev}(a) \;=\; \mathrm{Res}_{z=0}\big(\Gamma(z)\,\mathrm{Tr}(a\,|D|^{-2z})\big)\,, \tag{2.9}$$

$$\tau^q_{\rm ev}(a_0,\ldots,a_q) \;=\; \sum_{\mathbf{k}} c_{q,\mathbf{k}} ⨍_D \gamma\, a_0\,[D,a_1]^{(k_1)}\cdots[D,a_q]^{(k_q)}\,|D|^{-2|\mathbf{k}|-q}\,.$$

Each component $\tau^q=\tau^q_{\rm ev/odd}$ has finitely many nonzero summands, and $\tau^q\equiv 0$ for any $q>p$.

Since the expressions $\tau^q(a_0,\ldots,a_q)$ are unaffected by the scaling $D\mapsto tD$, $t\in\mathbb{R}$, we can write them in terms of the scale-invariant operators

$$\alpha^k(a) \;:=\; D^k a D^{-k}\,, \qquad a\in\mathcal{A}.$$

Indeed, using the obvious identity

$$[D,a]^{(k+1)}\,D^{-2k-3} \;=\; D^2\left([D,a]^{(k)}\,D^{-2k-1}\right)D^{-2} \;-\; [D,a]^{(k)}\,D^{-2k-1}\,,$$

one verifies by induction that for any $k\geq 0$,

$$[D,a]^{(k)}\,D^{-2k-1} \;=\; \sum_{j=0}^{k}(-1)^j\binom{k}{j}\left(\alpha^{2(k-j)+1}(a) \;-\; \alpha^{2(k-j)}(a)\right).$$

Therefore, for any $\ell\in\mathbb{Z}$,

$$\begin{aligned}[D,a]^{(k)}\,D^{-\ell} \;&=\; \sum_{j=0}^{k}(-1)^j\binom{k}{j}\left(\alpha^{2(k-j)+1}(a) \;-\; \alpha^{2(k-j)}(a)\right)D^{2k+1-\ell}\\ &=\; D^{2k+1-\ell}\sum_{j=0}^{k}(-1)^j\binom{k}{j}\left(\alpha^{\ell-2j}(a) \;-\; \alpha^{\ell-2j-1}(a)\right).\end{aligned}$$

With the abbreviated notation

$$\Sigma^{(k,\ell)}(a) \;:=\; \sum_{j=0}^{k}(-1)^j\binom{k}{j}\left(\alpha^{\ell-2j}(a) \;-\; \alpha^{\ell-2j-1}(a)\right)\,, \tag{2.10}$$

the above equality takes the form

$$[D,a]^{(k)}\,D^{-\ell} \;=\; D^{2k+1-\ell}\,\Sigma^{(k,\ell)}(a). \tag{2.11}$$

Successive application of the identity (2.11) brings the components τ^q with $q>0$ to the form

$$\tau^q(a_0,\ldots,a_q) = \sum_{\mathbf{k}} c_{q,\mathbf{k}} ⨍_D \gamma\, a_0\, \Sigma^{(k_1,\,2k_1+1)}(a_1)\cdots\Sigma^{(k_q,\,2(k_1+\cdots+k_q)+q)}(a_q)\,. \tag{2.12}$$

This formula covers the case of either parity, provided that in the odd case we define $\gamma := F = D|D|^{-1}$, and incorporate the factor $\sqrt{2i}$ in the expression (2.8) of the coefficients $c_{q,\mathbf{k}}$ for q odd.

2.2. Ansatz for the twisted case. Assume now that $(\mathcal{A},\mathfrak{H},D,\sigma)$ is a (p,∞)-summable (invertible) twisted σ-spectral triple. The twisted analogue of the usual bimodule of *gauge potentials* or *noncommutative differential forms* is obviously the linear subspace $\Omega^1_{D,\sigma}(\mathcal{A})\subset\mathcal{L}(\mathfrak{H})$ consisting of operators of the form

$$A = \sum_i a_i\big(D\,b_i - \sigma(b_i)\,D\big)\,, \qquad a_i, b_i\in\mathcal{A}\,,$$

which is a bimodule for the action

$$a\cdot\omega\cdot b = \sigma(a)\,\omega\,b\,, \quad \forall\, a,b\in\mathcal{A}\,, \quad \forall\,\omega\in\Omega^1_{D,\sigma}(\mathcal{A})\,.$$

In the presence of the Lipschitz regularity axiom (1.8), one can similarly define a bimodule $|\Omega|^1_{D,\sigma}(\mathcal{A})\subset\mathcal{L}(\mathfrak{H})$, by simply replacing D with $|D|$. Furthermore, as noted before *cf.* (1.21), the map

$$a\mapsto d_\sigma(a) = \,D\,a - \sigma(a)\,D$$

is a σ-derivation of $\mathcal{A}$ with values in $\Omega^1_{D,\sigma}(\mathcal{A})$, and clearly, so is the map

$$a\mapsto \delta_\sigma(a) = \,|D|\,a - \sigma(a)\,|D|\,.$$

However, in order for the analogue of the smoothness condition (1.15) to make sense, one needs to postulate the existence of an extension of the automorphism $\sigma\in\mathrm{Aut}\,(\mathcal{A})$, and consequently of the σ-derivation δ_σ, to a larger subalgebra of $\mathcal{L}(\mathfrak{H})$, which should contain $\Omega^1_{D,\sigma}(\mathcal{A})$ as well as its higher δ_σ-iterations.

Thus, the formulation of a twisted version for the pseudodifferential calculus is *not* canonical. Ignoring this aspect for now, let us pretend that an adequate analogue $\Psi^\bullet(\mathcal{A},\mathfrak{H},D,\sigma)$ of the algebra of pseudodifferential operators has already been constructed, and assume that the twisted σ-spectral triple $(\mathcal{A},\mathfrak{H},D,\sigma)$ admits a simple discrete dimension spectrum $\Sigma\subset\mathbb{C}$. We can then focus on finding an appropriate candidate for the local character cocycle. Denoting

$$\alpha^k_\sigma(a) \;:=\; D^k\,\sigma^{-k}(a)\,D^{-k}\,, \qquad a\in\mathcal{A}, \tag{2.13}$$

and

$$\Sigma^{(k,\ell)}_\sigma(a) \;:=\; \sum_{j=0}^{k}(-1)^j\binom{k}{j}\,\left(\alpha^{\ell-2j}_\sigma(a) \;-\; \alpha^{\ell-2j-1}_\sigma(a)\right)\,, \tag{2.14}$$

the analogues of the summands in Eq. (2.12) are the 'residue integrals'

$$⨍_D \gamma\, a_0\, \Sigma^{(k_1,\,2k_1+1)}_\sigma(a_1)\cdots\Sigma^{(k_q,\,2(k_1+\cdots+k_q)+q)}_\sigma(a_q)\,. \tag{2.15}$$

We next define the twisted version of the higher commutators as follows:

$$(a)_\sigma^{(k)} := \sum_{j=0}^{k} (-1)^j \binom{k}{j} D^{2(k-j)} \sigma^{2j}(a) D^{2j}\,, \tag{2.16}$$

respectively

$$[D,a]_\sigma^{(k)} := \sum_{j=0}^{k} (-1)^j \binom{k}{j} \left(D^{2(k-j)+1} \sigma^{2j}(a) D^{2j} - D^{2(k-j)} \sigma^{2j+1}(a) D^{2j+1} \right) . \tag{2.17}$$

REMARK 2.1. Noting that

$$\begin{aligned} (a)_\sigma^{(k+1)} &= D^2\, (a)_\sigma^{(k)} - (\sigma^2(a))_\sigma^{(k)}\, D^2\,, \\ [D,a]_\sigma^{(k+1)} &= D^2\, [D,a]_\sigma^{(k)} - [D,\sigma^2(a)]_\sigma^{(k)}\, D^2\,, \end{aligned}$$

one could be tempted to regard the expressions (2.16), (2.17) as genuine iterated twisted commutators with D^2. However, that would not be correct, because there is no guarantee that if, for instance, $[D,a]_\sigma = 0$ then $[D,a]_\sigma^{(k)} = 0$ for all $k \geq 1$. A counterexample can be easily obtained in the setting of §3.1, using Eq. (3.12).

At any rate, with the above notation, we can now put (2.15) in a form similar to Eq. (2.9). Indeed, the counterpart of the identity (2.11) is

$$D^{2k+1-\ell}\, \Sigma_\sigma^{(k,\ell)}(a) = [D, \sigma^{-\ell}(a)]_\sigma^{(k)}\, D^{-\ell},$$

which we then employ to reverse the process by which the expression (2.12) was obtained from (2.7). Keeping the same notational conventions used in (2.12), one thus arrives at the following Ansatz for the twisted version of the local character cocycle:

$$\tau_\sigma^q(a_0, \ldots, a_q) = \sum_{\mathbf{k}} c_{q,\mathbf{k}} ⨍_D \gamma\, a_0\, [D, \sigma^{-2k_1-1}(a_1)]_\sigma^{(k_1)} \cdots [D, \sigma^{-2(k_1+\cdots+k_q)-q}(a_q)]_\sigma^{(k_q)}\, |D|^{-2|\mathbf{k}|-q}. \tag{2.18}$$

There is an immediate obstruction for this formula to define a (b,B)-cocycle, which arises from the B-coboundary of τ_{odd}^1. Indeed,

$$\begin{aligned} B\tau_\sigma^1(a) &= \sum_{k\geq 0} c_{1,k} ⨍_D F\, [D, \sigma^{-2k-1}(a)]_\sigma^{(k)}\, D^{-2k-1} = \sum_{k\geq 0} c_{1,k} ⨍_D F\, \Sigma_\sigma^{(k,\,2k+1)}(a) \\ &= \sum_{k\geq 0} c_{1,k} \sum_{j=0}^{k} (-1)^j \binom{k}{j} \left(⨍_D F\, a_\sigma^{2(k-j)+1}(a) - ⨍_D F\, \alpha_\sigma^{2(k-j)}(a) \right) \\ &= \sum_{k\geq 0} c_{1,k} \sum_{j=0}^{k} (-1)^j \binom{k}{j} \left(⨍_D F\, \sigma^{-2(k-j)-1}(a) - ⨍_D F\, \sigma^{-2(k-j)}(a) \right). \end{aligned}$$

This expression vanishes if

$$⨍_D F\, \sigma(a) = ⨍_D F\, a\,, \qquad \forall\, a \in \mathcal{A}, \tag{2.19}$$

or equivalently

$$⨍_D [D,a]_\sigma\, |D|^{-1} = 0\,, \qquad \forall\, a \in \mathcal{A}\,,$$

As will become apparent in the next section, Eq. (2.19) has something in common with the Selberg principle for orbital integrals of reductive Lie groups. We shall show later that for a special class of conformally twisted spectral triples there are no higher obstructions to the validity of the Ansatz.

3. Conformal geometry and twisted spectral triples

In order to shed some light on the nature and plausibility of the above setup for the Ansatz, we examine in this section some authentic examples of twisted spectral triples arising in conformal geometry.

3.1. Transversely conformal spectral triple. Let M be a smooth connected closed spin manifold of dimension n. To each Riemannian metric g on M one can canonically associate a Dirac operator $D = \not{D}_g$ acting on the Hilbert space $\mathfrak{H} = \mathfrak{H}_g := L^2(M, \mathcal{S}_g)$ of L^2-sections of the spin bundle $\mathcal{S} = \mathcal{S}_g$, and thus a corresponding spectral triple $(C^\infty(M), \mathfrak{H}, D)$ over the algebra $C^\infty(M)$. Assume now that M is endowed with a conformal structure $[g]$, consisting of all Riemannian metrics conformally equivalent to a given Riemannian metric g. Let $SCO(M, [g])$ be the group of diffeomorphisms of M that preserve the conformal structure, the orientation and the spin structure. It is a Lie group, and we denote by $G = SCO(M, [g])_0$ its connected component of the identity. We then form the discrete crossed product algebra $\mathcal{A}_G = C^\infty(M) \rtimes G$. This algebra consists of finite sums of the form

$$a = \sum_\Gamma f_\phi \, v_\phi \,, \qquad f_\phi \in C^\infty(M), \quad \phi \in G,$$

with the product rule determined by

$$v_\phi \, f = (f \circ \phi^{-1}) \, v_\phi \,, \quad v_\phi \, v_\psi = v_{\phi \, \psi} \,.$$

It can be represented by bounded linear operators on the Hilbert space $\mathfrak{H} = L^2(M, \mathcal{S})$ of L^2-sections of the spin bundle $\mathcal{S}$, by letting a function $f \in C^\infty(M)$ act as the multiplication operator

$$\pi(f)\,(u) = f\,u \,, \qquad u \in L^2(M, \mathcal{S}) \,, \tag{3.1}$$

and the diffeomorphisms $\phi \in G$ act as translation operators

$$\pi(v_\phi)\,(u) \equiv V_\phi\,(u) := \tilde{\phi} \circ u \circ \phi^{-1} \,, \qquad u \in L^2(M, \mathcal{S}) \,, \tag{3.2}$$

where $\tilde{\phi}$ is the canonical lift of ϕ to an automorphism of $\mathcal{S}$; such a lift is well-defined, not just modulo $\mathbb{Z}/2\mathbb{Z}$, for any $\phi \in SCO(M, [g])_0$. To make G act by unitary operators, one needs to replace each operator V_ϕ^{-1}, $\phi \in G$, by the operator

$$U_\phi^{-1}\,(u) = e^{-nh_\phi}\, V_\phi^{-1}\,(u) = e^{-nh_\phi}\, \tilde{\phi}^{-1} \circ u \circ \phi \,, \qquad u \in L^2(M, \mathcal{S}) \,, \tag{3.3}$$

where $h_\phi \in C^\infty(M)$ is determined by the conformal factor via the equation

$$\phi^*(g) = e^{-4h_\phi}\, g \,. \tag{3.4}$$

Indeed, using the fact that the Riemannian volume forms are related by the equality

$$\mathrm{vol}_{\phi^*(g)} = e^{-2nh_\phi}\, \mathrm{vol}_g \,, \tag{3.5}$$

and denoting the fiberwise norm by $|\cdot|$, one easily checks that U_ϕ^{-1} is unitary:

$$\begin{aligned} \|U_\phi^{-1}(u)\|^2 &= \int_M e^{-2nh_\phi}\,|\tilde\phi^{-1}(u\circ\phi)|^2\,\mathrm{vol}_g = \int_M |\tilde\phi^{-1}(u\circ\phi)|^2\,\phi^*(\mathrm{vol}_g) \\ &= \int_M |u|^2\,\mathrm{vol}_g = \|u\|^2, \qquad \forall\, u\in L^2(M,\mathcal{S})\,. \end{aligned}$$

LEMMA 3.1. *For any $\phi\in G = SCO(M,[g])_0$ and with $D = D\!\!\!/_g$, one has*

$$U_\phi^*\circ D\circ U_\phi = e^{h_\phi}\circ D\circ e^{h_\phi}\,. \tag{3.6}$$

PROOF. Via the natural identification $\beta_g^{\phi^*(g)}$ corresponding to the change of metric, defined as in [**2**], the Dirac operator $D\!\!\!/_{\phi^*(g)}$ can be implemented as an operator

$$D_{\phi^*(g),\,g} = \left(\beta_g^{\phi^*(g)}\right)^{-1}\circ D\!\!\!/_{\phi^*(g)}\circ\beta_g^{\phi^*(g)}$$

acting on the sections of the bundle $\mathcal{S}_g$. It is explicitly given by the formula

$$D_{\phi^*(g),\,g} = e^{(n+1)h}\,D\!\!\!/_g\,e^{(-n+1)h}\,. \tag{3.7}$$

On the other hand, as differential operator,

$$D_{\phi^*(g),\,g} = V_\phi^{-1}\circ D\!\!\!/_g\circ V_\phi\,. \tag{3.8}$$

Combining (3.7) and (3.8) one obtains

$$V_\phi^{-1}\circ D\!\!\!/_g\circ V_\phi = e^{(n+1)h}\,D\!\!\!/_g\,e^{(-n+1)h}\,,$$

or equivalently

$$e^{-nh}\circ V_\phi^{-1}\circ D\!\!\!/_g\circ V_\phi\circ e^{nh} = e^{h}\circ D\!\!\!/_g\circ e^{h}\,.$$

□

Let σ be the algebra automorphism of $\mathcal{A}_G$ defined on generators by

$$\sigma(f\,v_\phi^{-1}) = e^{-2h_\phi}\,f\,v_\phi^{-1}\,, \qquad f\in C^\infty(M), \quad \phi\in G. \tag{3.9}$$

LEMMA 3.2. *The twisted commutators*

$$[D,\pi(a)]_\sigma := D\circ\pi(a) - \pi(\sigma(a))\circ D, \qquad a\in\mathcal{A}_G\,,$$

are bounded.

PROOF. It suffices to check the claimed property for $a = e^{-nh_\phi}\,v_\phi^{-1}$. In that case one has

$$[D,\pi(a)]_\sigma = D\circ U_\phi^* - e^{-2h_\phi}\circ U_\phi^*\circ D = \left(D - e^{-2h_\phi}\circ U_\phi^*\circ D\circ U_\phi\right)\circ U_\phi^*\,.$$

In view of Eq. (3.6), it follows that

$$\begin{aligned} [D,U_\phi^*]_\sigma &= \left(D - e^{-h_\phi}\circ D\circ e^{h_\phi}\right)U_\phi^* == -\,e^{-h_\phi}\,[D,e^{h_\phi}]\circ U_\phi^* \\ &= -c(dh_\phi)\,U_\phi^*. \end{aligned} \tag{3.10}$$

□

For further reference, we note that as a consequence of (3.10) one has

$$[D, f\, U_\phi^*]_\sigma \,=\, c(df \,-\, f\, dh_\phi)\, U_\phi^*\,, \tag{3.11}$$

and in particular

$$[D, e^{h_\phi}\, U_\phi^*]_\sigma \,=\, 0\,. \tag{3.12}$$

Proposition 3.3. *The algebra $\mathcal{A}_G = C^\infty(M) \rtimes G$, endowed with the automorphism $\sigma \in \operatorname{Aut}\mathcal{A}_G$ and the representation π on the Hilbert space $\mathfrak{H} = L^2(M, \mathcal{S})$, together with the Dirac operator $D = \not{D}_g$, defines an (n,∞)-summable σ-spectral triple $(\mathcal{A}_G, \mathfrak{H}, D, \sigma)$, which moreover satisfies the strong Lipschitz-regularity property*

$$|D|^{-t}\left(|D|^t\, a \,-\, \sigma^t(a)\,|D|^t\right) \,\in\, \mathcal{L}^{(n,\infty)}(\mathfrak{H})\,, \qquad \forall\, t \in \mathbb{R}. \tag{3.13}$$

Proof. The boundedness property (1.2) was verified in Lemma 3.2. To prove Lipschitz regularity, one notes that, when viewed as a pseudodifferential operator, $U_\phi^* \circ D \circ U_\phi$ has principal symbol

$$\sigma_{\rm pr}(U_\phi^* \circ D \circ U_\phi)(x,\xi) \,=\, e^{2h_\phi}\, c(\xi)\,, \quad \xi \in T_x^*M \tag{3.14}$$

where $c(\xi) \in \operatorname{End}(\mathcal{S}_x)$ stands for the Clifford multiplication by ξ. Furthermore, for any $t \in \mathbb{R}$, the principal symbol of $U_\phi^* \circ |D|^t \circ U_\phi$ is

$$\sigma_{\rm pr}(U_\phi^* \circ |D|^t \circ U_\phi)(x,\xi) \,=\, e^{2th_\phi}\,||\xi||^t\,, \tag{3.15}$$

since

$$U_\phi^* \circ |D|^t \circ U_\phi \,=\, |U_\phi^* \circ D \circ U_\phi|^t\,.$$

Now

$$|D|^t \circ U_\phi^* \,-\, \sigma^t(U_\phi^*) \circ |D|^t \,=\, \left(|D|^t \,-\, e^{-2th_\phi}\, U_\phi^* \circ |D|^t \circ U_\phi\right) \circ U_\phi^*\,,$$

and by Eq. (3.15),

$$\sigma_{\rm pr}\left(|D|^t \,-\, e^{-2th_\phi}\, U_\phi^* \circ |D|^t \circ U_\phi\right) \,=\, ||\xi||^t \,-\, e^{-2th_\phi}\, e^{2th_\phi}\,||\xi||^t \,=\, 0.$$

Thus, the operator $|D|^t - e^{-2th_\phi}\, U_\phi^* \circ |D|^t \circ U_\phi$ is pseudodifferential of order $t-1$, hence its product by $|D|^{-t}$ is of order -1 and therefore in $\mathcal{L}^{(n,\infty)}$. □

Remark 3.4. *In the same fashion, $\mathbb{R}^n$ with its standard metric g_0, together with the flat Dirac operator $D_0 = \not{D}_{g_0}$, gives rise to the (n,∞)-summable non-unital σ-spectral triple $(\mathcal{A}_{G_0}, \mathfrak{H}_0, D_0, \sigma)$, where $G_0 = CO(\mathbb{R}^n, g_0)$, and $\mathcal{A}_{G_0} = C_c^\infty(\mathbb{R}^n) \rtimes G_0$.*

According to the Ferrand-Obata theorem (*cf.* [**12**] for a complete proof), the conformal group $CO(M,[g])$ of a (not necessarily closed) manifold M of dimension $n \geq 2$ is *inessential*, *i.e.* reduces to the group of isometries for a metric in the conformal class $[g]$, except when M^n is conformally equivalent to the standard sphere S^n or to the standard Euclidean space $\mathbb{R}^n$. Correspondingly, the only twisted spectral triples arising from the above construction which are not isomorphic to ordinary spectral triples are those associated to the n-sphere and to the flat n-space.

3.2. Transverse noncommutative residue. With the same assumptions as in the preceding subsection, let $\Psi^\bullet(M;\mathcal{S})$ denote the algebra of classical pseudodifferential operators acting on the sections of the spin bundle. The group G acts on $\Psi^\bullet(M;\mathcal{S})$ in the natural fashion:

$$\phi \cdot P := V_\phi P V_\phi^{-1}, \qquad \phi \in G, \quad P \in \Psi^\bullet(M;\mathcal{S}). \tag{3.16}$$

One can thus form the crossed product algebra $\Psi^\bullet(M;\mathcal{S}) \rtimes G$. The representation $\pi : \mathcal{A}_G \to \mathcal{L}(\mathfrak{H})$ extends in a tautological manner to a representation of the enlarged algebra $\Psi^\bullet(M;\mathcal{S}) \rtimes G$ by densely defined linear operators on $\mathfrak{H} = L^2(M,\mathcal{S})$, which will still be denoted by π:

$$\pi(P\, v_\phi)\,(u) := P\big(V_\phi(u)\big) = P(\tilde{\phi} \circ u \circ \phi^{-1}), \qquad u \in C^\infty(M,\mathcal{S}). \tag{3.17}$$

We set out to show that given any $\mathcal{P} \in \Psi^N(\mathcal{A}_G)$ the zeta function

$$\zeta_{\mathcal{P}}(z) := \mathrm{Tr}(\mathcal{P}\,|D|^{-z}), \qquad \Re z > n+N \tag{3.18}$$

can be meromorphically continued to the whole complex plane; by linearity, it suffices to take $\mathcal{P} = P\,V_\phi$, with $P \in \Psi^N(M;\mathcal{S})$ and $\phi \in G$.

If G is inessential, ϕ is an isometry and the statement can be proved via the Mellin transform and heat kernel asymptotics (see [**1**, §6.3]). A more direct proof, given in [**11**], relies on the stationary phase method (*cf. e.g.* [**17**, Thm. 7.7.5]), applied to a phase function whose expression in local charts U covering a tubular neighborhood of the fixed point set M_ϕ is of the form

$$f(x,\xi) = \langle x - \phi(x), \xi\rangle, \qquad x \in U, \quad \xi \in \mathbb{R}^n,\ ||\xi|| = 1. \tag{3.19}$$

By restriction to the fibers of the normal bundle to M_ϕ, this function gives rise to a family of fiberwise phase functions, each having a single non-degenerate stationary point. Using the stationary phase for this family, it is shown in [**11**, Prop. 2.4] that the zeta function $\zeta_{V_\phi\, P}$ has a meromorphic extension to $\mathbb{C}$ whose poles are at most simple and located at the points $z_k = N + n_\phi - k$, $k \in \mathbb{Z}^+$, where $n_\phi = \dim M_\phi$.

In the sphere case, by Liouville's theorem the group of conformal automorphisms $CO(S^n,[g])$ coincides with the group $M(n) \cong PO(n+1,1)$ of Möbius transformations in dimension n, and $G = SCO(S^n,[g])$ is its connected component. Now if $\phi \in G$ is *elliptic*, *i.e.* conjugate to an element in the maximal compact subgroup $O(n+1)$, by replacing D in formula (3.18) with a conjugate by a unitary operator we can reduce to the isometric case.

The non-elliptic diffeomorphisms $\phi \in M(n)$ fall into two classes: *hyperbolic* and *parabolic* (see [**19**, §2]).

A *hyperbolic* transformation $\phi \in M(n)$ has two distinct fixed points, say x^+ and x^-, and its tangent map at each of these points $d\phi_{x^\pm} : T_{x^\pm}S^n \to T_{x^\pm}S^n$ is represented by an element of $O(n)\times\mathbb{R}^+$, with multiplier $\mu^\pm$, with $\mu > 1$. Because of the nontrivial multiplier, the phase function (3.19) has no critical points away from the zero section of the cotangent bundle T^*S^n. The stationary phase principle, in its most basic form (*cf.* [**17**, Thm. 7.7.1]) and utilized in the same manner as in [**11**], implies then that the zeta function $\zeta_{V_\phi\, P}$ extends to an entire function.

A *parabolic* transformation $\phi \in M(n)$ has a single fixed point $x_0 \in S^n$, however, $\det(\mathrm{Id} - d\phi_{x_0}) = 0$. Accordingly, the phase function has 0 as its only critical value, and the corresponding critical set is

$$C_\phi = \{(x_0,\xi) \,|\, \xi \in \mathbb{R}^k, \quad d\phi_{x_0}(\xi) = \xi\}. \tag{3.20}$$

The generalized stationary phase gives an asymptotic expansion

$$\int_{||\xi||=1} e^{ir\langle x-\phi(x),\,\xi\rangle}\, a(x,\xi)\, d^{n-1}\xi\, d^n x \sim \sum_{\alpha}\sum_{j=0}^{\infty}\sum_{k=0}^{2n-1} \delta_{j,k}(a)\, r^{\alpha-j}\, \log^k r\,, \tag{3.21}$$

where $\alpha = \frac{1}{2}\dim C_\phi - n$ and the distributions $\delta_{j,k}$ are supported in C_ϕ. Following the same line of arguments as in [**11**], but using stereographic coordinates instead of normal coordinates, one obtains the desired meromorphic continuation of the zeta function $\zeta_{V_\phi\, P}$, with (at most) simple poles at the points

$$z_k = N + \frac{1}{2}\dim C_\phi - k, \qquad \text{where} \qquad k \in \mathbb{Z}^+.$$

We summarize the conclusion of the preceding discussion as follows.

THEOREM 3.5. *For any $P \in \Psi^N(M^n;\mathcal{S})$ and any $\phi \in G$, the associated zeta function $\zeta_{P\, V_\phi}$ has a meromorphic extension to $\mathbb{C}$. Moreover,*

1^0 *if $\phi \in G$ is elliptic, then the poles of $\zeta_{V_\phi\, P}$ are at most simple and are located at the points $z_k = N + \dim M_\phi - k, \quad k \in \mathbb{Z}^+$;*

2^0 *if $\phi \in G$ is hyperbolic, then $\zeta_{V_\phi\, P}$ is entire;*

3^0 *if $\phi \in G$ is parabolic, then the poles of $\zeta_{V_\phi\, P}$ are at most simple and are located at the points $z_k = N + \frac{1}{2}\dim C_\phi - k, \quad k \in \mathbb{Z}^+$.*

This provides the *transverse noncommutative residue* functional

$$⨍_D \mathcal{P} \,=\, \mathrm{Res}_{z=0}\, \zeta_{\mathcal{P}}(2z)\,, \qquad \mathcal{P} \in \Psi^\bullet(M;\mathcal{S}) \rtimes G,$$

which satisfies a property analogous to the Selberg Principle.

COROLLARY 3.6. *For any hyperbolic transformation $\phi \in G$ and any $P \in \Psi^N(S^n;\mathcal{S})$,*

$$⨍_D P\, V_\phi \,=\, 0\,.$$

As another consequence, one can explicitly compute the candidate for the Hochschild character given by Eq. (1.20), and thus directly verify that it gives the expected result.

PROPOSITION 3.7. *The local Hochschild cocycle of the transversely conformal σ-spectral triple $(\mathcal{A}_G, \mathfrak{H}_g, \not{D}_g, \sigma)$ associated to a closed spin manifold M^n modulo the conformal group $G = SCO(M,[g])$ is a cyclic cocycle whose periodic cyclic cohomology class coincides with the transverse fundamental class $[M/G]$.*

PROOF. Let $a^k = f_k\, U^*_{\phi_k} \in \mathcal{A}_G$, $k = 0, 1, \ldots, n$. The integrand in the formula (1.20),

$$a^0\, [D, \sigma^{-1}(a^1)]_\sigma \, \cdots \, [D, \sigma^{-n}(a^n)]_\sigma \, |D|^{-n}\,,$$

can be put in the form $P\, V_\phi$ with $P \in \Psi^{-n}(S^n;\mathcal{S})$ and $\phi^{-1} = \phi_n \circ \cdots \circ \phi_0$. It follows from Prop. 3.5, specialized to the case when $N = -n$, that the zeta function $\zeta_{P\, V_\phi}(z)$ has no pole at $z = 0$. Therefore

$$⨍_D P\, V_\phi \,=\, 0\,, \qquad \text{unless} \qquad \phi = \mathrm{Id},$$

i.e. the cocycle (1.20) is localized at the identity. Employing Getzler's symbol calculus for asymptotic operators as indicated in [**8**, Remark II.1], and using the expression (3.11) of the twisted commutators, the Hochschild cocycle (1.20) can be explicitly computed. The end result is a cyclic cocycle, which is easily seen to differ by a coboundary from the standard transverse fundamental cocycle (*cf.* [**10**, Thm 3.11])

$$\tau_{M/G}(f_0 U^*_{\phi_0}, \dots, f_n U^*_{\phi_n}) = \begin{cases} \int_M f_0\, d(f_1 \circ \phi_0) \wedge \cdots \wedge d(f_n \circ \phi_{n-1} \circ \cdots \circ \phi_0), \\ \qquad\qquad \text{if} \qquad \phi_n \circ \cdots \circ \phi_0 = \mathrm{Id}; \\ \qquad\qquad 0 \qquad \text{otherwise}. \end{cases}$$

□

REMARK 3.8. If M is closed and G inessential, hence compact, the corresponding σ-spectral triple is a conformal perturbation, *cf.* §1.6, of an equivariant spectral triple. Its Connes-Chern character, given by (1.14), can be explicitly computed from the local index formula of [**8**] by employing an equivariant version of the Getzler symbol calculus, or the equivariant heat kernel techniques in [**1**].

3.3. Transverse similarities. Endow $\mathbb{R}^n$ with the Euclidean metric $g_0 = \sum_{i=1}^n dx^i \otimes dx^i$. The group $G = \mathrm{Sim}(n)$ of conformal (or similarity) transformations of the Euclidean n-space is generated by rotations, translations by vectors $y \in \mathbb{R}^n$,

$$\tau_y(x) \,=\, x - y, \qquad \forall\, x \in \mathbb{R}^n,$$

and homotheties $\rho_\lambda,\ \lambda > 0$,

$$\rho_\lambda(x) \,=\, \lambda^{-1} x, \qquad \forall\, x \in \mathbb{R}^n.$$

The only non-isometries are the homotheties,

$$\rho_\lambda^* g_0 \,=\, \lambda^{-2} g_0\,.$$

i.e. in the notation of §3.1 the corresponding conformal factor is

$$e^{-4h_{\rho_\lambda}}(x) \,=\, \lambda^{-2}\,, \quad \forall\, x \in \mathbb{R}^n.$$

With $\mathcal{A}_G := C_c^\infty(\mathbb{R}^n) \rtimes G$, the definition (3.9) of the automorphism $\sigma \in \mathrm{Aut}\,\mathcal{A}_G$ specializes to

$$\sigma(f\, U_\phi) \,=\, \mu(\phi)\; f\, U_\phi\,, \qquad \phi \in G\,, \tag{3.22}$$

where $\mu : G \to \mathbb{R}^+$ is the character determined by the *multiplier* of the similarity transformation:

$$\mu(\phi) = 1 \quad \text{if} \quad \phi \in O(n), \quad \mu(\tau_y) = 1, \forall\, y \in \mathbb{R}^n, \quad \text{and} \quad \mu(\rho_\lambda) \,=\, \lambda, \forall\, \lambda > 0.$$

Also, the covariace relation (3.6) becomes

$$U_\phi^{-1} \circ D \circ U_\phi \,=\, \mu(\phi)\, D\,, \qquad \phi \in G\,. \tag{3.23}$$

Let $\Psi_c^\bullet(\mathbb{R}^n; \mathcal{S})$ denote the algebra of classical pseudodifferential operators with x-compact support. Since the conformal factors are constant, one can easily extend σ to an automorphism of $\Psi_c^\bullet(\mathbb{R}^n; \mathcal{S}) \rtimes G$, by simply setting

$$\sigma(P\, U_\phi) \,=\, \mu(\phi)\; P\, U_\phi\,, \qquad \phi \in G, \quad P \in \Psi_c^\bullet(\mathbb{R}^n, \mathcal{S}). \tag{3.24}$$

Indeed,

$$\sigma(P\,U_\phi)\cdot\sigma(Q\,U_\psi)\;=\;\big(\mu(\phi)\;P\,U_\phi\big)\cdot\big(\mu(\psi)\;Q\,U_\psi\big)\;= \tag{3.25}$$
$$=\;\mu(\phi\psi)\,P\cdot(U_\phi Q U_\phi^{-1})\,U_{\phi\psi}\;=\;\sigma\big(P\cdot(U_\phi Q U_\phi^{-1})\,U_{\phi\psi}\big)\;=\;\sigma(P\,U_\phi\cdot Q\,U_\psi)\,.$$

PROPOSITION 3.9. *The residue functional* $⨍_D : \Psi_c^\bullet(\mathbb{R}^n;\mathcal{S})\rtimes G\to\mathbb{C}$ *is a σ-invariant trace.*

PROOF. The σ-invariance of the residue is a consequence of Corollary 3.6 (Selberg Principle). Indeed, let $\mathcal{P}=PU_\phi\in\Psi_c^\bullet(\mathbb{R}^n;\mathcal{S})\rtimes G$. If $\mu(\phi)\neq 1$ then ϕ has a unique fixed point $x_0\in\mathbb{R}^n$, at which $d\phi_{x_0}=\mu(\phi)\mathrm{Id}$. Thus, there are no fixed points on $T^*\mathbb{R}^n$, hence the zeta function $\zeta_{\mathcal{P}}(z)$ is entire (see §3.2). On the other hand, if $\mu(\phi)=1$ then $\sigma(\mathcal{P})=\mathcal{P}$.

To prove that the residue functional is a trace, let $PU_\phi,\,QU_\psi\in\Psi_c^\bullet(\mathbb{R}^n;\mathcal{S})\rtimes G$. One has

$$\begin{aligned}\mathrm{Tr}\big(PU_\phi\,QU_\psi\,|D|^{-z}\big)\;&=\;\mathrm{Tr}\big(QU_\psi\,|D|^{-z}P\,U_\phi\big)=\\ &=\;\mathrm{Tr}\big(QU_\psi\,P|D|^{-z}\,U_\phi\big)+\mathrm{Tr}\big(QU_\psi\,[|D|^{-z},P]\,U_\phi\big)\end{aligned}$$

Using the identity Eq. (2.2) to express the commutator $[|D|^{-z},P]$, one sees that

$$\mathrm{Res}_{z=0}\mathrm{Tr}\big(QU_\psi\,[|D|^{-z},P]\,U_\phi\big)\;=0\,,$$

hence,

$$⨍_D PU_\phi\,QU_\psi=\mathrm{Res}_{z=0}\mathrm{Tr}\big(QU_\psi\,P|D|^{-z}\,U_\phi\big)=\mathrm{Res}_{z=0}\mathrm{Tr}\big(QU_\psi\,PU_\phi\,U_\phi^{-1}|D|^{-z}\,U_\phi\big).$$

By Eq. (3.23), $\quad U_\phi^{-1}|D|^{-z}\,U_\phi=\mu(\phi)^{-z}|D|^{-z}$, therefore

$$⨍_D PU_\phi\,QU_\psi=\mathrm{Res}_{z=0}\,\big(\mu(\phi)^{-z}\,\mathrm{Tr}\big(QU_\psi\,PU_\phi\,|D|^{-z}\big)\big)=⨍_D QU_\psi\,PU_\phi.$$

□

REMARK 3.10. *One can explicitly verify in this specific case that the cochain* (2.18) *of the Ansatz does satisfy the cocycle identity*

$$b\,\tau_\sigma^{q-1}\;+\;B\,\tau_\sigma^{q+1}\;=\;0.$$

Indeed, the direct, albeit lengthy, computations by which the cocycle identity is checked in the beginning of the proof in [**8**, Theorem II.1] can be reproduced almost *verbatim*. Once the commutator with D^2 is substituted by the twisted commutator

$$\nabla_\sigma(\mathcal{P})\;=\;D^2\,\mathcal{P}\;-\;\sigma^2(\mathcal{P})\,D^2\,, \tag{3.26}$$

and usual iterated Leibniz rule is replaced with its twisted version

$$\nabla_\sigma^m(\mathcal{P}_1\mathcal{P}_2\cdots\mathcal{P}_q)\;=\;\sum_{m_1+\ldots+m_q=m}\frac{m!}{m_1!\cdots m_q!}\cdot \tag{3.27}$$
$$\nabla_\sigma^{m_1}(\sigma^{2(m_2+\ldots+m_q)}(\mathcal{P}_1))\,\nabla_\sigma^{m_2}(\sigma^{2(m_3+\ldots+m_q)}(\mathcal{P}_2))\cdots\,\nabla_\sigma^{m_q}(\mathcal{P}_q)\,,$$

the "integration by parts" property, which is repeatedly used in those calculations, becomes a consequence of Proposition 3.9. As a simple illustration,

$$⨍_D\nabla_\sigma(\mathcal{P})\,D^{-2}\;=\;⨍_D D^2\,\mathcal{P}\,D^{-2}\;-\;⨍_D\sigma^2(\mathcal{P})\;=\;⨍_D\mathcal{P}\;-\;⨍_D\sigma^2(\mathcal{P})\;=\;0\,.$$

A different approach, which provides a complete verification of the Ansatz in greater generality, constitutes the contents of the section that follows.

4. Twisting by scaling automorphisms and the local index formula

By abstracting the essential features of the preceding example, we define in this section a general class of spectral triples twisted by scaling automorphisms, for which we shall prove the validity of the Ansatz in its entirety.

4.1. Scaling automorphisms. Motivated by the Euclidean similarities in §3.3, we introduce the following abstract version of a spectral triple twisted by similarities.

DEFINITION 4.1. *Let $(\mathcal{A}, \mathfrak{H}, D)$ be a spectral triple over the (non-unital) involutive algebra $\mathcal{A}$. A* scaling automorphism *of $(\mathcal{A}, \mathfrak{H}, D)$ is defined by a unitary operator $U \in \mathcal{U}(\mathfrak{H})$ such that*

$$U\,\mathcal{A}\,U^* \,=\, \mathcal{A}\,, \quad \textit{and} \quad U\,D\,U^* \,=\, \mu(U)\,D, \quad \textit{with} \quad \mu(U) > 0. \tag{4.1}$$

Scaling automorphisms form a group $\mathrm{Sim}(\mathcal{A}, \mathfrak{H}, D)$, endowed by construction with a *scaling character* $\mu : \mathrm{Sim}(\mathcal{A}, \mathfrak{H}, D) \to \mathbb{R}^+$. Its subgroup $\mathrm{Ker}\,\mu$ consists of the *isometries* of $(\mathcal{A}, \mathfrak{H}, D)$, and will be denoted $\mathrm{Isom}(\mathcal{A}, \mathfrak{H}, D)$.

For the clarity of the exposition it will be convenient to assume D invertible. This can always be achieved by passage to the invertible double, *cf.* §1.3,

$$\widetilde{D} \,=\, D\hat{\otimes}\mathrm{Id} \,+\, \mathrm{Id}\hat{\otimes}F_1\,.$$

However, in doing so the similarity condition (4.1) cannot be exactly reproduced. Instead, it takes the modified form

$$U\,\widetilde{D}\,U^* \,=\, \mu(U)\,\widetilde{D}\hat{\otimes}\mathrm{Id} \,+\, (1-\mu(U))\,\mathrm{Id}\hat{\otimes}F_1\,. \tag{4.2}$$

We will explain at the end of the paper the minor modifications needed to handle the perturbed similarity condition.

In the remainder of the paper we fix a group of scaling automorphisms $G \subset \mathrm{Sim}(\mathcal{A}, \mathfrak{H}, D)$, and let $\mathcal{A}_G = \mathcal{A} \rtimes G$. We shall also denote by G_0 the subgroup of isometries in G.

PROPOSITION 4.2. *The formula*

$$\sigma(a\,U) \,=\, \mu(U)^{-1}\,a\,U, \quad \forall\, U \in G,\, a \in \mathcal{A}, \tag{4.3}$$

defines an automorphism $\sigma : \mathcal{A}_G \to \mathcal{A}_G$, and $(\mathcal{A}_G, \mathfrak{H}, D, \sigma)$ is a σ-spectral triple. Moreover,

$$[D, aU]_\sigma \,=\, [D, a]\,U. \tag{4.4}$$

PROOF. Indeed, one has for any monomials $a\,U,\ b\,V \in \mathcal{A}_G = \mathcal{A} \ltimes G$,

$$\begin{aligned}\sigma(a\,U)\,\sigma(b\,V) \,&=\, \big(\mu(U)^{-1}\,a\,U\big)\,\big(\mu(V)^{-1}\,b\,V\big) \,=\\ &=\, \mu(UV)^{-1}\,a\,(U\,b\,U^*)\,UV \,=\, \sigma\big(a\,U(b))\,UV\big) \,=\, \sigma(a\,U\,b\,V)\,.\end{aligned}$$

Furthermore, in view of (4.1),

$$[D, a\,U]_\sigma \quad=\quad [D,a]\,U \,+\, a\Big(D \,-\, \mu(U)^{-1}\,U\,D\,U^*\Big)\,U \,=\, [D,a]\,U \in \mathcal{L}(\mathfrak{H}).$$

∎

The resulting σ-spectral triple $(\mathcal{A}_G, \mathfrak{H}, D, \sigma)$ will be called *twisted by scaling automorphisms*. With the goal of establishing the validity of the Ansatz for twisted spectral triples of this form, we start with the assumption that the base spectral triple $(\mathcal{A}, \mathfrak{H}, D)$ is (p, ∞)-summable, and satisfies the smoothness condition (1.15). We recall that the corresponding algebra of pseudodifferential operators $\Psi^\bullet(\mathcal{A}, \mathfrak{H}, D)$ is $\mathbb{Z}$-filtered (by the order) and $\mathbb{Z}_2$-graded (even/odd), and that it includes the subalgebra of differential operators $\mathcal{D}(\mathcal{A}, \mathfrak{H}, D)$.

Proposition 4.3. *With the above notation and hypotheses,*

(1) *the action of G extends to an action by automorphisms on $\Psi(\mathcal{A}, \mathfrak{H}, D)$,*

$$P \mapsto U \triangleright P := U\, P\, U^*, \qquad \forall\, P \in \Psi(\mathcal{A}, \mathfrak{H}, D), \quad U \in G, \tag{4.5}$$

which respects both the order filtration and the even/odd grading;

(2) *the automorphism $\sigma \in \operatorname{Aut}(\mathcal{A}_G)$ extends to an automorphism σ of the crossed product algebra $\Psi(\mathcal{A}_G, \mathfrak{H}, D) := \Psi(\mathcal{A}, \mathfrak{H}, D) \rtimes G$, by setting*

$$\sigma(P\, U) \,=\, \mu(U)^{-1}\, P\, U, \quad \forall\, U \in G,\, P \in \Psi(\mathcal{A}, \mathfrak{H}, D), \tag{4.6}$$

(3) *the twisted commutators by D, $|D|$ and D^2 define twisted derivations d_σ, δ_σ, resp. ∇_σ, of the algebra $\Psi(\mathcal{A}_G, \mathfrak{H}, D)$.*

Proof. The condition (4.1) ensures that $\mathcal{D}(\mathcal{A}, \mathfrak{H}, D)$ remains invariant under conjugation by $U \in G$, and also implies that

$$U\, |D|^z\, U^* \,=\, \mu(U)^z\, |D|^z, \qquad \forall\, z \in \mathbb{C}. \tag{4.7}$$

The verification of the other claims is straightforward. $\square$

We now add the *extended simple dimension spectrum* hypothesis: *there exists a discrete set $\Sigma_G \subset \mathbb{C}$, such that the holomorphic functions*

$$\zeta_B(z) \,=\, \operatorname{Tr}(B\, |D|^{-z})\,, \qquad \Re z > p\,, \qquad \forall\, B \in \mathcal{B}_G := \mathcal{B} \rtimes G, \tag{4.8}$$

admit meromorphic extensions to $\mathbb{C}$ with simple poles in Σ_G, and *the functions $\Gamma(z)\, \zeta_B(z)$ decay rapidly on finite vertical strips.*

The proof of Proposition 3.9 applies verbatim and shows that the residue functional

$$⨍_D \mathcal{P} \,:=\, \operatorname{Res}_{z=0} \zeta_{\mathcal{P}}(2z)\,, \qquad \mathcal{P} \in \Psi(\mathcal{A}_G, \mathfrak{H}, D)$$

is automatically a trace. We require it to be σ-invariant:

$$⨍_D \sigma(\mathcal{P}) \,=\, ⨍_D \mathcal{P}. \tag{4.9}$$

This axiom *de facto* enforces the Selberg Principle, since it implies

$$⨍_D P\, U \,=\, 0, \qquad \text{if} \quad \mu(U) \,\neq\, 1, \qquad P \in \Psi(\mathcal{A}, \mathfrak{H}, D), \quad U \in G\,; \tag{4.10}$$

in particular, the residue functional is necessarily supported on $\Psi(\mathcal{A}_{G_0}, \mathfrak{H}, D)$, where $G_0 := \operatorname{Isom}(\mathcal{A}, \mathfrak{H}, D)$.

4.2. Twisted JLO brackets. We define the *twisted JLO bracket of order q* as the $(q+1)$-linear form on $\Psi(\mathcal{A}_G, \mathfrak{H}, D)$ which for $\alpha_0, \ldots, \alpha_q \in \Psi(\mathcal{A}, \mathfrak{H}, D)$ and $U_0, \ldots, U_q \in G$ has the expression

$$\begin{aligned}(4.11)\quad \langle \alpha_0 U_0^*, \ldots, \alpha_q U_q^* \rangle_D = \int_{\Delta_q} \mathrm{Tr}\big(\gamma\, \alpha_0 U_0^*\, e^{-s_0 \mu(U_0)^2 D^2}\, \alpha_1 U_1^*\, e^{-s_1 \mu(U_0 U_1)^2 D^2} \cdots \\ \cdots \alpha_q U_q^*\, e^{-s_q \mu(U_0 \cdots U_q)^2 D^2}\big),\end{aligned}$$

where the integration is over the q-simplex

$$\Delta_q := \big\{ s = (s_0, \cdots, s_q) \in \mathbb{R}^{q+1} \ \big|\ s_j \geq 0, \quad s_0 + \cdots + s_q = 1 \big\}.$$

Throughout the rest of this subsection, we shall assume that $\alpha_0, \ldots, \alpha_q$ are polynomial expressions in D and the elements of $\mathcal{A}$, $[D, \mathcal{A}]$, and are homogeneous in λ when D is replaced by λD. Given a JLO bracket as in (4.11), for any $\varepsilon > 0$ we denote by $\langle \alpha_0 U_0^*, \ldots, \alpha_q U_q^* \rangle_D(\varepsilon)$ the expression obtained by replacing every D occurring in each $\alpha_0, \ldots, \alpha_q$ by $\varepsilon^{1/2} D$. Equivalently,

$$(4.12)\qquad \langle \alpha_0 U_0^*, \ldots, \alpha_q U_q^* \rangle_D(\varepsilon) = \varepsilon^{\frac{m}{2}} \langle \alpha_0 U_0^*, \ldots, \alpha_q U_q^* \rangle_{\varepsilon^{1/2} D},$$

where m is the total degree in λ after replacing every D by λD in the product $\alpha_0 \cdots \alpha_q$. As before, by $U_0, \ldots, U_q$ we denote arbitrary elements in G.

Proposition 4.4. *Let $\alpha_0 \in \mathcal{A}$, and $\alpha_1, \ldots, \alpha_q \in [D, \mathcal{A}]$. There is an asymptotic expansion of the form*

$$(4.13)\qquad \langle \alpha_0 U_0^*, \ldots, \alpha_q U_q^* \rangle_D(\varepsilon) \sim_{\varepsilon \searrow 0} \sum_{j=0}^{m} (c_j + c_j' \log \varepsilon)\, \varepsilon^{\frac{q}{2} - \rho_j} + O(1),$$

with $\rho_0, \ldots, \rho_m$ a finite set of points in the half-plane $\Re z \geq \frac{q}{2}$.

Proof. Moving all the unitaries U_i to the rightmost position,

$$\begin{aligned}&\mathrm{Tr}\Big(\gamma\, \alpha_0 U_0^*\, e^{-s_0 \mu(U_0)^2 D^2} U_0 (U_0^* U_1^*)\, \alpha_1\, (U_1 U_0) U_0^* U_1^* e^{-s_1 \mu(U_0 U_1)^2 D^2} (U_1 U_0) \cdots \\ &(U_0^* \cdots U_{q-1}^*)\, \alpha_q\, (U_{q-1} \cdots U_0)(U_0^* \cdots U_q^*) e^{-s_q \mu(U_0 \cdots U_q)^2 D^2} (U_q \cdots U_0) U_0^* \cdots U_q^* \Big),\end{aligned}$$

the twisted JLO bracket relative to $\varepsilon^{1/2} D$ takes the form

$$\begin{aligned}(4.14)\quad &\langle \alpha_0 U_0^*, \ldots, \alpha_q U_q^* \rangle_{\varepsilon^{1/2} D} = \\ &= \int_{\Delta_q} \mathrm{Tr}\big(\gamma\, \alpha_0\, e^{-s_1 \varepsilon D^2}\, \alpha_1'\, e^{-(s_2 - s_1)\varepsilon D^2} \cdots \alpha_q'\, e^{-(1 - s_q)\varepsilon D^2}\, U_0^* \cdots U_q^*\big),\end{aligned}$$

where $\alpha_1' = U_0^* U_1^* \triangleright \alpha_1, \ldots, \alpha_q' = U_0^* \cdots U_{q-1}^* \triangleright \alpha_q$.

We next use the expansion

$$(4.15)\qquad e^{-\varepsilon D^2} \alpha \sim_{\varepsilon \searrow 0} \sum_{n=0}^{\infty} \frac{(-1)^n \varepsilon^n}{n!} \nabla^n(\alpha)\, e^{-\varepsilon D^2}, \qquad \alpha \in \mathcal{D}(\mathcal{A}, \mathfrak{H}, D),$$

which is the heat operator analogue of the expansion (2.3) (*cf.* also (4.21) *infra*, for its twisted version), to move the heat operators in Eq. (4.14) to the right and bring them all to the last position. One obtains

$$\langle \alpha_0 U_0^*, \dots, \alpha_q U_q^* \rangle_{\varepsilon^{1/2} D} \sim_{\varepsilon \searrow 0} \frac{1}{q!} \sum_{N \geq 0} \sum_{n_1 + \ldots + n_q = N} \frac{(-1)^N \varepsilon^N}{n_1! \cdots n_q!} \tag{4.16}$$
$$\cdot \int_{0 \leq s_1 \leq \ldots \leq s_q \leq 1} s_1^{n_1} \cdots s_q^{n_q} \cdot \mathrm{Tr}\big(\gamma \alpha_0 \nabla^{n_1}(\alpha_1') \cdots \nabla^{n_q}(\alpha_q') e^{-\varepsilon D^2} U_0^* \cdots U_q^*\big)$$
$$= \frac{1}{q!} \sum_{N \geq 0} \sum_{n_1 + \ldots + n_q = N} \frac{(-1)^N \varepsilon^N}{n_1! \cdots n_q! (n_1 + 1) \cdots (n_1 + \cdots n_q + q)} \cdot$$
$$\cdot \mathrm{Tr}\big(\gamma \alpha_0 \nabla^{n_1}(\alpha_1') \cdots \nabla^{n_q}(\alpha_q') e^{-\varepsilon D^2} U_0^* \cdots U_q^*\big).$$

The *extended simple dimension spectrum* hypothesis ensures that the zeta functions

$$\zeta_N(z) = \mathrm{Tr}\big(\gamma\, U_0^* \cdots U_q^*\, \alpha_0\, \nabla^{n_1}(\alpha_1') \cdots \nabla^{n_q}(\alpha_q')\, |D|^{-2z-2N}\big), \qquad \Re z > \frac{p}{2},$$

have meromorphic continuations with simple poles. One has

$$\zeta_N(z) = \frac{1}{\Gamma(z+N)} \int_0^\infty t^{z+N-1} \mathrm{Tr}\big(\gamma U_0^* \cdots U_q^*\, \alpha_0 \nabla^{n_1}(\alpha_1') \cdots \nabla^{n_q}(\alpha_q') e^{-tD^2}\big) dt.$$

Proceeding as in the proof of [**8**, Theorem II.1], one establishes by means of the inverse Mellin transform the existence of an asymptotic expansion

$$\varepsilon^{N+\frac{q}{2}}\, \mathrm{Tr}\big(\gamma\, U_0^* \cdots U_q^*\, \alpha_0\, \nabla^{n_1}(\alpha_1') \cdots \nabla^{n_q}(\alpha_q')\, e^{-\varepsilon D^2}\big) \sim_{\varepsilon \searrow 0} \tag{4.17}$$
$$\sum_j (c_{N,j} + c_{N,j}' \log \varepsilon)\, \varepsilon^{\frac{q}{2} - \rho_{N,j}} + O(1),$$

where the exponents $\rho_{N,j}$ are the poles of $\zeta_N(z)$ whose real parts are in the half-plane $\Re z \geq \frac{q}{2}$. □

Definition 4.5. *We define the* constant term $\langle \alpha_0 U_0^*, \dots, \alpha_q U_q^* \rangle_D|_0$ *as the* finite part Pf_0 *in the asymptotic expansion* $\langle \alpha_0 U_0^*, \dots, \alpha_q U_q^* \rangle_D(\varepsilon)$*; it is given by the coefficient* c_0 *when* $\rho_0 = \dfrac{q}{2}$*, and is* 0 *otherwise.*

Proposition 4.6. *The constant term* $\langle \alpha_0 U_0^*, \dots, \alpha_q U_q^* \rangle_D|_0$ *satisfies*

$$\langle \alpha_0 U_0^*, \dots, \alpha_q U_q^* \rangle_D|_0 = 0, \quad \textit{unless} \quad \mu(U_0 \cdots U_q) = 1; \tag{4.18}$$
$$\langle \sigma(\alpha_0 U_0^*), \dots, \sigma(\alpha_q U_q^*) \rangle_D|_0 = \langle \alpha_0 U_0^*, \dots, \alpha_q U_q^* \rangle_D|_0. \tag{4.19}$$

Proof. Up to a numerical factor, $\langle \alpha_0 U_0^*, \dots, \alpha_q U_q^* \rangle_D|_0$ coincides with the residue $\mathrm{Res}_{z=0}\, \zeta_N$. In view of the axiom (4.10),

$$\text{if} \quad \mu(U_0 \cdots U_q) \neq 1 \quad \text{then} \quad \mathrm{Res}_{z=0}\, \zeta_N = 0, \qquad \forall\, N \geq 0.$$

This proves the property (4.18), which in turn readily implies (4.19). □

In order to compute the constant term, we shall employ the elementary Duhamel-type commutator formula

$$e^{-(\beta-\alpha)\lambda^2 D^2} A - A\, e^{-(\beta-\alpha)\lambda^2 \mu^2 D^2} = \tag{4.20}$$
$$- \int_\alpha^\beta e^{-(s-\alpha)\lambda^2 D^2} \lambda^2 (D^2 A - \mu^2 A D^2) e^{-(\beta-s)\lambda^2 \mu^2 D^2}\, ds,$$

where $A \in \mathcal{A}_G$, $\lambda, \mu > 0$ and $[\alpha, \beta] \subset \mathbb{R}$. It is obtained by integrating the identity

$$\frac{d}{ds}\left(e^{-(s-\alpha)D^2} A\, e^{-(\beta-s)\mu^2 D^2}\right) = -e^{-(s-\alpha)D^2}(D^2A - \mu^2 AD^2)e^{-(\beta-s)\mu^2 D^2}$$

and then replacing D by λD.

By iterating (4.20), and using the abbreviation $\nabla_\mu(A) = D^2A - \mu^2 AD^2$, one obtains for any $N \in \mathbb{N}$,

$$e^{-t^2D^2}A = \sum_{k=0}^{N-1} \frac{(-1)^k t^{2k}}{k!} \nabla_\mu^k(A)\, e^{-t^2\mu^2D^2} + R_N(D, A, \mu, t). \tag{4.21}$$

The remainder is given by the formula

$$\begin{aligned} R_N(D, A, \mu, t) &= (-1)^N t^{2N} \int_{\Delta_N} e^{-s_1t^2D^2} \nabla_\mu^N(A) e^{-(1-s_1)t^2\mu^2D^2} ds_1 \cdots ds_N \\ &= \frac{(-1)^N t^{2N}}{(N-1)!} \int_0^1 (1-s)^{N-1} e^{-st^2D^2} \nabla_\mu^N(A) e^{-(1-s)t^2\mu^2D^2} ds, \end{aligned} \tag{4.22}$$

Using the finite-summability assumption, it is easy to estimate the above expression and thus show that Eq. (4.21) does provide an asymptotic expansion as $t \searrow 0$.

Applying this expansion for a twisted bracket as in Proposition 4.4 one obtains

$$\begin{aligned} \langle \alpha_0 U_0^*, \ldots, \alpha_q U_q^* \rangle_{tD} = \int_{\Delta_q} &\operatorname{Tr}\Big(\gamma\, \alpha_0 U_0^*\, e^{-s_1 \mu(U_0)^2 t^2 D^2} \alpha_1 U_1^* e^{-(s_2-s_1)\mu(U_0U_1)^2t^2D^2} \\ &\cdots e^{-(s_q-s_{q-1})\mu(U_0\cdots U_{q-1})^2t^2D^2} \alpha_q U_q^*\, e^{-(1-s_q)\mu(U_0\cdots U_q)^2t^2D^2}\Big) \end{aligned}$$

$$\sim_{t\searrow 0} \sum_{N\geq 0} (-1)^N t^{2N} \sum_{n_1+\ldots+n_q=N} \frac{\mu(U_0)^{2(n_1+\ldots+n_q)} \cdots \mu(U_{q-1})^{2n_q}}{n_1! \cdots n_q!} \cdot$$

$$\operatorname{Tr}\Big(\gamma\, \alpha_0 U_0^*\, \nabla_\sigma^{n_1}(\alpha_1 U_1^*) \cdots \nabla_\sigma^{n_q}(\alpha_q U_q^*)\, e^{-\mu(U_0\cdots U_q)^2t^2D^2}\Big) \int_{0\leq s_1\leq \ldots\leq s_q\leq 1} s_1^{n_1} \cdots s_q^{n_q}$$

$$= \sum_{N\geq 0} (-1)^N t^{2N} \sum_{n_1+\ldots+n_q=N} \frac{\mu(U_0)^{2(n_1+\ldots+n_q)} \cdots \mu(U_{q-1})^{2n_q}}{n_1! \cdots n_q!(n_1+1)\cdots(n_1+\cdots n_q+q)} \cdot$$

$$\operatorname{Tr}\Big(\gamma\, \alpha_0 U_0^*\, \nabla_\sigma^{n_1}(\alpha_1 U_1^*) \cdots \nabla_\sigma^{n_q}(\alpha_q U_q^*)\, e^{-t^2D^2}\Big).$$

In view of (4.18), we may assume $\mu(U_0 \cdots U_q)^{-2(n_1+\ldots+n_q)} = 1$; multiplying by $\mu(U_0 \cdots U_q)^{-2(n_1+\ldots+n_q)}$, we can continue by

$$= \sum_{N\geq 0} (-1)^N t^{2N} \sum_{n_1+\ldots+n_q=N} \frac{\mu(U_1)^{-2n_1} \cdots \mu(U_q)^{-2(n_1+\ldots+n_q)}}{n_1! \cdots n_q!(n_1+1)\cdots(n_1+\cdots n_q+q)} \cdot$$

$$\operatorname{Tr}\Big(\gamma\, \alpha_0 U_0^*\, \nabla_\sigma^{n_1}(\alpha_1 U_1^*) \cdots \nabla_\sigma^{n_q}(\alpha_q U_q^*)\, e^{-t^2D^2}\Big)$$

$$= \sum_{N\geq 0} (-1)^N t^{2N} \sum_{n_1+\ldots+n_q=N} \frac{1}{n_1! \cdots n_q!(n_1+1)\cdots(n_1+\cdots n_q+q)} \cdot$$

$$\operatorname{Tr}\Big(\gamma\, \alpha_0 U_0^*\, \nabla_\sigma^{n_1}\big(\sigma^{-2n_1}(\alpha_1 U_1^*)\big) \cdots \nabla_\sigma^{n_q}\big(\sigma^{-2(n_1+\ldots+n_q)}(\alpha_q U_q^*)\big)\, e^{-t^2D^2}\Big).$$

Comparing with the expansion obtained in the proof of Proposition 4.4, and converting the result into a residue via the Mellin transform, we arrive at the following conclusion.

Proposition 4.7. *The constant term has the expression*

$$\langle \alpha_0 U_0^*, \dots, \alpha_q U_q^* \rangle_D |_0 = \frac{1}{q!} \sum_{n_1, \dots, n_q \geq 0} \frac{(-1)^{n_1+\ldots+n_q} \Gamma\big(n_1 + \ldots + n_q + \frac{q}{2}\big)}{n_1! \cdots n_q! (n_1+1) \cdots (n_1 + \cdots n_q + q)} \cdot$$
$$\int\!\!\!\!\!\!- _D \gamma\, \alpha_0 U_0^* \nabla_\sigma^{n_1}\big(\sigma^{-2n_1}(\alpha_1 U_1^*)\big) \cdots \nabla_\sigma^{n_q}\big(\sigma^{-2(n_1+\ldots+n_q)}(\alpha_q U_q^*)\big) \, |D|^{-2(n_1+\ldots+n_q)-q}.$$

4.3. The cocycle identity. To compute coboundaries of the twisted brackets in the cyclic bicomplex, one needs to establish identities similar to those satisfied by the usual JLO brackets, *cf.* [**15**, Lemma 2.2].

Lemma 4.8. *With the same notation as in the preceding subsection, one has*

$$\langle \alpha_0 U_0^*, \dots, \alpha_q U_q^* \rangle_D \; = \sum_{k=0}^{q} \langle \alpha_0 U_0^*, \dots, 1, \alpha_k U_k^*, \dots, a_q U_q^* \rangle_D .$$

Proof. Proceeding as in [**15**, loc. cit.], one writes

$$\langle \alpha_0 U_0^*, \dots, \alpha_q U_q^* \rangle_D \; = \int_0^1 \langle \alpha_0 U_0^*, \dots, \alpha_q U_q^* \rangle_D \, ds \; =$$
$$= \int_0^1 ds \int_{0 \leq s_1 \leq \ldots \leq s_q \leq 1} \mathrm{Tr}\big(\gamma\, \alpha_0 U_0^* \, e^{-s_1 \mu(U_0)^2 D^2} \alpha_1 U_1^* \cdot$$
$$\cdot\, e^{-(s_2 - s_1)\mu(U_0 U_1)^2 D^2} \cdots \alpha_q U_q^* \, e^{-(1-s_q)\mu(U_0 \cdots U_q)^2 D^2}\big) \; =$$
$$= \int_{0 \leq s \leq s_1 \leq \ldots \leq s_q \leq 1} \mathrm{Tr}\big(\gamma\, \alpha_0 U_0^* \, e^{-s\mu(U_0)^2 D^2} \cdot 1 \cdot e^{-(s_1 - s)\mu(U_0)^2 D^2} \alpha_1 U_1^* \cdot$$
$$\cdot\, e^{-(s_2 - s)\mu(U_0 U_1)^2 D^2} \cdots \alpha_q U_q^* \, e^{-(1-s_q)\mu(U_0 \cdots U_q)^2 D^2}\big) \; +$$
$$+ \int_{0 \leq s_1 \leq s \leq \ldots \leq s_q \leq 1} \mathrm{Tr}\big(\gamma\, \alpha_0 U_0^* \, e^{-s_1 \mu(U_0)^2 D^2} \alpha_1 U_1^* \, e^{-(s - s_1)\mu(U_0 U_1)^2 D^2}$$
$$\cdot \quad 1 \cdot e^{-(s_2 - s)\mu(U_0 U_1)^2 D^2} \cdots \alpha_q U_q^* \, e^{-(1-s_q)\mu(U_0 \cdots U_q)^2 D^2}\big) \quad + \quad \ldots$$
$$\ldots \quad + \int_{0 \leq s_1 \leq \ldots \leq s_q \leq s \leq 1} \mathrm{Tr}\big(\gamma\, \alpha_0 U_0^* \, e^{-s_1 \mu(U_0)^2 D^2} \cdots \alpha_q U_q^* \cdot$$
$$\cdot\, e^{-(s - s_q)\mu(U_0 \cdots U_q)^2 D^2} \cdot 1 \cdot e^{-(1-s_q)\mu(U_0 \cdots U_q)^2 D^2}\big).$$

Lemma 4.9. *For $j = 1, \dots, q-1$, one has*

$$\langle \alpha_0 U_0^*, \dots, \alpha_{j-1} U_{j-1}^* \cdot \alpha_j U_j^*, \alpha_{j+1} U_{j+1}^*, \dots, \alpha_q U_q^* \rangle_D$$
$$- \langle \alpha_0 U_0^*, \dots, \alpha_{j-1} U_{j-1}^*, \alpha_j U_j^* \cdot \alpha_{j+1} U_{j+1}^*, \dots, \alpha_q U_q^* \rangle_D$$
$$= \langle \sigma^2(\alpha_0 U_0^*), \dots, \sigma^2(\alpha_{j-1} U_{j-1}^*), [D^2, \alpha_j U_j^*]_\sigma, \alpha_{j+1} U_{j+1}^*, \dots, \alpha_q U_q^* \rangle_D .$$

PROOF. Making use of the commutator formula (4.20), one can write

$$\begin{aligned}
&\langle \alpha_0 U_0^*, \dots, \alpha_{j-1} U_{j-1}^* \cdot \alpha_j U_j^*, \alpha_{j+1} U_{j+1}^*, \dots, \alpha_q U_q^* \rangle_D - \langle \alpha_0 U_0^*, \dots, \alpha_{j-1} U_{j-1}^*, \\
&\qquad\qquad\qquad \alpha_j U_j^* \cdot \alpha_{j+1} U_{j+1}^*, \dots, \alpha_q U_q^* \rangle_D \\
&= \int_{\Delta_q} \mathrm{Tr}\Big(\gamma \alpha_0 U_0^* \, e^{-s_1 \mu(U_0)^2 D^2} \cdots \alpha_{j-1} U_{j-1}^* \big(\int_{t_{j-1}}^{t_{j+1}} e^{-(s_j - s_{j-1}) \mu(U_0 \cdots U_{j-1})^2 D^2} \cdot \\
&\qquad \mu(U_0 \cdots U_{j-1})^2 [D^2, \alpha_j U_j^*]_\sigma e^{-(s_{j+1} - s_j) \mu(U_0 \cdots U_j)^2 D^2} ds_j \big) \cdot \\
&\qquad\qquad \alpha_{j+1} U_{j+1}^* \cdots \alpha_q U_q^* \, e^{-(1-s_q) \mu(U_0 \cdots U_q)^2 D^2} \Big) \\
&= \int_{\Delta_q} \mathrm{Tr}\Big(\gamma \sigma^2(\alpha_0 U_0^*) \, e^{-s_1 \mu(U_0)^2 D^2} \cdots \sigma^2(\alpha_{j-1} U_{j-1}^*) \\
&\big(\int_{t_{j-1}}^{t_{j+1}} e^{-(s_j - s_{j-1}) \mu(U_0 \cdots U_{j-1})^2 D^2} [D^2, \alpha_j U_j^*]_\sigma e^{-(s_{j+1} - s_j) \mu(U_0 \cdots U_j)^2 D^2} ds_j \big) \cdot \\
&\qquad \alpha_{j+1} U_{j+1}^* \cdots \alpha_q U_q^* \, e^{-(1-s_q) \mu(U_0 \cdots U_q)^2 D^2} \Big).
\end{aligned}$$

In contrast with the untwisted case, the cyclic symmetry property is no longer exactly satisfied. It only subsists in a weaker form.

LEMMA 4.10. *With m denoting the degree in D of the product $\alpha_0 \cdots \alpha_q$, one has*

$$\langle \alpha_0 U_0^*, \dots, \alpha_q U_q^* \rangle_D |_0 = \langle \alpha_1 U_1^*, \dots, \alpha_q U_q^*, \sigma^{-m}(\alpha_0 U_0^*) \rangle_D |_0. \tag{4.23}$$

Moreover, if $\mu(U_0 \cdots U_q) = 1$, then

$$\langle \alpha_0 U_0^*, \dots, \alpha_q U_q^* \rangle_D(\varepsilon) = \langle \alpha_1 U_1^*, \dots, \alpha_q U_q^*, \sigma^{-m}(\alpha_0 U_0^*) \rangle_D(\mu(U_0)^2 \varepsilon). \tag{4.24}$$

PROOF. Indeed,

$$\begin{aligned}
&\langle \alpha_0 U_0^*, \dots, \alpha_q U_q^* \rangle_D(\mu(U_0)^{-2} \varepsilon) = \\
&= \mu(U_0)^{-m} \varepsilon^{\frac{m}{2}} \int_{\Delta_q} \mathrm{Tr}\big(\gamma \alpha_0 U_0^* e^{-s_{q+1} \varepsilon D^2} \alpha_1 U_1^* e^{-s_1 \mu(U_1)^2 \varepsilon D^2} \cdots \\
&\cdots \alpha_q U_q^* \, e^{-s_q \mu(U_1 \cdots U_q)^2 \varepsilon D^2}\big) = \varepsilon^{\frac{m}{2}} \int_{\Delta_q} \mathrm{Tr}\big(\gamma \alpha_1 U_1^* e^{-s_1 \mu(U_1)^2 \varepsilon D^2} \cdots \\
&\cdots \alpha_q U_q^* \, e^{-s_q \mu(U_1 \cdots U_q)^2 \varepsilon D^2} \sigma^{-m}(\alpha_0 U_0^*) e^{-s_{q+1} \varepsilon D^2}\big),
\end{aligned}$$

which under the assumption $\mu(U_0 \cdots U_q) = 1$ equals

$$\begin{aligned}
= \quad & \varepsilon^{\frac{m}{2}} \int_{\Delta_q} \mathrm{Tr}\big(\gamma \alpha_1 U_1^* e^{-s_1 \mu(U_1)^2 \varepsilon D^2} \cdots \alpha_q U_q^* \, e^{-s_q \mu(U_1 \cdots U_q)^2 \varepsilon D^2} \cdot \\
& \cdot \sigma^{-m}(\alpha_0 U_0^*) \, e^{-s_{q+1} \mu(U_0 \cdots U_q)^2 \varepsilon D^2}\big) = \langle \alpha_1 U_1^*, \dots, \alpha_q U_q^*, \sigma^{-m}(\alpha_0 U_0^*) \rangle_D(\varepsilon).
\end{aligned}$$

This proves (4.24), and also implies the equality of the their constant terms. On the other hand, if $\mu(U_0 \cdots U_q) \neq 1$, then both sides of (4.23) vanish, *cf.* (4.18). $\square$

We now introduce the twisted version of the JLO cocycles by defining, for any $q+1$ elements $A_0, \dots, A_q \in \mathcal{A}_G$,

$$J^q(D)(A_0, \dots, A_q) = \langle A_0, [D, \sigma^{-1}(A_1)]_\sigma, \dots, [D, \sigma^{-q}(A_q)]_\sigma \rangle_D. \tag{4.25}$$

The collection $\{J^q(D)\}_{q=0,2,4,\ldots}$, resp. $\{J^q(D)\}_{q=1,3,5,\ldots}$, is a cochain in the entire cyclic cohomology bicomplex of $\mathcal{A}_G$ but, because of the failure of cyclic symmetry pointed out above, this cochain is not a cocycle. Instead, we form the 1-parameter family

$$J^q(\varepsilon^{1/2}D)(A_0,\ldots,A_q) := \varepsilon^{\frac{q}{2}}\langle A_0,[D,\sigma^{-1}(A_1)]_\sigma,\ldots,[D,\sigma^{-q}(A_q)]_\sigma\rangle_{\varepsilon^{1/2}D}\,, \tag{4.26}$$

where $\varepsilon\in\mathbb{R}^+$, and passing to the constant term we define

$$\mathcal{J}^q(D)(A_0,\ldots,A_q) \;:=\; \langle A_0,[D,\sigma^{-1}(A_1)]_\sigma,\ldots,[D,\sigma^{-q}(A_q)]_\sigma\rangle_D|_0\,. \tag{4.27}$$

According to Proposition 4.7, it has the explicit form predicted by the Ansatz

$$\begin{aligned}\mathcal{J}^q(D)(A_0,\ldots,A_q) \;&=\; \sum_{\mathbf{k}} c_{q,\mathbf{k}} {\int\!\!\!\!\!\!-}_{D}\,\gamma\, A_0\,[D,\sigma^{-2k_1-1}(A_1)]_\sigma^{(k_1)}\cdots\\ &\qquad\cdots[D,\sigma^{-2(k_1+\ldots+k_q)-q}(A_q)]_\sigma^{(k_q)}\,|D|^{-2|\mathbf{k}|-q},\end{aligned} \tag{4.28}$$

which in particular implies that

$$\mathcal{J}^q(D)=0,\qquad \text{for any}\quad q>p\,; \tag{4.29}$$

also, by Proposition 4.4

$$\mathcal{J}^q(D)(a_0U_0^*,\ldots,a_qU_q^*)\;=\;0,\qquad\text{if}\qquad \mu(U_0^*\cdots U_q^*)\neq 1\,. \tag{4.30}$$

Thus, $\{\mathcal{J}^q(D)\}_{q=0,2,4,\ldots}$, resp. $\{\mathcal{J}^q(D)\}_{q=1,3,5,\ldots}$, defines a cochain in the (b,B)-bicomplex of $\mathcal{A}_G$, which is supported on the conjugacy classes from G_0.

THEOREM 4.11. *The cochain $\mathcal{J}^\bullet(D)$ satisfies the cocycle identity*

$$b\mathcal{J}^{q-1}(D)(a_0U_0^*,\ldots,a_qU_q^*)\;+\;B\mathcal{J}^{q+1}(D)(a_0U_0^*,\ldots,a_qU_q^*)\;=\;0\,. \tag{4.31}$$

PROOF. The first stage of the proof will consist in establishing the identity

$$\begin{aligned}b\mathcal{J}^{q-1}(D)(a_0U_0^*,\ldots,a_qU_q^*)\;&=\;\sum_{j=1}^{q}(-1)^{j-1}\langle\sigma(a_0U_0^*),\ldots,\\ &[D,\sigma^{-(j-2)}(a_{j-1}U_{j-1}^*)]_\sigma,[D^2,\sigma^{-j}(a_jU_j^*)]_\sigma,[D,\sigma^{-j-1}(a_{j+1}U_{j+1}^*)]_\sigma,\ldots,\\ &\qquad\ldots,[D,\sigma^{-q}(a_qU_q^*)]_\sigma\rangle_D|_0.\end{aligned} \tag{4.32}$$

To this end, we compute

$$
\begin{aligned}
bJ^{q-1}(D)(a_0U_0^*,\dots,a_qU_q^*) \;=\; &\langle a_0U_0^*\cdot a_1U_1^*,\dots,[D,\sigma^{-(q-1)}(a_qU_q^*)]_\sigma\rangle_D \;+\\
&+\sum_{j=1}^{q-1}(-1)^j\langle a_0U_0^*,\dots,[D,\sigma^{-j}(a_jU_j^*\cdot a_{j+1}U_{j+1}^*)]_\sigma,\dots\rangle_D+\\
&+(-1)^q\langle a_qU_q^*\cdot a_0U_0^*,\dots,[D,\sigma^{-(q-1)}(a_{q-1}U_{q-1}^*)]_\sigma\rangle_D\\
=\;&\langle a_0U_0^*\cdot a_1U_1^*,\dots,[D,\sigma^{-(q-1)}(a_qU_q^*]_\sigma\rangle_D\;+\\
&-\langle a_0U_0^*\,,\,a_1U_1^*\cdot[D,\sigma^{-1}(a_2U_2^*)]_\sigma,\dots,[D,\sigma^{-(q-1)}(a_qU_q^*)]_\sigma\rangle_D\;+\\
&+\sum_{j=2}^{q-1}(-1)^{j-1}\langle a_0U_0^*,\dots\\
&\qquad\dots,[D,\sigma^{-(j-1)}(a_{j-1}U_{j-1}^*)]_\sigma\cdot\sigma^{-(j-1)}(a_jU_j^*),\dots\rangle_D\;+\\
&+\sum_{j=2}^{q-1}(-1)^j\langle a_0U_0^*,\dots,\sigma^{-(j-1)}(a_jU_j^*)\cdot[D,\sigma^{-j}(a_{j+1}U_{j+1}^*)]_\sigma,\dots\rangle_D\;+\\
&+(-1)^{(q-1)}\langle a_0U_0^*,[D,\sigma^{-1}(a_1U_1^*)]_\sigma,\dots\\
&\qquad\dots,[D,\sigma^{-(q-1)}(a_{q-1}U_{q-1}^*)]_\sigma\cdot\sigma^{-(q-1)}(a_qU_q^*)\rangle_D\\
&+(-1)^q\langle a_qU_q^*\cdot a_0U_0^*,\dots,[D,\sigma^{-(q-1)}(a_{q-1}U_{q-1}^*)]_\sigma\rangle_D,
\end{aligned}
$$

which by Lemma 4.9 is equal to

$$
\begin{aligned}
&\langle\sigma^2(a_0U_0^*),[D^2,a_1U_1^*]_\sigma,\dots,[D,\sigma^{-(q-1)}(a_qU_q^*)]_\sigma\rangle_D\;+\\
&+\sum_{j=2}^{q-1}(-1)^{j-1}\langle\sigma^2(a_0U_0^*),\dots,[D,\sigma^{-(j-3)}(a_{j-1}U_{j-1}^*)]_\sigma,[D^2,\sigma^{-(j-1)}(a_jU_j^*)]_\sigma,\\
&\qquad[D,\sigma^{-j}(a_{j+1}U_{j+1}^*)]_\sigma,\dots,[D,\sigma^{-(q-1)}(a_qU_q^*)]_\sigma\rangle_D\\
&+(-1)^{(q-1)}\langle a_0U_0^*,[D,\sigma^{-1}(a_1U_1^*)]_\sigma,\dots\\
&\qquad\dots,[D,\sigma^{-(q-1)}(a_{q-1}U_{q-1}^*)]_\sigma\cdot\sigma^{-(q-1)}(a_qU_q^*)\rangle_D\;+\\
&\qquad+(-1)^q\langle a_qU_q^*\cdot a_0U_0^*,\dots,[D,\sigma^{-(q-1)}(a_{q-1}U_{q-1}^*)]_\sigma\rangle_D.
\end{aligned}
$$

At this point we pass to the constant terms and use Eq. (4.23) for the last two terms, to replace them by the sum

$$
\begin{aligned}
&(-1)^{(q-1)}\langle[D,\sigma^{-1}(a_1U_1^*)]_\sigma,\dots\\
&\dots,[D,\sigma^{-(q-1)}(a_{q-1}U_{q-1}^*)]_\sigma\cdot\sigma^{-(q-1)}(a_qU_q^*),\sigma^{-(q-1)}(a_0U_0^*)\rangle_D|_0\\
&\qquad+(-1)^q\langle[D,\sigma^{-1}(a_1U_1^*)]_\sigma,\dots,[D,\sigma^{-(q-1)}(a_{q-1}U_{q-1}^*)]_\sigma,\\
&\qquad\qquad\sigma^{-(q-1)}(a_qU_q^*)\cdot\sigma^{-(q-1)}(a_0U_0^*)\rangle_D|_0\,.
\end{aligned}
$$

In turn, by Lemma 4.9 this equals

$$
\begin{aligned}
=(-1)^{(q-1)}\langle&[D,\sigma(a_1U_1^*)]_\sigma,\dots,[D,\sigma^{-(q-3)}(a_{q-1}U_{q-1}^*)]_\sigma,\\
&[D^2,\sigma^{-(q-1)}(a_qU_q^*)]_\sigma,\sigma^{-(q-1)}(a_0U_0^*)\rangle_D|_0.
\end{aligned}
$$

Applying once again Eq. (4.23), and taking into account that the homogeneity degree in D is $q+1$, the above expression becomes

$$= (-1)^{(q-1)}\langle \sigma^2(a_0U_0^*), [D, \sigma(a_1U_1^*)]_\sigma, \dots, [D, \sigma^{-(q-3)}(a_{q-1}U_{q-1}^*)]_\sigma, \\ [D^2, \sigma^{-(q-1)}(a_qU_q^*)]_\sigma\rangle_D|_0.$$

Summing up, one obtains

$$b\mathcal{J}^{q-1}(D)(a_0U_0^*, \dots, a_qU_q^*) = \\ \sum_{j=1}^{q}(-1)^{j-1}\langle \sigma^2(a_0U_0^*), \dots, [D, \sigma^{-(j-3)}(a_{j-1}U_{j-1}^*)]_\sigma, [D, [D, \sigma^{-(j-1)}(a_jU_j^*)]_\sigma]_\sigma, \\ [D, \sigma^{-j}(a_{j+1}U_{j+1}^*)]_\sigma, \dots, [D, \sigma^{-(q-1)}(a_qU_q^*)]_\sigma\rangle_D|_0.$$

Combined with the σ-invariance property (4.19), this completes the proof of Eq. (4.32).

The second stage of the proof will show that the B-boundary

$$B\mathcal{J}^{q+1}(D)(a_0U_0^*, \dots, a_qU_q^*) = \\ = \sum_{k=0}^{q}(-1)^{kq}\,\mathcal{J}^{q+1}(D)(1, a_kU_k^*, \dots a_qU_q^*, a_0U_0^*, \dots, a_{k-1}U_{k-1}^*),$$

satisfies the identity

$$B\mathcal{J}^{q+1}(D)(a_0U_0^*, \dots, a_qU_q^*) = \langle [D, \sigma^{-1}(a_0U_0^*)]_\sigma, \dots \\ \dots, [D, \sigma^{-(q+1)}(a_qU_q^*)]_\sigma\rangle_D|_0. \tag{4.33}$$

To this end, we note that by Lemma 4.8,

$$\langle [D, \sigma^{-1}(a_0U_0^*)]_\sigma, \dots, [D, \sigma^{-(q+1)}(a_qU_q^*)]_\sigma\rangle_D = \\ = \sum_{k=0}^{q}\langle [D, \sigma^{-1}(a_0U_0^*)]_\sigma, \dots, 1, [D, \sigma^{-(k+1)}(a_kU_k^*)]_\sigma, \dots, [D, \sigma^{-(q+1)}(a_qU_q^*)]_\sigma\rangle_D.$$

Passing to the constant term, we apply Eq. (4.23) k times to the k-th term of the sum and rewrite it in the form

$$\langle [D, \sigma^{-1}(a_0U_0^*)]_\sigma, \dots, 1, [D, \sigma^{-(k+1)}(a_kU_k^*)]_\sigma, \dots, [D, \sigma^{-(q+1)}(a_qU_q^*)]_\sigma\rangle_D|_0 \\ = (-1)^{kq}\langle 1, [D, \sigma^{-(k+1)}(a_kU_k^*)]_\sigma, \dots [D, \sigma^{-(q+1)}(a_qU_q^*)]_\sigma, [D, \sigma^{-(q+2)}(a_0U_0^*)]_\sigma, \\ \dots, [D, \sigma^{-(q+k+1)}(a_{k-1}U_{k-1}^*)]_\sigma\rangle_D|_0.$$

Summing up, one obtains

$$\langle [D, \sigma^{-1}(a_0U_0^*)]_\sigma, \dots, [D, \sigma^{-(q+1)}(a_qU_q^*)]_\sigma\rangle_D|_0 \\ = \sum_{k=0}^{q}(-1)^{kq}\langle 1, [D, \sigma^{-(k+1)}(a_kU_k^*)]_\sigma, \dots, [D, \sigma^{-(q+1)}(a_qU_q^*)]_\sigma, [D, \sigma^{-(q+2)}(a_0U_0^*)]_\sigma, \\ \dots, [D, \sigma^{-(q+k+1)}(a_{k-1}U_{k-1}^*)]_\sigma\rangle_D|_0 \\ = \sum_{k=0}^{q}(-1)^{kq}\,\mathcal{J}^{q+1}(D)(1, a_kU_k^*, \dots, a_qU_q^*, a_0U_0^*, \dots, a_{k-1}U_{k-1}^*),$$

which proves Eq. (4.33).

To relate the two identities satisfied by the coboundary operators, we use the generalized Leibniz rule (3.27) and write

$$
\begin{aligned}
&[D, a_0 U_0^* \, e^{-s_0 \mu(U_0)^2 D^2} \cdots [D, \sigma^{-q}(a_q U_q^*)]_\sigma \, e^{-s_q \mu(U_0 \cdots U_q)^2 D^2}]_\sigma \\
&= [D, a_0 U_0^*]_\sigma \, e^{-s_0 \mu(U_0)^2 D^2} \cdots [D, \sigma^{-q}(a_q U_q^*)]_\sigma \, e^{-s_q \mu(U_0 \cdots U_q)^2 D^2} + \\
&\sigma(a_0 U_0^*) \, e^{-s_0 \mu(U_0)^2 D^2} [D^2, \sigma^{-1}(a_1 U_1^*)]_\sigma \, e^{-s_1 \mu(U_0 U_1)^2 D^2} \cdots \\
&\qquad\qquad \cdots [D, \sigma^{-q}(a_q U_q^*)]_\sigma \, e^{-s_q \mu(U_0 \cdots U_q)^2 D^2} \; + \cdots + \\
&(-1)^{q-1} \sigma(a_0 U_0^*) e^{-s_0 \mu(U_0)^2 D^2} [D, a_1 U_1^*]_\sigma e^{-s_1 \mu(U_0 U_1)^2 D^2} \cdots \\
&\qquad\qquad\qquad \cdots [D^2, \sigma^{-q}(a_q U_q^*)]_\sigma e^{-s_q \mu(U_0 \cdots U_q)^2 D^2} .
\end{aligned}
$$

Since, in view of the Selberg property (4.10), $\displaystyle⨍_D$ vanishes on twisted graded commutators, one obtains

$$
\begin{aligned}
&-\langle [D, a_0 U_0^*]_\sigma, [D, \sigma^{-1}(a_1 U_1^*)]_\sigma, \ldots, [D, \sigma^{-q}(a_q U_q^*)]_\sigma \rangle_D |_0 \; = \\
&= \sum_{j=1}^{q} (-1)^{j-1} \langle \sigma(a_0 U_0^*), \ldots, [D, \sigma^{-(j-2)}(a_{j-1} U_{j-1}^*)]_\sigma, [D^2, \sigma^{-j}(a_j U_j^*)]_\sigma, \\
&\qquad\qquad\qquad [D, \sigma^{-j-1}(a_{j+1} U_{j+1}^*)]_\sigma, \ldots, [D, \sigma^{-q}(a_q U_q^*)]_\sigma \rangle_D |_0 .
\end{aligned}
$$

We can now rewrite Eq. (4.32) in the form

$$
\begin{aligned}
&b\mathcal{J}^{q-1}(D)(a_0 U_0^*, \ldots, a_q U_q^*) \; = \\
&\qquad - \langle [D, a_0 U_0^*]_\sigma, [D, \sigma^{-1}(a_1 U_1^*)]_\sigma, \ldots, [D, \sigma^{-q}(a_q U_q^*)]_\sigma \rangle_D |_0 \, ,
\end{aligned}
$$

Using once more the invariance property (4.19) and comparing with Eq. (4.33) one obtains the desired cocycle identity. □

4.4. Transgression and proof of the Ansatz. In view of the the property (4.30), we can restrict our considerations to the (b, B)-subcomplex $CC^*_{G_0}(\mathcal{A}_G)$ of cochains supported by the conjugacy classes in G_0. Far from being a mere convenience, this restriction is actually essential for the validity of the ensuing calculations.

We denote by $\iota_s(V)$ the *twisted contraction* operator on $CC^*(\Psi_G)$,

$$
\begin{aligned}
&\iota_\sigma(V)\langle A_0, \ldots, A_q \rangle_D \; = \\
&\qquad = \sum_{k=0}^{q} (-1)^{(\# A_0 + \cdots + \# A_k)\# V} \langle \sigma^2(A_0), \ldots, \sigma^2(A_k), V, A_{k+1}, \ldots, A_q \rangle_D ,
\end{aligned}
$$

where $\#A$ stands for the degree of A, and extend it to cochain-valued functions by setting

$$
(\iota_\sigma(V)\langle A_0, \ldots, A_q \rangle_D)(\varepsilon) := \iota_s(V)\big(\langle A_0, \ldots, A_q \rangle_D(\varepsilon)\big), \qquad \varepsilon \in \mathbb{R}^+ .
$$

In what follows, we shall denote by $\tau \mapsto D_\tau$ one of the following two families of operators $D_t = tD$, $t \in \mathbb{R}^+$ and $D_u = D|D|^{-u}$, $u \in [0,1]$, and will denote by $\dot{D}$ the corresponding derivative. In each case, we define the cochains $\mathcal{I}^q(D_\tau, V) \in CC^*_{G_0}(\mathcal{A}_G)$ by the formula

$$
\begin{aligned}
(4.34) \quad &\mathcal{I}^q(D_\tau, V)(a_0 U_0^*, \ldots, a_q U_q^*) \; = \\
&\qquad = \; \iota_\sigma(V)\langle a_0 U_0^*, [D_\tau, \sigma^{-1}(a_1 U_1^*)]_\sigma, \ldots, [D_\tau, \sigma^{-q}(a_q U_q^*)]_\sigma \rangle_{D_\tau} \, ,
\end{aligned}
$$

where V will be either the (odd) operator $\dot{D}_\tau$ or the (even) operator $[D_\tau, \dot{D}_\tau]$.

We would like to evaluate the expression

$$\frac{d}{d\tau} J^q(D_\tau) \,+\, b\slashed{J}^{q-1}(D_\tau, \dot{D}_\tau) + B\slashed{J}^{q+1}(D_\tau, \dot{D}_\tau)\,, \tag{4.35}$$

which vanishes in the untwisted case (*cf. e.g.* [**14**, Prop. 10.12]). The derivative

$$\frac{d}{d\tau} J^q(D_\tau)(a_0U_0^*, \ldots, a_qU_q^*) \;=\; \int_{\Delta_q} \frac{d}{d\tau}\mathrm{Tr}\Big(\gamma a_0 U_0^* e^{-s_1\mu(U_0)^2 D_\tau^2}[D_\tau, \sigma^{-1}(a_1U_1^*)]_\sigma$$
$$e^{-(s_2-s_1)\mu(U_0U_1)^2D_\tau^2}\cdots[D_\tau, \sigma^{-q}(a_qU_q^*)]_\sigma e^{-(1-s_q)\mu(U_0\cdots U_q)^2D_\tau^2}\Big).$$

splits into two sums of terms. The first sum simply consists of the derivatives of the twisted commutators

$$K^q(D_\tau)(a_0U_0^*, \ldots, a_qU_q^*) \,:=\, \sum_{j=1}^{q}\langle a_0U_0^*, \ldots, [\dot{D}_\tau, \sigma^{-j}(a_jU_j^*)]_\sigma, \ldots\rangle_{D_\tau}. \tag{4.36}$$

To evaluate the second sum one relies, as in the standard case, on the Duhamel formula

$$\frac{d}{d\tau} e^{-D_\tau^2} \;=\; -\int_0^1 e^{-sD_\tau^2}[D_\tau, \dot{D}_\tau]\, e^{-(1-s)D_\tau^2}\, ds.$$

By applying it in the form

$$\frac{d}{d\tau} e^{-(s_{j+1}-s_j)\mu^2D_\tau^2} \;=\; -\mu^2 \int_{s_j}^{s_{j+1}} e^{-(s-s_j)\mu^2D_\tau^2}[D_\tau, \dot{D}_\tau]\, e^{-(s_{j+1}-s)\mu^2D_\tau^2}\, ds\,, \tag{4.37}$$

one obtains

$$\sum_{j=0}^{q}\int_{\Delta_q}\mathrm{Tr}\Big(\gamma\cdots[D_\tau, \sigma^{-j}(a_jU_j^*)]_\sigma \frac{d}{d\tau} e^{-(s_{j+1}-s_j)\mu(U_0\cdots U_j)^2D_\tau^2}\cdots\Big) =$$
$$= -\sum_{j=0}^{q}\mu(U_0\cdots U_j)^2\int_{\Delta_{q+1}}\mathrm{Tr}\Big(\gamma\cdots[D_\tau, \sigma^{-j}(a_jU_j^*)]_\sigma$$
$$e^{-(s-s_j)\mu(U_0\cdots U_j)^2D_\tau}[D_\tau, \dot{D}_\tau]\, e^{-(s_{j+1}-s)\mu(U_0\cdots U_j)^2D_\tau^2}\cdots\Big) =$$
$$= -\sum_{j=0}^{q}\int_{\Delta_{q+1}}\langle\sigma^2(a_0U_0^*), \ldots, [D_\tau, \sigma^{-(j-2)}(a_jU_j^*)]_\sigma, [D_\tau, \dot{D}_\tau], \ldots,$$
$$\ldots, [D_\tau, \sigma^{-q}(a_qU_q^*)]_\sigma\rangle_{D_\tau} \qquad = \qquad -\slashed{J}^q(D_\tau, [D_\tau, \dot{D}_\tau])(a_0U_0^*, \ldots, a_qU_q^*)\,.$$

This gives the identity

$$\frac{d}{d\tau} J^q(D_\tau) \;=\; K^q(D_\tau) \,-\, \slashed{J}^q(D_\tau, [D_\tau, \dot{D}_\tau]). \tag{4.38}$$

On the other hand, in order to evaluate the coboundary of $\slashed{J}^\bullet(D_\tau, \dot{D}_\tau)$, as in the proof of Theorem 4.11, we apply the Leibniz rule to the integrand for the expression of $\slashed{J}^q(D_\tau, \dot{D}_\tau)(a_0U_0^*, \ldots, a_qU_q^*)$. By abuse of notation, we write this bracket operation in the form

$$[D_\tau,\, \slashed{J}^q(D_\tau, \dot{D}_\tau)(a_0U_0^*, \ldots, a_qU_q^*)]_\sigma\,,$$

and compute it as follows

$$
\begin{aligned}
&[D_\tau, \iota_\sigma(\dot{D}_\tau)\langle a_0U_0^*, [D_\tau, \sigma^{-1}(a_1U_1^*)]_\sigma, \dots, [D_\tau, \sigma^{-q}(a_qU_q^*)]_\sigma\rangle_{D_\tau}]_\sigma = \\
&= [D_\tau, \langle\sigma^2(a_0U_0^*), \dot{D}_\tau, [D_\tau, \sigma^{-1}(a_1U_1^*)]_\sigma, \dots, [D_\tau, \sigma^{-q}(a_qU_q^*)]_\sigma\rangle_{D_\tau}]_\sigma + \cdots \\
&= \langle[D_\tau, \sigma^2(a_0U_0^*)]_\sigma, \dot{D}_\tau, [D_\tau, \sigma(a_1U_1^*)]_\sigma, \dots, [D_\tau, \sigma^{-q}(a_qU_q^*)]_\sigma\rangle_{D_\tau} \\
&\ +\langle\sigma^4(a_0U_0^*), [D_\tau, \dot{D}_\tau], [D_\tau, \sigma^{-1}(a_1U_1^*)]_\sigma, \dots, [D_\tau, \sigma^{-q}(a_qU_q^*)]_\sigma\rangle_{D_\tau} + \\
&+\langle\sigma^4(a_0U_0^*), \dot{D}_\tau, [D_\tau, \sigma(a_1U_1^*)]_\sigma, [D_\tau, \sigma^{-2}(a_2U_2^*)]_\sigma, \dots, [D_\tau, \sigma^{-q}(a_qU_q^*)]_\sigma\rangle_{D_\tau} \\
&\qquad\qquad + \cdots \quad \text{and so on.}
\end{aligned}
$$

There are two kinds of terms appearing in this sum. Those which contain the term $[D_\tau, \sigma^2(a_0U_0^*)]_\sigma$ come from

$$
\iota(\dot{D}_\tau)\,\langle[D_\tau, a_0U_0^*]_\sigma, [D_\tau, \sigma^{-1}(a_1U_1^*)]_\sigma, \dots, [D_\tau, \sigma^{-q}(a_qU_q^*)]_\sigma\rangle_{D_\tau},
$$

and they closely resemble those appearing in $B\mathcal{Y}^{q+1}(D_\tau, \dot{D}_\tau)$. The remaining terms are of the form

$$
\begin{aligned}
&\langle\sigma^4(a_0U_0^*), \dots, \dot{D}_\tau, \dots, [D_\tau, \sigma^{-(j-2)}(a_{j-1}U_{j-1}^*)]_\sigma, [D_\tau^2, \sigma^{-j}(a_jU_j^*)]_\sigma, \\
&\qquad\qquad [D_\tau, \sigma^{-j-1}(a_{j+1}U_{j+1}^*)]_\sigma, \dots, [D_\tau, \sigma^{-q}(a_qU_q^*)]_\sigma\rangle_{D_s},
\end{aligned}
$$

or

$$
\begin{aligned}
&\langle\sigma^4(a_0U_0^*), \dots, [D_\tau, \sigma^{-(j-2)}(a_{j-1}U_{j-1}^*)]_\sigma, [D_\tau^2, \sigma^{-j}(a_jU_j^*)]_\sigma, \dots \\
&\qquad \dots, \dot{D}_\tau, \dots, [D_\tau, \sigma^{-j-1}(a_{j+1}U_{j+1}^*)]_\sigma, \dots, [D_\tau, \sigma^{-q}(a_qU_q^*)]_\sigma\rangle_{D_\tau},
\end{aligned}
$$

and they match those occurring in $b\mathcal{Y}^{q-1}(D_\tau, \dot{D}_\tau) + K^q$, plus terms which contain $[D_\tau, \dot{D}_\tau]$ and account for $\mathcal{Y}^{q-1}(D_\tau, [D_\tau, \dot{D}_\tau])$.

Indeed, for the b-coboundary of $\mathcal{Y}^\bullet(D_\tau, \dot{D}_\tau)$, we write

$$
\begin{aligned}
&b\mathcal{Y}^{q-1}(D_\tau, \dot{D}_\tau)(a_0U_0^*, \dots, a_qU_q^*) = \\
&\quad = \iota_\sigma(\dot{D}_\tau)\langle a_0U_0^* \cdot a_1U_1^*, [D_\tau, \sigma^{-1}(a_2U_2^*)]_\sigma, \dots, [D_\tau, \sigma^{-q+1}(a_qU_q^*)]_\sigma\rangle_{D_\tau} \\
&\quad - \iota_\sigma(\dot{D}_\tau)\langle a_0U_0^*\,,\, a_1U_1^* \cdot [D_\tau, \sigma^{-1}(a_2U_2^*)]_\sigma, \dots, [D_\tau, \sigma^{-(q-1)}(a_qU_q^*)]_\sigma\rangle_{D_\tau} + \\
&\sum_{j=2}^{q-1}(-1)^{j-1}\iota_\sigma(\dot{D}_\tau)\langle a_0U_0^*, \dots, [D_\tau, \sigma^{-(j-1)}(a_{j-1}U_{j-1}^*)]_\sigma \cdot \sigma^{-(j-1)}(a_jU_j^*), \dots\rangle_{D_\tau} \\
&+\sum_{j=2}^{q-1}(-1)^{j}\iota_\sigma(\dot{D}_\tau)\langle a_0U_0^*, \dots, \sigma^{-(j-1)}(a_jU_j^*) \cdot [D_\tau, \sigma^{-j}(a_{j+1}U_{j+1}^*)]_\sigma, \dots\rangle_{D_\tau} + \\
&+(-1)^{(q-1)}\iota_\sigma(\dot{D}_\tau)\langle a_0U_0^*, [D_\tau, \sigma^{-1}(a_1U_1^*)]_\sigma, \dots \\
&\qquad\qquad \dots, [D_\tau, \sigma^{-(q-1)}(a_{q-1}U_{q-1}^*)]_\sigma \cdot \sigma^{-(q-1)}(a_qU_q^*)\rangle_{D_\tau} + \\
&\qquad\qquad + (-1)^q\iota_\sigma(\dot{D}_\tau)\langle a_qU_q^* \cdot a_0U_0^*, \dots, [D_\tau, \sigma^{-(q-1)}(a_{q-1}U_{q-1}^*)]_\sigma\rangle_{D_\tau}\,.
\end{aligned}
$$

Let us take a closer look at the first two terms, and expand $\iota_\sigma(\dot{D}_\tau)$. One has

$$\begin{aligned}
&\iota_\sigma(\dot{D}_\tau)\langle a_0U_0^* \cdot a_1U_1^*, [D_\tau, \sigma^{-1}(a_2U_2^*)]_\sigma, \dots, [D_\tau, \sigma^{-q+1}(a_qU_q^*)]_\sigma\rangle_{D_\tau} \\
&\qquad - \iota_\sigma(\dot{D}_\tau)\langle a_0U_0^*\, ,\, a_1U_1^* \cdot [D, \sigma^{-1}(a_2U_2^*)]_\sigma, \dots, [D_\tau, \sigma^{-(q-1)}(a_qU_q^*)]_\sigma\rangle_{D_\tau} \\
&= \langle \sigma^2(a_0U_0^* \cdot a_1U_1^*),\, \dot{D}_\tau,\, [D_\tau, \sigma^{-1}(a_2U_2^*)]_\sigma, \dots, [D_\tau, \sigma^{-q+1}(a_qU_q^*)]_\sigma\rangle_{D_\tau} \\
&\qquad - \langle \sigma^2(a_0U_0^*),\, \dot{D}_\tau,\, a_1U_1^* \cdot [D_\tau, \sigma^{-1}(a_2U_2^*)]_\sigma, \dots, [D_\tau, \sigma^{-(q-1)}(a_qU_q^*)]_\sigma\rangle_{D_\tau} \\
&+ \sum_{k=2}^{q} (-1)^k \langle \sigma^2(a_0U_0^* \cdot a_1U_1^*), [D_\tau, \sigma(a_2U_2^*)]_\sigma, \dots, \\
&[D_\tau, \sigma^{-(k-3)}(a_kU_k^*)]_\sigma, \dot{D}_\tau, [D_\tau, \sigma^{-k}(a_{k+1}U_{k+1}^*)]_\sigma, \dots, [D_\tau, \sigma^{-(q-1)}(a_qU_q^*)]_\sigma\rangle_{D_\tau} \\
&- \sum_{k=2}^{q} (-1)^k \langle \sigma^2(a_0U_0^*), \sigma^2(a_1U_1^*) \cdot [D_\tau, \sigma(a_2U_2^*)]_\sigma, \dots, \\
&[D_\tau, \sigma^{-(k-3)}(a_kU_k^*)]_\sigma, \dot{D}_\tau, [D_\tau, \sigma^{-k}(a_{k+1}U_{k+1}^*)]_\sigma, \dots, [D_\tau, \sigma^{-(q-1)}(a_qU_q^*)]_\sigma\rangle_{D_\tau} \\
&= \langle \sigma^2(a_0U_0^* \cdot a_1U_1^*),\, \dot{D}_\tau,\, [D_\tau, \sigma^{-1}(a_2U_2^*)]_\sigma, \dots, [D_\tau, \sigma^{-q+1}(a_qU_q^*)]_\sigma\rangle_{D_\tau} \\
&\qquad - \langle \sigma^2(a_0U_0^*),\, \dot{D}_\tau,\, a_1U_1^* \cdot [D_\tau, \sigma^{-1}(a_2U_2^*)]_\sigma, \dots, [D_\tau, \sigma^{-(q-1)}(a_qU_q^*)]_\sigma\rangle_{D_\tau} \\
&+ \sum_{k=2}^{q} (-1)^k \langle \sigma^4(a_0U_0^*), [D_\tau^2, \sigma^2(a_1U_1^*)]_\sigma, [D_\tau, \sigma(a_2U_2^*)]_\sigma, \dots, \\
&[D_\tau, \sigma^{-(k-3)}(a_kU_k^*)]_\sigma, \dot{D}_\tau, [D_\tau, \sigma^{-k}(a_{k+1}U_{k+1}^*)]_\sigma, \dots, [D_\tau, \sigma^{-(q-1)}(a_qU_q^*)]_\sigma\rangle_{D_\tau}
\end{aligned}$$

where we have used Lemma 4.9 after the first pair of terms.

We now look at pairs of terms indexed by the same $j = 2, \dots, q-1$,

$$\begin{aligned}
&\iota_\sigma(\dot{D}_\tau)\langle a_0U_0^*, \dots, [D_\tau, \sigma^{-(j-1)}(a_{j-1}U_{j-1}^*)]_\sigma \cdot \sigma^{-(j-1)}(a_jU_j^*), \dots, \rangle_{D_\tau} \\
&- \iota_\sigma(\dot{D}_\tau)\langle a_0U_0^*, \dots, \sigma^{-(j-1)}(a_jU_j^*) \cdot [D_\tau, \sigma^{-j}(a_{j+1}U_{j+1}^*)]_\sigma, \dots, \rangle_{D_\tau} = \\
&= \sum_{k \le j-2} (-1)^k \langle \sigma^4(a_0U_0^*), \dots, \dot{D}_\tau, \dots, [D_\tau, \sigma^{-(j-3)}(a_{j-1}U_{j-1}^*)]_\sigma, \\
&\qquad\qquad\qquad [D_\tau^2, \sigma^{-(j-1)}(a_jU_j^*)]_\sigma, [D_\tau, \sigma^{-j}(a_{j+1}U_{j+1}^*)]_\sigma, \dots, \rangle_{D_\tau} \\
&+ (-1)^{j-1} \langle \sigma^2(a_0U_0^*), \dots, [D_\tau, \sigma^{-(j-3)}(a_{j-1}U_{j-1}^*)]_\sigma \cdot \sigma^{-(j-3)}(a_jU_j^*), \dot{D}_\tau, \dots, \rangle_{D_\tau} \\
&- (-1)^{j-1} \langle \sigma^2(a_0U_0^*), \dots, \sigma^{-(j-3)}(a_{j-1}U_{j-1}^*), \dot{D}_\tau, \\
&\qquad\qquad\qquad \sigma^{-(j-1)}(a_jU_j^*) \cdot [D_\tau, \sigma^{-j}(a_{j+1}U_{j+1}^*)]_\sigma, \dots, \rangle_{D_\tau} + \\
&+ \sum_{k \ge j} (-1)^k \langle \sigma^4(a_0U_0^*), \dots, [D_\tau, \sigma^{-(j-5)}(a_{j-1}U_{j-1}^*)]_\sigma, [D_\tau^2, \sigma^{-(j-3)}(a_jU_j^*)]_\sigma, \dots, \\
&\qquad\qquad\qquad\qquad \dots, \dot{D}_\tau, \dots, \rangle_{D_\tau}\,,
\end{aligned}$$

where we have again applied Lemma 4.9. We focus on the last two terms,

$$\begin{aligned}
&\iota(\dot{D}_\tau)\langle a_0U_0^*, [D_\tau, \sigma^{-1}(a_1U_1^*)]_\sigma, \dots, [D_\tau, \sigma^{-(q-1)}(a_{q-1}U_{q-1}^*)]_\sigma \cdot \sigma^{-(q-1)}(a_qU_q^*)\rangle_{D_\tau} \\
&\qquad\qquad - \iota(\dot{D}_\tau)\langle a_qU_q^* \cdot a_0U_0^*, \dots, [D_\tau, \sigma^{-(q-1)}(a_{q-1}U_{q-1}^*)]_\sigma\rangle_{D_\tau} \\
&= \langle \sigma^2(a_0U_0^*), \dot{D}_\tau, [D_\tau, \sigma^{-1}(a_1U_1^*)]_\sigma, \dots \\
&\qquad\qquad\qquad \dots, [D_\tau, \sigma^{-(q-1)}(a_{q-1}U_{q-1}^*)]_\sigma \cdot \sigma^{-(q-1)}(a_qU_q^*)\rangle_{D_\tau} \\
&\qquad - \langle \sigma^2(a_qU_q^* \cdot a_0U_0^*), \dot{D}_\tau, \dots, [D_\tau, \sigma^{-(q-1)}(a_{q-1}U_{q-1}^*)]_\sigma\rangle_{D_\tau} + \cdots
\end{aligned}$$

because one has to use Eq. (4.24) to prepare them for the application of Lemma 4.9, as follows:

$$
\begin{aligned}
&= (-1)^{(q-1)}\langle \dot{D}_\tau, [D_\tau, \sigma^{-1}(a_1U_1^*)]_\sigma, \ldots,\\
&\ldots, [D_\tau, \sigma^{-(q-1)}(a_{q-1}U_{q-1}^*)]_\sigma \cdot \sigma^{-(q-1)}(a_qU_q^*), \sigma^{-(q-1)}(a_0U_0^*)\rangle_{\mu(U_0)D_\tau}\\
&\qquad + (-1)^q\langle \dot{D}_\tau, [D_\tau, \sigma^{-1}(a_1U_1^*)]_\sigma, \ldots [D_\tau, \sigma^{-(q-1)}(a_{q-1}U_{q-1}^*)]_\sigma,\\
&\qquad\qquad \sigma^{-(q-1)}(a_qU_q^*) \cdot \sigma^{-(q-1)}(a_0U_0^*)\rangle_{\mu(U_qU_0)D_\tau}.
\end{aligned}
$$

In doing so, we have rescaled the operator D_τ. This kind of rescaling, which appears every time we need to make cyclic rearrangements, prevents the expression (4.35) from vanishing.

However, in the special case of the scaling family $\tau = t \mapsto D_t = tD$, one can integrate from 0 to ∞, replacing the ordinary integral near 0 with its finite part as in [**7**, §4]. Also, since D is invertible, the proof of Lemma 2 in *op.cit.* can be easily replicated to produce the necessary estimates for the behavior of $J^\bullet(tD)$ and $\slashed{Y}^\bullet(tD, D)$ as $t \nearrow \infty$. After integrating all the expressions involved in the above calculation, the mismatching disappears and all cancellations that take place in the untwisted case do occur in this case too. Indeed, taking as an example the first of the two terms above, one has

$$
\begin{aligned}
&\mathrm{Pf}_0 \int_\varepsilon^\infty \langle a_0U_0^*, D, [D_t, \sigma^{-1}(a_1U_1^*)]_\sigma, \ldots,\\
&\qquad\qquad \ldots, [D_t, \sigma^{-(q-1)}(a_{q-1}U_{q-1}^*)]_\sigma \cdot \sigma^{-(q-1)}(a_qU_q^*)\rangle_{tD}\, dt\\
&= \mathrm{Pf}_0 \int_\varepsilon^\infty \langle a_0U_0^*, D, [D, \sigma^{-1}(a_1U_1^*)]_\sigma, \ldots,\\
&\qquad\qquad \ldots, [D, \sigma^{-(q-1)}(a_{q-1}U_{q-1}^*)]_\sigma \cdot \sigma^{-(q-1)}(a_qU_q^*)\rangle_D(t)\, \frac{dt}{t}.
\end{aligned}
$$

By Eq. (4.24) this equals

$$
\begin{aligned}
&= \mathrm{Pf}_0 \int_\varepsilon^\infty \langle D, [D, \sigma^{-1}(a_1U_1^*)]_\sigma, \ldots,\\
&\ldots, [D, \sigma^{-(q-1)}(a_{q-1}U_{q-1}^*)]_\sigma \cdot \sigma^{-(q-1)}(a_qU_q^*),\ \sigma^{-q}(a_0U_0^*)\rangle_D(\mu(U_0)^2t)\, \frac{dt}{t},
\end{aligned}
$$

which after the substitution $t \mapsto \mu(U_0)^{-2}t$ becomes

$$
\begin{aligned}
&= \mathrm{Pf}_0 \int_{\mu(U_0)^2\varepsilon}^\infty \langle D, [D, \sigma^{-1}(a_1U_1^*)]_\sigma, \ldots,\\
&\ldots, [D, \sigma^{-(q-1)}(a_{q-1}U_{q-1}^*)]_\sigma \cdot \sigma^{-(q-1)}(a_qU_q^*),\ \sigma^{-q}(a_0U_0^*)\rangle_D(t)\, \frac{dt}{t}.
\end{aligned}
$$

In this way one obtains the following *transgression formula*:

LEMMA 4.12. *For any $q \geq 0$, one has*

$$
\begin{aligned}
&\mathrm{Pf}_0 J^q(tD) - \lim_{t\nearrow\infty} J^q(tD) = \\
&= b\left(\mathrm{Pf}_0 \int_\varepsilon^\infty \slashed{Y}^{q-1}(tD, D)\, dt\right) + B\left(\mathrm{Pf}_0 \int_\varepsilon^\infty \slashed{Y}^{q+1}(tD, D)\, dt\right).
\end{aligned}
\tag{4.39}
$$

On the other hand, with virtually identical arguments as in the proof of [**7**, Proposition 2], one establishes the similar vanishing result:

$$\lim_{t\nearrow\infty} J^q(tD) \,=\, 0\,. \tag{4.40}$$

In particular, for $q=p$ (summability dimension), applying the b-boundary to Eq. (4.39) and using the cocycle identity (4.31) together with the vanishing property (4.29), one obtains

$$b\,B\left(\mathrm{Pf}_0 \int_\varepsilon^\infty \mathcal{Y}^{p+1}(tD,D)\,dt\right) \,=\, b\mathcal{J}^p(D)) \,=\, -B\mathcal{J}^{p+2}(D)) \,=\, 0\,.$$

This shows that

$$\mathcal{T}^p(D) \,:=\, B\left(\mathrm{Pf}_0 \int_\varepsilon^\infty \mathcal{Y}^{p+1}(tD,D)\,dt\right) \tag{4.41}$$

is a cyclic cocycle.

LEMMA 4.13. *The (b,B)-cocycle $\mathcal{J}^\bullet(D)$ is cohomologous to the cyclic cocycle $\mathcal{T}^p(D)$.*

PROOF. In view of Eqs. (4.39) and (4.40), the difference between the two cochains is a total coboundary in the periodic cyclic complex:

$$\mathcal{J}^\bullet(D) - \mathcal{T}^p(D) \,=\, (b+B)\left(\mathrm{Pf}_0 \int_\varepsilon^\infty \mathcal{Y}^\bullet(tD,D)\,dt\right). \tag{4.42}$$

□

We are now ready to conclude the proof of the Ansatz for spectral triples twisted by scaling automorphisms.

THEOREM 4.14. *The periodic cyclic cohomology class in $HP^*(\mathcal{A}_G)$ of the cocycle $\mathcal{J}^\bullet(D)$ coincides with the Connes-Chern character $Ch^*(\mathcal{A}_G,\mathfrak{H},D)$.*

PROOF. The strategy for the proof remains the same as in [**7, 8**], and relies on employing the family $D_u = D|D|^{-u}$, $u\in[0,1]$ in order to construct a homotopy between the cocycle $\mathcal{T}^p(D)$ and the global cocycle τ_F^p. Using the fact that each D_u defines its own spectral triple twisted by scaling automorphisms (with character μ^{1-u}), and with similar analytic estimates and algebraic manipulations as above, one establishes the analogue of [**8**, Proposition 3] in the form

$$\mathcal{T}^p(D_0) - \mathcal{T}^p(D_1) \,=\, (b+B)\left(\int_0^1 \mathrm{Pf}_0 \int_\varepsilon^\infty \mathcal{Y}^\bullet(tD_u,D_u)\,dt\;du\right).$$

Since $D_1 = F$ and $F^2 = \mathrm{Id}$, the cyclic cocycle $\mathcal{T}^p(D_1)$ can be easily seen to coincide, up to the constant factor $\frac{\Gamma(\frac{p}{2}+1)}{2p!}$, with the very cocycle τ_F^p (*cf.* Eq. (1.9)) that defines the Connes-Chern character. □

4.5. The non-invertible case. As noted after Definition 4.1, the passage from $(\mathcal{A},\mathfrak{H},D)$ to the invertible double $(\widetilde{\mathcal{A}},\widetilde{\mathfrak{H}},\widetilde{D})$ necessitates the replacement of the exact similarity condition (4.1) by the perturbed version (4.2). This does affect the twisted commutators, but only up to higher order in the asymptotic expansion. More precisely,

$$[\varepsilon^{\frac12}\widetilde{D},\tilde{a}\,U^*]_\sigma \,=\, [\varepsilon^{\frac12}\widetilde{D},\tilde{a}]\,U^* \,+\, \varepsilon^{\frac12}\,a\,U^*\hat{\otimes}e_1F_1. \tag{4.43}$$

At the same time,

$$\widetilde{D}^2 = (D^2 + \mathrm{Id})\hat{\otimes}\mathrm{Id}\,, \qquad \text{hence} \qquad U\,\widetilde{D}^2\,U^* = \big(\mu(U)^2\,D^2 + \mathrm{Id}\big)\hat{\otimes}\mathrm{Id}\,.$$

Retracing the arguments leading to the expansion Eq. (4.17), one sees that the constant term remains unaffected by the perturbation.

Alternatively, one could proceed as in [**16**, §6.1], and add a compact operator 'mass' to the Dirac Hamiltonian. Specifically, in the construction of the invertible double, one takes

$$\widetilde{D} = D\hat{\otimes}\mathrm{Id} + K\hat{\otimes}F_1\,,$$

where $K \in \mathrm{OP}^{-\infty}$ is a smoothing operator such that

$$[K, D] = [D, K] \quad \text{and} \quad D^2 + K^2 \quad \text{is invertible.}$$

The similarity condition is again perturbed, this time by a smoothing operator. Since the residue integral factors through the complete symbols, the constant term remains of the same form as in Proposition 4.7, only with D replaced by $\widetilde{D}$.

4.6. Application to foliations with transverse similarity structure. We conclude by briefly indicating how one can use the above result in order to compute the index pairing for the leaf space of a foliation with transverse similarity structure.

A codimension n foliation $\mathcal{F}$ of an N-dimensional manifold V is said to have a *transverse similarity structure* if there exist an open cover $\{U_i\}_{i\in I}$ of V and a family $\{h_i : U_i \to \mathbb{R}^n\}_{i\in I}$ of submersions such that $\mathcal{F}|U_i = \{h_i^{-1}(y);\, y \in h_i(U_i)\}$ and the covering transformations $g_{ji} : h_i|U_i \cap U_j \to h_j|U_j \cap U_j$ are given by similarities in $\mathrm{Sim}(n)$. Concrete examples of such foliations can be found in [**20**], where for the case $n = N-1$ all nonsingular flows which admit a closed transversal (satisfying an additional property) are in fact classified. When $N = 3$ the notion of a transverse similarity structure to a nonsingular flow coincides with that of a *complex affine structure*, treated in [**13**] without the requirement for the existence of a closed transversal.

Given a foliation $\mathcal{F}$ with a transverse similarity structure, let $\mathcal{G}$ denote the smooth étale groupoid associated to a complete transversal M and let $\mathcal{A}_{\mathcal{G}} = C_c^\infty(\mathcal{G})$ (see [**4**, II, §§8-10]). The Dirac operator D on M defines a spectral triple twisted by similarities over the algebra $\mathcal{A}_{\mathcal{G}}$, whose Connes-Chern character is given by the cocycle $\mathcal{J}^\bullet(D) \in CC^\bullet(\mathcal{A}_{\mathcal{G}})$. On the other hand, let P be a proper $\mathcal{G}$-manifold [**4**, II, §10] with compact quotient $P/\mathcal{G}$, and let $\mathcal{D}$ be a $\mathcal{G}$-invariant elliptic differential operator on P. By a construction explained in [**6**, §5] for discrete groups and in [**4**, III, 7.γ] for étale groupoids, one associates to $\mathcal{D}$ a well-defined K-theory class

$$\mathrm{Ind}(\mathcal{D}) \in K_*(\mathcal{A}_{\mathcal{G}} \otimes \mathcal{R})\,,$$

where $\mathcal{R}$ is the algebra of infinite matrices with rapidly decaying entries. In the even case, this class can be represented by a difference idempotent

$$E_{\mathcal{D}} - E_0 \in M_k(\mathcal{A}_{\mathcal{G}} \otimes \mathcal{R})\,.$$

The index pairing between the K-homology class of D and the K-theory class of $\mathrm{Ind}(\mathcal{D})$ is then computed by the pairing of their explicitly expressed Chern characters:

$$\langle D, \mathrm{Ind}(\mathcal{D})\rangle = \langle \mathcal{J}^\bullet(D), ch_\bullet(E_{\mathcal{D}} - E_0)\rangle\,.$$

References

[1] N. Berline, E. Getzler, M. Vergne, **Heat kernels and Dirac operators**, Grundlehren der Mathematischen Wissenschaften, 298. Springer-Verlag, Berlin, 1992.

[2] J.-P. Bourguignon, P. Goduchon, Spineurs, opérateurs de Dirac et variations, *Commun. Math. Phys.* **144** (1992), 581-599.

[3] A. Connes, Noncommutative differential geometry, *Inst. Hautes Etudes Sci. Publ. Math.* **62** (1985), 257-360.

[4] A. Connes, **Noncommutative geometry**, Academic Press, 1994. ftp://ftp.alainconnes.org/book94bigpdf.pdf

[5] A. Connes, H. Moscovici, Transgression du caractère de Chern et cohomologie cyclique, *C.R. Acad. Sci. Paris* **303** (1986), 913-918.

[6] A. Connes, H. Moscovici, Cyclic cohomology, the Novikov conjecture and hyperbolic groups, *Topology* **29** (1990), 345–388.

[7] A. Connes, H. Moscovici, Transgression of the Chern character of finite-dimensional K-cycles, *Commun. Math. Phys.* **155** (1993), 103-122.

[8] A. Connes, H. Moscovici, The local index formula in noncommutative geometry, *Geom. Funct. Anal.* **5** (1995), 174-243.

[9] A. Connes, H. Moscovici, Hopf algebras, cyclic cohomology and the transverse index theorem, *Commun. Math. Phys.* **198** (1998), 199-246.

[10] A. Connes, H. Moscovici, Type III and spectral triples, **Traces in Geometry, Number Theory and Quantum Fields**, Aspects of Mathematics **E 38**, Vieweg Verlag 2008, pp. 57–71. http://www.arxiv.org/abs/math.OA/0609703

[11] S. Dave, An equivariant noncommutative residue, math.AP/061037. http://www.arxiv.org/abs/math.AP/0610371

[12] J. Ferrand, The action of conformal transformations on a Riemannian manifold, *Math. Ann.* **304** (1996), 277-291.

[13] E. Ghys, Flots transversalement affines et tissus feuilletés, *Mém. Soc. Math. France* **46** (1991), 123-150.

[14] J. M. Gracia-Bondia, J. C. Várilly, H. Figueroa, **Elements of Noncommutative Geometry**, *Birkhäuser Advanced Texts*, Birkhäuser Boston Inc., Boston,MA, 2001.

[15] E. Getzler, A. Szenes, On the Chern character of a theta-summable Fredholm module, *J. Func. Anal.* **84** (1989), 343-357.

[16] N. Higson, The residue index theorem of Connes and Moscovici. *Surveys in noncommutative geometry*, 71–126, **Clay Math. Proc., 6**, Amer. Math. Soc., Providence, RI, 2006.

[17] L. Hörmander, **The analysis of linear partial differential operators. I**, *Distribution theory and Fourier analysis.* Classics in Mathematics, Springer-Verlag, Berlin, 1983.

[18] A. Jaffe, A. Lesniewski, K. Osterwalder, Quantum K-theory: I. The Chern character, *Commun. Math. Phys.* **118** (1988), 1-14.

[19] R. Kulkarni, Conjugacy classes in $M(n)$, in: **Conformal Geometry**, R. Kulkarni and U. Pinkall (Eds.), Aspects Math. E 12, Vieweg, p. 41-64.

[20] T. Nishimori, A note on the classification of nonsingular flows with transverse similarity structures, *Hokkaido Math. J.* **21** (1992), 381-393.

Department of mathematics, The Ohio State University, Columbus, OH 43210, USA
E-mail address: henri@math.ohio-state.edu

Clay Mathematics Proceedings
Volume 11, 2010

Observables in Quantum Gravity

Alejandro Perez and Carlo Rovelli

Abstract. We study a family of physical observable quantities in quantum gravity. We call them W functions, or n-net functions. They represent transition amplitudes between quantum states of the geometry, are analogous to the n-point functions in quantum field theory, but depend on spin networks with n connected components. In particular, they include the three-geometry to three-geometry transition amplitude. The W functions are scalar under four-dimensional diffeomorphisms, and fully gauge invariant. They capture the physical content of the quantum gravitational theory. In particular, they can be used to compute scattering amplitudes between particle-like field quanta.

We show that W functions are the natural n-point functions of the field theoretical formulation of the gravitational spinfoam models. They can be computed from a perturbation expansion, which can be interpreted as a sum-over-four-geometries. Therefore the W functions bridge between the canonical (loop) and the covariant (spinfoam) formulations of quantum gravity. Following Wightman, the physical Hilbert space of the theory can be reconstructed from the W functions, if a suitable positivity condition is satisfied.

We compute explicitly the W functions in a "free" model in which the interaction giving the gravitational vertex is shut off, and we show that, in this simple case, we have positivity, the physical Hilbert space of the theory can be constructed explicitly and the theory admits a well-defined interpretation in terms of diffeomorphism invariant transition amplitudes between quantized geometries.

Preface, by C.R.

I met Alain in Cambridge, at the Newton Institute, in July 94. I was immediately deeply impressed by his astonishing intelligence and by the contagious force of his burning and almost childlike intellectual passion. I had the privilege of finding right away a common ground in our interests, and I found that Alain had much to teach me even in the subjects that I thought were more specific to my own work. Alain is a deep thinker, with an immense and youthful courage, and the rare capacity of finding genuine new ways for thinking about reality. In my opinion, he is among

2010 *Mathematics Subject Classification.* Primary: 83C45. Secondary: 81T05, 83C27.

Unité mixte de recherche (UMR 6207) du CNRS et des Universités de Provence (Aix-Marseille I), de la Méditerranée (Aix-Marseille II) et du Sud (Toulon-Var); laboratoire affilié à la FRUMAM (FR 2291).

the extremely rare thinkers that are today truly opening new paths for all of us. I admire him profoundly, as a scientist and as a human being.

The article that follows does not concern the topic on which Alain and I have collaborated (which is the nature of physical time [**1**]). But it concerns a problem on which Alain has long reflected: what are the natural observable quantities in a theory where space and time are not primary notions? The article reports work by Alejandro Perez and myself, which we did some time ago, but was not published. This work is at the basis of numerous subsequent developments in quantum gravity, and in particular of most of the recent calculations of graviton scattering amplitudes in the context of loop quantum gravity [**2**]. In particular, the relation between geometry transition amplitudes (and W functions) and particle scattering amplitudes was clarified in [**3**]. It is a pleasure to dedicate this work to Alain, as a sign of my admiration, and my pride and gratitude for his friendship.

1. Introduction

One of the hard problems in non-perturbative quantum gravity [**5**] is to construct a full set of physically meaningful observable quantities [**6**]. In this paper, we point out that there is a natural set of quantities that one can define in quantum general relativity, which are gauge invariant, have a natural physical interpretation, and could play the role played by the n-point functions in quantum field theory. We denote these quantities as W functions, or n-net functions. As the n-point functions in a quantum field theory, these quantities are not natural quantities of the corresponding classical field theory, namely in general relativity. Nevertheless, they capture the physical content of the quantum theory and are related to the classical theory.

The W functions are closely related to the three-geometry to three-geometry transition amplitude studied by Hawking [**7**]. However, they are not transition amplitudes between states in which the classical three-geometry has an arbitrary sharp value, but rather transition amplitudes between *eigenstates* of the three-geometry. In loop quantum gravity [**8, 9**], these eigenstates are characterized by discretized geometries and are labelled by abstract spin networks [**10, 11**], or s-knots. Thus the W functions are rather transition amplitudes between states with fixed amounts of "quanta of geometry". This is analogous to the n-point functions in field theory, which are not transition amplitudes between field configurations, but rather transition amplitudes between states characterized by a fixed number of "quanta of field" – that is, particles. Furthermore, the W functions generalize the three-geometry to three-geometry amplitude (a 2-point function) to arbitrary n-point functions; more precisely, we define the W functions as a functional $W(s)$ over an *algebra* $\mathcal{A}$ of abstract (not necessarily connected) spin networks. In this respect, the W functions are analogous to the Wightman distributions [**12**] (hence the choice of the letter W).

We start from a general definition of the W functions, based on canonical quantum gravity. We show that the W functions are well defined diffeomorphism invariant observable quantities and we clarify their physical interpretation. In this paper we focus on the case in which the dynamics is "real", in a sense defined below. The physical meaning of this reality and the extension of the formalism to the general case are discussed at the end of the paper.

A crucial property of the Wightman functions is the possibility of reconstructing the quantum field theory from them – a subtle application of the beautiful Gelfand-Naimark-Segal (GNS) representation theorem in the theory of C^*-algebras. We show here that the W functions have the same property: under appropriate conditions –in particular, a positivity condition– the physical Hilbert space of the theory and a suitable operator algebra can be reconstructed from the W functions, using the GNS construction. In other words, we explore the extension to the generally covariant context of Wightman's remarkable intuition that the content of a quantum field theory is encoded in its n-point functions. To this end, we need to strip Wightman's theory from all the "details" that follow from the existence of a background Minkowski space (positivity of the energy, uniqueness of the vacuum, microlocality. . .) and show that the core idea remains valid even in the absence of a background spacetime. For a line of investigation similar in spirit, see [**13**].

Since the diffeomorphism invariant quantum field theory can be characterized by its W functions, the way is open for defining a quantum theory of gravity by directly constructing its W functions. Remarkably, spinfoam models [**14, 15**] provide a natural perturbative definition of W functions. In particular, it has been recently shown that general spinfoam models can be obtained as the Feynman expansion of certain peculiar field theories over a group [**16, 17, 18, 19**]. We show here that the gauge invariant n-point functions of these field theories are precisely W functions. This construction provides a direct link between the field theoretical formulation of the spinfoam models and canonical quantum gravity. The link is similar in spirit to the link between the operator definition of quantum field theory and the construction of its n-point functions via a functional integral [**12**]. In particular, given the field theoretical formulation of a spinfoam model, we can construct a quantum gravity physical Hilbert space from its W functions. On the one hand, loop quantum gravity provides the general framework and, in particular, the physical interpretation of the W functions; on the other hand, the field theory over the group provides an indirect but complete definition of the dynamics. This is especially interesting in light, in particular, of the construction of *Lorentzian* spinfoam models [**20, 21**]. In turn, the perturbative expansion of the field theory defines a sum over spinfoams which can be directly interpreted as a sum over the 4-geometries formulation of quantum gravity. Some of the the ideas presented here are independently derived in [**22**].

We give an example of reconstruction in Section 5. We consider a simple "free" model, obtained by dropping the interaction term which gives the quantum gravity vertex. We prove positivity for this case, and thus the existence of a Hilbert space of spin networks for this quantum theory. Finally, in Section 6 we discuss the meaning of the reality assumption and the extension to the complex case.

All together, we obtain an attractive unified picture, in which canonical loop quantum gravity, covariant spinfoam models, and a family of diffeomorphism-invariant physical observables for quantum gravity fit into a unified scheme.

The ideas described in this paper were first presented in the second conference on Quantum Gravity in Warsaw, in June 1999.

2. The 2-net function $W(s, s')$ in canonical quantum gravity

In loop quantum gravity [**8, 9**], the Hilbert space $\mathcal{H}_{diff}$ of the states invariant under three-dimensional diffeomorphisms admits a discrete [**26**] basis of states $|s\rangle$, labelled by abstract spin networks (or s-knots) s. Let us define this space precisely. An abstract spin network s is an abstract graph (not necessarily connected) Γ_s, with links labelled by (nontrivial) $SU(2)$ representations, and nodes labelled by $SU(2)$ intertwiners [**25, 26**]. If the graphs of s and s' are different, the two spin network states $|s\rangle$ and $|s'\rangle$ are orthogonal; if the graphs are the same, the scalar product $\langle s|s'\rangle$ is the same as that given by the spin network basis of an $SU(2)$ lattice gauge theory [**9**].

For what follows, it is important to notice that there is a natural union operation $\cup$ defined on the abstract spin networks: $s \cup s'$ is the spin network defined by the graph formed by the two disconnected components Γ_s and $\Gamma_{s'}$.

The *extended* Hilbert space $\mathcal{H}_{ex}$ formed by the unconstrained states is spanned by a basis of *embedded* spin networks [**9**]. The relation between $\mathcal{H}_{diff}$ and the extended Hilbert space $\mathcal{H}_{ex}$ can be formally expressed in terms of a "projection" operator $P_{diff} : \mathcal{H}_{ex} \to \mathcal{H}_{diff}$ which sends an embedded spin network state of a suitable basis of $\mathcal{H}_{ex}$ into an abstract spin network state, or s-knot, in $\mathcal{H}_{diff}$ [**15**]. Due to the infinite volume of the group of diffeomorphisms (or to the fact that the zero eigenvalue of the diffeomorphism constraint is in the continuous spectrum) $\mathcal{H}_{diff}$ is not a proper subspace of $\mathcal{H}_{ex}$ and P_{diff} is not a true projection operator (hence the quotation marks), but several techniques for taking care of these technicalities are known, and the space $\mathcal{H}_{diff}$ and the operator P_{diff} are well defined.

The states $|s\rangle$ have a straightforward physical interpretation, which follows from the fact that they are projections on $\mathcal{H}_{diff}$ of eigenstates of the area and volume operators [**10, 11**]. The interpretation is the following. A state $|s\rangle$ represents a three-geometry. A three-geometry is an equivalence class of three-metrics under diffeomorphisms. The geometry represented by $|s\rangle$ is quantized, in the sense that it is formed by regions and surfaces having quantized values of volume and area. Intuitively, each node of s represents a "chunk" of space, whose (quantized) volume is determined by the intertwiner associated to the node. Two such chunks of space are adjacent if there is a link between the corresponding nodes. The (quantized) area of the surface that separates them is determined by the representation j associated to this link, according to the now well known relation [**11**]

$$A = 8\pi\gamma\hbar G\ \sqrt{j(j+1)} \tag{1}$$

where $\hbar, G, \gamma$ are the reduced Planck constant, the Newton constant and the Immirzi parameter (the dimensionless free parameter in the theory). This interpretation of the states $|s\rangle$ follows from the study of the area and volume operators on the Hilbert space of non-diffeomorphism invariant states. Notice that the states $|s\rangle$ are not gauge invariant either, and do not represent physical gauge invariant notions. The same is true for the corresponding classical notion of three-geometry: a three-geometry is determined by an ADM surface, which is a non-gauge-invariant notion in general relativity.

The dynamics of the theory is given by the Hamiltonian constraint $H(x)$, which is the operator that encodes the dynamics of quantum general relativity [**9**]. We assume here $H(x)$ to be a symmetric operator. The space of solutions of this constraint is the physical Hilbert space of the theory H_{Ph}. Instead of using the Hamiltonian constraint, we can work with the linear operator $P : H_{diff} \to H_{Ph}$ that projects onto the kernel of $H(x)$. (A suitable extension of H_{diff} to its generalized states –or any other of the many techniques developed for this purpose– should be used in order to take care of the technical complications in defining the Hilbert eigenspace corresponding to an eigenvalue in the continuous spectrum.) For more details on this operator, and, in particular, a more precise definition as a three-dimensional diffeomorphism invariant object, see [**15**]. Instead of worrying about the explicit construction of P, we assume here that the operator $P : H_{diff} \to H_{Ph}$ is given, and we consider the quantity

$$W(s, s') := \langle s|P|s'\rangle. \tag{2}$$

Our key observation is that this is a well-defined fully gauge invariant quantity, which represents a physical observable in quantum gravity and has a precise and well-understood physical interpretation.

The gauge invariance of $W(s, s')$ is immediate. All the objects on the r.h.s. of (2) are invariant under three-dimensional diffeomorphisms, therefore we only need to check invariance under time reparametrizations. An infinitesimal coordinate-time shift is generated by the Hamiltonian constraint. If we gauge transform (say) the bra state $\langle s|$ we obtain

$$\delta W(s, s') = \langle Hs|P|s'\rangle = \langle s|HP|s'\rangle = 0, \tag{3}$$

because P is precisely the projection on the kernel of H. Therefore $W(s, s')$ represents a gauge-invariant transition amplitude. In fact, this is precisely the physical three-geometry to three-geometry transition amplitude.

To clarify why the three-geometry to three-geometry transition amplitude is a physical gauge-invariant quantity, consider a simple analogy with a well known system. Consider a free relativistic particle in three spatial dimensions. Its physical description is given by its position $\vec{x}(t)$ at each time t. To have explicit Lorentz invariance in the formalism, the dynamics can be represented as a constrained reparametrization invariant dynamical system, by promoting the time variable t to the role of dynamical variable $x^0 = t$, and introducing an unphysical parameter "time" τ. The dynamics is then entirely determined by the constraints $p^2 - m^2 = 0$ and $p^0 > 0$. The corresponding constraints in the quantum theory are the Klein-Gordon equation and the restriction to its positive frequency solutions. The Hilbert space $\mathcal{H}_{ex}$ of the unconstrained states is formed by the square integrable functions on Minkowski space. The physical Hilbert space $\mathcal{H}_{Ph}$ of the physical states is formed by the positive frequency solutions of the Klein-Gordon equation. There is a well defined projection operator P, which restricts any state in $\mathcal{H}_{Ph}$ (more precisely, in the extension of $\mathcal{H}_{Ph}$ which includes its generalized states) to its mass shell, positive frequency, component. Now, consider the (generalized) state $|\vec{x}, x^0\rangle$ in $\mathcal{H}$. This is the eigenstate of both the position $\vec{x}$ and the time x^0 operators, which are well defined self-adjoint operators on $\mathcal{H}$. The interpretation of $|\vec{x}, x^0\rangle$ is clear: it is a particle at the Minkowski spacetime point $(\vec{x}, x^0)$. On the other hand, this is clearly *not* a physical state: there is no physical particle that can "stay" in a single point of spacetime (where is it after a second?). It is a state that does not satisfy the dynamics. Notice also that in $\mathcal{H}$ two such states at two different points of Minkowski space are orthogonal. However, given the state $|\vec{x}, x^0\rangle$ in $\mathcal{H}$, we can project it down to H_{Ph} and define the *physical* state

$$|\vec{x}, x^0\rangle_{Ph} = P\,|\vec{x}, x^0\rangle. \tag{4}$$

In momentum space, this amounts to restricting it to its mass shell positive frequency components. In coordinate space, this amounts to spreading out the delta function to a full solution of the Klein-Gordon equation, which –as its happens– at time x^0 is concentrated around $\vec{x}$, but at other times is spread around the future and past light cones of $(\vec{x}, x^0)$. The state $|\vec{x}, x^0\rangle_{Ph}$ is a physical state, and has a physical interpretation consistent with the dynamics: it is a (Heisenberg) state in which the particle is in $\vec{x}$ at time x^0, and has appropriately moved around in space at other times. The transition amplitude between two such states is a physically meaningful quantity. Indeed, it is nothing else than the familiar propagator in Minkowski space. But notice that

$$W(\vec{x}, x^0; \vec{x}', x^{0\prime}) = {}_{Ph}\langle\vec{x}, x^0|\vec{x}', x^{0\prime}\rangle_{Ph} = \langle\vec{x}, x^0|P|\vec{x}', x^{0\prime}\rangle, \tag{5}$$

Namely, the propagator is nothing but the matrix element of the projection operator P between the *unphysical* states $|\vec{x}, x^0\rangle$!

It is clear that the structure illustrated is the precisely the same as in quantum gravity. A classical three-geometry is determined by three degrees of freedom per space point. Two of these correspond to physical degrees of freedom of the gravitational field, in analogy with the dependent variable $\vec{x}$ above. The third is the independent temporal variable, in analogy with the x^0 variable in the example above.[1] Therefore s, precisely as $(\vec{x}, x^0)$, includes the dependent as well as the independent (time) variables. The states $|s\rangle$ are quantum states concentrated at a single three-geometry. Precisely as the states $|\vec{x}, x^0\rangle$, these are unphysical, because spacetime cannot be concentrated on a unique three-geometry, in the very same sense in which a particle cannot be at a unique point of Minkowski space. The projection P projects a state $|s\rangle$ into a physical state which spreads across three-geometries, and the transition amplitude (2) gives the amplitude of measuring the three-geometry corresponding to s after we have measured the three geometry corresponding to s'. This amplitude is well defined and diffeomorphism invariant.

3. Reality of P and W functions

Let us now return to the gravitational theory. We assume in this section that P has the following property, which we call (for reason that will become clear later on) "reality"

$$\langle s_1 \cup s_3|P|s_2\rangle = \langle s_1|P|s_2 \cup s_3\rangle. \tag{6}$$

The physical meaning of this property, as well as the extension of the formalism to the case in which this property does not hold, is discussed in Section 6.

Consider the vector in $\mathcal{H}_{Ph}$

$$|0\rangle_{Ph} \equiv P|0\rangle. \tag{7}$$

and, in general,

$$|s\rangle_{Ph} \equiv P|s\rangle. \tag{8}$$

(See the particle analogy discussed at the end of last section.) The 2-net function $W(s, s')$, defined in (2), can then be written also as

$$W(s, s') = {}_{Ph}\langle s|s'\rangle_{Ph}. \tag{9}$$

[1] Of course, there is no a priori physical distinction between the two sets. This is because the dynamics of general relativity is relational: it provides relations between quantities which are on an equal footing, not a preferred temporal variable. The advantage of the formalism we are considering here is that it does not require such a distinction to be made. It does not require one to single out a preferred time variable.

Clearly the states $|s\rangle_{Ph}$ form an overcomplete basis of $\mathcal{H}_{Ph}$. In particular, there will be relations between them, of the form

$$\sum_s c_s|s\rangle_{Ph} = 0 \tag{10}$$

for appropriate complex numbers c_s. Notice that (10) is equivalent to $P\sum_s c_s|s\rangle = 0$, or $\langle s'P\sum_s c_s|s\rangle = 0, \forall s'$. This can also be rewritten as $\langle s'\cup s''P\sum_s c_s|s\rangle = 0, \forall s'$ and, because of the reality (6) of P, as $\langle s'P\sum_s c_s|s\cup s"\rangle = 0, \forall s'$. Therefore

$$\sum_s c_s|s\cup s''\rangle_{Ph} = 0 \tag{11}$$

for all s'', whenever (10) holds. Using this fact, we define on $\mathcal{H}_{Ph}$ the operator

$$\hat{\phi}_s|s'\rangle_{Ph} = |s'\cup s\rangle_{Ph}. \tag{12}$$

This definition is well posed, in spite of the overcompleteness of the vectors $|s\rangle_{Ph}$, because of (11), that is, ϕ_s sends the vanishing linear combinations (10) of states into the linear combinations (11) which are still vanishing. Also, notice that $\hat{\phi}_s$ is self-adjoint, again because of the reality of P,

$$\begin{aligned} {}_{Ph}\langle s_1|\phi_s^\dagger|s_2\rangle_{Ph} &= {}_{Ph}\langle \phi_s s_1|s_2\rangle_{Ph} = {}_{Ph}\langle s_1\cup s|s_2\rangle_{Ph} = \langle s_1\cup s|P|s_2\rangle \\ &= \langle s_1|P|s_2\cup s\rangle \ = {}_{Ph}\langle s_1|s_2\cup s\rangle_{Ph} = {}_{Ph}\langle s_1|\phi_s|s_2\rangle_{Ph}, \end{aligned} \tag{13}$$

and it commutes with itself,

$$[\hat{\phi}_s, \hat{\phi}_{s'}] = 0, \tag{14}$$

since

$$\hat{\phi}_s\hat{\phi}_{s'} = \hat{\phi}_{s'}\hat{\phi}_s = \hat{\phi}_{s\cup s'}. \tag{15}$$

The 2-net function $W(s,s')$, defined in (2), can now be written as

$$W(s,s') = {}_{Ph}\langle 0|\hat{\phi}_s\hat{\phi}_{s'}|0\rangle_{Ph}. \tag{16}$$

More generally, we can define

$$W(s) = {}_{Ph}\langle 0|\hat{\phi}_s|0\rangle_{Ph}. \tag{17}$$

so that

$$W(s,s') = W(s\cup s'). \tag{18}$$

Now, consider the linear space $\mathcal{A}$ formed by the (formal) linear combinations of spin networks, with complex coefficients

$$A = \sum_s c_s s. \tag{19}$$

There is a natural product defined on $\mathcal{A}$ by $s\cdot s' = s\cup s'$, and a natural star operation defined by $s^* = s$ (Here we refer to spin networks labeled by $SU(2)$ representations and each representation of $SU(2)$ is conjugate to itself. When spin networks are labeled by representations of groups which are not self-conjugate the star operation should replace representations with dual representations.) We define the norm $||A|| = \sup_s|c_s|$. We obtain in this way a C^*-algebra structure on $\mathcal{A}$. The quantity $W(s)$, defined in (17), defines a linear functional on $\mathcal{A}$. A straightforward calculation shows that the functional is positive

$$W(A^*A) \geq 0. \tag{20}$$

We can thus apply the Gelfand-Naimark-Segal construction to the C^*-algebra $\mathcal{A}$ and the positive linear functional W, obtaining a Hilbert space $\mathcal{H}$, a "vacuum" state $|0)$ and a representation ϕ of $\mathcal{A}$ in the Hilbert space, such that

$$W(s) = (0|\hat{\phi}(s)|0). \tag{21}$$

But it is clear that in doing so we have simply reconstructed the Hilbert space $\mathcal{H}_{Ph}$, the "vacuum" state $|0) = |0\rangle_{Ph}$ and the algebra of the operators $\hat{\phi}(s) = \hat{\phi}_s$. In other words, the content of the canonical theory of quantum gravity can be coded, in the spirit of Wightman, in the positive linear functional $W(s)$ over the algebra $\mathcal{A}$ of spin networks.

We can thus determine the dynamics of the theory by giving $W(s)$, instead of explicitly giving the projection P, or the Hamiltonian constraint, and reconstruct the physical Hilbert space from $W(s)$. In particular, the main physical gauge-invariant observable, namely the three-geometry to three-geometry transition amplitude, is simply the value of $W(s)$ on the spin networks s formed by two disjoint components.

We close this section with a comment about locality. The sense in which general relativity is a local theory is far more subtle than in ordinary field theory. For a detailed discussion of this issue see for instance [**6**]. In particular, physical gauge invariant observables are independent of the spacetime coordinates $\vec{x}, t$, and therefore they are not localized on the spacetime manifold, which is coordinatized by $\vec{x}, t$. Nevertheless, the dynamics of general relativity is still local in an appropriate sense. This locality should be reflected in a general property of the W functions. Roughly, we expect that if a spin network s can be cut into two parts (connected to each other) s_{ext} and s_{in}, and a second spin network s' can be cut into two parts (connected to each other) s'_{ext} and s'_{in}, and if $s_{ext} = s'_{ext}$, then $W(s, s')$ should be independent of s_{ext}. In other words, the local evolution in a part of the spin network should be independent of what happens elsewhere on the spin network. A precise formulation of this property and its consequences deserve to be studied.

4. $W(s)$ in field theories over a group

In the last few years, intriguing developments in quantum gravity have been obtained using the spinfoam [**14**] formalism. Recently, it has been shown that any spinfoam model can be derived from an auxiliary field theory over a group manifold [**16, 17**]. Several spinfoam models defined from auxiliary theories defined over a group have been developed. They are covariant, have remarkable finiteness properties [**18**], exist in Lorentzian form [**21**] and represent intriguing covariant models for a quantum theory of the gravitational field. In this section, we illustrate the emergence of a $W(s)$ functional over $\mathcal{A}$ in the context of these field theories over a group manifold. For a similar derivation see [**22**].

For concreteness, and simplicity of the presentation, let us consider a specific model. Consider a real field theory for a scalar field defined over a group manifold $\phi(g^i) = \phi(g^1, g^2, g^3, g^4)$, where $g^i \in G$ (that is $\phi : G^4 \to \mathbb{C}$), which we choose for the moment to be a compact Lie group [**16, 18**]. The field $\phi(g^1, g^2, g^3, g^4)$ is defined to be symmetric under permutation of its four arguments, G-invariant in the sense that it satisfies $P_g\phi = \phi$, where the operator P_g is defined by

$$P_g\phi(g^1, g^2, g^3, g^4) \equiv \int_G dg\ \phi(g^1g, g^2g, g^3g, g^4g), \tag{22}$$

and finally to be zero-mode free, namely

$$\int_G dg\ \phi(g^1 g, g^2, g^3, g^4) = 0. \tag{23}$$

This last property is necessary to make the relationship between spin network states and field operator products (described below) a one-to-one relationship. The dynamics is given by an action $S[\phi]$, which we do not specify for the moment. The n-point functions of the theory are functions of n four-tuples of group elements; they have the form

$$W(g_1^{i_1}, \ldots, g_n^{i_n}) = \int [D\phi]\ \phi(g_1^{i_1}) \cdots \phi(g_n^{i_n})\ e^{iS[\phi]}. \tag{24}$$

Let us work in momentum space. Using the Peter-Weyl theorem, we expand the field in terms of the matrix elements of the irreducible representations $D^{(N)}_{\alpha\beta}$ of G.

$$\phi(g_1, \ldots, g_4) = \sum_{N_1 \ldots N_4} \Phi^{\alpha_1 \ldots \alpha_4}_{(N_1 \ldots N_4)\beta_1 \ldots \beta_4}\ D^{(N_1)\beta_1}_{\alpha_1}(g_1) \cdots D^{(N_4)\beta_4}_{\alpha_4}(g_4). \tag{25}$$

We denote by $C^{N_1 \ldots N_4\,\Lambda}_{\alpha_1 \ldots \alpha_4}$ a normalized basis in the space of the intertwiners between the representations $N_1, \ldots, N_4$. Imposing the G-invariance of the field on the momentum space components, and using the relation

$$\int_G dg\ D^{(N_1)}_{\alpha_1\beta_1}(g) \cdots D^{(N_4)}_{\alpha_4\beta_4}(g) = \sum_\Lambda\ C^{N_1 \ldots N_4\,\Lambda}_{\alpha_1 \ldots \alpha_4}\ \ C^{N_1 \ldots N_4\,\Lambda}_{\beta_1 \ldots \beta_4} \tag{26}$$

we can write the field as

$$\phi(g_1, \ldots, g_4) = \sum_{N_1, \ldots, N_4} \Phi^{\alpha_1 \ldots \alpha_4}_{(N_1 \ldots N_4)\beta_1 \ldots \beta_4}\ D^{(N_1)\gamma_1}_{\alpha_1}(g_1) \cdots D^{(N_4)\gamma_4}_{\alpha_4}(g_4) \sum_\Lambda C^{N_1 \ldots N_4, \Lambda}_{\gamma_1 \ldots \gamma_4} C^{\beta_1 \ldots \beta_4}_{N_1 \ldots N_4, \Lambda}, \tag{27}$$

or, defining (for later convenience)

$$\phi^{\alpha_1 \ldots \alpha_4}_{N_1 \ldots N_4, \Lambda} := \frac{\Phi^{\alpha_1 \ldots \alpha_4}_{(N_1 \ldots N_4)\beta_1 \ldots \beta_4} C^{N_1 \ldots N_4\,\Lambda}_{\beta_1 \ldots \beta_4}}{\Delta_{N_1}\Delta_{N_2}\Delta_{N_3}\Delta_{N_4}}, \tag{28}$$

where Δ_N is the dimension of the representation N, by

$$\phi(g_1, \ldots, g_4) = \sum_{N_1, \ldots, N_4, \Lambda} \phi^{\alpha_1 \ldots \alpha_4}_{N_1 \ldots N_4, \Lambda} \left(\Delta_{N_1} \ldots \Delta_{N_4} D^{(N_1)\gamma_1}_{\alpha_1}(g_1) \cdots D^{(N_4)\gamma_4}_{\alpha_4}(g_4) C^{N_1 \ldots N_4, \Lambda}_{\gamma_1 \ldots \gamma_4} \right). \tag{29}$$

We can take the quantities $\phi^{\alpha_1 \ldots \alpha_4}_{N_1 \ldots N_4, \Lambda}$ as the independent "Fourier components" of the field, and therefore write the W functions, in momentum space, as

$$W^{\alpha_1^1\alpha_2^1\alpha_3^1\alpha_4^1}_{N_1^1N_2^1N_3^1N_4^1,\Lambda^1} \cdots {}^{\alpha_1^n\alpha_2^n\alpha_3^n\alpha_4^n}_{N_1^nN_2^nN_3^nN_4^n,\Lambda_n} = \int [D\phi]\ \phi^{\alpha_1^1\alpha_2^1\alpha_3^1\alpha_4^1}_{N_1^1N_2^1N_3^1N_4^1,\Lambda^1} \cdots \phi^{\alpha_1^n\alpha_2^n\alpha_3^n\alpha_4^n}_{N_1^nN_2^nN_3^nN_4^n,\Lambda_n}\ \ e^{iS[\phi]}. \tag{30}$$

However, the measure and the action are G-invariant. Therefore the only nontrivial independent W functions are given by G-invariant combinations of fields, where G acts on each index α_i^n by the representation N_i^n . There is only one way of obtaining G-singlets: to have the indices α_i^n all paired –with the two indices of the pair sitting in the same representation– and to sum over the paired indices. Each independent W function is determined by a choice of indices and their pairing.

In order to describe these index choices and pairings, let us associate to each field $\phi^{\alpha_1\alpha_2\alpha_3\alpha_4}_{N_1N_2N_3N_4,\Lambda}$ in the integrand a four-valent node; we associate to this node the intertwiner Λ_n, and to each of its four links a representation N_i^n. We then connect the links between two fields with paired indices. We obtain a graph, with nodes labelled by intertwiners and links labelled by representations (satisfying Clebsch-Gordan-like relations), namely a spin network s (in the group G). Thus, independent W functions are labelled by spin networks!

In other words, to each spin network s, with nodes n labelled by intertwiners Λ_n and links l labelled by representations N_l, we can associate a gauge invariant product of field operators ϕ_s

$$\phi_s = \sum_{\alpha_l} \prod_n \phi^{\alpha_1^n\alpha_2^n\alpha_3^n\alpha_4^n}_{N_1^nN_2^nN_3^nN_4^n,\Lambda_n} \tag{31}$$

where α_i^n is the index associated to the link l which is the i-th link of the node n. And we define

$$W(s) = \int [D\phi]\, \phi_s\, e^{iS[\phi]}. \tag{32}$$

Therefore the field theory over the group defines a W functional over the spin network algebra $\mathcal{A}$. If $W(s)$ is positive, we then have immediately, thanks to the GNS theorem, a Hilbert space and an algebra of field operators whose vacuum expectation value is $W(s)$. Under suitable conditions, this could be identified with the physical Hilbert space of quantum gravity. For this, the group G has to be $SU(2)$, or, alternatively, the representations and the intertwiners should be in correspondence with those of $SU(2)$. This is the case in particular for the gravitational $SO(4)$ and $SO(3,1)$ models [**20, 21**] in which the dynamics restricts the representations to the simple, or balanced, representations, which can be identified with the irreducible $SU(2)$ representations.

The relation between field theory and quantum gravity becomes much more transparent by expressing $W(s)$ explicitly as a perturbation expansion. Indeed, as shown in Ref. [**16, 17, 18**], the standard field theoretical perturbation expansion of $W(s)$ in Feynman graphs turns out to be a sum over spinfoams. In particular $W(s,s') = W(s \cup s')$ is given by a sum over all spinfoams σ bounded by the spin networks s and s'

$$W(s,s') = \sum_{\sigma,\ \partial\sigma = s\cup s'} A(\sigma) \tag{33}$$

where $A(\sigma)$ is a complex amplitude associated to the spinfoam σ (see Ref. [**16, 17, 18, 19**]). A spinfoam σ admits an interpretation as a (discretized) 4-geometry. In particular, σ can be the complex dual to a four-dimensional cellular complex, and the representations and intertwiners are naturally related to areas and volumes of the elementary 2- and 3-cells. Therefore (33) is a (precise) implementation of the representation of quantum gravity as a sum over geometries, introduced by Wheeler and Misner [**23**], and developed by Hawking and collaborators [**7**]. The generation of the spacetimes summed over as Feynman diagrams is a four-dimensional analog of the two-dimensional quantum gravity models developed some time ago in the context of string theory in zero dimensions [**28**].

5. A "free" theory

As a simple example, we sketch here the structure of a very simple model in which the action $S[\phi]$ contains only a kinetic part and no interaction part.

$$S[\phi] = (i/2) \int dg^1 \cdots dg^4 \ \phi^2(g^1, \cdots, g^4). \tag{34}$$

A straightforward calculation yields the action in momentum space,

$$S[\phi]=\frac{i}{2}\phi^{\alpha_1\ldots\alpha_4}_{N_1\ldots N_4,\Lambda}\phi^{\beta_1\ldots\beta_4}_{M_1\ldots M_4,\tilde{\Lambda}} \left((\Delta_{N_1}\cdots\Delta_{N_4})\delta_{\alpha_1\beta_1}\cdots\delta_{\alpha_4\beta_4}\delta^{N_1M_1}\cdots\delta^{N_4M_4}\delta^{\Lambda\tilde{\Lambda}}\right), \tag{35}$$

where sum over repeated indices is understood.

Every n-point function of the field theory can be calculated as functional derivatives of the generating function $\mathcal{W}(J)$ defined as

$$\mathcal{W}(J) = \int \mathcal{D}\phi \ \exp\left(iS[\phi] + J^{\alpha_1\ldots\alpha_4}_{N_1\ldots N_4,\Lambda}\phi^{N_1\ldots N_4,\Lambda}_{\alpha_1\ldots\alpha_4}\right), \tag{36}$$

which can be easily computed using standard Gaussian integration:

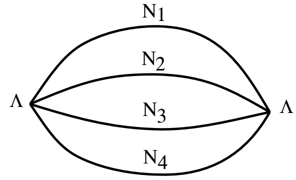
$$W(J) = \mathcal{C}\exp\left(\frac{1}{2}\frac{J^{\alpha_1\ldots\alpha_4}_{N_1\ldots N_4,\Lambda}\ J^{N_1\ldots N_4,\Lambda}_{\alpha_1\ldots\alpha_4}}{\Delta_{N_1}\cdots\Delta_{N_4}}\right), \tag{37}$$

The n-point function depends on the boundary 4-valent spin network defined by the action of the $J^{\alpha_1\ldots\alpha_4}_{N_1\ldots N_4,\Lambda}$'s, which can be represented by a 4-valent node carrying the representations N_1, N_2, N_3, N_4 respectively and an intertwiner colored by Λ. Contraction of their indices α_i is represented by the connection of the corresponding links. Let us illustrate with an example the computation of the Wightman functions for this free theory. The two-point function $W(s_1, s_2)$ is given by

$$W(s_1,s_2) = \left\{\frac{\delta}{\delta J^{\alpha_1\ldots\alpha_4}_{N_1\ldots N_4,\Lambda}}\frac{\delta}{J^{\alpha_1\ldots\alpha_4}_{N_1\ldots N_4,\Lambda}}\frac{\delta}{\delta J^{\beta_1\ldots\beta_4}_{M_1\ldots M_4,\Gamma}}\frac{\delta}{\delta J^{\beta_1\ldots\beta_4}_{M_1\ldots M_4,\Gamma}}W(J)\right\}_{J=0}, \tag{38}$$

where the boundary spin networks s_1, and s_2 are given by the corresponding contraction of the J's functional derivatives, namely:

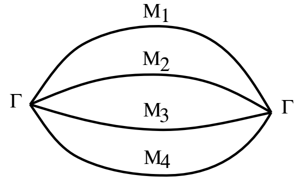
$$\frac{\delta}{\delta J^{\alpha_1\ldots\alpha_4}_{N_1\ldots N_4,\Lambda}}\frac{\delta}{J^{\alpha_1\ldots\alpha_4}_{N_1\ldots N_4,\Lambda}} \rightarrow \tag{39}$$



,

and

$$\frac{\delta}{\delta J^{\beta_1\ldots\beta_4}_{M_1\ldots M_4,\Gamma}}\frac{\delta}{\delta J^{\beta_1\ldots\beta_4}_{M_1\ldots M_4,\Gamma}} \rightarrow \tag{40}$$



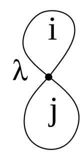
.

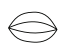
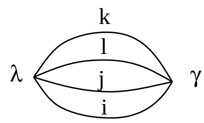
A straightforward calculation gives

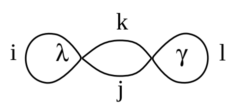
$$W(S_1, S_2) = 1 + \delta_{N_1M_1}\cdots\delta_{N_4M_4}\delta_{\Lambda\tilde{\Lambda}}. \tag{41}$$

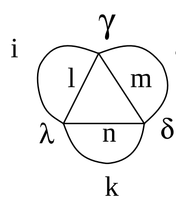
The C^*-algebra $\mathcal{A}$ is defined as the algebra generated by the free sum of (not necessarily connected) 4-valent spin networks over $SO(4)$ as in (19). The $*$ operation is simply defined by complex conjugation of the components of A. We define the functional W over the algebra by means of the corresponding Wightman functions of our spinfoam model. The fact that the positivity condition holds for W (namely, $W(A^*A) \geq 0$) can be easily seen from the form of the functional measure. We can explicitly construct an orthonormal basis in $\mathcal{H}_{Ph}$ as follows. There are two kind of situations: spin networks which do not interfere with the vacuum $|0\rangle$ (the empty spin network), and those which do. In the first case the projection is trivial and the elements of the physical Hilbert space $\mathcal{H}_{Ph}$ are the simply the original spin network states. Some examples are the following:

$$|0\rangle = 1,$$

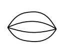
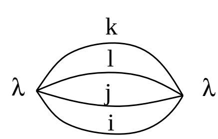
$|ij, \lambda\rangle$  $=$  ,

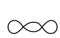
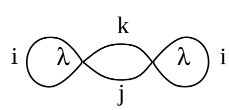
$|ijkl, \lambda\gamma\rangle$  $=$  for $\lambda \neq \gamma$,

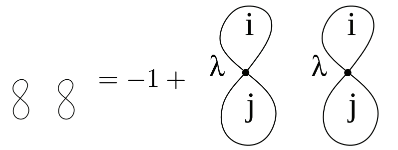
$|ijkl, \lambda\gamma\rangle$  $=$  for $\lambda \neq \gamma$ or $i \neq l$,

$|ijklmn, \lambda\gamma\delta\rangle$  $=$  .

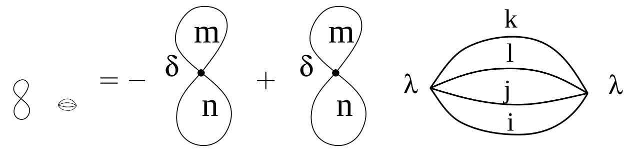
The states which interfere with the vacuum are those for which there are closed bubble diagrams from the given spin network to 'nothing'. In those orthonormal states the physical state can be constructed by simply subtracting the vacuum part using the standard Gram-Schmidt procedure. For example

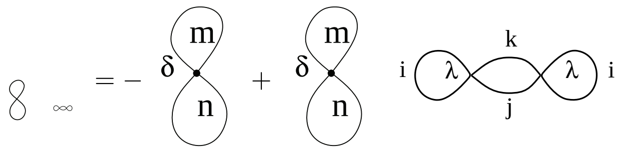
$|ijkl, \lambda\rangle$  $= -1 +$  ,

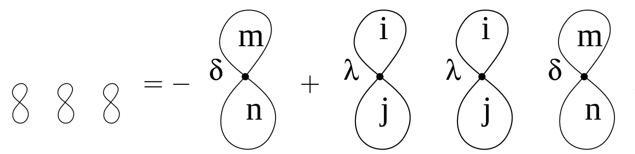
$|ijk, \lambda\rangle$  $= -1 +$  ,

$|ij,\lambda\rangle$ 

and so on. Other states turn out to be just the tensor product of the previous ones, namely

$|ijklmn,\lambda\delta\rangle$ 

$|ijkmn,\lambda\delta\rangle$ 

$|ijmn,\lambda\delta\rangle$ 

and so on. The procedure can be clearly continued to construct an orthonormal basis of $\mathcal{H}_{Ph}$.

6. Complex P

An important ingredient in the above construction is the assumption that P is real, equation (6). This assumption greatly simplifies the construction, allowing a simple definition of the ϕ_s operators. Here we discuss the meaning of this assumption and the extension of the formalism to the case in which P is non-real, or "complex".

To clarify the meaning of the reality condition, let us represent graphically the 2-net function $W(s,s')$ as in Figure 1, when s and s' are connected. If s' is formed by two connected components s_1 and s_2, we represent it as in Figure 2. Then the reality of P is expressed by the equality in Figure 3. That is, it represents an a priori indistinguishability between past and future boundaries of spacetime. This property is closely connected with crossing symmetry [**15**], which is essentially the analogous property at the level of the Hamiltonian constraint. The property is natural from the perspective of Atiyah's topological quantum field theory axiomatic framework [**27**]. It is perhaps natural to expect this property for Euclidean quantum gravity. Whether we should expect Lorentzian quantum gravity to have the same property, on the other hand, is not clear to us. On the one hand, the causal structure of the Lorentzian four-geometries seems to suggest that one should distinguish past and future boundaries. Notice also that (6) implies that the transition amplitudes $W(s,s')$ are real, because

$$W(s,s') = \langle s|P|s'\rangle = \langle 0|P|s'\cup s\rangle = \langle s'|P|s\rangle = \overline{\langle s|P|s'\rangle} = \overline{W(s,s')}, \tag{42}$$

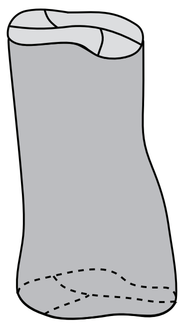


FIGURE 1. The three-geometry to three-geometry transition amplitude $W(s, s')$ between two connected three-geometries.

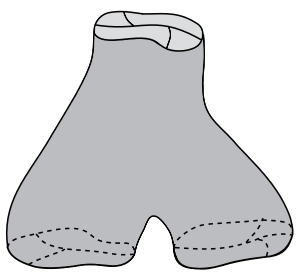


FIGURE 2. The transition amplitude between a doubly connected and a connected three-geometry.

which would prevent quantum mechanical interference between the $|s\rangle$ basis states (but not between generic states). On the other hand, however, temporal relations between boundaries may be induced a posteriori by the dynamics, instead of being a priori given in the structure of the formalism itself.

If we drop the reality condition on P, the main difficulty is that the definition (12) of the field operator becomes inconsistent. However, one can still retain a (partial) characterization of the field operator ϕ_s by requiring only

$$\hat{\phi}_s|0\rangle_{Ph} = |s\rangle_{Ph}. \tag{43}$$

This is certainly consistent. Notice that in general we have then

$$\phi_s^\dagger \neq \phi_s. \tag{44}$$

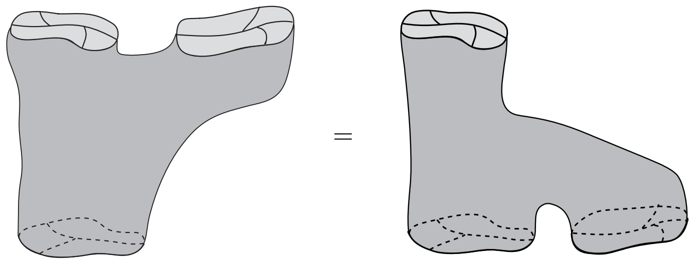


FIGURE 3. The reality condition on P.

And the 2-net function is now given by

$$W(s, s') = {}_{Ph}\langle 0|\hat{\phi}_s^\dagger \hat{\phi}_{s'}|0\rangle_{Ph}. \tag{45}$$

That is, we have to add the adjoint operation to equation (16).

The relevant abstract C^*-algebra now has a non-trivial star operation, different from $s^* = s$. If $\hat{\phi}_s^\dagger$ and $\hat{\phi}_{s'}$ are independent, the C^*-algebra is generated by products of s's and s^*'s, and

$$W(s, s') = W(s^* \cup s'). \tag{46}$$

Starting from the field theory, we may generate a W functional on the complex algebra $\mathcal{A}$ by using a complex field, instead of a real one. The resulting structure will be explored in detail elsewhere.

A closely related problem is whether the n-net W functions should be thought as analogous to the Wightman distributions (the vacuum expectation values of products of field operators), to the Feynman distributions (the vacuum expectation values of time-ordered products of field operators) or rather to the Schwinger functions (the appropriate analytic continuations of the Wightman distributions to imaginary time). In the case of a conventional quantum field theory defined on Minkowski space, one can apply the GNS reconstruction theorem to the Wightman distributions. On the other hand, one obtains directly the Feynman distributions as functional integrals of products of fields (with suitable "prescriptions" at the poles); while one can obtain the Schwinger functions as momenta of a stochastic process [**12**]. The Osterwalder-Schrader reconstruction theorem [**29**] that allows the reconstruction of the Wightman distributions from the Schwinger functions requires a duality "star" operation to be defined, corresponding to the inversion of the time variable. Presumably, the distinction between these different families of n-point functions makes no sense in the generally covariant context. The peculiar analytic structure of the n-point functions of field theory on Minkowski space is a consequence of the positivity of the energy (the Fourier transform of a function with support on the positive numbers is analytic in the upper complex plane.), while there is no notion of positivity of the energy in quantum gravity – indeed, there is no notion of energy at all– and one should be careful in trying to generalize standard quantum field theoretical prejudices to the generally covariant context.

7. Conclusion

We have studied the family of quantities $W(s)$, which we propose as the main physical observables of a quantum theory of gravity. We have proposed a general framework, based on these quantities, that ties the canonical (loop) and the covariant (spinfoam) approaches to quantum gravity. The connection between the two formalisms is provided by the GNS reconstruction theorem, and parallels the connection between the Hilbert space and the functional formulations of conventional quantum field theory, which one obtains from the properties of the n-point functions.

Many issues deserve to be clarified. Among these are the reality of P and the complex algebra $\mathcal{A}$; the locality property of $W(s)$ mentioned at the end of Section 3; and the connection between $W(s)$ and the S-matrix when spacetime admits asymptotic regions. An explicit construction of the C^*-algebra and its GNS construction in the case of the 2-dimensional theory [**30**] would also be of great interest and presumably not too hard to do.

References

[1] A. Connes and C. Rovelli, "Von Neumann algebra automorphisms and time thermodynamics relation in general covariant quantum theories," Class. Quant. Grav. 11 (1994) 2899.

[2] L Modesto, C Rovelli: "Particle scattering in loop quantum gravity", Phys Rev Lett 95 (2005) 191301 C Rovelli: "Graviton propagator from background-independent quantum gravity", Phys.Rev.Lett.97:151301,2006. E Bianchi, L Modesto, C Rovelli, S Speziale: "Graviton propagator in loop quantum gravity", Class.Quant.Grav. 23 (2006) 6989-7028. J Engle, R Pereira, C Rovelli: "The loop-quantum-gravity vertex-amplitude", Phys.Rev.Lett.99:161301,2007. E Alesci, C Rovelli: "The complete LQG propagator: I. Difficulties with the Barrett-Crane vertex", Phys.Rev.D76:104012, 2007. J Engle, R Pereira, C Rovelli: "Flipped spinfoam vertex and loop gravity", Nucl.Phys.B798:251-290, 2008. J Engle, R Pereira, C Rovelli: "LQG vertex with finite Immirzi parameter", Nucl.Phys.B799:136-149, 2008. E Alesci, C Rovelli: "The complete LQG propagator: II. Asymptotic behavior of the vertex", Phys.Rev.D77:044024, 2008.

[3] F Mattei, C Rovelli, S Speziale, M Testa, "From 3-geometry transition amplitudes to graviton states," Nucl Phys B739 (2006) 234–253,

[4] C Perini, C Rovelli, S Speziale: "Self-energy and vertex radiative corrections in LQG", arXiv:08101714

[5] For an overview of the present approaches to the problem of the description of quantum spacetime, see for instance: C Rovelli, *Strings, loops and the others: a critical survey on the present approaches to quantum gravity* in *Gravitation and Relativity: At the turn of the Millennium*, N Dadhich J Narlikar eds, pg 281-331 (Inter-University centre for Astronomy and Astrophysics, Pune 1998); gr-qc/9803024.

[6] For a general discussion, see: C Rovelli, *The century of the incomplete revolution: searching for general relativistic quantum field theory*, J. Math. Phys. 41, 3776 (2000), hep-th/9910131; *What is observable in classical and quantum gravity?*, Classical and Quantum Gravity 8, 297 (1991)

[7] SW Hawking, *The path-integral approach to quantum gravity*, in *General Relativity: An Einstein Centenary Survey*, SW Hawking and W Israel eds. (Cambridge University Press, Cambridge 1979).

[8] C Rovelli, L Smolin, *Knot theory and quantum gravity*, Phys Rev Lett 61 (1988) 1155. C Rovelli, L Smolin, *Loop space representation of quantum general relativity*, Nucl Phys, B331 (1990) 80. For a review, see: C Rovelli, *Loop Quantum Gravity*, Living Reviews in Relativity (refereed electronic journal), invited review paper. http://www.livingreviews.org/Articles/Volume1/1998-1rovelli; gr-qc/9709008. For an introduction, see: M Gaul, C Rovelli,

Loop quantum gravity and the meaning of diffeomorphism invariance, in *Towards Quantum Gravity. Proceedings of the XXXV Karpacz International Winter School on Theoretical Physics* J Kowalski–Glikman ed (Springer Verlag, Heidelberg, 1999). R Gambini and J Pullin, "Loops, knots, gauge theories and quantum gravity," (Cambridge, UK: Cambridge University Press 1996).

[9] C Rovelli, *Quantum Gravity*, (Cambridge University Press, Cambridge, 2004.) T Thiemann, "Modern Canonical Quantum General Relativity," (Cambridge, UK: Cambridge University Press 2007).

[10] C Rovelli, L Smolin, *Spin networks and quantum gravity*, Phys Rev D53 (1995) 5743–5759. JC Baez, Adv in Math Phys 117 (1996) 253; gr-qc/9411007. JC Baez, in *The Interface of Knots and Physics*, LH Kauffman ed, (American Mathematical Society, Providence, Rhode Island, 1996), gr-qc/9504036.

[11] C Rovelli and L Smolin, *Discreteness of area and volume in quantum gravity*, Nucl Phys B442 (1995) 593, gr-qc/9411005; Erratum: Nucl Phys B456 (1995) 734. Ashtekar A, Lewandowski J 1997, "Quantum theory of geometry I: area operators" *Class and Quantum Grav* 14(1997) A55; "II : volume operators", *Adv Theo Math Phys* 1(1997) 388-429.

[12] A S Wightman, "Quantum field theory in terms of vacuum expectation values," Phys Rev 101 (1956) 860. R F Streater, A S Wightman, "PCT, spin and statistics, and all that", Mathematical Physics Monograph Series, Benjamin-Cummings, (1964).

[13] A. Ashtekar, D. Marolf, J. Mourao, T. Thiemann, *Osterwalder-Schrader Reconstruction and Diffeomorphism Invariance*, quant-ph/9904094.

[14] M Reisenberger, *Worldsheet formulations of gauge theories and gravity*, talk given at the 7th Marcel Grossmann Meeting Stanford, July 1994; gr-qc/9412035. J Iwasaki, *A definition of the Ponzano-Regge quantum gravity model in terms of surfaces*, J Math Phys 36 (1995) 6288. J Baez, *spinfoam Models*, Class Quant Grav 15 (1998) 1827-1858; gr-qc/9709052. *An introduction to spinfoam models of BF theory and quantum gravity*, in Geometry and Quantum Physics, eds. Helmut Gausterer and Harald Grosse, Lecture Notes in Physics, Springer, Berlin, 2000, pp. 25-93, gr-qc/9905087. J Iwasaki *A surface theoretic model of quantum gravity*, gr-qc/9903112. J Iwasaki,*A lattice quantum gravity model with surface-like excitations in 4-dimensional spacetime*, gr-qc/0006088. L Freidel, K Krasnov *spinfoam Models and the classical action principle*, Adv Theor Math Phys 2 (1999) 1183-1247. R De Pietri, *Canonical* loop *quantum gravity and spinfoam models*, Proceeding of the XXIII SIGRAV conference, Monopoli (ITALY), September 1998. F Markopoulou, L Smolin, *Causal evolution of spin networks*, Nucl Phys B508 (1997) 409-430.

[15] M Reisenberger, C Rovelli, *Sum over Surfaces form of loop quantum gravity*, Phys Rev D56 (1997) 3490-3508. C Rovelli, *Quantum gravity as a sum over surfaces*, Nucl Phys B57 (1997) 28-43. C Rovelli, *The projector on physical states in loop quantum gravity*, gr-qc/9806121.

[16] R De Pietri, L Freidel, K Krasnov, C Rovelli, *Barrett-Crane model from a Boulatov-Ooguri field theory over a homogeneous space*, Nuclear Physics, to appear, hep-th/9907154.

[17] M Reisenberger, C Rovelli, *Spinfoam models as Feynman diagrams*, gr-qc/0002083. M Reisenberger, C Rovelli, *Spacetime as a Feynman diagram: the connection formulation*, gr-qc/0002095.

[18] A Perez, C Rovelli, *A spinfoam model without bubble divergences*, Nuclear Physics B599 (2001) 255-282. A Perez, *Finiteness of a spinfoam model for euclidean quantum general relativity*, gr-qc/0011058. L Crane, A Perez, C Rovelli: *Finiteness in spinfoam quantum gravity*, Physical Review Letters 87 (2001) 181301.

[19] L Freidel, *Group Field Theory: an overview*, Int Journ Theor Phys 44 (2005) 1769-1783. D Oriti, *The group field theory approach to quantum gravity*, arXiv:gr-qc/0607032.

[20] JW Barrett, L Crane, *Relativistic spin networks and quantum gravity*, J Math Phys 39 (1998) 3296.

[21] JW Barrett, L Crane, *A Lorentzian Signature Model for Quantum General Relativity*, Class. Quantum Grav. 17 (2000) 3101-3118; gr-qc/9904025. A Perez, C Rovelli, *spinfoam model for Lorentzian General Relativity*, Phys Rev D63 (2001) 04150, gr-qc/0009021; *3+1 spinfoam model of quantum gravity with spacelike and timelike components* gr-qc/0011037. R. Pereira, "Lorentzian LQG vertex amplitude", arXiv:0710.5043.

[22] A Mikovic, *Quantum field theory of spin networks*, gr-qc/0102110.

[23] C Misner, *Feynman quantization of General relativity*, Rev Mod Phys 29 (1957) 497.

[24] C Rovelli, "Ashtekar formulation of general relativity and loop space non-perturbative Quantum Gravity: a report" Class Quant Grav 8 (1991) 1613-1675.

[25] A Ashtekar, J Lewandowski, D Marolf, J Mourao, T Thiemann, *Quantization of diffeomorphism invariant theories of connections with local degrees of freedom*, J Math Phys 36 (1995) 6456.

[26] W. Fairbairn and C. Rovelli, "Separable Hilbert space in loop quantum gravity," J. Math. Phys. 45, 2802 (2004)

[27] M Atiyah, *Quantum theory and geometry* J Math Phys 36:6069-6072,1995. "The Geometry and Physics of Knots", Cambridge Univ Press, (1990).

[28] E Brézin, C Itzykson, G Parisi, J Zuber, Commun Math Phys 59 (1978) 35. F David, Nucl Phys B257 (1985) 45. J Ambjorn, B Durhuus, J Fröhlich, Nucl Phys B257 (1985) 433. VA Kazakov, IK Kostov, AA Migdal, Phys Lett 157 (1985) 295. D V Boulatov, V A Kazakov, I K Kostov, A A Migdal, Nucl Phys B275 (1986) 641. M Douglas, S Shenker, Nucl Phys B335(1990) 635. D Gross, A Migdal, Phys Rev Lett 64 (1990) 635; E Brézin, V A Kazakov, Phys Lett B236 (1990) 144. For the passage from 2d models to field theories on a group, see DV Boulatov, Mod Phys Lett A7 (1992) 1629.

[29] K Osterwalder, R Schrader, Comm Math Phys 31 (1973) 83; 42 (1975) 281.

[30] R Livine, A Perez, C Rovelli, *2d manifold-independent spinfoam theory*, Classical and Quantum Gravity 20 (2003), 4425. gr-qc/0102051.

Centre de Physique Théorique de Luminy, Case 907, F-13288 Marseille, France
E-mail address: perez@cpt.univ-mrs.fr

Centre de Physique Théorique de Luminy, Case 907, F-13288 Marseille, France
E-mail address: rovelli@cpt.univ-mrs.fr

Clay Mathematics Proceedings
Volume 11, 2010

On the Fundamental Group of II_1 Factors and Equivalence Relations Arising from Group Actions

Sorin Popa and Stefaan Vaes

Dedicated to Alain Connes at the occasion of his 60th birthday

Abstract. Given a countable group G, we consider the sets $\mathcal{S}_{\mathrm{factor}}(G)$, $\mathcal{S}_{\mathrm{eqrel}}(G)$, of subgroups $\mathcal{F} \subset \mathbb{R}_+$ for which there exists a free ergodic probability measure preserving action $G \curvearrowright X$ such that the fundamental group of the associated II_1 factor $\mathrm{L}^\infty(X) \rtimes G$, respectively orbit equivalence relation $\mathcal{R}(G \curvearrowright X)$, equals $\mathcal{F}$. We prove that if $G = \Gamma^{*\infty} * \mathbb{Z}$, with $\Gamma \neq 1$, then $\mathcal{S}_{\mathrm{factor}}(G)$ and $\mathcal{S}_{\mathrm{eqrel}}(G)$ contain $\mathbb{R}_+$ itself, all of its countable subgroups, as well as uncountable subgroups that can have any Hausdorff dimension $\alpha \in (0,1)$. We deduce that there exist II_1 factors of the form $M = \mathrm{L}^\infty(X) \rtimes \mathbb{F}_\infty$ such that the fundamental group of M is $\mathbb{R}_+$, but $M \overline{\otimes} \mathrm{B}(\ell^2(\mathbb{N}))$ admits no continuous trace scaling action of $\mathbb{R}_+$. We then prove that if $G = \Gamma * \Lambda$, with Γ, Λ finitely generated ICC groups, one of which has property (T), then $\mathcal{S}_{\mathrm{factor}}(G) = \mathcal{S}_{\mathrm{eqrel}}(G) = \{\{1\}\}$.

1. Introduction

Some of the most intriguing phenomena concerning group measure space II_1 factors $M = \mathrm{L}^\infty(X) \rtimes G$ and orbit equivalence relations $\mathcal{R} = \mathcal{R}(G \curvearrowright X)$, arising from free ergodic probability measure preserving (p.m.p.) actions $G \curvearrowright X$ of countable groups G on probability spaces (X, μ), pertain to their *fundamental groups* $\mathcal{F}(M)$, $\mathcal{F}(\mathcal{R})$ ([**20**]). Although much progress has been made in understanding and calculating these invariants, many natural questions on how the group G may affect the behavior of $\mathcal{F}(M)$, $\mathcal{F}(\mathcal{R})$ remain open.

A first indication that certain properties of G can impact the invariants independently of the way it acts appeared in Connes' groundbreaking work on the classification and the structure of von Neumann factors, from the 1970's. Thus, a side effect of the uniqueness of the amenable II_1 factor [**5**] and of the amenable II_1 equivalence relation [**7**], is that $\mathcal{F}(M) = \mathcal{F}(\mathcal{R}) = \mathbb{R}_+$ whenever the group G is amenable. On the other hand, arguments from Connes' rigidity paper [**4**] were

2010 *Mathematics Subject Classification.* Primary: 37A20. Secondary: 22D40, 46L10, 46LS, 46L55.

The first author was partially supported by NSF Grant, DMS-0601082.

The second author was partially supported by ERC Starting Grant VNALG-200749 and Research Programme G.0231.07 of the Research Foundation – Flanders (FWO).

used to show that if G is infinite conjugacy class (ICC) and has the property (T) of Kazhdan, then $\mathcal{F}(M), \mathcal{F}(\mathcal{R})$ are countable for any free ergodic p.m.p. action of G ([**25**], [**17**]).

Then, in the late 1990's, Gaboriau discovered that certain groups G, such as the free groups with finitely many generators, $\mathbb{F}_n, 2 \leq n < \infty$, give rise to orbit equivalence relations $\mathcal{R} = \mathcal{R}(G \curvearrowright X)$ with $\mathcal{F}(\mathcal{R}) = \{1\}$, for *any* free ergodic p.m.p. action $F \curvearrowright X$ [**15**]. Moreover, many factors of the form $\mathrm{L}^\infty(X) \rtimes \mathbb{F}_n$ were shown to have trivial fundamental group as well (cf. [**29**], [**24**]) and it is strongly believed that, in fact, this holds true for all $\mathbb{F}_n \curvearrowright X$.

In turn, a completely new type of phenomena emerged in the case $G = \mathbb{F}_\infty$, where it was shown that there exist free ergodic p.m.p. actions $\mathbb{F}_\infty \curvearrowright X$ with the fundamental group of the associated II_1 factors and equivalence relations, $\mathcal{F}(M), \mathcal{F}(\mathcal{R})$, ranging over a "large" family of subgroups $\mathcal{F} \subset \mathbb{R}_+$, containing $\mathbb{R}_+$ itself, all its countable subgroups, as well as "many" uncountable subgroups $\neq \mathbb{R}_+$ [**31**]. In fact, it was conjectured in [**31**] that any group $\mathcal{F}$ that can be realized as a fundamental group of a II_1 factor or equivalence relation can also be realised as $\mathcal{F}(\mathrm{L}^\infty(X) \rtimes \mathbb{F}_\infty)$, $\mathcal{F}(\mathcal{R}(\mathbb{F}_\infty \curvearrowright X))$, for some free ergodic p.m.p. action $\mathbb{F}_\infty \curvearrowright X$.

Related to all these phenomena, we introduced in [**31**] the sets $\mathcal{S}_{\mathrm{factor}}(G)$, $\mathcal{S}_{\mathrm{eqrel}}(G)$, of subgroups $\mathcal{F} \subset \mathbb{R}_+$ for which there exists a free ergodic m.p. action $G \curvearrowright X$ such that $\mathcal{F}(\mathrm{L}^\infty(X) \rtimes G) = \mathcal{F}$, respectively $\mathcal{F}(\mathcal{R}(G \curvearrowright X)) = \mathcal{F}$. Using this notation, the result in [**31**] shows, more precisely, that $\mathcal{S}_{\mathrm{factor}}(\mathbb{F}_\infty) \cap \mathcal{S}_{\mathrm{eqrel}}(\mathbb{F}_\infty)$ contains the set $\mathcal{S}_{\mathrm{centr}}$ of all subgroups $\mathcal{F} \subset \mathbb{R}_+$ for which there exists a free ergodic action of an amenable group Λ on an infinite measure space (Y, ν), such that the set of scalars $t > 0$ that can appear as scaling constants of non-singular automorphisms θ of (Y, ν) commuting with $\Lambda \curvearrowright Y$ equals $\mathcal{F}$. In turn, $\mathcal{S}_{\mathrm{centr}}$ is shown to contain $\mathbb{R}_+$, all its countable subgroups and uncountable subgroups $\mathcal{F} \subset \mathbb{R}_+$ with arbitrary Hausdorff dimension in the interval $(0, 1)$ ([**31**]). While an abstract characterization of $\mathcal{S}_{\mathrm{factor}}(\mathbb{F}_\infty)$, $\mathcal{S}_{\mathrm{eqrel}}(\mathbb{F}_\infty)$ remains elusive, it was noticed in [**31**] that subgroups in either set, as in fact subgroups in $\mathcal{S}_{\mathrm{factor}}(G)$, $\mathcal{S}_{\mathrm{eqrel}}(G)$ for any G, must be Borel sets and Polishable.

Our purpose in this paper is to estimate (or even completely calculate) the invariants $\mathcal{S}_{\mathrm{factor}}(G)$, $\mathcal{S}_{\mathrm{eqrel}}(G)$ for other classes of groups G. We target two types of results: on the one hand, detecting classes of groups G for which $\mathcal{S}_{\mathrm{factor}}(G)$, $\mathcal{S}_{\mathrm{eqrel}}(G)$ are "large", containing for instance the set $\mathcal{S}_{\mathrm{centr}}$ defined above (like in the case case $G = \mathbb{F}_\infty$); on the other hand, detecting classes of groups G for which $\mathcal{S}_{\mathrm{factor}}(G)$, $\mathcal{S}_{\mathrm{eqrel}}(G)$ contain only "small" subgroups of $\mathbb{R}_+$ (e.g. countable, or just $\{1\}$).

Thus, our first result enlarges considerably the class of groups G for which we can show that the set $\mathcal{S}_{\mathrm{centr}}$ is contained in both $\mathcal{S}_{\mathrm{factor}}(G)$ and $\mathcal{S}_{\mathrm{eqrel}}(G)$.

THEOREM 1.1. *Let Γ be a non-trivial group, Σ an infinite amenable group and denote $G = \Gamma^{*\infty} * \Sigma$. Then,*

$$\mathcal{S}_{\mathrm{centr}} \subset \mathcal{S}_{\mathrm{factor}}(G) \qquad \text{and} \qquad \mathcal{S}_{\mathrm{centr}} \subset \mathcal{S}_{\mathrm{eqrel}}(G) .$$

Moreover, there exist free ergodic p.m.p. actions $G \curvearrowright (X, \mu)$ such that the II_1 factor $M = \mathrm{L}^\infty(X) \rtimes G$ has fundamental group $\mathcal{F}(M) = \mathbb{R}_+$, but the II_∞ factor $M \overline{\otimes} \mathrm{B}(\ell^2(\mathbb{N}))$ admits no trace scaling action of $\mathbb{R}_+$.

In Section 6, we will show that if the full group of an equivalence relation $\mathcal{R}$ on a probability space (X, μ) contains a property (T) group acting ergodically on X, then $\mathcal{F}(\mathcal{R})$ is countable. Thus, if a group Γ appearing in Theorem 1.1 contains an infinite subgroup Λ with the property (T) and if $G = \Gamma^{*\infty} * \Sigma \curvearrowright X$ is a free ergodic p.m.p. action such that $\mathcal{R}_G$ has fundamental group equal to an uncountable group in $\mathcal{S}_{\mathrm{centr}}$, then the restriction of $G \curvearrowright X$ to Λ cannot be ergodic.

Note that the last part of Theorem 1.1 provides group measure space II_1 factors $M = \mathrm{L}^\infty(X) \rtimes G$ which do have fundamental group equal to $\mathbb{R}_+$ yet cannot appear in the continuous decomposition of a type III_1 factor. The problem of whether such II_1 factors exist was posed over the years by several people, including Connes, Takesaki, and more recently Shlyakhtenko. The fact that there are even factors of the form $\mathrm{L}^\infty(X) \rtimes \mathbb{F}_\infty$ satisfying this property (by simply taking $\Gamma = \Sigma = \mathbb{Z}$ in 1.1) should be contrasted with the fact that the II_∞ factor associated with $L(\mathbb{F}_\infty)$ does admit a trace scaling action of $\mathbb{R}_+$, by [**34**].

Note that all groups of the form $G = \Gamma^{*\infty} * \Sigma$, covered by the above theorem, have infinite first ℓ^2-Betti number, $\beta_1^{(2)}(G) = \infty$, and in fact $\beta_n^{(2)}(G) = \infty, 0$, $\forall n \geq 2$. On the other hand, by Gaboriau's scaling formula for ℓ^2-Betti numbers [**15**], any free ergodic p.m.p. action of a group G with $\beta_n^{(2)}(G) \neq 0, \infty$, for some n, gives rise to an equivalence relation $\mathcal{R}_G$ with trivial fundamental group, $\mathcal{F}(\mathcal{R}_G) = \{1\}$. In other words, $\mathcal{S}_{\mathrm{eqrel}}(G) = \{\{1\}\}$. While it is still an open question whether the corresponding II_1 factors $M = \mathrm{L}^\infty(X) \rtimes G$ satisfy $\mathcal{F}(M) = \{1\}$ as well (i.e. $\mathcal{S}_{\mathrm{factor}}(G) = \{\{1\}\}$), our next result provides a large class of groups G for which this is indeed the case.

THEOREM 1.2. *Let Γ and Λ be infinite, finitely generated groups. Assume that Γ is ICC and that one of the following conditions holds.*

a) $\Gamma = \Gamma_1 \times \Gamma_2$, *with Γ_1 non-trivial and Γ_2 non-amenable,*
b) Γ *admits a non-virtually abelian, normal subgroup Γ_1 with the relative property (T).*

Then, $\mathcal{S}_{\mathrm{factor}}(\Gamma * \Lambda) = \mathcal{S}_{\mathrm{eqrel}}(\Gamma * \Lambda) = \{\{1\}\}$.

When viewed from the perspective of Connes' discrete decomposition of type III_λ factors with $0 < \lambda < 1$ ([**6**]) and respectively Connes-Takesaki continuous decomposition of type III_1 factors ([**10**]), the above result provides a large class of groups G with the property that no II_1 factor M arising from an arbitrary free ergodic p.m.p. action of G can appear in the decomposition of a type III factor (i.e., as Connes puts it, no such M can appear as the "shadow" of a type III factor).

While II_1 factors $M = \mathrm{L}^\infty(X) \rtimes \Gamma$ arising from free ergodic p.m.p. actions $\Gamma \curvearrowright X$ of ICC property (T) groups always have countable fundamental group (cf. [**4**], [**25**], [**17**]), it was not known whether there exist cases when $\mathcal{F}(M) \neq \{1\}$. Our next result gives the first such examples. It also provides the first "concrete" examples of free ergodic p.m.p. actions $\Gamma \curvearrowright X$ with the associated II_1 factors M having fundamental group $\neq \{1\}, \mathbb{R}_+$. Indeed, the actions in Theorem 1.1 above and in [**31**] are shown to exist by using a Baire-category argument, at some point, while in 1.3 below they are specific G-actions, obtained as diagonal products of Bernoulli and profinite actions.

THEOREM 1.3. *Let $\mathcal{F} \subset \mathbb{Q}_+$ be a subgroup generated by a subset of the prime numbers. Let $G = \mathbb{Z}^n \rtimes \mathrm{SL}(n, \mathbb{Z})$ with $n \geq 3$. Then G admits a free ergodic p.m.p. action $G \curvearrowright (X, \mu)$ such that the fundamental group of $\mathrm{L}^\infty(X) \rtimes G$ and of $\mathcal{R}(G \curvearrowright X)$ equals $\mathcal{F}$.*

We in fact believe that any subgroup of $\mathbb{Q}_+$ can be realized as the fundamental group of a factor or equivalence relation arising from a free ergodic p.m.p. action of $\mathbb{Z}^n \rtimes \mathrm{SL}(n,\mathbb{Z})$, $n \geq 3$. The question of whether there exist free ergodic p.m.p. actions of an ICC property (T) group $G \curvearrowright X$ such that $\mathcal{F}(\mathcal{R}_G)$ or $\mathcal{F}(\mathrm{L}^\infty(X) \rtimes G)$ contains irrational numbers remains open. In fact, it is not even known whether the union of all the fundamental groups of II_1 factors and equivalence relations arising from free ergodic p.m.p. actions of a fixed ICC property (T) group G is necessarily countable or not.

Finally, noticing that for a large number of groups G it is known that $\{1\} \in \mathcal{S}_{\mathrm{factor}}(G)$ (see e.g. [**27**], [**29**], [**30**]), we conjecture that this is in fact the case for all non-amenable groups G. If true, this would also show that the only possibilities for $\mathcal{S}_{\mathrm{factor}}(G), \mathcal{S}_{\mathrm{eqrel}}(G)$ to be single point sets are $\mathcal{S}_{\mathrm{factor}}(G) = \mathcal{S}_{\mathrm{eqrel}}(G) = \{\mathbb{R}_+\}$, $\mathcal{S}_{\mathrm{factor}}(G) = \mathcal{S}_{\mathrm{eqrel}}(G) = \{\{1\}\}$, the first situation corresponding to G being amenable. This would provide a new, interesting facet of the dichotomy amenable/non-amenable for groups.

2. Preliminaries

The *fundamental group* $\mathcal{F}(M)$ of a II_1 factor M, introduced in [**20**], is defined as the following subgroup of $\mathbb{R}_+$.

$$\mathcal{F}(M) = \{\tau(p)/\tau(q) \mid p, q \text{ are non-zero projections in } M \text{ such that } pMp \cong qMq\} .$$

We call II_1 equivalence relation on a standard probability space (X,μ) any ergodic probability measure preserving (p.m.p.) measurable equivalence relation with countable equivalence classes. The fundamental group $\mathcal{F}(\mathcal{R})$ of a II_1 equivalence relation $\mathcal{R}$ is defined as

$$\mathcal{F}(\mathcal{R}) = \{\mu(Y)/\mu(Z) \mid \mathcal{R}|_Y \cong \mathcal{R}|_Z\} .$$

Whenever $\Gamma \curvearrowright (X,\mu)$ is a free ergodic p.m.p. action, we denote by $\mathcal{R}(\Gamma \curvearrowright X)$ the associated II_1 orbit equivalence (OE) relation and by $\mathrm{L}^\infty(X) \rtimes \Gamma$ the associated group measure space II_1 factor [**20**].

Definition 2.1. A free ergodic p.m.p. action $\Gamma \curvearrowright (X,\mu)$ is called *rigid* if the corresponding inclusion $\mathrm{L}^\infty(X) \subset \mathrm{L}^\infty(X) \rtimes \Gamma$ is rigid in the sense of [**29**, Proposition 4.1].

Some sets of subgroups of $\mathbb{R}$ and ergodic measures. Given a countable group Γ, we are interested in

$$\begin{aligned} \mathcal{S}_{\mathrm{factor}}(\Gamma) := \{\mathcal{F} \subset \mathbb{R}_+ \mid\ & \text{there exists a free ergodic p.m.p. action } \Gamma \curvearrowright (X,\mu) \\ & \text{such that } \mathcal{F}(\mathrm{L}^\infty(X) \rtimes \Gamma) = \mathcal{F}\} , \\ \mathcal{S}_{\mathrm{eqrel}}(\Gamma) := \{\mathcal{F} \subset \mathbb{R}_+ \mid\ & \text{there exists a free ergodic p.m.p. action } \Gamma \curvearrowright (X,\mu) \\ & \text{such that } \mathcal{F}(\mathcal{R}(\Gamma \curvearrowright X)) = \mathcal{F}\} . \end{aligned}$$

In [**31**, Theorem 5.3 and formula (2.2)], we have shown that both $\mathcal{S}_{\mathrm{factor}}(\mathbb{F}_\infty)$ and $\mathcal{S}_{\mathrm{eqrel}}(\mathbb{F}_\infty)$ contain $\mathcal{S}_{\mathrm{centr}}$, defined as

$$\begin{aligned} \mathcal{S}_{\mathrm{centr}} := \{\mathcal{F} \subset \mathbb{R}_+ \mid\ & \text{there exists } \Lambda \curvearrowright (Y,\eta) \text{ , a free ergodic m.p. action,} \\ & \text{with } \Lambda \text{ amenable and } \operatorname{mod}(\mathrm{Centr}_\Lambda(Y)) = \mathcal{F} \} . \end{aligned}$$

Following [**1**, Section 4], we call *ergodic measure on* $\mathbb{R}$ any σ-finite measure ν on the Borel sets of $\mathbb{R}$ having the following properties, where we denote $\lambda_x(y) = x + y$.

- For all $x \in \mathbb{R}$, either $\nu \circ \lambda_x = \nu$ or $\nu \circ \lambda_x \perp \nu$.
- There exists a countable subgroup $Q \subset \mathbb{R}$ such that $\nu \circ \lambda_x = \nu$ for all $x \in Q$ and such that every Q-invariant Borel function on $\mathbb{R}$ is ν-almost everywhere constant.

For every ergodic measure ν on $\mathbb{R}$, one defines

$$H_\nu := \{x \in \mathbb{R} \mid \nu \circ \lambda_x = \nu\} .$$

As shown in [**1**], the groups H_ν can have arbitrary Hausdorff dimension and all $\exp(H_\nu)$ belong to $\mathcal{S}_{\mathrm{centr}}$. We refer to [**31**, Section 2 and the proof of Theorem 5.3] for a detailed exposition.

Intertwining by bimodules and the notation $A \underset{M}{\prec} B$. In Sections 3 and 4, we use the method of intertwining by bimodules, introduced by the first author in [**29**]. Let (M, τ) be a von Neumann algebra with faithful normal tracial state τ. We use the notation $M^n = \mathrm{M}_n(\mathbb{C}) \otimes M$. When $A, B \subset M^n$ are possibly non-unital embeddings, we write $A \underset{M}{\prec} B$ if there exists a non-zero partial isometry $v \in 1_A(\mathrm{M}_{n,m}(\mathbb{C}) \otimes M)$ and a, possibly non-unital, normal *-homomorphism $\rho : A \to B^m$ satisfying $av = v\rho(a)$ for all $a \in A$. Several equivalent formulations of this property can be given; see [**26**, Theorem 2.1] (see also [**37**, Theorem C.3]).

Suppose that A and B are Cartan subalgebras of the II_1 factor M. Let $A_0 \subset A$ be a von Neumann subalgebra such that $A_0' \cap M = A$. By [**29**, Theorem A.1], $A_0 \underset{M}{\prec} B$ if and only if there exists a unitary $u \in M$ such that $uAu^* = B$.

3. Groups G for which $\mathcal{S}_{\mathrm{factor}}(G)$ contains uncountable groups

The following theorem, whose proof is given at the end of the section, provides a large family of groups G such that $\mathcal{S}_{\mathrm{factor}}(G)$ and $\mathcal{S}_{\mathrm{eqrel}}(G)$ is large, in the sense that both contain $\mathcal{S}_{\mathrm{centr}}$. Moreover, we prove that G admits free ergodic p.m.p. actions $G \curvearrowright (X, \mu)$ such that the II_1 factor $M := \mathrm{L}^\infty(X) \rtimes G$ has fundamental group $\mathbb{R}_+$, but nevertheless, the II_∞ factor $M \overline{\otimes} \mathrm{B}(\ell^2(\mathbb{N}))$ admits no strongly continuous trace scaling action of $\mathbb{R}_+$.

The groups G involved are infinite free product groups and should be contrasted with the groups G treated in Theorem 4.1, for which $\mathcal{S}_{\mathrm{factor}}(G)$ is trivial (cf. Remark 4.2).

THEOREM 3.1. *Let Γ be a non-trivial group, Σ an infinite amenable group and denote $G = \Gamma^{*\infty} * \Sigma$. Then,*

$$\mathcal{S}_{\mathrm{centr}} \subset \mathcal{S}_{\mathrm{factor}}(G) \qquad \text{and} \qquad \mathcal{S}_{\mathrm{centr}} \subset \mathcal{S}_{\mathrm{eqrel}}(G) .$$

Moreover, there exist free ergodic p.m.p. actions $G \curvearrowright (X, \mu)$ such that the II_1 factor $M = \mathrm{L}^\infty(X) \rtimes G$ has fundamental group $\mathcal{F}(M) = \mathbb{R}_+$, but the II_∞ factor $M \overline{\otimes} \mathrm{B}(\ell^2(\mathbb{N}))$ admits no trace scaling action of $\mathbb{R}_+$.

In the course of the proof of Theorem 3.1, we will also obtain the following result.

THEOREM 3.2. *There exist II_1 factors M_1 and M_2 such that $\mathcal{F}(M_1) \neq \mathbb{R}_+ \neq \mathcal{F}(M_2)$, but nevertheless $\mathcal{F}(M_1 \overline{\otimes} M_2) = \mathbb{R}_+$.*

Let G be a countable group with subgroup Γ. Suppose that $G \curvearrowright (X, \mu)$ is a free p.m.p. action such that the restriction to Γ is ergodic. Slightly changing notations compared to [**31**, Section 2], denote by $\mathrm{Emb}(\Gamma, G)$ the set of non-singular partial automorphisms ϕ of (X, μ) satisfying $\phi(g \cdot x) \in G \cdot \phi(x)$ for all $g \in \Gamma$ and almost all $x \in X$ with $x, g \cdot x \in D(\phi)$. Denote by $[[G]]$ the full pseudogroup of the OE relation $\mathcal{R}(G \curvearrowright X)$, i.e. the set of a partial automorphisms ϕ of (X, μ) satisfying $\phi(x) \in G \cdot x$ for almost all $x \in D(\phi)$.

The following lemma generalizes [**31**, Theorem 4.1].

LEMMA 3.3. *Let Γ be an infinite group, Λ an arbitrary group, both acting freely and p.m.p. on (X, μ). There exists a free p.m.p. action $\Gamma^{*\infty} * \Lambda \overset{\alpha}{\curvearrowright} (X, \mu)$ with the following properties.*

- *The restriction of α to $\Gamma^{*\infty}$ is ergodic and rigid (in the sense of Definition 2.1).*
- $\mathrm{Emb}(\Gamma^{*\infty}, \Gamma^{*\infty} * \Lambda) = [[\Gamma^{*\infty} * \Lambda]]$.
- *The restriction of α to any of the copies of Γ, resp. to Λ, is conjugate to the originally given action.*

PROOF. Denote the given actions by $\Gamma \overset{\beta}{\curvearrowright} (X, \mu)$ and $\Lambda \overset{\rho}{\curvearrowright} (X, \mu)$. We introduce the following notations:

$$\begin{aligned} \Gamma^{*\infty} &= \mathop{\ast}_{n=-1}^{\infty} G_n \quad \text{with all } \ G_n \cong \Gamma \ , \\ \Gamma_n &:= \mathop{\ast}_{k=-1}^{n} G_k \ , \\ \Gamma_E &:= G_{-1} * G_0 * \mathop{\ast}_{n \in E} G_n \quad \text{whenever } \ E \subset \mathbb{N} \ . \end{aligned}$$

By [**16**, Theorem 1.2], take a free ergodic p.m.p. action $\Gamma_0 \overset{\alpha_0}{\curvearrowright} (X, \mu)$ such that α_0 is a rigid action and such that the restrictions of α_0 to G_{-1} and G_0 are conjugate to the action β. By [**36**, Category Lemma] and [**19**, Lemma A.1], extend α_0 to a free action of $\Gamma_0 * \Lambda$ on (X, μ), still denoted by α_0, whose restriction to Λ is conjugate to the action ρ.

Extend the action α_0 inductively to free actions $\Gamma_n * \Lambda \overset{\alpha_n}{\curvearrowright} (X, \mu)$ following the procedure in [**31**, Section 3] and such that the restriction of α_n to $G_k \subset \Gamma_n$ is conjugate to β for all $k \leq n$. We end up with the free action $\Gamma^{*\infty} * \Lambda \overset{\alpha_\infty}{\curvearrowright} (X, \mu)$. For every infinite subset $E \subset \mathbb{N}$, we denote by α_E the restriction of α_∞ to $\Gamma_E * \Lambda$. Following the proof of [**31**, Theorem 4.1], there exists an infinite subset $E \subset \mathbb{N}$ such that $\mathrm{Emb}(\Gamma_E, \Gamma_E * \Lambda) = [[\Gamma_E * \Lambda]]$. Since $\Gamma_E \cong \Gamma^{*\infty}$, the lemma is proved. □

REMARK 3.4. Using the methods of [**16**, Section 2.3], Lemma 3.3 can be shown for $\Gamma_1 * \Gamma_2$ instead of the infinite free product $\Gamma^{*\infty}$, for arbitrary infinite groups Γ_1, Γ_2 with given free p.m.p. actions on (X, μ). Such a generalization does not provide a refinement for Theorem 3.1 though, since the proof of Theorem 3.1 involves taking once more an infinite free product.

For the formulation of the following theorem, recall that the automorphism group $\mathrm{Aut}(N)$ of a von Neumann algebra with separable predual is a Polish group under the topology making the maps $\mathrm{Aut}(N) \to N_* : \alpha \mapsto \omega \circ \alpha$ continuous for all $\omega \in N_*$. Similarly, the group $\mathrm{Aut}(Y, \eta)$ of non-singular isomorphisms of (Y, η)

(up to equality almost everywhere) is a Polish group and $\mathrm{Centr}_{\mathrm{Aut}\,Y}(\Gamma_2)$ is a closed subgroup whenever $\Gamma_2 \curvearrowright (Y,\eta)$ is a non-singular action.

THEOREM 3.5. *Let* $\Gamma_1 * \Gamma_2 \overset{\alpha}{\curvearrowright} (X,\mu)$ *be a free p.m.p. action. Let* $\Gamma_2 \curvearrowright (Y,\eta)$ *be a free ergodic action preserving the infinite standard measure* η. *Consider the action* $\Gamma_1 * \Gamma_2 \curvearrowright X \times Y$ *given by*

$$g \cdot (x,y) = (g\cdot x, y) \;\; \forall g \in \Gamma_1 \;\;,\;\; h\cdot(x,y) = (h\cdot x, h\cdot y) \;\; \forall h \in \Gamma_2\;. \tag{3.1}$$

Make the following assumptions.

- *The restriction of* α *to* Γ_1 *is ergodic and rigid.*
- *We have* $\mathrm{Emb}(\Gamma_1, \Gamma_1 * \Gamma_2) = [[\Gamma_1 * \Gamma_2]]$.
- Γ_2 *is amenable.*

Then, the following holds.

(1) *The map*

$$\Theta : \mathrm{Centr}_{\mathrm{Aut}\,Y}(\Gamma_2) \to \mathrm{Aut}(\mathcal{R}(\Gamma_1 * \Gamma_2 \curvearrowright X \times Y)) : \Delta \mapsto \Theta_\Delta$$
$$\text{where } \;\Theta_\Delta(x,y) = (x, \Delta(y))$$

induces an onto group isomorphism between $\mathrm{Centr}_{\mathrm{Aut}\,Y}(\Gamma_2)$ *and* $\mathrm{Out}(\mathcal{R}(\Gamma_1 * \Gamma_2 \curvearrowright X \times Y))$.

(2) *Define the* II_∞ *factor* $N := \mathrm{L}^\infty(X \times Y) \rtimes (\Gamma_1 * \Gamma_2)$. *Denote for every* $\Delta \in \mathrm{Centr}_{\mathrm{Aut}\,Y}(\Gamma_2)$, *by* θ_Δ *the corresponding automorphism of* N.

(a) *The group* $\mathrm{Aut}(N)$ *is generated by the three subgroups* $\{\theta_\Delta \mid \Delta \in \mathrm{Centr}_{\mathrm{Aut}\,Y}(\Gamma_2)\}$, *the inner automorphism group* $\mathrm{Inn}(N) = \{\mathrm{Ad}\, u \mid u \in \mathcal{U}(N)\}$ *and the group of automorphisms*$^{(1)}$ $H := \{\theta \in \mathrm{Aut}(N) \mid \theta(a) = a$ *for all* $a \in \mathrm{L}^\infty(X \times Y)\}$.

(b) *The subgroup* $\mathrm{Inn}(N)\cdot H$ *of* $\mathrm{Aut}(N)$ *is closed and normal in* $\mathrm{Aut}(N)$ *and the map*

$$\mathrm{Centr}_{\mathrm{Aut}\,Y}(\Gamma_2) \to \frac{\mathrm{Aut}(N)}{\mathrm{Inn}(N)\cdot H} : \Delta \mapsto \theta_\Delta$$

is an isomorphism and homeomorphism of Polish groups.

PROOF. The proof of (1) is identical to [**31**, Lemma 5.1]. It remains to prove (2).

Write $A = \mathrm{L}^\infty(X)$ and $B = \mathrm{L}^\infty(Y)$. We first prove that every automorphism of N preserves the Cartan subalgebra $A \overline{\otimes} B$ up to unitary conjugacy. Together with point 1, this implies 2(a). So, let θ be an automorphism of $N := (A\overline{\otimes}B) \rtimes (\Gamma_1 * \Gamma_2)$. Take a projection $p \in A\overline{\otimes}B$ of finite trace and put $q = \theta(p)$. After unitary conjugacy, we may assume that $q \in A\overline{\otimes}B$. By [**29**, Theorem A.1], it is sufficient to prove that $\theta(Ap) \underset{qNq}{\prec} (A\overline{\otimes}B)q$.

Since $\theta(Ap) \subset qNq$ is rigid, [**19**, Theorem 5.1] implies that

$$\theta(Ap) \underset{qNq}{\prec} q\big((A\overline{\otimes}B) \rtimes \Gamma_i\big)q \quad \text{for some } \; i = 1,2.$$

Since $\theta(Ap)$ is quasi-regular in qNq, [**19**, Theorem 1.1] implies that $\theta(Ap) \underset{qNq}{\prec} (A\overline{\otimes}B)q$.

$^{(1)}$Note that H is isomorphic to the group of S^1-valued 1-cocycles for the action $\Gamma_1 * \Gamma_2 \curvearrowright X \times Y$.

We finally prove 2(b). Observe that $\mathcal{U}(N)$ is a Polish group in a natural way and that the map $\mathcal{U}(N) \to \operatorname{Aut}(N) : u \mapsto \operatorname{Ad} u$ is a continuous group morphism. Define H as in the formulation of the theorem and note that H is a closed subgroup of $\operatorname{Aut}(N)$. We form the semidirect product Polish group $\mathcal{U}(N) \rtimes H$ in such a way that $\pi : \mathcal{U}(N) \rtimes H \to \operatorname{Aut}(N) : \pi(u,\theta) = (\operatorname{Ad} u) \circ \theta$ is a group morphism. Note that π is continuous and denote $K := (\mathcal{U}(N) \rtimes H)/\operatorname{Ker}\pi$. Again, K is a Polish group. We form the semidirect product Polish group $K \rtimes \operatorname{Centr}_{\operatorname{Aut} Y}(\Gamma_2)$ in such a way that

$$\rho : K \rtimes \operatorname{Centr}_{\operatorname{Aut} Y}(\Gamma_2) \to \operatorname{Aut}(N) : \rho(k,\Delta) = \pi(k)\theta_\Delta$$

is a group morphism. Then ρ is a continuous and injective group morphism between Polish groups. Moreover, by (2a), ρ is onto. So, ρ is a homeomorphism. Hence, $\operatorname{Inn}(N) \cdot H = \rho(K)$ is closed and normal in $\operatorname{Aut}(N)$ and the map $\Delta \mapsto \theta_\Delta$ provides an isomorphism and homeomorphism between $\operatorname{Centr}_{\operatorname{Aut} Y}(\Gamma_2)$ and $\operatorname{Aut}(N)/(\operatorname{Inn}(N) \cdot H)$. $\square$

LEMMA 3.6. *Let $\Gamma * \Lambda \curvearrowright (X,\mu)$ be a free p.m.p. action with the restriction to Γ being ergodic. Let $\Lambda \curvearrowright (Y,\eta)$ be a free ergodic action preserving the infinite standard measure η. Assume that Λ is amenable. Consider $\Gamma * \Lambda \curvearrowright X \times Y$ as in* (3.1)*. Let $Z \subset Y$ be a subset of finite measure and define the II_1 equivalence relation $\mathcal{R}$ as the restriction of $\mathcal{R}(\Gamma * \Lambda \curvearrowright X \times Y)$ to Z.*

Whenever Σ is an infinite amenable group, there exists a free ergodic p.m.p. action $\Gamma^{*\infty} * \Sigma \curvearrowright X \times Z$ such that $\mathcal{R} = \mathcal{R}(\Gamma^{*\infty} * \Sigma \curvearrowright X \times Z)$.

PROOF. Denote by $\mathcal{R}_1$ the equivalence relation given by the restriction of $\mathcal{R}(\Lambda \curvearrowright X \times Y)$ to $X \times Z$. Note that $\mathcal{R}_1$ need not be ergodic. Since Λ is amenable and almost every equivalence class of $\mathcal{R}_1$ is infinite, the results in [**7**] and [**22**] allow one to take a free p.m.p. action $\Sigma \curvearrowright X \times Z$ whose OE relation is precisely $\mathcal{R}_1$.

Since the action of Λ on (Y,η) is ergodic, take $\phi_n \in [[\Lambda]]$ with $\operatorname{dom}(\phi_n) = X \times Z$ and $\operatorname{range}(\phi_n) = X \times Z_n$, where Z_n, $n \in \mathbb{N}$, forms a partition of Y (up to measure zero). Since the action of Γ leaves every $X \times Z_n$ globally invariant, we can view $\phi_n^{-1}\Gamma\phi_n$ as a group of automorphisms of $X \times Z$. It is now an exercise to check that $\mathcal{R}$ is freely generated by the OE relations of $\phi_n^{-1}\Gamma\phi_n$, $n \in \mathbb{N}$, together with $\Sigma \curvearrowright X \times Z$. This provides us with the required free action of $\Gamma^{*\infty} * \Sigma \curvearrowright X \times Z$. $\square$

The following is the final ingredient in the proof of Theorem 3.1.

LEMMA 3.7. *There exist ergodic measures ν, ν' on $\mathbb{R}$ such that $H_\nu \neq \mathbb{R} \neq H_{\nu'}$ and $H_\nu + H_{\nu'} = \mathbb{R}$.*

PROOF. As explained in [**31**, Section 2], an ergodic measure ν on $\mathbb{R}$ can be associated to any pair $(a_n), (b_n)$ of sequences in $\mathbb{N}$ satisfying $\sum_{n=1}^\infty b_n^{-1} < \infty$ and $b_n < a_n/2$ for all n, in such a way that

$$H_\nu = \Big\{ x \in \mathbb{R} \;\Big|\; \sum_{n=1}^\infty \frac{a_n}{b_n} \|a_1 \cdots a_{n-1} x\| < \infty \Big\}$$

where $\|x\|$ denotes the distance of $x \in \mathbb{R}$ to $\mathbb{Z} \subset \mathbb{R}$. Take $a_n = 2^{2n+2}$, $b_n = 2^{2n}$ and associate with it the ergodic measure ν. Take $a'_n = 2^{2n+1}$, $b'_n = 2^{2n-1}$ and associate with it the ergodic measure ν'. First of all,

$$\sum_{n=1}^\infty \frac{b_n}{a_1 \cdots a_n} \notin H_\nu \quad \text{and} \quad \sum_{n=1}^\infty \frac{b'_n}{a'_1 \cdots a'_n} \notin H_{\nu'}$$

proving that $H_\nu \neq \mathbb{R} \neq H_{\nu'}$.

Let now $x \in \mathbb{R}$ and write

$$x = x_0 + \sum_{n=1}^{\infty} \frac{x_n}{a_1 \cdots a_n} \quad \text{with} \quad x_n \in \{0, \ldots, a_n - 1\} \; .$$

Write for every $n \in \mathbb{N}$, $x_n = y_n + \sqrt{a_n} z_n$ with $y_n, z_n \in \{0, \ldots, \sqrt{a_n}\}$. Define

$$y = x_0 + \sum_{n=1}^{\infty} \frac{y_n}{a_1 \cdots a_n} \quad \text{and} \quad z = \sum_{n=1}^{\infty} \frac{2z_n}{a_1' \cdots a_n'} \; .$$

One checks that $y \in H_\nu$, $z \in H_{\nu'}$ and $x = y + z$. So, $H_\nu + H_{\nu'} = \mathbb{R}$. $\square$

Proof of Theorem 3.1. Since $\Gamma^{*\infty} = (\Gamma * \Gamma)^{*\infty}$, we may assume that Γ is an infinite group. Let Σ, Λ be infinite amenable groups and $\Lambda \curvearrowright (Y, \eta)$ a free ergodic action preserving the infinite standard measure η. Set $G := \Gamma^{*\infty} * \Sigma$. We prove the existence of a free ergodic p.m.p. action of $G \curvearrowright Z$ such that the associated II$_1$ factor $M := \mathrm{L}^\infty(Z) \rtimes G$ and equivalence relation $\mathcal{R} := \mathcal{R}(G \curvearrowright Z)$ have the following properties.

(1) The fundamental group of M and the fundamental group of $\mathcal{R}$ equal $\operatorname{mod}(\operatorname{Centr}_{\operatorname{Aut} Y}(\Lambda))$.

(2) The II$_\infty$ factor $M \overline{\otimes} \mathrm{B}(\ell^2(\mathbb{N}))$ admits a strongly continuous trace-scaling action of $\mathbb{R}_+$ if and only if the group morphism

$$\operatorname{mod} : \operatorname{Centr}_{\operatorname{Aut} Y}(\Lambda) \to \mathbb{R}_+ \tag{3.2}$$

is onto and splits continuously.

Choose any free p.m.p. actions $\Gamma \curvearrowright (X, \mu)$ and $\Lambda \curvearrowright (X, \mu)$. Take a free p.m.p. action $\Gamma^{*\infty} * \Lambda \curvearrowright (X, \mu)$ satisfying the conclusions of Lemma 3.3. Define $\Gamma^{*\infty} * \Lambda \curvearrowright X \times Y$ by (3.1), with $\Gamma_1 = \Gamma^{*\infty}$ and $\Gamma_2 = \Lambda$. Define the II$_1$ equivalence relation $\mathcal{R}$ by restricting $\mathcal{R}(\Gamma^{*\infty} * \Lambda \curvearrowright X \times Y)$ to a subset Z of finite measure. By Lemma 3.6, we can take a free ergodic p.m.p. action $G \curvearrowright Z$ whose OE relation equals $\mathcal{R}$. By (1) of Theorem 3.5,

$$\mathcal{F}(\mathcal{R}) = \operatorname{mod}(\operatorname{Centr}_{\operatorname{Aut} Y}(\Lambda)) \; .$$

Put $M := \mathrm{L}^\infty(Z) \rtimes G$ and note that $M \overline{\otimes} \mathrm{B}(\ell^2(\mathbb{N})) \cong \mathrm{L}^\infty(X \times Y) \rtimes (\Gamma^{*\infty} * \Lambda)$. By (2a) of Theorem 3.5, also

$$\mathcal{F}(M) = \operatorname{mod}(\operatorname{Aut}(N)) = \operatorname{mod}(\operatorname{Centr}_{\operatorname{Aut} Y}(\Lambda)) \; .$$

If the group morphism (3.2) splits continuously, it is clear that N admits a strongly continuous trace scaling action. The converse follows from (2b) of Theorem 3.5.

In order to conclude the proof of Theorem 3.1, we have to construct an action $\Lambda \curvearrowright (Y, \eta)$ such that $\operatorname{mod}(\operatorname{Centr}_{\operatorname{Aut} Y}(\Lambda)) = \mathbb{R}_+$, but the morphism (3.2) does not split continuously. By Lemma 3.7, we can take ergodic measures ν_1, ν_2 on $\mathbb{R}$ such that $H_{\nu_1} \neq \mathbb{R} \neq H_{\nu_2}$, while $H_{\nu_1} + H_{\nu_2} = \mathbb{R}$. By formula (2.2) in [**31**], we can take amenable groups Λ_1, Λ_2 and free ergodic infinite measure preserving actions $\Lambda_i \curvearrowright (Y_i, \eta_i)$ such that $\operatorname{mod}(\operatorname{Centr}_{\operatorname{Aut} Y_i}(\Lambda_i)) = \exp(H_{\nu_i})$. Since the homomorphism mod is continuous, we equip $\exp(H_{\nu_i})$ with the (Polish) quotient topology. In this way, the H_{ν_i} become Polish groups and the embedding $H_{\nu_i} \hookrightarrow \mathbb{R}$ continuous.

We prove that

$$\operatorname{mod} : \operatorname{Centr}_{\operatorname{Aut}(Y_1 \times Y_2)}(\Lambda_1 \times \Lambda_2) \to \mathbb{R}_+$$

admits no continuous splitting. Assume that it does. Since the left-hand side of the previous formula equals $\mathrm{Centr}_{\mathrm{Aut}\,Y_1}(\Lambda_1) \times \mathrm{Centr}_{\mathrm{Aut}\,Y_2}(\Lambda_2)$, the homomorphism

$$H_{\nu_1} \times H_{\nu_2} \to \mathbb{R} : (x, y) \mapsto x + y$$

admits a continuous splitting. We then find continuous homomorphisms $\theta_i : \mathbb{R} \to H_{\nu_i}$ such that $x = \theta_1(x) + \theta_2(x)$ for all $x \in \mathbb{R}$. Since the embedding $H_{\nu_i} \hookrightarrow \mathbb{R}$ is continuous, there exist $\lambda_i \in \mathbb{R}$ such that $\theta_i(x) = \lambda_i x$ for all $x \in \mathbb{R}$. But, $H_{\nu_i} \neq \mathbb{R}$, forcing $\lambda_i = 0$ for $i = 1, 2$, a contradiction. □

PROOF OF THEOREM 3.2. By Lemma 3.7, we can take ergodic measures ν_1, ν_2 on $\mathbb{R}$ such that $H_{\nu_1} \neq \mathbb{R} \neq H_{\nu_2}$, while $H_{\nu_1} + H_{\nu_2} = \mathbb{R}$. By formula (2.2) in [**31**], $\exp(H_{\nu_i}) \in \mathcal{S}_{\mathrm{centr}}$, so that by Theorem 3.1, we can take II_1 factors M_i with $\mathcal{F}(M_i) = \exp(H_{\nu_i})$. Then the fundamental group of $M_1 \overline{\otimes} M_2$ contains $\exp(H_{\nu_1} + H_{\nu_2})$ and hence equals $\mathbb{R}_+$. □

4. Groups G for which $\mathcal{S}_{\mathrm{factor}}(G)$ is trivial

Combining results from [**3, 15, 19**], we prove that for the following groups G, $\mathcal{S}_{\mathrm{factor}}(G)$ is trivial.

THEOREM 4.1. *Let Γ and Λ be infinite, finitely generated groups. Assume that Γ is ICC and that one of the following conditions holds.*

a) $\Gamma = \Gamma_1 \times \Gamma_2$ *is a non-trivial direct product with Γ_2 being non-amenable,*
b) Γ *admits a non-virtually abelian, normal subgroup Γ_1 with the relative property (T).*

Then $\mathcal{S}_{\mathrm{factor}}(\Gamma * \Lambda) = \mathcal{S}_{\mathrm{eqrel}}(\Gamma * \Lambda) = \{\{1\}\}$.

REMARK 4.2. Observe that Theorem 4.1 implies that, in general, Theorem 3.1 is false if we only take a finite free product.

PROOF OF THEOREM 4.1. Let $\Gamma * \Lambda \curvearrowright (X, \mu)$ be a free ergodic p.m.p. action. Note first that by [**15**, Propriétés 1.5], we have $0 < \beta_1^{(2)}(\Gamma * \Lambda) < \infty$. Hence, by [**15**, Corollaire 5.7], the fundamental group of the OE relation $\mathcal{R}(\Gamma * \Lambda \curvearrowright X)$ is trivial.

Write $A = \mathrm{L}^\infty(X)$, $M_1 = A \rtimes \Gamma$, $M_2 = A \rtimes \Lambda$. Finally, set $M = A \rtimes (\Gamma * \Lambda) = M_1 *_A M_2$. Suppose that $p \in A$ is a projection and $\theta : M \to pMp$ a *-isomorphism. It remains to prove that $\theta(A)$ and pA are unitarily conjugate, since this implies that $\mathcal{F}(M) = \mathcal{F}(\mathcal{R}(\Gamma * \Lambda \curvearrowright X))$.

Under assumption a), we invoke [**3**, Theorem 4.2] and under assumption b), we invoke [**19**, Theorem 5.1] and conclude in both cases that $\theta(L(\Gamma_1)) \underset{M}{\prec} M_i$ for some $i = 1, 2$. Take a projection $q \in M_i^n$, a non-zero partial isometry $v \in p(\mathrm{M}_{1,n}(\mathbb{C}) \otimes M)q$ and a unital *-homomorphism $\rho : L(\Gamma_1) \to qM_i^n q$ satisfying $\theta(a)v = v\rho(a)$ for all $a \in L(\Gamma_1)$. In both cases a) and b), the group Γ_1 is not virtually abelian. Hence, $\rho(L(\Gamma_1)) \underset{M_i}{\not\prec} A$. By [**19**, Theorem 1.1], the normalizer of $\rho(L(\Gamma_1))$ inside $qM^n q$ is contained in $qM_i^n q$. Since v^*v commutes with $\rho(L(\Gamma_1))$, we may first of all assume that $q = v^*v$. Next, it follows that $v^*\theta(L(\Gamma))v \subset qM_i^n q$. Hence, $\theta(L(\Gamma)) \underset{M}{\prec} M_i$.

Repeating the previous paragraph, we may assume that $\rho : L(\Gamma) \to qM_i^n q$, $\theta(a)v = v\rho(a)$ for all $a \in L(\Gamma)$ and $v^*v = q$. Since Γ is an ICC group, we get that

$$M \cap L(\Gamma)' = M_1 \cap L(\Gamma)' = A^\Gamma \ .$$

So, $vv^* \in \theta(A^\Gamma)$. It follows that $v^*\theta(A)v$ is a Cartan subalgebra of qM^nq. Moreover, for all $g \in \Gamma$, the unitary $v^*\theta(u_g)v = \rho(u_g)$ belongs to qM_i^nq and normalizes $v^*\theta(A)v$. Then, [**19**, Theorem 1.8] implies that there exists $w \in qM^n$ such that $ww^* = q$, $w^*w \in A^n$ and $w^*v^*\theta(A)vw = w^*wA^n$. It follows that $\theta(A)$ and pA are unitarily conjugate. $\square$

5. $\mathcal{S}_{\mathrm{factor}}(\mathbb{Z}^n \rtimes \mathrm{SL}(n,\mathbb{Z}))$ is non-trivial, for all $n \geq 3$

When Γ is an ICC property (T) group, all groups in $\mathcal{S}_{\mathrm{factor}}(\Gamma)$ are countable (cf. [**17**, Proof of Theorem 1.7], or [**25**, Theorem 4.5.1]). Nevertheless, $\mathcal{S}_{\mathrm{factor}}(\Gamma)$ can be non-trivial, as shown by the next theorem, in which we show that if $\Gamma = \mathbb{Z}^n \rtimes \mathrm{SL}(n,\mathbb{Z})$, $n \geq 3$, then $\mathcal{S}_{\mathrm{factor}}(\Gamma)$ contains "many" subgroups of $\mathbb{Q}_+$. It is unclear though whether there exists a free ergodic p.m.p. action $\Gamma \curvearrowright (X,\mu)$ of an ICC property (T) group Γ such that $\mathcal{F}(\mathcal{R}(\Gamma \curvearrowright X)) \not\subset \mathbb{Q}_+$.

THEOREM 5.1. *Let $\mathcal{F} \subset \mathbb{Q}_+$ be a subgroup generated by a subset of the prime numbers. Let $\Gamma = \mathbb{Z}^n \rtimes \mathrm{SL}(n,\mathbb{Z})$ with $n \geq 3$. Then Γ admits a "concrete" free ergodic p.m.p. action $\Gamma \curvearrowright (X,\mu)$ such that both the fundamental group of $\mathrm{L}^\infty(X) \rtimes \Gamma$ and of $\mathcal{R}(\Gamma \curvearrowright X)$ equal $\mathcal{F}$.*

We prove Theorem 5.1 as a consequence of the following more general result.

THEOREM 5.2. *Let Γ be a group having a normal, non-virtually abelian subgroup Σ with the relative property (T) and with Γ/Σ being finitely generated. Let $\Gamma \supset \Gamma_1 \supset \Gamma_2 \supset \cdots$ be a decreasing sequence of finite index subgroups such that the action $\Gamma \curvearrowright (X,\mu) := \varprojlim \Gamma/\Gamma_n$ is essentially free. Consider the diagonal product action $\Gamma \curvearrowright X \times [0,1]^\Gamma$ of $\Gamma \curvearrowright X$ and the Bernoulli action $\Gamma \curvearrowright [0,1]^\Gamma$.*

Then the fundamental groups of the associated II_1 factor and II_1 equivalence relation are both equal to

$$\left\{ \frac{[\Gamma:\Lambda_1]}{[\Gamma:\Lambda_2]} \;\middle|\; \Gamma_n \subset \Lambda_1 \cap \Lambda_2 \ \textit{for large enough } n, \textit{ and } \Lambda_1 \curvearrowright \varprojlim \Lambda_1/\Gamma_n \ \textit{is conjugate with} \ \Lambda_2 \curvearrowright \varprojlim \Lambda_2/\Gamma_n \right\}. \tag{5.1}$$

Conjugacy of two profinite actions can be expressed in purely group-theoretic terms; see e.g. [**18**, Proposition 1.8].

Before proving Theorem 5.2, we introduce some terminology and an auxiliary result. Recall that a 1-*cocycle* $\omega : \Gamma \times X \to \Lambda$ for an action $\Gamma \curvearrowright (X,\mu)$ with values in a countable group Λ is a measurable map satisfying

$$\omega(gh,x) = \omega(g,h\cdot x)\,\omega(h,x) \quad \text{for all } g,h \in \Gamma \text{ and almost all } x \in X\,.$$

The 1-cocycles $\omega, \omega' : \Gamma \times X \to \Lambda$ are called *cohomologous* if there exists a measurable map $\varphi : X \to \Lambda$ satisfying $\omega'(g,x) = \varphi(g\cdot x)\omega(g,x)\varphi(x)^{-1}$ almost everywhere. We identify homomorphisms from Γ to Λ with 1-cocyles ω that are independent of the x-variable.

DEFINITION 5.3. Let $\Gamma \curvearrowright (X,\mu)$ be a free ergodic p.m.p. action. We say that a 1-cocycle $\omega : \Gamma \times X \to \Lambda$ *virtually untwists* if there exists

- a finite index subgroup $\Gamma_0 < \Gamma$ and a measurable map $\pi : X \to \Gamma/\Gamma_0$ satisfying $\pi(g\cdot x) = g\pi(x)$ almost everywhere,
- a 1-cocycle $\omega' : \Gamma \times \Gamma/\Gamma_0 \to \Lambda$ for the action $\Gamma \curvearrowright \Gamma/\Gamma_0$,

such that ω is cohomologous to the 1-cocycle $(g, x) \mapsto \omega'(g, \pi(x))$.

We call $\Gamma \curvearrowright (X, \mu)$ *virtually cocycle superrigid* (with countable target groups) if every 1-cocycle with values in a countable group Λ virtually untwists.

A *stable orbit equivalence* between free ergodic p.m.p. actions $\Gamma \curvearrowright (X, \mu)$ and $\Lambda \curvearrowright (X', \mu')$ is a map $\Delta : X \to X'$ satisfying the following properties.

- For almost every $x \in X$, we have $\Delta(\Gamma \cdot x) = \Lambda \cdot \Delta(x)$.
- There exists a partition $X = \bigsqcup_n X_n$ of X into measurable subsets $X_n \subset X$ and there exist measurable subsets $X'_n \subset X'$ such that for every $n \in \mathbb{N}$, the restriction of Δ to X_n is a non-singular isomorphism between X_n and X'_n.

By ergodicity, all of these non-singular isomorphisms $\Delta|_{X_n} : X_n \to X'_n$ are measure scaling, with the scaling being independent of n. This scaling factor is called the *compression constant* of the stable OE Δ and denoted by $c(\Delta)$.

The *Zimmer* 1-*cocycle* $\omega : \Gamma \times X \to \Lambda$ associated with the stable OE Δ is defined by

$$\Delta(g \cdot x) = \omega(g, x) \cdot \Delta(x) \quad \text{almost everywhere.}$$

Two stable OEs $\Delta_1, \Delta_2 : X \to X'$ are called *similar* if $\Delta_1(x) \in \Lambda \cdot \Delta_2(x)$ for almost all $x \in X$. Note that similar stable OEs give rise to cohomologous 1-cocycles.

Whenever $X_0 \subset X$ and $X'_0 \subset X'$ are non-negligible measurable subsets and $\Delta_0 : X_0 \to X'_0$ is a non-singular isomorphism satisfying $\Delta_0(\Gamma \cdot x \cap X_0) = \Lambda \cdot \Delta_0(x) \cap X'_0$ for almost all $x \in X_0$, ergodicity allows one to choose a measurable map $p : X \to X_0$ with $p(x) \in \Gamma \cdot x$ for almost all $x \in X$ and then, $\Delta := \Delta_0 \circ p$ defines a stable OE. Another choice of p gives rise to a similar stable OE. It follows that

$$\mathcal{F}(\mathcal{R}(\Gamma \curvearrowright X)) = \{c(\Delta) \mid \Delta \text{ is a stable OE between } \Gamma \curvearrowright X \text{ and } \Gamma \curvearrowright X\} .$$

Let $\Gamma \curvearrowright (X, \mu)$. We say that the action $\Gamma \curvearrowright X$ is *induced* from $\Gamma_1 \curvearrowright X_1$ if X_1 is a non-neglible measurable subset of X and $\Gamma_1 < \Gamma$ is a finite index subgroup such that $g \cdot X_1 = X_1$ for all $g \in \Gamma_1$ and $\mu(g \cdot X_1 \cap X_1) = 0$ if $g \in \Gamma - \Gamma_1$. Obviously, in this situation $\Gamma \curvearrowright X$ is stably orbit equivalent to $\Gamma_1 \curvearrowright X_1$ with compression constant $[\Gamma : \Gamma_1]^{-1}$.

The following provides one more instance of a general principle going back to [**38**, Proposition 4.2.11]. For other versions of this, see [**28**, Proposition 5.11] and [**37**, Lemma 4.7].

PROPOSITION 5.4. *Let $\Delta : X \to X'$ be a stable OE between the free ergodic p.m.p. actions $\Gamma \curvearrowright (X, \mu)$ and $\Lambda \curvearrowright (X', \mu')$. If the associated Zimmer* 1*-cocycle virtually untwists (see Definition 5.3), there exist finite index subgroups $\Gamma_1 < \Gamma$, $\Lambda_1 < \Lambda$, non-negligible measurable subsets $X_1 \subset X$, $X'_1 \subset X'$ and a finite normal subgroup $H \lhd \Gamma_1$ such that*

1. *$\Gamma \curvearrowright X$ is induced from $\Gamma_1 \curvearrowright X_1$,*
2. *$\Lambda \curvearrowright Y$ is induced from $\Lambda_1 \curvearrowright Y_1$,*
3. *the actions $\Gamma_1/H \curvearrowright X_1/H$ and $\Lambda_1 \curvearrowright Y_1$ are conjugate,*

and such that the stable OE Δ is similar to the composition of the canonical stable OEs given by (1), (3) and (2). In particular, the compression constant of Δ equals

$$c(\Delta) = \frac{[\Lambda : \Lambda_1]}{[\Gamma : \Gamma_1]|H|} .$$

PROOF. Let $\Delta(g \cdot x) = \omega(g,x) \cdot \Delta(x)$ almost everywhere. By our assumption, take a finite index subgroup Γ_1, a quotient map $\pi : X \to \Gamma/\Gamma_1$ and a 1-cocycle $\omega' : \Gamma \times \Gamma/\Gamma_1 \to \Lambda$ such that $\pi(g \cdot x) = g\pi(x)$ almost everywhere and such that ω is cohomologous to the 1-cocycle $(g,x) \mapsto \omega'(g,\pi(x))$. Define $X_1 = \pi^{-1}(e\Gamma_1)$. By construction, $\Gamma \curvearrowright X$ is induced from $\Gamma_1 \curvearrowright X_1$.

Denote by Δ_1 the restriction of Δ to X_1. Then Δ_1 is a stable OE between $\Gamma_1 \curvearrowright X_1$ and $\Lambda \curvearrowright Y$. By construction, the 1-cocycle associated with Δ_1 is cohomologous to a homomorphism from Γ_1 to Λ. The conclusion of the proposition now follows from [**37**, Lemma 4.7]. $\square$

In order to show the equality of the fundamental groups of the II_1 factor and the II_1 equivalence relation associated with the Γ-actions defined in Theorem 5.2, we need the following result about automatic preservation of Cartan subalgebras.

PROPOSITION 5.5. *Let Γ be a countable group having a normal, non-virtually abelian subgroup Σ with the relative property (T). Let $\Gamma \curvearrowright (X,\mu)$ be a free ergodic p.m.p. action and assume that this action admits a free and profinite quotient: there exists a free profinite p.m.p. action $\Gamma \curvearrowright (X_1,\mu_1)$ and a quotient map $\pi : X \to X_1$ satisfying $\pi(g \cdot x) = g \cdot \pi(x)$ almost everywhere.*

Let (Y_0,η_0) be a non-trivial standard probability space and set $(Y,\eta) = (Y_0,\eta_0)^\Gamma$. Consider the diagonal action $\Gamma \curvearrowright (X \times Y, \mu \times \eta)$. Set $M = \mathrm{L}^\infty(X \times Y) \rtimes \Gamma$.

Then every isomorphism $\theta : M \to pMp$ preserves, up to unitary conjugacy, the natural Cartan subalgebras of M, pMp.

PROOF. Set $A = \mathrm{L}^\infty(X)$ and $B = \mathrm{L}^\infty(Y)$. Let $\theta : M \to pMp$ be an isomorphism. Denote $A_1 = \mathrm{L}^\infty(X_1)$ and view A_1 as a globally Γ-invariant von Neumann subalgebra of A.

Almost literally repeating [**26**, Theorem 4.1] (see also [**37**, Lemma 6.1]), we find that $\theta(L(\Sigma)) \underset{M}{\prec} A \rtimes \Gamma$. Take a projection $q \in (A \rtimes \Gamma)^n$, a non-zero partial isometry $v \in p(\mathrm{M}_{1,n}(\mathbb{C}) \otimes M)q$ and a unital *-homomorphism $\alpha : L(\Sigma) \to q(A \rtimes \Gamma)^n q$ such that $\theta(a)v = v\alpha(a)$ for all $a \in L(\Sigma)$.

Since Σ is normal in Γ and since $\Sigma \curvearrowright X_1$ is profinite, the quasi-normalizer of $L(\Sigma)$ inside M contains $A_1 \rtimes \Gamma$. Since Σ is non-virtually abelian, $L(\Sigma)$ cannot be embedded in an amplification of A. So, by [**37**, Proposition D.5],

$$v^*\theta(A_1 \rtimes \Gamma)v \subset (A \rtimes \Gamma)^n .$$

It follows in particular that $\theta(A_1) \underset{M}{\prec} A \rtimes \Gamma$. We claim that in fact $\theta(A_1) \underset{M}{\prec} A$. Indeed, if this were not the case, applying once more [**37**, Proposition D.5] (and using the regularity of $A_1 \subset (A \overline{\otimes} B) \rtimes \Gamma$) would yield $M \underset{M}{\prec} A \rtimes \Gamma$, a contradiction. This proves the claim.

Since $\Gamma \curvearrowright X_1$ is free, we have $A_1' \cap M = A \overline{\otimes} B$. Hence, the proposition follows from [**29**, Theorem A.1]. $\square$

We are now ready to prove Theorem 5.2.

PROOF OF THEOREM 5.2. Put $(X,\mu) = \varprojlim \Gamma/\Gamma_n$ as in the formulation of the theorem. We assume $\Gamma \curvearrowright X$ to be essentially free. Let $Y = [0,1]^\Gamma$ and denote by η the infinite product of the Lebesgue measure on $[0,1]$. We consider the diagonal action $\Gamma \curvearrowright X \times Y$.

By Proposition 5.5, we have

$$\mathcal{F}(\mathrm{L}^\infty(X \times Y) \rtimes \Gamma) = \mathcal{F}(\mathcal{R}(\Gamma \curvearrowright X \times Y)) \; .$$

Whenever $\Lambda_1 < \Gamma$ is a subgroup containing Γ_n for large enough n, the action $\Gamma \curvearrowright (X,\mu)$ is induced from the action $\Lambda_1 \curvearrowright X_1 := \varprojlim \Lambda_1/\Gamma_n$, and hence $\Gamma \curvearrowright X \times Y$ is induced from $\Lambda_1 \curvearrowright X_1 \times Y$. Since $\Lambda_1 \curvearrowright [0,1]^\Gamma$ and $\Lambda_1 \curvearrowright [0,1]^{\Lambda_1}$ are isomorphic actions, it follows that the set defined by (5.1) is part of the fundamental group $\mathcal{F}(\mathcal{R}(\Gamma \curvearrowright X \times Y))$.

A combination of [**28**, Theorem 0.1] and [**18**, Theorem B] yields that the diagonal action $\Gamma \curvearrowright X \times Y$ is virtually cocycle superrigid in the sense of Definition 5.3. So, we can apply Proposition 5.4.

Let $\Delta : X \times Y \to X \times Y$ be a stable OE between $\Gamma \curvearrowright X \times Y$ and itself. We have to prove that $c(\Delta)$ belongs to the set defined in (5.1). Proposition 5.4 provides us with finite index subgroups G_1, G_2 of Γ, non-negligible measurable subsets $Z_1, Z_2 \subset X \times Y$ and a finite normal subgroup H of G_1 such that $\Gamma \curvearrowright X \times Y$ is induced from $G_i \curvearrowright Z_i$ and such that $G_1/H \curvearrowright Z_1/H$ is conjugate to $G_2 \curvearrowright Z_2$, say through the isomorphism $\Delta : Z_1/H \to Z_2$ and the group isomorphism $\delta : G_1 \to G_2$. Finally, $c(\Delta) = \frac{[\Gamma:G_2]}{[\Gamma:G_1]|H|}$.

Since the Bernoulli action $G_i \curvearrowright Y$ is mixing, we have $Z_i = X_i \times Y$, with $\Gamma \curvearrowright X$ being induced from $G_i \curvearrowright X_i$. Moreover, still because the Bernoulli action is mixing, $\Delta(x,y) = (\Delta_0(x), \ldots)$, where $\Delta_0 : X_1/H \to X_2$ is an isomorphism conjugating the actions $G_1/H \curvearrowright X_1/H$ and $G_2 \curvearrowright X_2$ through the group isomorphism $\delta : G_1/H \to G_2$.

Denote by $\pi_n : X \to \Gamma/\Gamma_n$ the natural quotient map. By [**18**, Lemma 4.1], we find $k \in \mathbb{N}$ and $g \in \Gamma$ such that $g\Gamma_k g^{-1} \subset G_1$ and $X_1 = \pi_k^{-1}(G_1 g \Gamma_k)$. Moreover, since $\Gamma \curvearrowright X$ is free, we can take k large enough and assume that $H \cap \Gamma_k = \{e\}$. Replacing G_1 by $g^{-1}G_1 g$ and X_1 by $g^{-1} \cdot X_1$, we may assume that $g = e$. Note that $G_1/H \curvearrowright X_1/H = \varprojlim G_1/(\Gamma_n H)$ is induced from $(\Gamma_k H)/H \curvearrowright \varprojlim (\Gamma_k H)/(\Gamma_n H)$ and that the latter is conjugate to $\Gamma_k \curvearrowright \varprojlim \Gamma_k/\Gamma_n$, because $\Gamma_k \cap H = \{e\}$.

It follows that $\Gamma \curvearrowright X$ is induced from $\delta((\Gamma_k H)/H) \curvearrowright \Delta_0(\varprojlim (\Gamma_k H)/(\Gamma_n H))$. Applying as above [**18**, Lemma 4.11], we find $h \in \Gamma$ such that, after replacing δ by $g \mapsto h\delta(g)h^{-1}$ and Δ_0 by $x \mapsto h \cdot \Delta_0(x)$, we have $\Gamma_n \subset \Lambda_2 := \delta((\Gamma_k H)/H)$ for n large enough and

$$\Delta_0(\varprojlim (\Gamma_k H)/(\Gamma_n H)) = \varprojlim \Lambda_2/\Gamma_n \; .$$

Denoting $\Lambda_1 = \Gamma_k$, we have constructed finite index subgroups $\Lambda_1, \Lambda_2 \subset \Gamma$ such that $\Gamma_n \subset \Lambda_1 \cap \Lambda_2$ for n large enough and such that the actions $\Lambda_i \curvearrowright \varprojlim \Lambda_i/\Gamma_n$ are conjugate for $i = 1, 2$. Tracing back the construction, we also have

$$c(\Delta) = \frac{[\Gamma : \Lambda_2]}{[\Gamma : \Lambda_1]}$$

concluding the proof of the theorem. □

Proof of Theorem 5.1. Let $\mathcal{F}$ be a subgroup of $\mathbb{Q}_+$ generated by a non-empty subset $\mathcal{P}$ of the prime numbers. The case $\mathcal{F} = \{1\}$ will be discussed at the end of the proof. Denote by R the subring of $\mathbb{Q}$ generated by $\mathcal{P}^{-1}$. Note that $R^* = \mathcal{F} \cup (-\mathcal{F})$. Set $G = R^n \rtimes \mathrm{GL}(n, R)$ and $\Gamma = \mathbb{Z}^n \rtimes \mathrm{SL}(n, \mathbb{Z})$. Let $G = \{g_1, g_2, \ldots\}$ and define the finite index subgroups $\Gamma_k < \Gamma$ as $\Gamma_k = \Gamma \cap \bigcap_{i=1}^k g_i \Gamma g_i^{-1}$. Define the profinite action $\Gamma \curvearrowright (X,\mu) := \varprojlim \Gamma/\Gamma_k$.

We first argue why $\Gamma \curvearrowright (X,\mu)$ is essentially free. Let $p \in \mathcal{P}$ and take $k_1 < k_2 < k_3 < \cdots$ such that $(0, p^l 1) \in \{g_1, \ldots, g_{k_l}\}$. One checks that $\Gamma_{k_l} \subset G_l := p^l \mathbb{Z}^n \rtimes \mathrm{SL}(n,\mathbb{Z})$, and hence it suffices to prove freeness of $\Gamma \curvearrowright \varprojlim \Gamma/G_l$. The latter has been shown in [**18**, discussion before Corollary 5.8].

Consider the diagonal action $\Gamma \curvearrowright X \times [0,1]^\Gamma$. Denote by $\mathcal{F}$ the set defined in (5.1). By Theorem 5.2, we have to prove that $\mathcal{F} = R^*_+$. It is more convenient to write $X = \varprojlim \Gamma/\Gamma_F$, where F runs through the finite subsets of G and $\Gamma_F := \Gamma \cap \bigcap_{g \in F} g\Gamma g^{-1}$. Whenever $g \in G$, the action $\Gamma \curvearrowright X$ is induced from

$$\Gamma \cap g\Gamma g^{-1} \curvearrowright \varprojlim_{g \in F \subset G} \frac{\Gamma \cap g\Gamma g^{-1}}{\Gamma_F}$$

and is induced from

$$g^{-1}\Gamma g \cap \Gamma \curvearrowright \varprojlim_{g^{-1} \in F \subset G} \frac{g^{-1}\Gamma g \cap \Gamma}{\Gamma_F} \; .$$

The two actions are conjugate by construction. If $g = (x, A)$, one checks that

$$\frac{[\Gamma : \Gamma \cap g\Gamma g^{-1}]}{[\Gamma : g^{-1}\Gamma g \cap \Gamma]} = |\det A| \; .$$

It follows that $R^*_+ \subset \mathcal{F}$.

Conversely, we claim that whenever $\Lambda_1, \Lambda_2 < \Gamma$ are isomorphic finite index subgroups of Γ such that $\Gamma_k \subset \Lambda_1 \cap \Lambda_2$ for some k, then $[\Gamma : \Lambda_1]/[\Gamma : \Lambda_2] \in R^*_+$. Once this claim is proven, we get the required equality $\mathcal{F} = R^*_+$. Let $\delta : \Lambda_1 \to \Lambda_2$ be an isomorphism and $\Gamma_k \subset \Lambda_1 \cap \Lambda_2$. We have $\delta(\Lambda_1 \cap \mathbb{Z}^n) = \Lambda_2 \cap \mathbb{Z}^n$. An elementary argument for this fact can be given by repeating the beginning of the proof of [**32**, Proposition 7.1]. Denoting by $\pi : \Gamma \to \mathrm{SL}(n,\mathbb{Z})$ the quotient map, $\pi(\Lambda_1)$ and $\pi(\Lambda_2)$ are isomorphic finite index subgroups of $\mathrm{SL}(n,\mathbb{Z})$. Using [**18**, Lemma 5.2], it follows that $[\mathrm{SL}(n,\mathbb{Z}) : \pi(\Lambda_1)] = [\mathrm{SL}(n,\mathbb{Z}) : \pi(\Lambda_2)]$. Hence, we get

$$\frac{[\Gamma : \Lambda_1]}{[\Gamma : \Lambda_2]} = \frac{[\mathrm{SL}(n,\mathbb{Z}) : \pi(\Lambda_1)]\,[\mathbb{Z}^n : \mathbb{Z}^n \cap \Lambda_1]}{[\mathrm{SL}(n,\mathbb{Z}) : \pi(\Lambda_2)]\,[\mathbb{Z}^n : \mathbb{Z}^n \cap \Lambda_2]} = \frac{[\mathbb{Z}^n : \mathbb{Z}^n \cap \Lambda_1]}{[\mathbb{Z}^n : \mathbb{Z}^n \cap \Lambda_2]} \; .$$

Being finite index subgroups of $\mathbb{Z}^n$, we have $\Lambda_i \cap \mathbb{Z}^n = B_i \mathbb{Z}^n$ for some $B_i \in \mathrm{M}_n(\mathbb{Z})$ with $\det B_i \neq 0$, $i = 1, 2$. It follows that there exists $A \in \mathrm{GL}(n,\mathbb{Q})$ such that $\delta(x,1) = (Ax, 1)$ for all $(x,1) \in \Lambda_1 \cap \mathbb{Z}^n$. Hence,

$$\frac{[\mathbb{Z}^n : \mathbb{Z}^n \cap \Lambda_1]}{[\mathbb{Z}^n : \mathbb{Z}^n \cap \Lambda_2]} = \frac{[\mathbb{Z}^n : \mathbb{Z}^n \cap A^{-1}\mathbb{Z}^n]}{[\mathbb{Z}^n : \mathbb{Z}^n \cap A\mathbb{Z}^n]} \; .$$

Since

$$\mathbb{Z}^n \cap \Gamma_k \subset \mathbb{Z}^n \cap \Lambda_1 \cap \Lambda_2 \subset \mathbb{Z}^n \cap A\mathbb{Z}^n \cap A^{-1}\mathbb{Z}^n \; ,$$

we find $\alpha \in R^* \cap (\mathbb{N} - \{0\})$ such that $\alpha \mathbb{Z}^n \subset \mathbb{Z}^n \cap A\mathbb{Z}^n \cap A^{-1}\mathbb{Z}^n$ for $i = 1, 2$. We conclude that $A \in \mathrm{GL}(n, R)$ and finally,

$$\frac{[\Gamma : \Lambda_1]}{[\Gamma : \Lambda_2]} = |\det A|^{-1} \in R^*_+ \; .$$

To conclude the proof of the theorem, we need to construct a free ergodic p.m.p. action $\Gamma \curvearrowright (X,\mu)$ such that the associated II_1 factor has trivial fundamental group. By [**27**, Corollary 0.2], the Bernoulli action $\Gamma \curvearrowright [0,1]^\Gamma$ has this property. Other examples can be given as follows. Let $p_1, p_2, \ldots$ be an enumeration of the prime numbers and set $\Gamma_k = p_1 \cdots p_k \mathbb{Z}^n \rtimes \mathrm{SL}(n,\mathbb{Z})$. By Theorem 5.2 and [**18**, Corollary

5.8], the diagonal product of the Bernoulli action $\Gamma \curvearrowright [0,1]^\Gamma$ and the profinite action $\Gamma \curvearrowright \varprojlim \Gamma/\Gamma_k$, provides a crossed product II_1 factor with trivial fundamental group. □

6. Property (T) and countability of the fundamental group

In his celebrated "rigidity paper" [**4**], Connes showed that II_1 factors arising from ICC groups with the property (T) of Kazhdan have countable fundamental group. Using the same ideas, it was later shown that, for a separable II_1 factor M to have countable $\mathcal{F}(M)$, it is in fact sufficient that M contains a subfactor with the property (T) in the sense of [**9**] and having trivial relative commutant, $N' \cap M = \mathbb{C}$ (cf. Theorem 4.6.1 in [**25**]; see also [**21**] for a more general statement). It was also shown that if M is a separable II_1 factor, then the family of subfactors $N_i \subset M, i \in I$, having property (T) and trivial relative commutant, is countable modulo conjugacy by unitaries in M (cf. Theorem 4.5.1 in [**25**]; see also [**23**] for a related result). In this section, we prove some analogous results for II_1 equivalence relations.

In particular, these results show that given any Kazhdan group Γ, $\mathcal{S}_{\mathrm{eqrel}}(\Gamma)$ can only contain countable subgroups of $\mathbb{R}_+$, and if in addition Γ is ICC then the same holds true for $\mathcal{S}_{\mathrm{factor}}(\Gamma)$. More generally, Part 1 of Theorem 6.4 below shows that this is still the case if the center of Γ "virtually coincides" with its *FC radical* (as defined before 6.4). However, if one drops this assumption on Γ, then the situation becomes quite complicated. Thus, Part 2 of Theorem 6.4 shows that if a property (T) group Γ is residually finite and has non-virtually abelian FC radical, then Γ admits free ergodic p.m.p. actions $\Gamma \curvearrowright (X,\mu)$ such that $\mathrm{L}^\infty(X) \rtimes \Gamma$ is McDuff and hence its fundamental group is equal to $\mathbb{R}_+$.

We first need some notation. Thus, if $\mathcal{R}$ is a II_1 equivalence relation on the standard probability space (X,μ), then we denote by $[\mathcal{R}]$ the *full group* of the equivalence relation $\mathcal{R}$, consisting of all non-singular isomorphisms $\phi : X \to X$ satisfying $(x,\phi(x)) \in \mathcal{R}$ for almost all $x \in X$. The *full pseudogroup* of $\mathcal{R}$ is denoted by $[[\mathcal{R}]]$ and consists of all non-singular partial automorphisms ϕ between measurable subsets $D(\phi), R(\phi) \subset X$, satisfying $(x,\phi(x)) \in \mathcal{R}$ for almost all $x \in D(\phi)$. Note that, since $\mathcal{R}$ is II_1, every $\phi \in [[\mathcal{R}]]$ is measure preserving. If $\Gamma \subset [[\mathcal{R}]]$ is a subgroup, we denote by $s(\Gamma) \subset X$ its *support*, i.e. the subset $Y \subset X$ with the property that $R(g) = D(g) = Y$, $\forall g \in \Gamma$. Two such subgroups $\Gamma, \Lambda \subset [[\mathcal{R}]]$ are *conjugate* by an element in $[[\mathcal{R}]]$ if there exists $\phi \in [[\mathcal{R}]]$ such that $D(\phi) = s(\Gamma)$, $R(\phi) = s(\Lambda)$ and $\phi\Gamma = \Lambda\phi$.

THEOREM 6.1. *Let $\mathcal{R}$ be a II_1 equivalence relation on the probability space (X,μ).*

(1) *If $[\mathcal{R}]$ contains a property (T) group Γ implementing an ergodic action on (X,μ), then $\mathcal{F}(\mathcal{R})$ is countable. More generally, if $[\mathcal{R}]$ contains a countable group Γ having a subgroup $H \subset \Gamma$ with the relative property (T) implementing an ergodic action on (X,μ), then $\mathcal{F}(\mathcal{R})$ is countable.*
(2) *Let $\mathcal{T}$ be the set of property (T) subgroups $\Gamma \subset [[\mathcal{R}]]$ acting ergodically on $s(\Gamma)$. Then $\mathcal{T}$ is countable, modulo conjugacy by elements in $[[\mathcal{R}]]$.*

Note that the ergodicity assumption of the action of Γ on (X,μ) in Part 1 of the above statement is crucial. Indeed, Theorem 3.1 provides examples of free ergodic p.m.p. actions $G \curvearrowright (X,\mu)$ such that $\mathcal{R}(G \curvearrowright X)$ has uncountable fundamental group, but nevertheless G contains a subgroup having property (T) (which, a fortiori, acts non-ergodically on (X,μ)). In turn, the existence of a property (T) subgroup of $[\mathcal{R}]$ acting freely and ergodically on X, does not ensure that the II_1 factor $L(\mathcal{R})$ has countable fundamental group. Indeed, by [**8**] there exist free ergodic p.m.p. actions of groups of the form $G = \Gamma \times \Sigma$, with Γ having property (T) and acting by Bernoulli shifts (thus ergodically), such that $M = \mathrm{L}^\infty(X) \rtimes G$ splits off the hyperfinite II_1 factor, and thus $\mathcal{F}(M) = \mathbb{R}_+$. In fact, as pointed out in [**28**], more than being countable, the fundamental group of $\mathcal{R}(G \curvearrowright X)$ is trivial.

Note also that in the case $\mathcal{R}$ comes from a free ergodic action of a property (T) group, $\Gamma \curvearrowright (X,\mu)$, Part 1 of Theorem 6.1 was already shown in [**17**, Corollary 1.8], in the case Γ is ICC, and in [**18**, Theorem 5.9], in the general case. We will use the above result in [**33**] to prove that the II_1 equivalence $\mathcal{R}$ obtained by restricting the II_∞ equivalence relation implemented by $\mathrm{SL}(n,\mathbb{Z}) \curvearrowright \mathbb{R}^n$, $n \geq 4$, to a subset of measure 1 has property (T) in the sense of Zimmer, yet cannot be implemented by an action (even non-free) of a property (T) group because $\mathcal{F}(\mathcal{R}) = \mathbb{R}_+$.

We will prove Theorem 6.1 by contradiction, using the property (T) of the subgroups and a "separability" argument, in the spirit of [**25**]. For more on this strategy of proofs, which grew out of Connes' rigidity paper [**4**], we send the interested reader to Section 4 in [**30**]. As a result of this argument, we obtain two copies $\Gamma_1, \Gamma_2 \subset [[\mathcal{R}]]$ of the same property (T) group, which are uniformly close one to the other. This in turn gives rise to a non-zero intertwiner $\phi \in [[\mathcal{R}]]$ between Γ_1, Γ_2. But if the Γ_i-actions are assumed ergodic, this forces $\mu(s(\Gamma_1)) = \mu(s(\Gamma_2))$ and the conjugacy of Γ_1, Γ_2.

The existence of an intertwiner between uniformly close subgroups in $[[\mathcal{R}]]$ is the subject of the next lemma. Recall that $[[\mathcal{R}]]$ has a natural metric space structure, inherited from the Hilbert-norm $\|\cdot\|_2$ of the underlying II_1 factor $L(\mathcal{R})$ associated with $\mathcal{R}$, by viewing every $\phi \in [[\mathcal{R}]]$ as a partial isometry in $L(\mathcal{R})$ and using the $\|\cdot\|_2$-norm on the latter. The metric can be concretely written as

$$d(\phi,\psi)^2 = \mu\big(D(\phi) \triangle\ D(\psi)\big) + 2\mu\big(\{x \in D(\phi) \cap D(\psi) \mid \phi(x) \neq \psi(x)\}\big)\ ,$$

where $\triangle$ denotes the symmetric difference of two sets. We will also need the natural σ-finite measure $\mu^{(1)}$ on $\mathcal{R} \subset X \times X$, defined by the formula

$$\mu^{(1)}(\mathcal{U}) = \int_X \#\{y \mid (x,y) \in \mathcal{U}\}\ d\mu(x) = \int_X \#\{x \mid (x,y) \in \mathcal{U}\}\ d\mu(y)$$

for all measurable subsets $\mathcal{U} \subset \mathcal{R}$.

Lemma 6.2. *Let $\mathcal{R}$ be a II_1 equivalence relation on the standard probability space (X,μ). Suppose that Γ is a countable group, $X_0, Y_0 \subset X$ and let*

$$\alpha : \Gamma \to [\mathcal{R}|_{X_0}] \quad \textit{and} \quad \beta : \Gamma \to [\mathcal{R}|_{Y_0}]$$

be group morphisms satisfying $d(\alpha_g,\beta_g) \leq 1/5$ for all $g \in \Gamma$ and $\mu(X_0), \mu(Y_0) \geq 3/4$. Then there exist non-negligible measurable subsets $X_1 \subset X_0$, $Y_1 \subset Y_0$ and $\phi \in [[\mathcal{R}]]$ with $D(\phi) = X_1$, $R(\phi) = Y_1$ such that

X_1 is globally $(\alpha_g)_{g\in\Gamma}$-invariant , Y_1 is globally $(\beta_g)_{g\in\Gamma}$-invariant, and

$\phi(\alpha_g(x)) = \beta_g(\phi(x))$ for almost all $x \in D(\phi)$.

Proof. Denote by Tr the normal faithful semi-finite trace on $\mathrm{L}^\infty(\mathcal{R})$ given by integration along $\mu^{(1)}$. Let $p \in \mathrm{L}^\infty(\mathcal{R})$ be the projection onto $\mathcal{R} \cap X_0 \times Y_0$ and

$e \leq p$ the projection onto $\{(z,z) \mid z \in X_0 \cap Y_0\}$. Set $B = \mathrm{L}^\infty(\mathcal{R})p$. The group Γ acts by automorphisms ρ_g of B given by

$$(\rho_g F)(x,y) = F(\alpha_{g^{-1}}(x), \beta_{g^{-1}}(y)) \quad \text{for almost all } (x,y) \in \mathcal{R} \cap X \times Y \; .$$

Since $\|\rho_g(e) - e\|_{2,\mathrm{Tr}}^2 = 2\mu\big(\{z \in X_0 \cap Y_0 \mid \alpha_{g^{-1}}(z) \neq \beta_{g^{-1}}(z)\}\big)$, we get

$$\|\rho_g(e) - e\|_{2,\mathrm{Tr}} \leq \frac{1}{5} \quad \text{for all } g \in \Gamma \; .$$

Define $a \in B^+$ as the unique element of minimal $\|\cdot\|_{2,\mathrm{Tr}}$ in the weakly closed convex hull $\overline{\mathrm{conv}}\{\rho_g(e) \mid g \in \Gamma\}$. It follows that $\|a - e\|_{2,\mathrm{Tr}} \leq 1/5$ and that $\rho_g(a) = a$ for all $g \in \Gamma$. Note that $0 \leq a \leq 1$. Defining f as the spectral projection $f = \chi_{[1/2,1]}(a)$, we find that $\|f - e\|_{2,\mathrm{Tr}} \leq 2/5$ and $\rho_g(f) = f$ for all $g \in \Gamma$. We write $f = \chi_W$, where $W \subset \mathcal{R} \cap X_0 \times Y_0$ is globally $(\rho_g)_{g\in\Gamma}$-invariant and satisfies

$$\mu^{(1)}\big(W \;\triangle\; \{(z,z) \mid z \in X_0 \cap Y_0\}\big) \leq \frac{4}{25} \; . \tag{6.1}$$

Denote ${}_xW := \{y \in Y_0 \mid (x,y) \in W\}$ and $W_y := \{x \in X_0 \mid (x,y) \in W\}$. Define

$$W_0 := \{(x,y) \in W \mid \; {}_xW \text{ and } W_y \text{ are singletons } \} \; .$$

Then W_0 is still globally $(\rho_g)_{g\in\Gamma}$-invariant. Since $\mu(X_0), \mu(Y_0) \geq 3/4$ and $\mu(X_0 \triangle\; Y_0) \leq 1/25$, we have $\mu(X_0 \cap Y_0) \geq 3/4 - 1/25$. By (6.1), the set of $x \in X_0$ such that ${}_xW$ is a singleton then has measure at least $3/4 - 1/25 - 4/25$. The same holds for the set of $y \in Y_0$ such that W_y is a singleton. So, W_0 has measure at least $1/10$. By construction, W_0 is the graph of a partial automorphism $\phi \in [[\mathcal{R}]]$ satisfying all the conclusions of the lemma. $\square$

Proof of Theorem 6.1. Let us first prove Part 1 of the theorem. By the relative property (T) of $H \subset \Gamma$, there exist $F \subset \Gamma$ and $0 < \varepsilon < 1/4$ such that whenever $\pi : \Gamma \to \mathcal{U}(\mathcal{H})$ is a unitary representation of Γ on a Hilbert space $\mathcal{H}$ and $\xi_0 \in \mathcal{H}$ a unit vector satisfying $\|\pi(g)\xi_0 - \xi_0\| \leq \varepsilon$ for all $g \in F$, then $\|\pi(h)\xi_0 - \xi_0\| \leq 1/8$ for all $h \in H$.

Choose for every $t \in (0,1)$ a measurable subset $Y_t \subset X$ with $\mu(Y_t) = t$ and such that $Y_s \subset Y_t$ if $s \leq t$. Assume that the fundamental group of $\mathcal{R}$ is uncountable. For every $t \in \mathcal{F}(\mathcal{R}) \cap (3/4, 1)$, choose an isomorphism $\Delta_t : X \to Y_t$ between $\mathcal{R}$ and $\mathcal{R}|_{Y_t}$. Note that Δ_t scales the measure μ by t. Define $\alpha_g^t = \Delta_t \circ \alpha_g \circ \Delta_t^{-1}$. Note that $\alpha_g^t \in [[\mathcal{R}]]$ with $D(\alpha_g^t) = R(\alpha_g^t) = Y_t$. Since $\mathcal{F}(\mathcal{R}) \cap (3/4,1)$ is uncountable, separability of the metric space $([[\mathcal{R}]], d)$ yields $s, t \in \mathcal{F}(\mathcal{R}) \cap (3/4, 1)$ with $s < t$ and $d(\alpha_g^s, \alpha_g^t) \leq \varepsilon/2$ for all $g \in F$.

Define the Hilbert space $\mathcal{H} = \mathrm{L}^2(\mathcal{R} \cap Y_s \times Y_t, \mu^{(1)})$ and the unitary representation

$$\pi : \Gamma \to \mathcal{U}(\mathcal{H}) : (\pi(g)\xi)(x,y) = \xi(\alpha_{g^{-1}}^s(x), \alpha_{g^{-1}}^t(y)) \; .$$

Set $\Delta_s := \{(y,y) \mid y \in Y_s\}$ and $\xi_0 := s^{-1/2}\chi_{\Delta_s}$. Then ξ_0 is a unit vector in $\mathcal{H}$ and, for all $g \in F$,

$$\|\pi(g)\xi_0 - \xi_0\|^2 = 2s^{-1}\mu\big(\{y \in Y_s \mid \alpha_{g^{-1}}^s(y) \neq \alpha_{g^{-1}}^t(y)\}\big) \leq s^{-1} d(\alpha_g^s, \alpha_g^t)^2 \leq \varepsilon^2 \; .$$

It follows that $\|\pi(h)\xi_0 - \xi_0\| \leq 1/8$ for all $h \in H$. So, for all $h \in H$, we have

$$2\mu\big(\{y \in Y_s \mid \alpha_h^s(y) \neq \alpha_h^t(y)\}\big) \leq \frac{s}{64} \leq \frac{1}{64} \; .$$

Since also, given $g \in F$,

$$\mu(Y_t \setminus Y_s) \leq d(\alpha_g^s, \alpha_g^t)^2 \leq \frac{\varepsilon^2}{4} \leq \frac{1}{64} ,$$

it follows that

$$d(\alpha_h^s, \alpha_h^t)^2 \leq \mu(Y_t \setminus Y_s) + \frac{1}{64} < \frac{1}{25}$$

for all $h \in H$. Since $(\alpha_h)_{h\in H}$ implements an ergodic action on (X, μ), the same holds for $(\alpha_h^s)_{h\in H}$, $(\alpha_h^t)_{h\in H}$ and so, Lemma 6.2 provides an element $\phi \in [[\mathcal{R}]]$ with $D(\phi) = Y_s$ and $R(\phi) = Y_t$. Since ϕ is a measure preserving isomorphism between Y_s and Y_t and $\mu(Y_s) = s < t = \mu(Y_t)$, we reached a contradiction.

To prove Part 2 of Theorem 6.1, assume by contradiction that there exist uncountably many subgroups $\{\Gamma_i \mid i \in I\}$ in $[[\mathcal{R}]]$ which have property (T) and are non-conjugate in $[[\mathcal{R}]]$. We continue to use the measurable subsets $Y_t \subset X$ with $\mu(Y_t) = t$ and $Y_s \subset Y_t$ whenever $s \leq t$. By the ergodicity of $\mathcal{R}$, we may assume that for every $i \in I$, the support of Γ_i is one of the Y_s.

By Shalom's theorem [**35**, Theorem 6.7], every property (T) group is the quotient of a finitely presented property (T) group. Since there are only countably many finitely presented groups, we may assume that all Γ_i's are quotients of the same property (T) group Γ through surjective homomorphisms $\alpha_i : \Gamma \to \Gamma_i$. Finally, we may assume that there exists $t \in (0, 1)$ such that $\mu(s(\Gamma_i)) \in (3t/4, t)$ for all $i \in I$. So, replacing $\mathcal{R}$ by $\mathcal{R}|_{Y_t}$, we may assume that $\mu(s(\Gamma_i)) \in (3/4, 1)$ for all $i \in I$.

By the property (T) of Γ, there exist $F \subset \Gamma$ and $0 < \varepsilon < 1/4$ such that whenever $\pi : \Gamma \to \mathcal{U}(\mathcal{H})$ is a unitary representation of Γ on a Hilbert space $\mathcal{H}$ and $\xi_0 \in \mathcal{H}$ a unit vector satisfying $\|\pi(g)\xi_0 - \xi_0\| \leq \varepsilon$ for all $g \in F$, then $\|\pi(g)\xi_0 - \xi_0\| \leq 1/8$ for all $g \in \Gamma$.

Now, by the separability of $([[\mathcal{R}]], d)$, there exist $i \neq j$ such that $d(\alpha_i(g), \alpha_j(g)) \leq \varepsilon/2$, $\forall g \in F$. Let $Y_i \subset X$, $Y_j \subset X$ be the support of Γ_i resp. Γ_j and assume $Y_i \subset Y_j$. We define $\mathcal{H}$, $\xi_0 \in \mathcal{H}$, $\pi : \Gamma \to \mathcal{U}(\mathcal{H})$ as before, but replacing α_g^s by $\alpha_i(g)$, α_g^t by $\alpha_j(g)$, Y_s by Y_i and Y_t by Y_j. The same estimates then show that ξ_0 is a unit vector satisfying $\|\pi_g(\xi_0) - \xi_0\| \leq \varepsilon$, $\forall g \in F$. Thus, $\|\pi(g)\xi_0 - \xi_0\| \leq 1/8$ for all $g \in \Gamma$. As before, this translates into $d(\alpha_i(g), \alpha_j(g)) \leq 1/4$, $\forall g \in \Gamma$. By Lemma 6.2, this implies Γ_i, Γ_j are conjugate by an element in $[[\mathcal{R}]]$, contradicting our initial assumption. □

Part 2 of Theorem 6.1 readily implies that the functor $\Gamma \mapsto \mathcal{R}_\Gamma$, from free ergodic p.m.p. actions of property (T) groups with morphisms given by conjugacy, to the associated equivalence relations with morphisms given by orbital isomorphism, is "countable to one". In other words, there are at most countably many non-conjugate free ergodic p.m.p. actions in each OE class of a free ergodic p.m.p. action of a property (T) group. In fact, even more is true: any free ergodic p.m.p. action of a property (T) group follows "orbit equivalent superrigid, modulo countable classes", in a sense made precise below.

Corollary 6.3. *Let $\Gamma \curvearrowright X$ be a free ergodic p.m.p. action of a property (T) group. Let $\Lambda_i \curvearrowright X_i$, $i \in I$, be a family of free ergodic p.m.p. actions such that $\mathcal{R}_\Gamma \simeq \mathcal{R}_{\Lambda_i}^{t_i}$, for some $t_i > 0$. Then the family I is countable, modulo conjugacy of actions.*

Proof. We may assume that $t_i \geq 1/c$ for all $i \in I$ and some $c > 0$. Setting $\mathcal{R} = (\mathcal{R}_\Gamma)^c$, we can view all Λ_i as subgroups of $[[\mathcal{R}]]$, with the action of Λ_i on $s(\Lambda_i) \subset X$ being conjugate to $\Lambda_i \curvearrowright X_i$. By [**14**, Corollary 1.4], all Λ_i have property (T). So, by Part 2 of Theorem 6.1, the family I is countable modulo conjugacy of actions. $\square$

When Γ is an ICC property (T) group, all groups in $\mathcal{S}_{\text{factor}}(\Gamma)$ are countable (cf. [**17**, Proof of Theorem 1.7], or [**25**, Theorem 4.5.1]). The next theorem generalizes this result to Kazhdan groups Γ with the property that the center $\mathcal{Z}(\Gamma)$ has finite index in the *FC-radical* Γ_f of Γ, defined by

$$\Gamma_f := \{g \in \Gamma \mid \ g \ \text{ has a finite conjugacy class } \} \; .$$

On the other hand, we prove in the second part of the theorem below that if Γ is a residually finite property (T) group such that Γ_f is not virtually abelian (i.e., $\mathcal{Z}(\Gamma_f) < \Gamma_f$ has infinite index), then Γ admits a free ergodic p.m.p. action on (X, μ) with $\mathrm{L}^\infty(X) \rtimes \Gamma$ being McDuff and hence, $\mathbb{R}_+ \in \mathcal{S}_{\text{factor}}(\Gamma)$.

At the time of finishing a first version of this article, the only known examples of Kazhdan groups Γ with infinite FC-radical Γ_f were such that $\mathcal{Z}(\Gamma)$ has finite index in Γ_f (see e.g. [**2**, Example 1.7.13] and [**11**, Definition 2.4]). While we were unable to show whether or not there exist residually finite Kazhdan groups Γ with non virtually abelian FC-radical Γ_f, after discussing this problem with several specialists, it was indicated to us by Mark Sapir and Denis Osin that such groups probably do exist. Very recently, this was confirmed by Mikhail Ershov [**13**] who showed that every Golod-Shafarevich group has a residually finite quotient whose FC-radical is not virtually abelian. Since he proved in [**12**] that there exist Golod-Shafarevich groups with property (T) and since property (T) passes to quotients, there indeed exist free ergodic p.m.p. actions $\Gamma \curvearrowright (X, \mu)$ of property (T) groups such that $\mathrm{L}^\infty(X) \rtimes \Gamma$ is McDuff.

Theorem 6.4. *Let Γ be a property (T) group.*

(1) *If $\mathcal{Z}(\Gamma)$ has finite index in the FC-radical Γ_f, then $\mathcal{S}_{\text{factor}}(\Gamma)$ only contains countable groups.*
(2) *If Γ is residually finite and $[\Gamma_f, \mathcal{Z}(\Gamma_f)] = \infty$, then Γ admits a free ergodic profinite p.m.p. action on (X, μ) such that $\mathrm{L}^\infty(X) \rtimes \Gamma$ is McDuff.*

Proof. Whenever $H \subset \Gamma$, denote by $\mathrm{C}_\Gamma(H)$ the centralizer of H inside Γ.

Assume first that $\mathcal{Z}(\Gamma)$ has finite index in Γ_f. Let $\Gamma \curvearrowright (X, \mu)$ be free ergodic p.m.p. Write $A := \mathrm{L}^\infty(X)$ and $M := A \rtimes \Gamma$. Define $\Lambda := \mathrm{C}_\Gamma(\Gamma_f)$. Since $\mathcal{Z}(\Gamma)$ has finite index in Γ_f, it follows that Λ has finite index in Γ. A fortiori, the subgroup $\Lambda_1 := \Lambda \cdot \Gamma_f$ has finite index in Γ. Also, the subalgebra A^Λ of Λ-invariant functions in A, is finite dimensional and globally Λ_1-invariant. Consider the subalgebra $B := A^\Lambda \rtimes \Lambda_1$ of M. Since $L(\Lambda_1) \subset B$ has finite index, it follows that B has property (T). On the other hand $M \cap B' \subset M \cap L(\Lambda)'$ and it is straightforward to check that $M \cap L(\Lambda)' \subset A^\Lambda \rtimes \Gamma_f$. So, we get $M \cap B' \subset B$. By [**21**, Theorem A.1], it follows that $\mathcal{F}(M)$ is countable.

Suppose from now on that Γ is residually finite and $[\Gamma_f, \mathcal{Z}(\Gamma_f)] = \infty$. Let $\Gamma_f = \{h_1, h_2, \ldots\}$ be an enumeration. Let $H_n \lhd \Gamma$ be a decreasing sequence of normal, finite index subgroups with $\bigcap_n H_n = \{e\}$. Define

$$\Gamma_n := H_n \cap \bigcap_{s \in \Gamma / \mathrm{C}_\Gamma(h_1, \ldots, h_n)} s \, \mathrm{C}_\Gamma(h_1, \ldots, h_n) s^{-1} \; .$$

By construction, Γ_n is a decreasing sequence of normal, finite index subgroups with $\bigcap_n \Gamma_n = \{e\}$ and such that for all $h \in \Gamma_f$, we have $\Gamma_n \subset \mathrm{C}_\Gamma(h)$ for all n large enough.

Denote $(X,\mu) = \varprojlim(\Gamma/\Gamma_n, \text{counting probability measure})$. Consider the natural free, ergodic, profinite, p.m.p. action $\Gamma \curvearrowright (X,\mu)$. Put $A = \mathrm{L}^\infty(X)$ and $M := A \rtimes \Gamma$. For every $s \in \Gamma$ and $n \in \mathbb{N}$, denote by $\chi_{s\Gamma_n}$ the function equal to 1 on $s\Gamma_n$ and zero elsewhere and interpret $\chi_{s\Gamma_n}$ as a projection in A.

For every $h \in \Gamma_f$, define the unitary $v_h \in M \cap \mathrm{L}(\Gamma)'$ by

$$v_h := \sum_{s \in \Gamma/\Gamma_n} \chi_{s\Gamma_n} u_{sh^{-1}s^{-1}} \quad \text{for } n \text{ large enough, meaning } \Gamma_n \subset \mathrm{C}_\Gamma(h).$$

It is straightforward to check that $\Gamma_f \to \mathcal{U}(M \cap \mathrm{L}(\Gamma)') : h \mapsto v_h$ is a group morphism and that $\tau(v_h) = 0$ whenever $h \neq e$.

Claim. If for all $n \in \mathbb{N}$, we have $h_n \in \Gamma_f \cap \Gamma_n$ with $h_n \neq e$, then (v_{h_n}) is a central sequence in M with $\tau(v_{h_n}) = 0$ for all n. For all n, we have $v_{h_n} \in \mathrm{L}(\Gamma)'$. So, to prove the claim, it suffices to take $k \in \mathbb{N}$, $g \in \Gamma$ and prove that

$$\lim_n \|[\chi_{g\Gamma_k}, v_{h_n}]\|_2 = 0 \; .$$

But, by construction, $\chi_{g\Gamma_k}$ and v_{h_n} commute when $n \geq k$.

Since $\mathcal{Z}(\Gamma_f) < \Gamma_f$ has infinite index and since Γ_f has finite conjugacy classes, it follows that Γ_f has no finite index abelian subgroups. So, for every n, the finite index subgroup $\Gamma_f \cap \Gamma_n$ of Γ_f is non-abelian. Therefore, we can choose $h_n, h'_n \in \Gamma_f \cap \Gamma_n$ such that $h_n h'_n h_n^{-1} {h'_n}^{-1} \neq e$. By the claim above, v_{h_n} and $v_{h'_n}$ are central sequences. By construction, $\tau(v_{h_n} v_{h'_n} v_{h_n}^* v_{h'_n}^*) = 0$ for all n. So, M is McDuff.

□

References

[1] J. Aaronson, The intrinsic normalising constants of transformations preserving infinite measures. *J. Analyse Math.* **49** (1987), 239-270.

[2] B. Bekka, P. de la Harpe and A. Valette, Kazhdan's property (T). *New Mathematical Monographs* **11**. Cambridge University Press, Cambridge, 2008.

[3] I. Chifan and C. Houdayer, Bass-Serre rigidity results in von Neumann algebras. *Duke Math. J.* **153** (2010), 23-54.

[4] A. Connes, A factor of type II_1 with countable fundamental group. *J. Operator Theory* **4** (1980), 151-153.

[5] A. Connes, Classification of injective factors. *Ann. Math.* **104** (1976), 73-115.

[6] A. Connes, Une classification des facteurs de type III, *Ann. Ec. Norm. Sup.* **6** (1973), 133-252.

[7] A. Connes, J. Feldman & B. Weiss, An amenable equivalence relation is generated by a single transformation. *Ergodic Theory Dynam. Systems* **1** (1981), 431-450.

[8] A. Connes and V.F.R. Jones, A II_1 factor with two non-conjugate Cartan subalgebras, *Bull. Amer. Math. Soc.* **6** (1982), 211-212.

[9] A. Connes and V.F.R. Jones, Property (T) for von Neumann algebras, *Bull. London Math. Soc.* **17** (1985), 57-62.

[10] A. Connes and M. Takesaki, The flow of weights of a factor of type III. *Tohoku Math. J.* **29** (1977), 473-575.

[11] Y. de Cornulier, Finitely presentable, non-Hopfian groups with Kazhdan's property (T) and infinite outer automorphism group. *Proc. Amer. Math. Soc.* **135** (2007), 951-959.

[12] M. Ershov, Golod-Shafarevich groups with property (T) and Kac-Moody groups. *Duke Math. J.* **145** (2008), 309-339.

[13] M. Ershov, Kazhdan groups whose FC-radical is not virtually abelian. *Preprint.* Available at `http://people.virginia.edu/~mve2x`

[14] A. Furman, Gromov's measure equivalence and rigidity of higher rank lattices. *Ann. of Math.* **150** (1999), 1059–1081.

[15] D. Gaboriau, Invariants l^2 de relations d'équivalence et de groupes. *Publ. Math. Inst. Hautes Études Sci.* **95** (2002), 93-150.

[16] D. Gaboriau, Relative property (T) actions and trivial outer automorphism groups. *Preprint.* `arXiv:0804.0358`

[17] S.L. Gefter and V.Ya. Golodets, Fundamental groups for ergodic actions and actions with unit fundamental groups. *Publ. Res. Inst. Math. Sci.* **24** (1988), 821-847.

[18] A. Ioana, Cocycle superrigidity for profinite actions of property (T) groups. *Duke Math. J.*, to appear. `arXiv:0805.2998`

[19] A. Ioana, J. Peterson and S. Popa, Amalgamated free products of w-rigid factors and calculation of their symmetry groups. *Acta Math.* **200** (2008), 85-153.

[20] F. Murray, J. von Neumann, Rings of operators IV, *Ann. Math.* **44** (1943), 716-808.

[21] R. Nicoara, S. Popa and R. Sasyk, On II_1 factors arising from 2-cocycles of w-rigid groups. *J. Funct. Anal.* **242** (2007), 230-246.

[22] D.S. Ornstein and B. Weiss, Ergodic theory of amenable group actions. *Bull. Amer. Math. Soc. (N.S.)* **2** (1980), 161-164.

[23] N. Ozawa, There is no separable universal II_1-factor, *Proc. Amer. Math. Soc.* **132** (2004), 487-490.

[24] N. Ozawa and S. Popa, On a class of II_1 factors with at most one Cartan subalgebra. *Ann. Math.* **172** (2010), 713-749.

[25] S. Popa, Correspondences, INCREST preprint No 56, 1986 (unpublished). http://www.math.ucla.edu/~popa/preprints.html

[26] S. Popa, Strong rigidity of II_1 factors arising from malleable actions of w-rigid groups, I. *Invent. Math.* **165** (2006), 369-408.

[27] S. Popa, Strong rigidity of II_1 factors arising from malleable actions of w-rigid groups, II. *Invent. Math.* **165** (2006), 409-452.

[28] S. Popa, Cocycle and orbit equivalence superrigidity for malleable actions of w-rigid groups. *Invent. Math.* **170** (2007), 243-295.

[29] S. Popa, On a class of type II_1 factors with Betti numbers invariants. *Ann. of Math.* **163** (2006), 809-899.

[30] S. Popa, Deformation and rigidity for group actions and von Neumann algebras, in "Proceedings of the International Congress of Mathematicians" (Madrid 2006), Volume I, EMS Publishing House, Zurich 2006/2007, pp. 445-479.

[31] S. Popa and S. Vaes, Actions of $\mathbb{F}_\infty$ whose II_1 factors and orbit equivalence relations have prescribed fundamental group. *J. Amer. Math. Soc.* **23** (2010), 383-403.

[32] S. Popa and S. Vaes, Strong rigidity of generalized Bernoulli actions and computations of their symmetry groups. *Adv. Math.* **217** (2008), 833–872.

[33] S. Popa and S. Vaes, Cocycle and orbit superrigidity for lattices in $\mathrm{SL}(n,\mathbb{R})$ acting on homogeneous spaces. In *Geometry, rigidity, and group actions,* Eds. B. Farb and D. Fisher. The University of Chicago Press, to appear. `arXiv:0810.3630`

[34] F. Radulescu, A one parameter group of automorphisms of $L(\mathbb{F}_\infty)$ scaling the trace. *Comptes Rendus Acad. Sci. Paris* **314** (1992), 1027-1032.

[35] Y. Shalom, Rigidity of commensurators and irreducible lattices. *Invent Math.* **141** (2000), 1-54.

[36] A. Törnquist, Orbit equivalence and actions of $\mathbb{F}_n$. *J. Symbolic Logic* **71** (2006), 265-282.

[37] S. Vaes, Rigidity results for Bernoulli actions and their von Neumann algebras (after Sorin Popa). Séminaire Bourbaki, exp. no. 961. *Astérisque* **311** (2007), 237-294.

[38] R.J. Zimmer, Ergodic theory and semisimple groups. Birkhäuser, Boston, 1984.

Mathematics Department; University of California at Los Angeles, CA 90095-1555, US

E-mail address: popa@math.ucla.edu

Department of Mathematics, K.U.Leuven, Celestijnenlaan 200B; B–3001 Leuven, Belgium

E-mail address: stefaan.vaes@wis.kuleuven.be

Clay Mathematics Proceedings
Volume 11, 2010

Leibniz Seminorms for "Matrix Algebras Converge to the Sphere"

Marc A. Rieffel

In celebration of the sixtieth birthday of Alain Connes

ABSTRACT. In an earlier paper of mine relating vector bundles and Gromov–Hausdorff distance for ordinary compact metric spaces, it was crucial that the Lipschitz seminorms from the metrics satisfy a strong Leibniz property. In the present paper, for the now non-commutative situation of matrix algebras converging to the sphere (or to other spaces) for quantum Gromov–Hausdorff distance, we show how to construct suitable seminorms that also satisfy the strong Leibniz property. This is in preparation for making precise certain statements in the literature of high-energy physics concerning "vector bundles" over matrix algebras that "correspond" to monopole bundles over the sphere. We show that a fairly general source of seminorms that satisfy the strong Leibniz property consists of derivations into normed bimodules. For matrix algebras our main technical tools are coherent states and Berezin symbols.

Introduction

In a previous paper [**30**] I showed how to give a precise meaning to statements in the literature of high-energy physics and string theory of the kind "Matrix algebras converge to the sphere". (See [**30**] for numerous references to the relevant physics literature.) I did this by introducing the concept of "compact quantum metric spaces", in which the metric data is given by a seminorm on the non-commutative "algebra of functions". This seminorm plays the role of the usual Lipschitz seminorm on the algebra of continuous functions on an ordinary compact metric space. However, I was somewhat puzzled by the fact that I needed virtually no algebraic conditions on the seminorm, only an important analytic condition. But when I later began trying to give precise meaning to further statements in the physics literature of the kind "here are the vector bundles over the matrix algebras that correspond to the monopole bundles over the sphere" (see [**32**] for many references), I found that

2010 *Mathematics Subject Classification.* Primary 46L87; Secondary 53C23, 58B34, 81R15, 81R30.

Key words and phrases. quantum metric space, Gromov–Hausdorff distance, Lipschitz, Leibniz seminorm, coadjoint orbits, coherent states, Berezin symbols.

The research reported here was supported in part by National Science Foundation grant DMS-0500501.

for ordinary metric spaces a strong form of the Leibniz inequality for the seminorm played a crucial role [**32**]. (See, for example, the proof of proposition 2.3 of [**32**].) However, on returning to the non-commutative case of matrix algebras converging to the sphere (or to other spaces), for some time I did not see how to construct useful seminorms that brought the matrix algebras and sphere close together while also having the strong Leibniz property. The main purpose of this paper is to show how to construct such seminorms. As in the earlier paper [**30**], the setting is that of coadjoint orbits of compact semisimple Lie groups, of which the 2-sphere is the simplest example. The main technical tools continue to be coherent states and Berezin symbols.

In the first four sections of this paper we show that a fairly general setting for obtaining seminorms that possess the strong Leibniz property that we need consists of derivations into normed bimodules, and we examine various aspects of this topic. The strong Leibniz property for a seminorm L on a normed unital algebra A consists of the usual Leibniz inequality together with the inequality

$$L(a^{-1}) \leq \|a^{-1}\|^2 L(a)$$

whenever a is invertible in A. I have not seen this latter inequality discussed in the literature, but it plays a crucial role in [**32**]. In Section 4 we put together the various conditions that we have found to be important, and thereby give a tentative definition for a "compact C^*-metric space".

In Section 5 we examine the use of seminorms with the strong Leibniz property in connection with quantum Gromov–Hausdorff distance. (I expect that many of the ideas and techniques developed in this paper will apply to many other classes of examples beyond "Matrix algebras converge to the sphere".) In Section 6 we extend to the case of strongly Leibniz seminorms the construction technique introduced in [**29**] that we called "bridges". Sections 7 and 8 contain those pieces of our development that can be carried out for certain homogeneous spaces of any compact group (including finite ones). Section 9 gives the statement of our main theorem for coadjoint orbits, while Sections 10 through 13 contain the detailed technical development needed to prove our main theorem. Finally, in Section 14 we relate our results to other variants of quantum Gromov–Hausdorff distance that have been developed by David Kerr, Hanfeng Li, and Wei Wu [**13, 14, 18, 19, 41, 42, 43**].

We can describe our basic setup and our main theorem somewhat more specifically as follows, where definitions for various terms are given in later sections. Let G be a compact semisimple Lie group, let $(U, \mathcal{H})$ be an irreducible unitary representation of G, and let P be the rank-one projection along a highest weight vector for $(U, \mathcal{H})$. Let α be the action of G on $\mathcal{L}(\mathcal{H})$ by conjugation by U, and let H be the α-stability group of P. Let $A = C(G/H)$. Let ω be the highest weight for U, and for each $n \in \mathbb{Z}_{>0}$ let $(U^n, \mathcal{H}^n)$ be the irreducible representation of G of highest weight $n\omega$. Let α also denote the action of G on $B^n = \mathcal{L}(\mathcal{H}^n)$ by conjugation by U^n.

Choose on G a continuous length-function ℓ. Then ℓ and the translation action of G on A, as well as the actions α of G on each B^n, determine seminorms L_A on A and L_{B^n} on B^n that make (A, L_A) and each (B^n, L_{B^n}) into compact C^*-metric spaces.

Main Theorem (sketchy statement of Theorem 9.1). *For any $\varepsilon > 0$ there exists an N such that for any $n \geq N$ we can explicitly construct a strongly Leibniz seminorm,*

L_n, on $A \oplus B^n$ making $A \oplus B^n$ into a compact C^*-metric space, such that the quotients of L_n on A and B^n are L_A and L_{B^n}, and for which the quantum Gromov-Hausdorff distance between A and B^n is no greater than ε.

I plan to apply the results of this paper in a future paper to discuss vector bundles over non-commutative spaces (e.g., monopole bundles), along the lines used for ordinary spaces in [**32**].

I developed part of the material presented here during a ten-week visit at the Isaac Newton Institute in Cambridge, England, in the Fall of 2006. I am very appreciative of the stimulating and enjoyable conditions provided by the Isaac Newton Institute.

I am grateful to Hanfeng Li for some important comments on the first version of this paper, which led to some substantial improvements given in the present version.

1. Strongly Leibniz seminorms

From my investigation of the relation between vector bundles on compact metric spaces that are close together, both for ordinary spaces [**32**] and for non-commutative spaces (a continuing investigation), I have found that the following properties are very important when considering the seminorms that play the role of the Lipschitz seminorms of ordinary metric spaces. Unless the contrary is stated, we allow our seminorms to take the value $+\infty$, but we require that they take value 0 at 0. We use the usual conventions for calculating with $+\infty$. The following definition is close to definition 2.1 of [**32**].

DEFINITION 1.1. Let A be a normed unital algebra over $\mathbb{R}$ or $\mathbb{C}$, and let L be a seminorm on A. We say that:

1) L is *Leibniz* if it satisfies the inequality
$$L(ab) \leq L(a)\|b\| + \|a\|L(b)$$
for all $a, b \in A$.
2) L is *strongly Leibniz* if it is Leibniz, and $L(1) = 0$, and if for any $a \in A$ that has an inverse in A, we have
$$L(a^{-1}) \leq \|a^{-1}\|^2 L(a).$$
3) L is *finite* if $L(a) < \infty$ for all $a \in A$.
4) L is *semifinite* if $\{a : L(a) < \infty\}$ is norm-dense in A.
5) L is *continuous* if it is norm-continuous.
6) L is *lower-semicontinuous* if for one $r \in \mathbb{R}_{>0}$, hence for all $r > 0$, the set
$$\{a \in A : L(a) \leq r\}$$
is norm-closed in A.

If, furthermore, A is a $*$-normed algebra (i.e., has an isometric involution), then we define L^* by $L^*(a) = L(a^*)$ for $a \in A$. We then say that L is a *$*$-seminorm* if $L = L^*$.

The proof of the following proposition is straightforward.

PROPOSITION 1.2. *Let A be a unital normed algebra.*

i. *Let L be a seminorm on A and let $r \in \mathbb{R}^+$. If L satisfies one of the properties 1–6 above then rL satisfies that same property.*

ii. *Let L_1 and L_2 be two seminorms on A. If they are both Leibniz, or strongly Leibniz, or finite, or continuous, or lower semicontinuous, then so is $L_1 + L_2$.*

iii. *Let $\{L_\alpha\}$ be a family of seminorms on A, possibly infinite, and let L be the supremum of this family. (That is, $L = \bigvee_\alpha L_\alpha$, defined by $L(a) = \sup_\alpha\{L_\alpha(a)\}$. For two seminorms, L and L', we will denote their maximum by $L \vee L'$.) Then L is a seminorm on A, and if each L_α is Leibniz, or strongly Leibniz, or lower-semicontinuous, then so is L.*

iv. *If A is a $*$-normed algebra and if L satisfies one of the properties 1–6 above, then L^* satisfies that same property.*

I have seen no discussion of the strong Leibniz property in the literature. I do not know of an example of a finite Leibniz seminorm which does not satisfy the inequality for $L(a^{-1})$ in the definition of "strongly Leibniz". But if we allow the value $+\infty$ then examples can be constructed in the following way. Let A be a unital normed algebra and let B be a unital subalgebra of A. Let L_0 be a finite Leibniz seminorm on B. Define a Leibniz seminorm, L, on A by $L(a) = L_0(a)$ if $a \in B$ and $L(a) = +\infty$ otherwise. If B contains an element that is invertible in A but not in B then L is not strongly Leibniz. For example, let $A = C([0,1])$ and let B be its subalgebra of polynomial functions, with $L_0(f) = \|f'\|$. (This example is not lower-semicontinuous.)

Let $A^f = \{a : L(a) < \infty\}$. It is clear that if L is Leibniz then A^f is a subalgebra of A. If L is in fact strongly Leibniz and $a \in A^f$, then clearly a is invertible in A^f if and only if it is invertible in A. It follows that for any $a \in A^f$ the spectrum of a in A^f will be the same as its spectrum in A. In stupid examples we may have $1 \notin A^f$, but with that understood, we see that:

Proposition 1.3. *If L is strongly Leibniz then A^f is a spectrally stable subalgebra of A.*

The importance of this proposition will be seen in Section 3. We also remark that if A has an involution and if L is a Leibniz seminorm that is also a $*$-seminorm, then A^f is a $*$-subalgebra of A.

Simple arguments prove the following two propositions.

Proposition 1.4. *Let A be a normed unital algebra, and let L be a seminorm on A. Let B be a unital subalgebra of A, equipped with the norm from A. If L is Leibniz, or strongly Leibniz, or finite, or continuous, or lower-semicontinuous, then so is the restriction of L to B as a seminorm on B. (But if L is semifinite, its restriction to B need not be semifinite.)*

Proposition 1.5. *Let A be a $*$-normed unital algebra and let L be a seminorm on A. Let $\tilde{L} = L \vee L^*$. Then $\tilde{L}$ is a $*$-seminorm. If L is Leibniz, or strongly Leibniz, or finite, or continuous, or lower-semicontinuous, then so is $\tilde{L}$. (But if L is semifinite, $\tilde{L}$ need not be semifinite.)*

So in this way we can usually arrange to work with $*$-seminorms when dealing with $*$-algebras.

Here is another way to combine seminorms:

Proposition 1.6. *Let $L_1, \dots, L_n$ be seminorms on a normed unital algebra A, and let $\|\cdot\|_0$ be a norm on $\mathbb{R}^n$ with the property that if $(r_j), (s_j) \in \mathbb{R}^n$, and if*

$0 \le r_j \le s_j$ *for all* j, $1 \le j \le n$, *then* $\|(r_j)\|_0 \le \|(s_j)\|_0$. *Define a seminorm,* N, *on* A *by* $N(a) = \|(L_j(a))\|_0$, *with the evident meaning if* $L_j(a) = \infty$ *for some* j. *If each* L_j *satisfies a particular one of properties* 1, 2, 3, 5, 6 *of Definition* 1.1 *then* N *satisfies that property too.*

Proof. If each L_j is Leibniz, then

$$\begin{aligned} N(ab) = \|(L_j(ab))\|_0 &\le \|(L_j(a)\|b\| + \|a\|L_j(b))\|_0 \\ &\le \|(L_j(a))\|_0\|b\| + \|a\|\|(L_j(b))\|_0 = N(a)\|b\| + \|a\|N(b), \end{aligned}$$

and if each L_j is strongly Leibniz, then also

$$N(a^{-1}) = \|(L_j(a^{-1}))\|_0 \le \|(\|a^{-1}\|^2 L_j(a))\|_0 = \|a^{-1}\|^2 N(a).$$

It is clear that if each L_j is finite, or continuous, then so is N.

Suppose instead that each L_j is lower-semicontinuous. Let (a_m) be a sequence in A which converges in norm to $a \in A$, and suppose that there is a constant, K, such that $N(a_m) \le K$ for each m. For each m let $p^m = (L_j(a_m)) \in \mathbb{R}^n$, so that $\|p^m\|_0 \le K$. Since the K-ball of $\mathbb{R}^n$ for $\|\cdot\|_0$ is compact, we can pass to a convergent subsequence, so we can assume that the sequence $\{p^m\}$ converges to a vector, p, in $\mathbb{R}^n$ such that $\|p\|_0 \le K$ and whose entries are non-negative. Let $\varepsilon > 0$ be given. Then there is an integer m_ε such that if $m \ge m_\varepsilon$ then $L_j(a_m) \le p_j + \varepsilon$ for each j. Since each L_j is lower-semicontinuous, it follows that $L_j(a) \le p_j + \varepsilon$ for each j. Then $N(a) = \|(L_j(a))\|_0 \le \|(p_j + \varepsilon)\|_0 \le \|p\|_0 + \varepsilon\|(1, \dots, 1)\|_0$. Thus $N(a) \le K$ since $\|p\|_0 \le K$ and ε is arbitrary. ☐

2. General sources of strongly Leibniz seminorms

We will now examine general methods for constructing strongly Leibniz seminorms. We recall first [**11**] that a *first-order differential calculus* over a unital algebra A is a pair (Ω, d) consisting of a bimodule Ω over A and a derivation d from A into Ω, that is, a linear map from A into Ω such that

$$d(ab) = (da)b + a(db)$$

for all $a, b \in \Omega$. (We will always assume that our bimodules are such that 1_A acts as the identity operator on both left and right.) It is common to assume that Ω is generated as a bimodule by the range of d, but we will not need to impose this requirement, though it can always be arranged by replacing Ω by its sub-bimodule generated by the range of d.

Suppose now that A is a normed unital algebra (with $\|1_A\| = 1$), and that (Ω, d) is a first-order differential calculus for A. Assume further that Ω is equipped with a norm that makes it into a normed A-bimodule, that is,

$$\|a\omega b\|_\Omega \le \|a\|\|\omega\|_\Omega\|b\|$$

for all $a, b \in A$ and $\omega \in \Omega$. We will then say that $(\Omega, d, \|\cdot\|_\Omega)$ is a *normed first-order differential calculus*. We do not require that d be continuous for the norms on A and Ω. We define L on A by

$$L(a) = \|da\|_\Omega.$$

Notice that L is finite, and that L is continuous if d is.

Proposition 2.1. *Let* L *on* A *be defined as above for a normed first-order differential calculus. Then* L *is strongly Leibniz.*

PROOF. That L is Leibniz follows immediately from the definitions of a derivation and of a normed bimodule. To see that L is strongly Leibniz, notice first that from the definition of a derivation we obtain $d(1_A) = 0$, so that $L(1_A) = 0$. Suppose now that a is an invertible element of A. Then

$$0 = d(1_A) = d(aa^{-1}) = (da)a^{-1} + a(d(a^{-1})).$$

Thus

$$d(a^{-1}) = -a^{-1}(da)a^{-1}.$$

On taking the norm we see that $L(a^{-1}) \leq \|a^{-1}\|^2 L(a)$. □

We remark that no effective characterization seems to be known as to which Leibniz seminorms come from normed first-order differential calculi (or of inner ones) as above. They fall within the scope of the "flat" differential seminorms defined in definition 4.3 of [**4**], and for which equivalent conditions are given in theorem 4.4 of [**4**]. Necessary conditions for a differential seminorm to be flat are given immediately after definition 4.3 and in proposition 4.7 of [**4**]. For Leibniz seminorms we see above that a further necessary condition is that of being strongly Leibniz.

Let us now give some examples.

EXAMPLE 2.2. Let (X, ρ) be a compact metric space. For given $x_0, x_1 \in X$ with $x_0 \neq x_1$ let Ω_{x_0,x_1} be $\mathbb{R}$ or $\mathbb{C}$ according to whether $A = C(X)$ is over $\mathbb{R}$ or $\mathbb{C}$, and define actions of A on Ω_{x_0,x_1} by

$$f \cdot \omega = f(x_0)\omega, \quad \omega \cdot f = \omega f(x_1).$$

Define d by

$$df = (f(x_1) - f(x_0))/\rho(x_1, x_0).$$

It is easily checked that (Ω_{x_0,x_1}, d) is a first-order differential calculus over A. Give $A = C(X)$ its supremum norm, $\|\cdot\|_\infty$, and give Ω_{x_0,x_1} the usual norm on $\mathbb{R}$ or $\mathbb{C}$. Then Ω_{x_0,x_1} is a normed A-bimodule. Clearly d is continuous. We set

$$L_{x_0,x_1}(f) = \|df\| = |f(x_1) - f(x_0)|/\rho(x_1, x_0).$$

Then from Proposition 2.1 it follows that L_{x_0,x_1} is strongly Leibniz (and continuous).

Now let L be the supremum of the L_{x_0,x_1} over all pairs (x_0, x_1) with $x_1 \neq x_0$. We obtain in this way the usual Lipschitz seminorm, L^ρ, on $C(X)$. From Proposition 1.2 it follows that L^ρ is strongly Leibniz and lower-semicontinuous. Of course L^ρ is not continuous in general. But L^ρ is semifinite, since it is finite on the functions $f_{x_0}(x) = \rho(x, x_0)$, and these already generate a dense subalgebra, as seen by means of the Stone–Weierstrass theorem.

This example can be recast in a quite familiar form in the following way. Let $Z = (X \times X) \setminus \Delta$ where Δ is the diagonal of $X \times X$. Thus Z is a locally compact space. Let $\Omega = C_b(Z)$, the linear space of bounded continuous functions on Z with the supremum norm. Then Ω is a normed $C(X)$-bimodule for the actions

$$(f\omega)(x_0, x_1) = f(x_0)\omega(x_0, x_1), \quad (\omega f)(x_0, x_1) = \omega(x_0, x_1)f(x_1).$$

Let A denote the subalgebra of $C(X)$ consisting of the Lipschitz functions, and define a derivation d from A to $C(Z)$ by

$$(df)(x_0, x_1) = (f(x_1) - f(x_0))/\rho(x_1, x_0).$$

Then the usual Lipschitz seminorm is given by $L^\rho(f) = \|df\|_\infty$. Alternatively, let $\Omega = C(Z)$, the space of all continuous, possibly unbounded, functions on Z, as a $C(X)$-bimodule in the above way. Then d can be defined on all of $C(X)$ by the above formula. We can now consider the supremum norm on $C(Z)$, taking value $+\infty$ on unbounded functions (so a bit beyond our definitions above), and again set $L^\rho(f) = \|df\|_\infty$.

Example 2.3. Let us now consider some examples in which the normed unital algebra A may be non-commutative. If Ω is an A-bimodule, one always has the corresponding inner derivations. That is, if $\omega \in \Omega$ we can set $d^\omega(a) = \omega a - a\omega$. If Ω is a normed A-bimodule then d^ω is continuous, with $\|d^\omega\| \leq 2\|\omega\|$. The corresponding seminorm, L^ω, defined by $L^\omega(a) = \|d^\omega(a)\|$, is then a continuous strongly Leibniz seminorm.

Suppose now that B is a unital normed algebra and that π is a unital homomorphism from A into B. Then we can view B as a bimodule over A in the evident way, and obtain inner derivations and corresponding strongly Leibniz seminorms, which are continuous if π is.

Example 2.4. Suppose now that π is a non-degenerate representation of A as operators on a normed space X, so that π can be viewed as a unital homomorphism from A into $B(X)$, the algebra of bounded operators on X. Then $B(X)$ can be viewed in the evident way as a bimodule over A, and any element, D, of $B(X)$ determines an inner derivation, and corresponding seminorm

$$L(a) = \|D\pi(a) - \pi(a)D\| = \|[D, \pi(a)]\|.$$

More generally, if one has two representations, π^1 and π^2 of A on X, then one can view $B(X)$ as an A-bimodule via

$$a \cdot T \cdot b = \pi^1(a)T\pi^2(a),$$

and again any element, D, of $B(X)$ will determine an inner derivation. (The twisted commutators in equation 2.4 and lemma 2.2 of [**8**] fit into this view, except that there D is usually an unbounded operator.) Alternately one can assemble π^1 and π^2 into one representation on $X \oplus X$, and use the operator $\left(\begin{smallmatrix} 0 & D \\ D & 0 \end{smallmatrix}\right)$ on $X \oplus X$.

As an important particular case, for X we can take A itself and let π be the left-regular representation of A on itself. As element of $B(X)$ we can take an isometric algebra automorphism, α, of A. Then

$$\begin{aligned}(\alpha \circ \pi(a) - \pi(a) \circ \alpha)(b) &= \alpha(ab) - a\alpha(b) \\ &= (\alpha(a) - a)\alpha(b).\end{aligned}$$

From this we see that

$$\|\alpha \circ \pi(a) - \pi(a) \circ \alpha\| = \|\alpha(a) - a\|,$$

so that if we set $L(a) = \|\alpha(a) - a\|$ then L will be a continuous strongly Leibniz seminorm. We can view this in another way. View A as a bimodule over A by

$$a \cdot b \cdot c = ab\alpha(c),$$

and set $d(a) = \alpha(a) - a$. It is easily checked that d is a (continuous) derivation, and so from Proposition 2.1 we see again that L is strongly Leibniz. (This does not require that α be isometric.)

EXAMPLE 2.5. Now let G be a group, and let α be an action of G on A, that is, a homomorphism from G into $\mathrm{Aut}(A)$. Let ℓ be a length-function on G. For each $x \in G$ with $x \neq e_G$ the map $a \mapsto \|\alpha_x(a) - a\|/\ell(x)$ is a continuous strongly Leibniz seminorm. Let L be the supremum over G of all of these seminorms, so that

$$L(a) = \sup\{\|\alpha_x(a) - a\|/\ell(x) : x \neq e_G\}.$$

By Proposition 1.2 we see that L is a lower-semicontinuous strongly-Leibniz seminorm. Of course L may not be semifinite. But if G is a locally compact group, if A is complete, so a Banach algebra, if α is a strongly continuous action by isometric automorphisms of A, and if ℓ is a continuous length-function, then the discussion before theorem 2.2 of [**28**] shows that L is semifinite. The discussion there is stated just for C^*-algebras, but it applies without change to Banach algebras.

EXAMPLE 2.6. Suppose now that G is a connected Lie group, and that α is a strongly continuous action of G on A by isometric automorphisms. Let $\mathfrak{g}$ denote the Lie algebra of G, and let A^∞ denote the dense subalgebra of smooth elements of A for the action α. We let α also denote the corresponding infinitesimal action of $\mathfrak{g}$ on A^∞, defined by

$$\alpha_X(a) = \left.\frac{d}{dt}\right|_{t=0} \alpha_{\exp(tX)}(a)$$

for $X \in \mathfrak{g}$ and $a \in A^\infty$. The argument in the proof of lemma 3.1 of [**28**] works here, and shows that

$$\|\alpha_X(a)\| = \sup\{\|\alpha_{\exp(tX)}(a) - a\|/|t| : t \neq 0\}.$$

It follows from Proposition 1.2 that the map $a \mapsto \|\alpha_X(a)\|$ is a finite lower-semicontinuous strongly-Leibniz seminorm on A^∞. Suppose further that we are given a norm on $\mathfrak{g}$, and that we set

$$L(a) = \sup\{\|\alpha_X(a)\| : \|X\| \leq 1\}.$$

It follows again from Proposition 1.2 that L is a lower-semicontinuous strongly-Leibniz seminorm on A^∞, which is easily seen to be finite, but which may well not be norm-continuous.

EXAMPLE 2.7. Suppose now that G is a Lie group and that $(U, \mathcal{H})$ is a strongly continuous representation of G on a Hilbert space $\mathcal{H}$. As discussed in section 3 of [**28**] we can define an action, α, of G on $\mathcal{B}(\mathcal{H})$ by $\alpha_x(T) = U_x T U_x^*$, and we can let B be the largest subalgebra of $\mathcal{B}(\mathcal{H})$ on which this action is strongly continuous. We can then apply the discussion of the previous example to obtain a seminorm L on B^∞. If A is a unital $*$-subalgebra of B^∞ (which need not be carried into itself by α), then according to Proposition 1.4 the restriction of L to A is a lower-semicontinuous strongly-Leibniz $*$-seminorm which is clearly finite.

EXAMPLE 2.8. Let us consider the above situation for the special case in which $\mathfrak{g} = \mathbb{R}$. Then U is generated by a self-adjoint (often unbounded) operator, D, on $\mathcal{H}$, that is, $U_t = e^{itD}$ for all $t \in \mathbb{R}$. Then it follows easily that for $T \in B^\infty$

$$L(T) = \|[D, T]\|,$$

and in particular that the commutator $[D, T]$ is a bounded operator. All of this will then be true for any $T \in A \subset B^\infty$. This applies in particular to the "Dirac" operators on which Connes [**6, 11**] bases his approach to metric non-commutative differential geometry.

3. Closed seminorms

We adapt here some definitions from section 4 of [**26**]. Let A be a normed unital algebra, and let $\bar{A}$ denote its completion. Let L be a seminorm on A (value $+\infty$ allowed) and let

$$\mathcal{L}_1 = \{a \in A : L(a) \leq 1\}.$$

Let $\bar{\mathcal{L}_1}$ be the closure of $\mathcal{L}_1$ in $\bar{A}$, and let $\bar{L}$ denote the corresponding "Minkowski functional" on $\bar{A}$, defined by setting, for $c \in \bar{A}$,

$$\bar{L}(c) = \inf\{r \in \mathbb{R}^+ : c \in r\bar{\mathcal{L}_1}\}.$$

The value $+\infty$ must be allowed. Then $\bar{L}$ is a seminorm on $\bar{A}$, and the proof of proposition 4.4 of [**26**] tells us that if L is lower-semicontinuous, then $\bar{L}$ is an extension of L. We call $\bar{L}$ the *closure* of L. We see that the set $\{c \in \bar{A} : \bar{L}(c) \leq 1\}$ is closed in $\bar{A}$. We say that the original seminorm L on A is *closed* if $\mathcal{L}_1$ is closed in $\bar{A}$, or, equivalently, is complete for the norm on A. Clearly if L is closed, then it is lower-semicontinuous. If L is closed and is not defined on all of $\bar{A}$, then $\bar{L}$ is obtained simply by giving it value $+\infty$ on all the elements of $\bar{A}$ that are not in A. It is clear that if L is semifinite then so is $\bar{L}$. We recall that a unital subalgebra B of a unital algebra A is said to be spectrally stable in A if for any $b \in B$ the spectrum of b as an element of B is the same as its spectrum as an element of A, or equivalently, that any b that is invertible in A is invertible in B.

PROPOSITION 3.1. *Let L be a Leibniz seminorm on a normed unital algebra A. Then $\bar{L}$ is Leibniz. Set*

$$\bar{A}^f = \{c \in \bar{A} : \bar{L}(c) < \infty\}.$$

If $L(1) < \infty$, then $\bar{A}^f$ is a unital spectrally-stable subalgebra of the norm closure of $\bar{A}^f$ in $\bar{A}$. If A is defined over $\mathbb{C}$, then $\bar{A}^f$ is stable under the holomorphic-function calculus of its closure.

PROOF. Let $c, d \in \bar{A}$. If $\bar{L}(c) = \infty$ or $\bar{L}(d) = \infty$ there is nothing to show for the Leibniz condition. Otherwise, we can find sequences $\{a_n\}$ and $\{b_n\}$ in A such that $\{a_n\}$ converge to c while $\{L(a_n)\}$ converges to $\bar{L}(c)$ and $L(a_n) \leq \bar{L}(c)$ for all n, and similarly for $\{b_n\}$ and d. Then $a_n b_n$ converges to cd and

$$L(a_n b_n) \leq L(a_n)\|b_n\| + \|a_n\|L(b_n) \leq \bar{L}(c)\|b_n\| + \|a_n\|\bar{L}(d),$$

and the right-hand side converges to $\bar{L}(c)\|d\| + \|c\|\bar{L}(d)$. Thus $\bar{L}$ is Leibniz.

If $L(1) < \infty$ so that $\bar{A}^f$ is a unital subalgebra of $\bar{A}$, then it follows from proposition 3.12 of [**4**] (or proposition 1.7 and theorem 1.17 of [**36**], or lemma 1.6.1 of [**9**]) that $\bar{A}^f$ is spectrally stable in its closure in $\bar{A}$, and in fact is stable under the holomorphic-function calculus there. We sketch the proof in our simpler setting. Define a new norm, M, on $\bar{A}^f$ by

$$M(c) = \|c\| + \bar{L}(c).$$

Then, as mentioned after definition 4.5 of [**26**], $\bar{A}^f$ will be complete for the norm M because $\bar{L}$ is closed. (See the proof of proposition 1.6.2 of [**39**].) Because $\bar{L}$ is Leibniz, M is easily seen to be an algebra norm, so that $\bar{A}^f$ becomes a Banach algebra. Let $c \in \bar{A}^f$. From the Leibniz rule we find that $\bar{L}(c^n) \leq n\|c\|^{n-1}\bar{L}(c)$, so that

$$M(c^n) \leq \|c\|^n + n\|c\|^{n-1}\bar{L}(c).$$

From this it follows that if $\|c\| < 1$ then the series $\sum_{n=0}^{\infty} c^n$ converges for M to an element of $\bar{A}^f$. Thus $1 - c$ is invertible in $\bar{A}^f$. It follows that if instead $\|1 - c\| < 1$ then c is invertible in $\bar{A}^f$. From this it is then easily seen (e.g. lemma 3.38 of [**11**]) that if $c \in \bar{A}^f$ and if c is invertible in the norm-closure of $\bar{A}^f$ in $\bar{A}$, then c is invertible in $\bar{A}^f$. Consequently $\bar{A}^f$ is spectrally stable in its closure in $\bar{A}$.

Assume now that A is defined over $\mathbb{C}$. For the definition and properties of the holomorphic-function (or "symbolic") calculus see [**12, 34**]. It is well-known that a dense subalgebra that is spectrally stable and is a Banach algebra for a norm stronger that the norm of the bigger algebra, is stable under the holomorphic-function calculus. (See the comments after definition 3.25 of [**11**].) We briefly recall the reason, for our context. For notational simplicity we assume that $\bar{A}^f$ is dense in $\bar{A}$. Let $c \in \bar{A}^f$, and let f be a $\mathbb{C}$-valued function defined and holomorphic on an open neighborhood $\mathcal{O}$ of the spectrum $\sigma_{\bar{A}}(c)$. Let γ be the union of a finite number of curves in $\mathcal{O}$ that surrounds $\sigma_{\bar{A}}(c)$ in the usual way such that the Cauchy integral formula using γ gives f on a neighborhood of $\sigma_{\bar{A}}(c)$. Since $\bar{A}^f$ is spectrally stable in $\bar{A}$, the function $z \mapsto (z - c)^{-1}$, well-defined on γ, has values in $\bar{A}^f$. Since $\bar{A}^f$ is a Banach algebra for M, this function is continuous for M, and the integral

$$f(c) = \frac{1}{2\pi i} \int_{\gamma} f(z)(z - c)^{-1} dz$$

is well defined in $\bar{A}^f$. Since the homomorphism from $\bar{A}^f$ with norm M to $\bar{A}$ with its original norm is clearly continuous, the image of $f(c)$ in $\bar{A}$ will be expressed by the same integral but now interpreted in $\bar{A}$. But $f(c) \in \bar{A}^f$. So the above integral, but interpreted in $\bar{A}$, gives an element of $\bar{A}^f$ as was to be shown. $\square$

For the use of the holomorphic-function calculus when dealing with algebras over $\mathbb{R}$ see proposition 2.4 of [**32**].

One reason that the property of being closed under the holomorphic-function calculus is important is that it implies that $\bar{A}^f$ and its closure, say B, in $\bar{A}$ have essentially the same finitely-generated projective modules ("vector bundles") in the sense that any such right module V for B is of the form $V = W \otimes_{\bar{A}^f} B$ for such a right module W over $\bar{A}^f$, unique up to isomorphism. This is crucial to [**32**], and to our proposed discussion of projective modules and quantum Gromov–Hausdorff distance for non-commutative C^*-algebras. The inclusion of $\bar{A}^f$ into B also gives an isomorphism of their K-groups. (See appendix IIIC of [**6**] and theorem 3.44 of [**11**].)

PROPOSITION 3.2. *Let L be a strongly-Leibniz seminorm on a normed algebra A. Assume that A^f is dense and spectrally stable in $\bar{A}$. Then the closure, $\bar{L}$, of L is strongly Leibniz.*

PROOF. It is clear that $\bar{L}(1) = 0$. From Proposition 3.1 we know that $\bar{L}$ is Leibniz. Thus we only need to verify the condition on inverses. Suppose now that $c \in \bar{A}$ and that c is invertible in $\bar{A}$. If $\bar{L}(c) = \infty$ there is nothing to show, so assume that $c \in \bar{A}^f$. Then there is a sequence $\{a_n\}$ in A that converges to c while $\{L(a_n)\}$ converges to $\bar{L}(c)$ with $L(a_n) \leq \bar{L}(c)$ for all n (so $a_n \in A^f$). Since c is invertible in $\bar{A}$, and the set of invertible elements of a unital Banach algebra is open, the elements a_n are eventually invertible in $\bar{A}$. Since A^f is assumed to be spectrally stable in $\bar{A}$ the elements a_n are eventually invertible in A^f. Thus we can adjust the sequence $\{a_n\}$ so that each element is invertible in A^f. Then the sequence $\{a_n^{-1}\}$

converges to c^{-1}, while for each n

$$L(a_n^{-1}) \leq \|a_n^{-1}\|^2 L(a_n) \leq \|a_n^{-1}\|^2 \bar{L}(c).$$

It follows easily that $\bar{L}(c^{-1}) \leq \|c^{-1}\|^2 \bar{L}(c)$. Thus $\bar{L}$ is strongly Leibniz. □

4. C^*-metrics

Up to this point we have ignored the crucial analytic property of the seminorms that define quantum metric spaces, i.e., Lip-norms. We recall this property here, for our special context of unital C^*-algebras. Let A be a unital $*$-algebra equipped with a C^*-norm (but not assumed to be complete). Let L be a seminorm on A such that $L(1) = 0$. Define a metric, ρ_L, on the state space, $S(A)$, of A by

$$\rho_L(\mu, \nu) = \sup\{|\mu(a) - \nu(a)| : a = a^* \text{ and } L(a) \leq 1\}.$$

(Without further hypotheses ρ_L might take the value $+\infty$.) We will say that L is a Lip-norm if the topology on $S(A)$ from ρ_L coincides with the weak-$*$ topology on $S(A)$. In our definition of Lip-norms in definition 2.1 of [**29**] we, in effect, assumed that our seminorms L were defined only on the self-adjoint part of A, but still defined ρ_L as above. The comments before definition 2.1 of [**29**] show that if L is a $*$-seminorm then ρ_L would not change if the condition "$a = a^*$" above were omitted.

We now come to the definition that seems to be dictated by our investigation of vector bundles and Gromov–Hausdorff distance, both for ordinary metric spaces [**32**] and for quantum ones. It should be viewed as tentative, since future experience may require additional hypotheses.

Definition 4.1. Let A be a unital C^*-normed algebra and let L be a seminorm on A (possibly taking value $+\infty$). We will say that L is a *C^*-metric* on A if

a) L is a lower-semicontinuous strongly-Leibniz $*$-seminorm,
b) L (restricted to A^{sa}) is a Lip-norm,
c) A^f is spectrally stable in the completion, $\bar{A}$, of A.

By a *compact C^*-metric space* we mean a pair (A, L) consisting of a unital C^*-normed algebra A and a C^*-metric L on A.

In using the word "space" above, we should logically be referring to objects in the dual to the category of unital C^*-algebras. But we will not make this distinction explicit during our discussions in this paper.

We need condition c) in Definition 4.1 so that we can apply Proposition 3.2 to conclude that the closure of a C^*-metric is strongly Leibniz and itself satisfies condition c). Hanfeng Li has pointed out to me that the subalgebra of polynomials in the algebra of continuous functions on the unit interval with the usual Lipschitz seminorm shows that condition c) is independent of conditions a) and b).

At this time it is not clear to me how best to define C^*-metric spaces that are locally compact but not compact, though some substantial indications can be gleaned from the results in [**16**].

Recall [**4**] that a $*$-subalgebra B of a C^*-algebra A is said to be stable under the C^2-function calculus for self-adjoint elements if for any $b \in B$ with $b^* = b$ and any twice continuously differentiable function f on $\mathbb{R}$, the element $f(b)$ of A, defined by the continuous-function calculus on self-adjoint elements of A, is in fact again in B.

Proposition 4.2. *Every C^*-metric on a unital C^*-normed algebra is semifinite. Let L be a C^*-metric on a unital C^*-normed algebra A, and let $\bar{L}$ be its closure on the completion $\bar{A}$ of A (so $\bar{L}$ is an extension of L). Then $\bar{L}$ is a C^*-metric. Let $\bar{A}^f$ be defined as earlier (so now $\bar{A}^f$ is dense in $\bar{A}$). Then $\bar{A}^f$ is stable both under the holomorphic-function calculus of $\bar{A}$ and the C^2-calculus on self-adjoint elements of $\bar{A}$.*

Proof. Let L be a C^*-metric on a unital C^*-normed algebra A, and let A^f be defined as above. Suppose that A^f is not dense in A. Then it is easily seen that there is an $a \in A$ with $a^* = a$ that is not in the closure of A^f. By the Hahn–Banach theorem there is a linear functional of norm 1 on the self-adjoint part of A that has value 0 on all of the self-adjoint part of A^f. From lemma 2.1 of [**26**] it then follows that there are two distinct states of A which agree on A^f. Then the distance between these two states for the metric ρ_L determined by L is 0, which contradicts the requirement that the topology on $S(A)$ determined by ρ_L coincides with the weak-$*$ topology.

The fact that $\bar{L}$ is a C^*-metric is seen as follows. By definition, $\bar{L}$ is closed, and so lower-semicontinuous. As remarked above, $\bar{L}$ is strongly Leibniz by condition c) and Proposition 3.2. The closure of a Lip-norm is again a Lip-norm, giving the same metric on the state-space, as seen in proposition 4.4 of [**26**]. That $\bar{L}$ satisfies condition c) follows from Proposition 3.1.

The fact that $\bar{A}^f$ is stable under the holomorphic-function calculus of $\bar{A}$ follows immediately from the semifiniteness of $\bar{L}$ and Proposition 3.1. The fact that $\bar{A}^f$ is stable for the C^2-function calculus on self-adjoint elements of $\bar{A}$ follows quickly from proposition 6.4 of [**4**], which actually gives a slightly stronger fact. □

The condition that L be a Lip-norm is often a difficult one to verify for various specific examples. But most of the Lip-norms that have been constructed on C^*-algebras so far are in fact C^*-metrics. We explain this now for several of the classes of examples described in sections 2 and 3 of [**28**].

Example 4.3. Let A be a unital C^*-algebra, let G be a compact group, and let α be an action of G on A that is ergodic in the sense that if an $a \in A$ satisfies $\alpha_x(a) = a$ for all $x \in G$ then $a \in \mathbb{C}1_A$. Let ℓ be a continuous length function on G, and define a seminorm L on A, as in Example 2.5, by

$$L(a) = \sup\{\|\alpha_x(a) - a\|/\ell(x) : x \notin e_G\}.$$

It is shown in [**28**] that L is a Lip-norm. But we saw in Example 2.5 that L is lower-semicontinuous and strongly Leibniz. Since L is defined on all of A, the spectral stability of $\bar{A}^f$ in A follows from Proposition 3.1.

The next several examples involve "Dirac" operators in various settings.

Example 4.4. This class of examples is the main class discussed in Connes's first paper [**5**] on metric aspects of non-commutative geometry. It is discussed briefly as example 3.6 of [**28**]. Let G be a discrete group and let $A = C_r^*(G)$ be its reduced group C^*-algebra acting on $\ell^2(G)$. Let ℓ be a length function on G. As Dirac operator take the operator $D = M_\ell$ of pointwise multiplication by ℓ on $\ell^2(G)$. The one-parameter unitary group generated by D simply sends t to the operator of pointwise multiplication by the function $e^{it\ell}$. We are then in the context of Examples 2.7 and 2.8. It is easily seen that the dense subalgebra $C_c(G)$ of functions of finite support is in the smooth algebra B^∞ for the action of G on $\mathcal{L}(\ell^2(G))$. As in Example 2.8 we thus obtain a lower-semicontinuous strongly-Leibniz semifinite $*$-seminorm on A, which for any $f \in C_c(G)$ is given by

$$L(f) = \|[D, f]\|.$$

From Proposition 3.1 it follows that A^f (for the closure of L) is stable for the holomorphic-function calculus. But for stupid length functions L can fail to be a Lip-norm, and it is not easy to see when it is a Lip-norm, and thus a C^*-metric. In [**27**], by means of a long and interesting argument, it is shown that L is a Lip-norm, and thus a C^*-metric, for $G = \mathbb{Z}^d$ (and even for the twisted group algebra $C^*(\mathbb{Z}^d, \gamma)$ where γ is a bicharacter on $\mathbb{Z}^d$) when ℓ is either a word-length function or the restriction to $\mathbb{Z}^d$ of a norm on $\mathbb{R}^d$. In [**22**] it is shown, by techniques entirely different from those used for the case of $\mathbb{Z}^d$, that if G is a hyperbolic group and ℓ is a word-length function on G then L is a Lip-norm, and thus a C^*-metric. For other classes of infinite discrete groups, e.g., nilpotent ones, it remains a mystery as to whether L is a Lip-norm if ℓ is a word-length function. Some related examples can be found in [**1**].

Example 4.5. Let α be an action of the d-dimensional torus $\mathbb{T}^d$, $d \geq 2$, on a unital C^*-algebra A. In [**24**] it is shown that for any skew-symmetric real $d \times d$ matrix θ one can deform the product on A to get a new C^*-algebra, A_θ. Connes and Landi [**7**] show that when M is a compact spin Riemannian manifold and α is a smooth action of $\mathbb{T}^d$ on M, so on $A = C(M)$, leaving the Riemannian metric invariant, and lifting to the spin bundle, then there is a natural Dirac operator for the (usually non-commutative) deformed algebra A_θ. As in Examples 2.8 and 4.3, this Dirac operator determines a $*$-seminorm, L, on A_θ which is lower semicontinuous, strongly Leibniz, and semifinite. Hanfeng Li [**17**] showed that L is a Lip*-norm. Thus L is a C^*-metric.

5. Quotient seminorms and proximity

We now try to modify the definition of quantum Gromov–Hausdorff distance so as to use the above definition of C^*-metrics. This involves quotient seminorms, so we begin by exploring them. There are at least three difficulties that confront us, namely that the quotient of a Leibniz seminorm may not be Leibniz, that the quotient of a strongly Leibniz seminorm, even if it is Leibniz, may not be strongly Leibniz, and that reasonable $*$-seminorms can agree on self-adjoint elements but still be distinct. We begin by considering the first difficulty.

Let L be a Leibniz seminorm on a unital normed algebra C, and let $\pi : C \twoheadrightarrow A$ be a unital homomorphism from C onto a unital normed algebra A. Let $\tilde{L}^A$ be the quotient seminorm on A, defined by

$$\tilde{L}^A(a) = \inf\{L(c) : c \in C \text{ and } \pi(c) = a\}.$$

It is known [**4**] that $\tilde{L}^A$ need not be Leibniz. (See also lemma 4.3 of [**19**] and the comments just before it.) But the situation can be partly rescued by the following definition.

Definition 5.1. Let C, A, π and L be as above, and assume that π is norm non-increasing. We say that L is *π-compatible* if for every $a \in A$ and every $\varepsilon > 0$ there is a $c \in C$ such that $\pi(c) = a$ and simultaneously

$$L(c) \leq \tilde{L}^A(a) + \varepsilon \qquad \text{and} \qquad \|c\| \leq \|a\| + \varepsilon.$$

PROPOSITION 5.2. *Let C, A, π and L be as above. If L is π-compatible then the norm on A coincides with the quotient norm from C, and $\tilde{L}^A$ is Leibniz.*

PROOF. The statement about the norms is easily verified. Suppose now that $a, b \in A$ and $\varepsilon > 0$ are given. Since L is π-compatible, we can find $c, d \in C$ such that $\pi(c) = a$ and $\pi(d) = b$ and the conditions of Definition 5.1 are satisfied. Then $\pi(cd) = \pi(ab)$, and so

$$\begin{aligned} \tilde{L}^A(ab) &\leq L(cd) \leq L(c)\|d\| + \|c\|L(d) \\ &\leq (\tilde{L}^A(a) + \varepsilon)(\|b\| + \varepsilon) + (\|a\| + \varepsilon)(\tilde{L}^A(b) + \varepsilon). \end{aligned}$$

Since ε is arbitrary, we see that $\tilde{L}^A$ is Leibniz. □

However, if L is strongly Leibniz and if $\tilde{L}^A$ is Leibniz, there seems to be no reason that $\tilde{L}^A$ need be strongly Leibniz (though I do not have an example showing this difficulty). I do not know of a useful way to partly rescue this difficulty.

We now consider the third difficulty. It is quite instructive to first consider ordinary metric spaces. For this purpose π-compatibility is useful.

PROPOSITION 5.3. *Let (Z, ρ) be a compact metric space, and let $C = C(Z)$ be its C^*-algebra of continuous complex-valued functions. Let X be a closed subset of Z, let $A = C(X)$, and let $\pi : C \to A$ be the usual restriction homomorphism. Then the Leibniz seminorm L^ρ for ρ is π-compatible.*

PROOF. Let $f \in A$. Let Q be the radial retraction of $\mathbb{C}$ onto its ball of radius $\|f\|_\infty$ centered at 0. It is easily seen that the Lipschitz constant of Q is 1. Then for any $h \in C$ with $\pi(h) = f$ we can set $g = Q \circ h$ and we will have $\pi(g) = f$ and $L(g) \leq L(h)$ while $\|g\| = \|f\|$. This quickly gives the desired result. □

We remark that the above argument does not work for matrix-valued functions, as employed in [**32**], since the radial retraction no longer has Lipschitz constant 1 [**31**].

While Proposition 5.3 appears favorable, the difficulty is that the quotient of L^ρ on A need not agree with the Lipschitz seminorm from the metric ρ_X on X coming from restricting ρ:

EXAMPLE 5.4. (See [**39, 31**].) Let (X, ρ_X) be the metric space containing exactly 3 points, at distance 2 from each other. We can ask what the Gromov–Hausdorff distance is from (X, ρ_X) to a metric space consisting of one point, say p. It is easily seen that the answer is 1, with the metric ρ on $Z = X \cup \{p\}$ that extends ρ_X giving p distance 1 to each point of X. Now let f be the function on X which sends the three points of X to the three different cube roots of 1 in $\mathbb{C}$. It is not difficult to see that the extension of f to Z that has the smallest Lipschitz norm is the extension g that sends p to 0. But $L^\rho(g)$ is easily seen to be substantially larger than $L^{\rho_X}(f)$. As remarked in [**32, 31**], this is possible because the metric on Z is somewhat hyperbolic.

On the other hand, for any compact metric space (Z, ρ), any closed subset X of Z, and for any $f \in C_{\mathbb{R}}(X)$, there is a $g \in C_{\mathbb{R}}(Z)$ with $g|_X = f$, $\|g\| = \|f\|$ and $L^\rho(g) = L^{\rho_X}(f)$ [**31**]. This shows in particular that here L^{ρ_X} does coincide with the quotient seminorm from L^ρ. It also means that for the situation of Example 5.4 we have two Leibniz seminorms on $C(X)$ which agree on real-valued functions but are nevertheless distinct. (For a related phenomenon see [**23**].) From the comments at the end of the first paragraph of Section 4 we see that these two seminorms will give the same metrics on the set of probability measures on X, and in particular the same metrics on X.

We now turn our attention to Gromov–Hausdorff distance. Let (A, L_A) and (B, L_B) be C^*-metric spaces. The evident way to adapt the definition of quantum Gromov–Hausdorff distance given in definition 4.2 of [**29**] is to require that the seminorms L considered on $A \oplus B$ be C^*-metrics. Example 5.4 shows that we cannot require the quotient of L on A to agree with L_A, except on self-adjoint elements (though for the main class of examples considered in later sections they will agree, so those examples are better behaved than Example 5.4). Then we do not know whether the quotient is Leibniz. We could impose π-compatibility to ensure this, but then we still may not have the strong Leibniz property, so it is not clear that it is useful to impose this.

Perhaps as our topic develops in the future it will become clearer what are the best conditions to impose. Anyway, guided by the above observations, we set, parallel to notation 4.1 of [**29**]:

Notation 5.5. Let (A, L_A) and (B, L_B) be compact C^*-metric spaces. We let $\mathcal{M}_C(L_A, L_B)$ denote the collection of all C^*-metrics, L, on $A \oplus B$ such that the quotient of L on A agrees with L_A on self-adjoint elements of A, and similarly for the quotient of L on B.

We want to modify the definition of quantum Gromov–Hausdorff distance, dist_q, given in definition 4.2 of [**29**] by requiring that the seminorms involved there are in $\mathcal{M}_C(L_A, L_B)$. But I am not able to show that the resulting notion satisfies the triangle inequality. When one tries to imitate the proof of the triangle inequality for dist_q given in theorem 4.3 of [**29**], one of the main obstacles is in showing that the Lip-norm L_{AC} of lemma 4.6, which is defined as a quotient seminorm, is a C^*-metric. I would not be surprised if the triangle inequality fails. So the term "distance" should not be used. I will use instead the term "proximity". Thus:

Definition 5.6. Let (A, L_A) and (B, L_B) be compact C^*-metric spaces. We define their proximity by

$$\mathrm{prox}(A, B) = \inf\{\mathrm{dist}_H^{\rho_L}(S(A), S(B)) : L \in \mathcal{M}_C(L_A, L_B)\}.$$

This definition makes sense in the following way. Both $S(A)$ and $S(B)$ are closed subsets of $S(A \oplus B)$. Much as at the beginning of Section 4, ρ_L is a metric on $S(A \oplus B)$, and $\mathrm{dist}_H^{\rho_L}$ is ordinary Hausdorff distance with respect to ρ_L. We note that the hypotheses in the definition of $\mathcal{M}_C(L_A, L_B)$ are such that proposition 3.1 of [**29**] applies, so that for any $L \in \mathcal{M}_C(L_A, L_B)$ the restrictions of ρ_L to $S(A)$ and $S(B)$ coincide with ρ_{L_A} and ρ_{L_B}. Put another way, when we associate to each $L \in \mathcal{M}_C(L_A, L_B)$ its restriction to the self-adjoint part of $A \oplus B$ we obtain a map from $\mathcal{M}_C(L_A, L_B)$ to $\mathcal{M}(L_A^s, L_B^s)$, where L_A^s denotes the restriction of L_A to the self-adjoint part of A, and similarly for L_B^s. This map need not be either injective or surjective.

It is clear that

$$\mathrm{dist}_q(A, B) \leq \mathrm{prox}(A, B),$$

since $\mathrm{prox}(A, B)$ is an infimum over a subset of the seminorms used to define $\mathrm{dist}_q(A, B)$. Thus if we have a sequence (B^n, L_{B^n}) of C^*-metric spaces for which the sequence $\mathrm{prox}(A, B^n)$ converges to 0, then it follows that (B^n, L_{B^n}) converges to (A, L_A) for quantum Gromov–Hausdorff distance. For this reason the absence of the triangle inequality will not be too serious a problem. The advantage of prox, as mentioned earlier, is that the use of seminorms L on $A \oplus B$ that are C^*-metrics permits one to try to generalize to C^*-metric spaces the results about vector bundles obtained in [**32**] for ordinary metric spaces. (We plan to discuss this in a future paper.)

6. Bimodule bridges

In the development of quantum Gromov–Hausdorff distance given in [**29**] and used in [**30**], a very convenient method for constructing suitable seminorms L on $A \oplus B$ involved suitable continuous seminorms N on $A \oplus B$ that we called "bridges", with L then defined as

$$L(a, b) = L_A(a) \vee L_B(b) \vee N(a, b).$$

Within the context of the present paper it is natural to require that N satisfy a suitable Leibniz condition. There is an evident condition to consider, coming from viewing N as a seminorm on the algebra $A \oplus B$. But it seems more appropriate to require the stronger condition

$$N((a, b)(a', b')) \leq N(a, b)\|b'\| \ + \ \|a\|N(a', b').$$

Examples show that this condition can be interpreted as indicating that N only provides metric data between A and B, and not within A or within B.

We will find it very useful to use bridges that come from normed bimodules. Such bridges will satisfy the Leibniz condition stated above. Let A and B be unital C^*-algebras, and let Ω be an A-B-bimodule. We say that Ω is a normed bimodule if it is equipped with a norm that satisfies, much as in Section 2,

$$\|a\omega b\| \leq \|a\|\|\omega\|\|b\|$$

for all $a \in A$, $b \in B$ and $\omega \in \Omega$. We assume that the identity elements of A and B both act as the identity operator on Ω.

Definition 6.1. Let (A, L_A) and (B, L_B) be C^*-metric spaces. By a *bimodule bridge* for (A, L_A) and (B, L_B) we mean a normed A-B-bimodule Ω together with a distinguished element $\omega_0 \neq 0$ such that when we form the seminorm N on $A \oplus B$ defined by

$$N(a, b) = \|a\omega_0 - \omega_0 b\|,$$

it has the property that for any $a \in A$ with $a = a^*$ and any $\varepsilon > 0$ there is a $b \in B$ with $b^* = b$ such that

$$L_B(b) \vee N(a, b) \leq L_A(a) + \varepsilon,$$

and similarly for A and B interchanged.

Theorem 6.2. *Let (Ω, ω_0) be a bimodule bridge for the C^*-metric spaces (A, L_A) and (B, L_B), and let N be defined as above in terms of (Ω, ω_0). Define L on $A \oplus B$ by*

$$L(a, b) = L_A(a) \vee L_B(b) \vee N(a, b) \vee N(a^*, b^*).$$

Then $L \in \mathcal{M}_C(L_A, L_B)$.

PROOF. One can show directly that N is strongly Leibniz, or view Ω as an $(A \oplus B)$-bimodule in the evident way and apply Proposition 2.1. Since N is also continuous, it follows from Proposition 1.2 that L is lower-semicontinuous and strongly Leibniz. Clearly L is a $*$-seminorm. Thus condition a) of Definition 4.1 is satisfied.

We now want to apply theorem 5.2 of [**29**] to show that L is a Lip*-norm. We must thus show that $N \vee N^*$, restricted to the self-adjoint part of $A \oplus B$, is a bridge as defined in definition 5.1 of [**29**]. From its bimodule source it is clear that $N(1_A, 1_B) = 0$, while $N(1_A, 0) \neq 0$ since $\omega_0 \neq 0$. Since also N is continuous, it follows that the first two conditions of definition 5.1 are satisfied. The main technical condition of Definition 6.1 directly implies that condition 3 of definition 5.1 of [**29**] is satisfied, so that $N \vee N^*$ is indeed a bridge, and so L, restricted to self-adjoint elements, is a Lip-norm. Thus L is a Lip*-norm, and so condition b) of Definition 4.1 is satisfied.

Because N is clearly finite, $(A \oplus B)^f$, as defined for L, coincides with $A^f \oplus B^f$. From the fact that A^f and B^f are by assumption spectrally stable in their completion it follows easily that $(A \oplus B)^f$ is spectrally stable in its completion. Thus L satisfies condition c) of Definition 4.1, so that L is a C^*-metric.

Suppose now that we are given $a \in A$ with $a = a^*$. From the formula for L it is clear that $L(a, b) \geq L_A(a)$ for all $b \in B$. Let $\varepsilon > 0$ be given. Then by Definition 6.1 there is a $b \in B$ with $b = b^*$ such that

$$L_B(b) \vee N(a, b) \leq L_A(a) + \varepsilon.$$

Since N and N^* agree on self-adjoint elements, it follows that $L(a, b) \leq L_A(a) + \varepsilon$. Since ε is arbitrary, it follows that the quotient of L on A applied to a gives $L_A(a)$. In the same way the quotient of L on B, restricted to self-adjoint elements, gives L_B on self-adjoint elements. Thus $L \in \mathcal{M}_C(L_A, L_B)$. $\square$

In the next sections we will see how to construct useful bimodule bridges for "matrix algebras converging to the sphere".

Hanfeng Li has pointed out to me that prox is dominated by the "nuclear Gromov-Hausdorff distance" dist_{nu} that he defined in remark 5.5 of [**19**] and studied further in section 5 of [**14**]. He gives a proof of this in the appendix of [**20**]. (He uses the term "nuclear" because this distance has favorable properties for nuclear C^*-algebras.) We sketch here how this works, so that it can be easily compared with what we have done above. The crux of Li's approach is that he restricts attention to bimodules of a quite special kind. Specifically, for unital C^*-algebras A and B let $\mathcal{H}(A, B)$ denote the collection of all triples (D, ι_A, ι_B) consisting of a unital C^*-algebra D and injective (so isometric) unital homomorphisms ι_A and ι_B from A and B into D. We can then view D as an A-B-bimodule in the evident way. For a C^*-metric L_A on A Li sets

$$\mathcal{E}(L_A) \;=\; \{a \in A^{sa} :\; L_A(a) \leq 1\},$$

the L_A-unit-ball in A^{sa}. Then for any $(D, \iota_A, \iota_B) \in \mathcal{H}(A, B)$ he considers

$$\mathrm{dist}_H(\iota_A(\mathcal{E}(L_A)), \iota_B(\mathcal{E}(L_B))),$$

the ordinary Hausdorff distance in D for the norm of D. Even though $\mathcal{E}(L_A)$ and $\mathcal{E}(L_B)$ are unbounded, this distance is finite, for the following reason. Let r_A be the radius of (A, L_A), as defined in section 2 of [**26**], so that $\|\tilde{a}\|^{\sim} \leq\; r_A L_A(a)$ for any $a \in A^{sa}$, where $\tilde{}$ denotes image in the quotient $A^{sa}/\mathbb{R}1_A$, with the quotient norm. Then if $a \in \mathcal{E}(L_A)$ so that $L_A(a) \leq 1$, it follows that $a = a' + t1_A$ for some $t \in \mathbb{R}$ and $a' \in A^{sa}$ with $\|a'\| \leq r_A$. Let $b = t1_B$, so that $b \in \mathcal{E}(L_B)$. Then

$$\|\iota_A(a) - \iota_B(b)\| = \|a'\| \leq r_A.$$

Thus $\iota_A(\mathcal{E}(L_A))$ is in the r_A-neighborhood of $\iota_B(\mathcal{E}(L_B))$. By also interchanging the roles of a and b we see that

$$\mathrm{dist}_H(\iota_A(\mathcal{E}(L_A)), \iota_B(\mathcal{E}(L_B))) \ \leq \ \max(r_A, r_B).$$

Then Li defines $\mathrm{dist}_{nu}(A, B)$ (or, more precisely, $\mathrm{dist}_{nu}(L_A, L_B)$) to be

$$\inf\{\mathrm{dist}_H(\iota_A(\mathcal{E}(L_A)), \iota_B(\mathcal{E}(L_B))) : (D, \iota_A, \iota_B) \in \mathcal{H}(A, B)\}.$$

Li shows as follows that dist_{nu} satisfies the triangle inequality. Suppose that a third compact C^*-metric space (C, L_C) is given. Let $d_{AB} = \mathrm{dist}_{nu}(A, B)$, and similarly for d_{BC} and d_{AC}. Given $\varepsilon > 0$ we can find $(D, \iota_A, \iota_B) \in \mathcal{H}(A, B)$ and $(E, \rho_B, \rho_C) \in \mathcal{H}(B, C)$ such that

$$\mathrm{dist}_H(\iota_A(\mathcal{E}(L_A)), \iota_B(\mathcal{E}(L_B)) \ \leq \ d_{AB} \ + \ \varepsilon,$$

and similarly for d_{BC}. Let $F = D *_B E$ be *an* amalgamated product of D and E over B (using the inclusions ι_B and ρ_B). This means that there are unital injective homomorphisms σ_D and σ_E of D and E into F such that $\sigma_D \circ \iota_B \ = \ \sigma_E \circ \rho_B$. (It is natural to cut down to the subalgebra generated by the images of D and E in F.)

Before continuing, we remark that it is easy to construct a universal amalgamated free product, $A *_C B$, if one does not insist that the homomorphisms into it from A and B are injective. One takes the quotient of the universal (i.e. full) free product $A * B$ by the ideal generated by the desired relations from C. See [**21**]. What is not as simple is to show that the evident homomorphisms of A and B into the universal $A *_C B$ are injective. This was first shown by Blackadar in [**3**]. In a comment added in proof in that paper, Blackadar says that John Phillips has shown him a preferable proof. Blackadar has shown me this proof of John Phillips, and since it seems not to have appeared in print up to now, we sketch it here. Hanfeng Li has pointed out to me that a version of the argument in a substantially more complicated situation appears in the proof of proposition 2.2 of [**2**].

To simplify notation we simply view C as a unital subalgebra of each of A and B. The crux of the matter is to show that there are faithful (non-degenerate) representations of A and B on the same Hilbert space whose restrictions to C are equal. We construct such representations as follows.

(1) Let $(\pi_1, \mathcal{H}_1)$ be a faithful representation of A. Form the restricted representation $(\pi_1|_C, \mathcal{H}_1)$ of C, and extend it to a representation $(\rho_1, \mathcal{H}_1 \oplus \mathcal{K}_1)$ of B. (This can be done by decomposing into cyclic representations and extending their states – see lemma 2.1 of [**2**] .)
(2) Notice that $\rho_1|_C$ carries $\mathcal{H}_1$ into itself and so carries $\mathcal{K}_1$ into itself. Extend $(\rho_1|_C, \mathcal{K}_1)$ to a representation $(\pi_2, \mathcal{K}_1 \oplus \mathcal{H}_2)$ of A.
(3) Extend $(\pi_2|_C, \mathcal{H}_2)$ to a representation $(\rho_2, \mathcal{H}_2 \oplus \mathcal{K}_2)$ of B.
(4) Continue this process through all the positive integers, and form $\mathcal{H} = \bigoplus_1^\infty(\mathcal{H}_j \oplus \mathcal{K}_j)$. The π_j's and ρ_j's combine to give representations π and ρ of A and B on $\mathcal{H}$ which can be checked to agree on C. Since π_1 was chosen to be a faithful representation of A, so is π. Thus the homomorphism from

A into $A *_C B$ must be injective. The situation is symmetric for A and B, so the homomorphism from B into $A *_C B$ must also be injective.

We return to demonstrating the triangle inequality for dist_{nu}. Let $\tau_A = \sigma_D \circ \iota_A$ and $\tau_C = \sigma_E \circ \rho_C$. Then $(F, \tau_A, \tau_C) \in \mathcal{H}(A, C)$. Furthermore, if $a \in \mathcal{E}(L_A)$ then there is a $b \in \mathcal{E}(L_B)$ such that $\|\iota_A(a) - \iota_B(b)\| \leq d_{AB} + \varepsilon$, so that $\|\tau_A(a) - \sigma_D(\iota_B(b))\| \leq d_{AB} + \varepsilon$. In the same way there exists $c \in \mathcal{E}(L_C)$ such that $\|\sigma_E(\rho_B(b)) - \tau_C(c)\| \leq d_{BC} + \varepsilon$. But $\sigma_D(\iota_B(b)) = \sigma_E(\rho_B(b))$, and so

$$\|\tau_A(a) - \tau_C(c)\| \leq d_{AB} + d_{BC} + 2\varepsilon.$$

In this way we find that

$$\mathrm{dist}_{nu}(L_A, L_C) \ \leq \ \mathrm{dist}_{nu}(L_A, L_B) \ + \ \mathrm{dist}_{nu}(L_B, L_C).$$

Further favorable properties of dist_{nu} are presented in [**19, 14**] that we will not discuss here.

Given $(D, \iota_A, \iota_B) \in \mathcal{H}(A, B)$, we can view D as a normed A-B-bimodule in the evident way, and as special element we can choose $\omega_0 = 1_D$. The corresponding bounded seminorm N_D on $A \oplus B$ is then simply defined by

$$N_D(a, b) = \|\iota_A(a) - \iota_B(b)\|.$$

Given C^*-metrics L_A and L_B on A and B, we can seek constants γ such that $\gamma^{-1} N_D$ is a bimodule bridge for L_A and L_B. Let $\delta = \mathrm{dist}_H(\iota_A(\mathcal{E}(L_A)), \iota_B(\mathcal{E}(L_B)))$. Given any $\varepsilon > 0$ we show that $\delta + \varepsilon$ is such a constant. Let $a \in A^{sa}$ with $L_A(a) = 1$. Then there is a $b \in B^{sa}$ such that $L_B(b) \leq 1$ and $\|\iota_A(a) - \iota_B(b)\| \leq \delta + \varepsilon$, so that

$$L_B(b) \vee (\delta + \varepsilon)^{-1} N_D(a, b) \leq 1 = L_A(a).$$

We can interchange the roles of A and B. Thus we see that $(\delta+\varepsilon)^{-1} N_D$ is indeed a bimodule bridge. Notice that for any $a \in A$ and $b \in B$ we have $N(a^*, b^*) = N(a, b)$. Thus when we define L on $A \oplus B$ by

$$L(a, b) = L_A(a) \vee L_B(b) \vee (\delta + \varepsilon)^{-1} N_D(a, b)$$

it follows from Theorem 6.2 that $L \in \mathcal{M}_C(L_A, L_B)$.

But even more is true. As suggested by Li, we will follow the argument in the last paragraph of the proof of proposition 4.7 of [**18**]. Let $\mu \in S(A)$. View A and B as subalgebras of D via ι_A and ι_B. By the Hahn-Banach theorem, extend μ to a state $\tilde{\nu}$ of D, and then restrict $\tilde{\nu}$ to B to get $\nu \in S(B)$. Then for $a \in A^{sa}$ and $b \in B^{sa}$ we have

$$|\mu(a) - \nu(b)| = |\tilde{\nu}(a) - \tilde{\nu}(b)| \leq \|\iota_A(a) - \iota_B(b)\| \leq (\delta + \varepsilon) L(a, b),$$

where L is defined as above. Consequently if $L(a, b) \leq 1$ then we have $|\mu(a) - \nu(b)| \leq (\delta + \varepsilon)$. Thus μ is in the $\delta + \varepsilon$-neighborhood of $S(B)$ for the metric ρ_L on $S(A \oplus B)$. The same argument works with the roles of A and B reversed. Since ε is arbitrary, we see from this that

$$\mathrm{prox}(A, B) \leq \mathrm{dist}_{nu}(A, B),$$

as asserted.

In [**19, 14**] Li indicates that dist_{nu} works very well with many of the classes of specific examples whose metric aspects have been studied. In particular, he pointed out to me that dist_{nu} can be used to give an alternate proof of our Main Theorem (in a qualitative way). This alternate proof is attractive because of its quite general approach. However, a proof via dist_{nu} appears to me to be less concrete and quantitative than that which we give in the next sections, both because the proof via dist_{nu} uses a somewhat deep theorem of Blanchard on the subtrivialization of continuous fields of nuclear C^*-algebras (as discussed in remark 5.5 of [**19**]), and because of its use of the Hahn-Banach theorem seen just above. The proof we will give provides specific estimates for the approximation, and provides a constructive way of finding a state for one of the algebras that is close to a given state of the other algebra.

Motivated by Li's definition of his nuclear distance I did find, for use in [**33**], some convenient bimodules that are C^*-algebras. But the corresponding injective maps are not unital, so these bimodules do not actually fit into Li's framework.

7. Matrix algebras and homogeneous spaces

In this section we begin the study of our main example. Our discussion will be fairly parallel to that in [**30**] but with some important differences. For the reader's convenience we will include here some fragments of [**30**] in order to make precise our setting. We will usually use the notation used in [**30**].

Let G be a compact group (perhaps even finite at first). Let U be an irreducible unitary representation of G on a Hilbert space $\mathcal{H}$. Let $B = \mathcal{L}(\mathcal{H})$ denote the C^*-algebra of linear operators on $\mathcal{H}$ (a "full matrix algebra"). There is a natural action, α, of G on B by conjugation by U. That is, $\alpha_x(T) = U_x T U_x^*$ for $x \in G$ and $T \in B$. We introduce metric data into the picture by choosing a continuous length-function, ℓ, on G. We require that ℓ satisfy the additional condition that $\ell(xyx^{-1}) = \ell(y)$ for $x, y \in G$. This ensures that the metric on G defined by ℓ is invariant under both left and right translations. As in Example 2.5 we define a seminorm, L_B, on B by

$$L_B(T) = \sup\{\|\alpha_x(T) - T\|/\ell(x) : x \neq e_G\}.$$

Then L_B is a C^*-metric on B for the reasons given in Example 4.3.

Let P be a rank-one projection in B. Let $H = \{x \in G : \alpha_x(P) = P\}$, the stability subgroup for P. Let $A = C(G/H)$, the C^*-algebra of continuous complex-valued functions on G/H. We let λ denote the usual action of G on G/H, and so on A, by translation. We define a seminorm, L_A, on A as in Example 2.5 by

$$L_A(f) = \sup\{\|\lambda_x(f) - f\|/\ell(x) : x \neq e_G\}.$$

Again, L_A is a C^*-metric for the reasons given in Example 4.3.

We can then ask for estimates of $\mathrm{prox}(A, B)$. To obtain such an estimate we need to construct a suitable C^*-metric on $A \oplus B$. We do this as follows. For any $T \in B$ its Berezin covariant symbol, σ_T, is defined by

$$\sigma_T(x) = \mathrm{tr}(T\alpha_x(P)),$$

for $x \in G$. Here tr is the usual unnormalized trace on B. Because of the definition of H we see that $\sigma_T \in C(G/H) = A$. When the $\alpha_x(P)$'s are viewed as giving states of B via tr as above, they form a "coherent state", assigning a pure state of B to each pure state of A. Once we note that tr is α-invariant, it is easy to see that σ is a unital, positive, norm-non-increasing α-λ-equivariant operator from B to A. However eventually one really wants also the property that if $\sigma_T = 0$ then $T = 0$. This is equivalent to the linear span of the $\alpha_x(P)$'s in B being all of B. It is an interesting question as to which representations U admit such a P, and how many such P's, even for finite groups.

We let $\Omega = \mathcal{L}(B, A)$, the Banach space of linear operators from B to A, equipped with the operator norm corresponding to the C^*-norms on A and B.

(Perhaps we should be using the space of completely bounded operators here.) We let M and Λ denote the left regular representations of A and B. Then Ω is an A-B-bimodule for the operations

$$f\omega = M_f \circ \omega \qquad \text{and} \qquad \omega T = \omega \circ \Lambda_T.$$

It is easily checked that Ω is a normed A-B-bimodule. Of course $\sigma \in \Omega$. We will take our bimodule bridge for (A, L_A) and (B, L_B) to be of the form $(\Omega, \gamma^{-1}\sigma)$ where γ is a positive real number that is yet to be determined. Set

$$N_\sigma(f, T) = \|M_f \circ \sigma - \sigma \circ \Lambda_T\|.$$

Then the seminorm N from $(\Omega, \gamma^{-1}\sigma)$ is defined by

$$N(f, T) = \gamma^{-1} N_\sigma(f, T).$$

We need to determine the values of γ for which $(\Omega, \gamma^{-1}\sigma)$ is a bimodule bridge so that, in particular, the corresponding seminorm L has L_A and L_B as quotients for self-adjoint elements. But, as a first step in showing what the implication for proximity will be, we have:

Proposition 7.1. *Suppose that γ is such that $(\Omega, \gamma^{-1}\sigma)$ is a bimodule bridge for L_A and L_B, and let N be the seminorm it determines. Let $L = L_A \vee L_B \vee N \vee N^*$ and let ρ_L be the metric on $S(A \oplus B)$ that L determines. Then $S(A)$ is in the γ-neighborhood of $S(B)$ for ρ_L.*

Proof. Let $\mu \in S(A)$. We must find a $\nu \in S(B)$ such that $\rho_L(\mu, \nu) \leq \gamma$. We choose $\nu = \mu \circ \sigma$. Let $(f, T) \in A \oplus B$ be such that $L(f, T) \leq 1$, so that $N(f, T) \leq 1$ and thus $N_\sigma(f, T) \leq \gamma$. Then

$$\begin{aligned}|\mu(f, T) - \nu(f, T)| &= |\mu(f) - \mu(\sigma_T)| \leq \|f - \sigma_T\| \\ &= \|(M_f \circ \sigma - \sigma \circ \Lambda_T)(I)\| \leq N_\sigma(f, T) \leq \gamma,\end{aligned}$$

where I is the identity element in B. From the definition of ρ_L it follows that $\rho_L(\mu, \nu) \leq \gamma$. ☐

We remark that in our earlier paper on "matrix algebras converge to the sphere" [**30**] the bridge N that we had used was $N(f, T) = \gamma^{-1}\|f - \sigma_T\|$. The above calculation reveals that this old N is related to our new one just by applying our $M_f \circ \sigma - \sigma \circ \Lambda_T$ to the identity operator. The old N is not Leibniz.

To proceed further we now obtain another expression for N_σ which will be more convenient for some purposes. We note that for $S, T \in B$ and $f \in A$ we have

$$(M_f \circ \sigma - \sigma \circ \Lambda_T)(S) = f\sigma_S - \sigma_{TS},$$

and that when this is evaluated at $x \in G/H$ we obtain

$$\begin{aligned}f(x)\sigma_S(x) - \sigma_{TS}(x) &= f(x)\,\mathrm{tr}(S\alpha_x(P)) - \mathrm{tr}(TS\alpha_x(P)) \\ &= \mathrm{tr}(\alpha_x(P)(f(x)I - T)S).\end{aligned}$$

The operator norm of $M_f \circ \sigma - \sigma \circ \Lambda_T$ is then the supremum of the absolute value of the above expression taken over all $x \in G/H$ and $S \in B$ with $\|S\| \leq 1$. But tr gives a pairing that expresses the dual of B with its operator norm as B with the trace-class norm, which we denote by $\|\cdot\|_1$. From this fact we see that

$$\|M_f \circ \sigma - \sigma \cdot \Lambda_T\| = \sup\{\|\alpha_x(P)(f(x)I - T)\|_1 : x \in G/H\}.$$

But if R is a rank-one operator then $R^*R = r^2Q$ for some rank-one projection Q and some $r \in \mathbb{R}^+$, so that

$$\|R\|_1 = \mathrm{tr}((R^*R)^{1/2}) = r = \|R^*R\|^{1/2} = \|R\|,$$

where the norm on the right is the operator norm. In this way we obtain:

PROPOSITION 7.2. *For $f \in A$ and $T \in B$ we have*

$$N_\sigma(f,T) = \sup\{N_x(f,T) : x \in G/H\}$$

where $N_x(f,T) = \|\alpha_x(P)(f(x)I - T)\|$.

We remark that $N_x(f,T)$ can easily be checked to be strongly Leibniz. Also, see the first section of [**33**] for a different, perhaps more convenient, normed bimodule that determines N_σ.

8. The choice of the constant γ

Let us first see what choices of γ ensure that L has L_A as a quotient. It suffices to choose γ such that for any $f \in A$ we can find $T \in B$ such that $L_B(T) \vee N(f,T) \leq L_A(f)$. On G/H let us momentarily use the G-invariant measure of mass 1 to give A the norm from $L^2(G/H)$. Similarly, on B we put the Hilbert–Schmidt norm from the *normalized* trace. Then σ has an adjoint operator, which we denote by $\breve{\sigma}$. It is easily computed [**30**] to be defined by

$$\breve{\sigma}_f = d\int_{G/H} f(x)\alpha_x(P)dx,$$

where d is the dimension of $\mathcal{H}$. One can easily verify that $\breve{\sigma}$ is a positive and λ-α-equivariant map from A to B. Furthermore, $\breve{\sigma}_1 = d\int \alpha_x(P)dx$, which is clearly α-invariant, and so is a scalar multiple of I since U is irreducible. But clearly the usual trace of $d\int \alpha_x(P)dx$ is d. Thus $\breve{\sigma}_1 = I$, that is, $\breve{\sigma}$ is unital. (This is why we used the normalized traces in defining $\breve{\sigma}$.) It follows that $\breve{\sigma}$ is also norm non-increasing.

Then, given $f \in A$, we will choose T to be $T = \breve{\sigma}_f$. It is easily seen (as in the proof of proposition 1.1 of [**30**]) that $L_B(\breve{\sigma}_f) \leq L_A(f)$. For any $x \in G/H$ we have by equivariance of $\breve{\sigma}$

$$N_x(f,\breve{\sigma}_f) = \|\alpha_x(P)(f(x)I - \breve{\sigma}_f)\| = \|P((\lambda_x^{-1}f)(e)I - \breve{\sigma}_{\lambda_x^{-1}f})\|.$$

Since f is arbitrary and L_A is λ-invariant, it clearly suffices for us to consider $\|P(f(e)I - \breve{\sigma}_f)\|$. But

$$\begin{aligned}\|P(f(e)I - \breve{\sigma}_f)\| &= \left\|P\left(f(e)d\int \alpha_y(P)dy - d\int f(y)\alpha_y(P)dy\right)\right\| \\ &= d\left\|\int (f(e) - f(y))P\alpha_y(P)dy\right\| \\ &\leq L_A(f)\ d\int \rho_{G/H}(e,y)\|P\alpha_y(P)\|dy,\end{aligned}$$

where $\rho_{G/H}$ is the ordinary metric on G/H from L_A. From all of this we obtain:

PROPOSITION 8.1. *Set $\gamma^A = d\int \rho_{G/H}(e,y)\|P\alpha_y(P)\|dy$. Then for any $\gamma \geq \gamma^A$ the seminorm $L = L_A \vee L_B \vee \gamma^{-1}(N_\sigma \vee N_\sigma^*)$ on $A \oplus B$ has L_A as its quotient on A.*

We remark that in the above proposition we do not have to restrict attention to self-adjoint elements, in contrast to the requirement in Definition 6.1. Note that $\breve{\sigma}_{\bar{f}} = (\breve{\sigma}_f)^*$. I do not know whether the above condition on γ is the best that can be obtained in the absence of further hypotheses on G, U, P and ℓ.

We now consider the quotient of L on B. Given $T \in B$ we seek $f \in A$ such that $L_A(f) \vee N(f,T) \le L_B(T)$. We choose $f = \sigma_T$, and seek what requirement this puts on γ. As above, it is easy to check that $L_A(\sigma_T) \le L_B(T)$. Again by equivariance we have

$$N_x(\sigma_T, T) = \|\alpha_x(P)(\mathrm{tr}(T\alpha_x(P))I - T)\| = \|P(\mathrm{tr}(P\alpha_x^{-1}(T))I - \alpha_x^{-1}(T))\|.$$

Since T is arbitrary and L_B is α-invariant, it suffices to choose γ large enough that $\|P\,\mathrm{tr}(PT) - PT\| \le \gamma L_B(T)$ for all $T \in B$. Notice that the left-hand side gives a seminorm (with value 0 for $T = P$ or I) on the quotient space $\tilde{B} = B/\mathbb{C}I$, while L_B gives a norm on $\tilde{B}$. Since B is finite-dimensional, there does exist a finite γ such that the above inequality is satisfied. Notice also that $\sigma_{T^*} = (\sigma_T)^-$. Thus we obtain:

Proposition 8.2. *Define γ^B by*

$$\gamma^B = \sup\{\|P\,\mathrm{tr}(PT) - PT\| : T \in B \text{ and } L_B(T) \le 1\}.$$

Then γ^B is finite, and for any $\gamma \ge \gamma^B$ the seminorm $L = L_A \vee L_B \vee \gamma^{-1}(N_\sigma \vee N_\sigma^*)$ on $A \oplus B$ has L_B as its quotient on B.

For later use we now express $\|P(\mathrm{tr}(PT)I - T)\|$ in a different form. Since taking adjoints is an isometry, and by the C^*-relation, and by the fact that if R is a positive operator then $\|PRP\| = \mathrm{tr}(PRP)$ because P is of rank 1, we have

$$\begin{aligned}\|P(\mathrm{tr}(PT)I - T)\|^2 &= \|P(\mathrm{tr}(PT)I - T)(\mathrm{tr}(PT)I - T)^*P\| \\ &= \mathrm{tr}\left(P(\mathrm{tr}(PT)I - T)(\mathrm{tr}(PT)I - T)^*P)\right) \\ &= |\,\mathrm{tr}(PT)|^2 - \mathrm{tr}(PTP)\overline{\mathrm{tr}}(PT) \\ &\quad - \mathrm{tr}(PT)\,\mathrm{tr}(PT^*P) + \mathrm{tr}(PTT^*P) \\ &= \mathrm{tr}(PTT^*P) - |\,\mathrm{tr}(PT)|^2.\end{aligned}$$

Thus:

Proposition 8.3. *For any $T \in B$ we have*

$$\|P(\mathrm{tr}(PT)I - T)\| = (\mathrm{tr}(PTT^*P) - |\,\mathrm{tr}(PT)|^2)^{1/2}.$$

We remark that if ξ is a unit vector in the range of P then

$$\mathrm{tr}(PTT^*P) - |\,\mathrm{tr}(PT)|^2 = \langle TT^*\xi, \xi\rangle - |\langle T^*\xi, \xi\rangle|^2.$$

When T is self-adjoint this is the "mean-square deviation" of T in the state determined by ξ [**37**].

We now need to consider how small a neighborhood of $S(A)$ contains $S(B)$. Let $\nu \in S(B)$ be given. We choose $\mu = \nu \circ \breve{\sigma}$, and observe that $\mu \in S(A)$. Let $(f,T) \in A \oplus B$ be such that $L(f,T) \le 1$, so that $N_\sigma(f,T) \le \gamma$. Then

$$\begin{aligned}|\mu(f,T) - \nu(f,T)| &= |\nu(\breve{\sigma}_f) - \nu(T)| \le \|\breve{\sigma}_f - T\| \\ &= \left\| d\int f(x)\alpha_x(P)dx - d\int \alpha_x(P)T dx \right\|\end{aligned}$$

$$= d\left\|\int \alpha_x(P)(f(x)I - T)dx\right\| \le d\int N_x(f,T)dx$$
$$\le dN_\sigma(f,T) \le d\gamma.$$

But the presence of d here causes us difficulties later, so we take another path, namely that used near the end of section 2 of [**30**]. We have

$$\|\breve{\sigma}_f - T\| \le \|\breve{\sigma}_f - \breve{\sigma}(\sigma_T)\| + \|\breve{\sigma}(\sigma_T) - T\|$$
$$\le \|f - \sigma_T\| + \|\breve{\sigma}(\sigma_T) - T\| \le \gamma^A + \|\breve{\sigma}(\sigma_T) - T\|,$$

where we have used that $\|f - \sigma_T\| \le N_\sigma(f,T)$, as seen in the proof of Proposition 7.1. Notice that $T \mapsto \|\breve{\sigma}(\sigma_T) - T\|$ is a seminorm on B which takes value 0 for $T = I$, and so descends to a seminorm on $\tilde{B} = B/\mathbb{C}I$, where L_B becomes a norm.

Notation 8.4. We set

$$\delta^B = \sup\{\|T - \breve{\sigma}(\sigma_T)\| : L_B(T) \le 1\}.$$

With this notation the above discussion gives:

Proposition 8.5. *Suppose that $\gamma \ge \gamma^A \vee \gamma^B$, so that L has L_A and L_B as quotients (where $L = L_A \vee L_B \vee \gamma^{-1}(N_\sigma \vee N_\sigma^*)$). Then $S(B)$ is in the $(\gamma^A + \delta^B)$-neighborhood of $S(A)$.*

9. The set-up for compact Lie groups

We now specialize to the case in which G is a compact connected semisimple Lie group. We use many of the techniques used in sections 6 and 7 of [**30**], and we usually use the notation established in sections 5 and 6 of [**30**]. We now review that notation. We let $\mathfrak{g}_0$ denote the Lie algebra of G, while $\mathfrak{g}$ denotes the complexification of $\mathfrak{g}_0$. We choose a maximal torus in G, with corresponding Cartan subalgebra of $\mathfrak{g}$, its set of roots, and a choice of positive roots. We let $(U, \mathcal{H})$ be an irreducible unitary representation of G, and we let U also denote the corresponding representation of $\mathfrak{g}$. We choose a highest weight vector, ξ, for $(U, \mathcal{H})$ with $\|\xi\| = 1$. For any $n \in \mathbb{Z}_{\ge 1}$ we set $\xi^n = \xi^{\otimes n}$ in $\mathcal{H}^{\otimes n}$, and we let $(U^n, \mathcal{H}^n)$ be the restriction of $U^{\otimes n}$ to the $U^{\otimes n}$-invariant subspace, $\mathcal{H}^n$, of $\mathcal{H}^{\otimes n}$ which is generated by ξ^n. Then $(U^n, \mathcal{H}^n)$ is an irreducible representation of G with highest weight vector ξ^n, and its highest weight is just n times the highest weight of $(U, \mathcal{H})$. We denote the dimension of $\mathcal{H}^n$ by d_n.

We let $B^n = \mathcal{L}(\mathcal{H}^n)$. The action of G on B^n by conjugation by U^n will be denoted simply by α. We assume that a continuous length function, ℓ, has been chosen for G, and we denote the corresponding C^*-metric on B^n by L_n^B. We let P^n denote the rank-one projection along ξ^n. Then the α-stability subgroup H for $P = P^1$ will also be the stability subgroup for each P^n. Let γ_n^A and γ_n^B be the constants defined in Propositions 8.1 and 8.2 but for P^n.

As done earlier, we let $A = C(G/H)$, and we let L_A be the seminorm on A for ℓ and the action of G. We can now state the main theorem of this paper.

Theorem 9.1. *Let notation be as above. Set $\gamma_n = \max\{\gamma_n^A, \gamma_n^B\}$ for each n, and let L_n be defined on $A \oplus B^n$ as in Proposition 7.1 but using γ_n. Then $L_n \in \mathcal{M}_C(L_A, L_B)$, and the sequence $\{L_n\}$ shows that the sequence $\{\mathrm{prox}(A, B^n)\}$ converges to 0 as n goes to ∞.*

The next three sections will be devoted to the proof of this theorem.

10. The proof that $\gamma_n^A \to 0$

Consistent with the notation of Proposition 8.1, we have set

$$\gamma_n^A = d_n \int \rho_{G/H}(e,x)\|P^n\alpha_x(P^n)\|dx.$$

PROPOSITION 10.1. *The sequence* $\{\gamma_n^A\}$ *converges to* 0.

PROOF. For any two vectors η, ζ we let $\langle \eta, \zeta\rangle_0$ denote the rank-one operator that they determine. Then for any n we have

$$\begin{aligned}\|P^n\alpha_x(P^n)\| &= \|\langle \xi^n, \xi^n\rangle_0 \langle U_x^n\xi^n, U_x^n\xi^n\rangle_0\| \\ &= |\langle U_x^n\xi^n, \xi^n\rangle| = |\langle U_x\xi, \xi\rangle|^n = \|P\alpha_x(P)\|^n.\end{aligned}$$

We use the analogous treatment given in lemma 3.3 and theorem 3.4 of [**30**], where it is shown that $d_n|\langle U_x\xi,\xi\rangle|^{2n}dx$ $(= d_n\|P^n\alpha_x(P^n)\|^2dx)$ is a probability measure on G/H, and that the sequence of these probability measures converges in the weak-$*$ topology to the δ-measure on G/H supported at eH. Since $\rho_{G/H}(e,e) = 0$, it follows that the sequence $d_n \int \rho_{G/H}(e,x)\|P^n\alpha_x(P^n)\|^2dx$ converges to 0 . Now

$$\begin{aligned}\gamma_{2n}^A =& d_{2n}\int \rho_{G/H}(e,x)\|P\alpha_x(P)\|^{2n}dx \\ &= (d_{2n}/d_n)d_n\int \rho_{G/H}(e,x)\|P^n\alpha_x(P^n)\|^2dx,\end{aligned}$$

and so if we can show that (d_{2n}/d_n) is bounded, then we find that the sequence $\{\gamma_{2n}\}$ converges to 0. We use the Weyl dimension formula, as presented for example in theorem 4.14.6 of [**38**], to show that $\{d_{2n}/d_n\}$ is bounded. We let ω be the highest weight of U for our choice $\mathcal{P}$ of positive roots. If one examines the dimension formula, it is evident that one only needs to use those positive roots α such that $\langle\omega,\alpha\rangle > 0$. We denote this set by $\mathcal{P}_\omega$, and we denote its cardinality by p. It is clear that for any $n \in \mathbb{Z}_{>0}$ we have $\mathcal{P}_{n\omega} = \mathcal{P}_\omega$. The Weyl dimension formula then tells us that

$$d_n = \Big(\prod\langle n\omega+\delta,\alpha\rangle\Big)\Big/\Big(\prod\langle\delta,\alpha\rangle\Big)$$

where both products are taken over $\mathcal{P}_\omega$, and δ is half the sum of the positive roots. Thus

$$\begin{aligned}d_{2n}/d_n &= \Big(\prod\langle 2n\omega+\delta,\alpha\rangle\Big)\Big/\Big(\prod\langle n\omega+\delta,\alpha\rangle\Big) \\ &= \prod(1+\langle n\omega,\alpha\rangle/\langle n\omega+\delta,\alpha\rangle) \le 2^p,\end{aligned}$$

so that the sequence d_{2n}/d_n is bounded as needed, and consequently the sequence $\{\gamma_{2n}^A\}$ converges to 0. In the same way, we find that $d_{n+1}/d_n \le (1+n^{-1})^p$. Since $0 \le \|P\alpha_x(P)\| \le 1$, we have $\|P\alpha_x(P)\|^n \ge \|P\alpha_x(P)\|^{n+1}$. Thus the integrals defining γ_n^A are non-increasing. It follows that $\gamma_{2n+1}^A \le (1+(2n)^{-1})^p\gamma_{2n}^A$. Since the sequence $\{\gamma_{2n}^A\}$ converges to 0, it follows that the sequence $\{\gamma_{2n+1}^A\}$ does also, so that the sequence $\{\gamma_n^A\}$ converges to 0. $\square$

11. Properties of Berezin symbols

We now need results related to those given in sections 4 and 5 of [**30**], leading to the proof of theorem 6.1 of [**30**], and we will shortly also need theorem 6.1 of [**30**] itself. But Jeremy Sain has found a substantial simplification of the proof of theorem 6.1 of [**30**]. He gives his argument in section 4.4 of [**35**] in the more complicated context of quantum groups. We will use his arguments here in our present context. This will in particular provide Sain's proof of theorem 6.1 of [**30**].

As in [**30**], we denote the Berezin symbol map from B^n to $A = C(G/H)$ by σ^n. From theorem 3.1 of [**30**] we find that σ^n is injective because ξ^n is a highest weight vector. Consistent with the notation defined near the beginning of Section 8, we denote the adjoint of σ^n by $\breve{\sigma}^n$. We let

$$\delta_n^A = \int_{G/H} \rho_{G/H}(e,x) d_n \mathrm{tr}(P^n \alpha_x(P^n))\, dx. \tag{11.1}$$

In section 3 of [**30**] δ_n^A was denoted by γ_n, and theorem 3.4 of [**30**] shows both that the sequence $\{\delta_n^A\}$ converges to 0, and that

$$\|f - \sigma^n(\breve{\sigma}^n(f))\|_\infty \ \le\ \delta_n^A L_A(f) \tag{11.2}$$

for all $f \in A$ and all n. We remark that $\sigma^n \circ \breve{\sigma}^n$ is often called the "Berezin transform" (for a given n).

As in section 4 of [**30**] we let $\hat{G}$ denote the set of equivalence classes of irreducible unitary representations of G. For any finite subset $\mathcal{S}$ of $\hat{G}$ we let $A_{\mathcal{S}}$ and $B^n_{\mathcal{S}}$ denote the direct sum of the isotypic components of A and B^n for the representations in $\mathcal{S}$ and for the actions of G on A and B^n (and similarly for actions on other Banach spaces). Since σ^n is equivariant, it carries $B^n_{\mathcal{S}}$ into $A_{\mathcal{S}}$. Since σ^n is injective, it follows that the dimension of $B^n_{\mathcal{S}}$ is no larger than that of $A_{\mathcal{S}}$, which is finite.

Since $\{\delta_n^A\}$ converges to 0, it follows from Inequality 11.2 that $\sigma^n \circ \breve{\sigma}^n$ converges strongly to the identity operator on the space of functions f for which $L_A(f) < \infty$. But $A_{\mathcal{S}}$ is contained in this space and is finite-dimensional, and $\sigma^n \circ \breve{\sigma}^n$ carries $A_{\mathcal{S}}$ into itself for each n. Consequently $\sigma^n \circ \breve{\sigma}^n$ restricted to $A_{\mathcal{S}}$ converges in norm to the identity operator on $A_{\mathcal{S}}$. It follows that there is an integer, $N_{\mathcal{S}}$, such that $\sigma^n \circ \breve{\sigma}^n$ on $A_{\mathcal{S}}$ is invertible and $\|(\sigma^n \circ \breve{\sigma}^n)^{-1}\| < 2$ for every $n > N_{\mathcal{S}}$. In particular, σ^n from $B^n_{\mathcal{S}}$ to $A_{\mathcal{S}}$ will be surjective for $n > N_{\mathcal{S}}$. Since, as mentioned above, σ^n is always injective, and $\|\sigma^n\| = 1 = \|\breve{\sigma}^n\|$ for all n, we can quickly see that:

Lemma 11.3. *(See corollary 4.17 of* [**35**]*.) For $n > N_{\mathcal{S}}$ both σ^n and $\breve{\sigma}^n$ going between $A_{\mathcal{S}}$ and $B^n_{\mathcal{S}}$ are invertible and their inverses have operator norm no bigger than 2.*

Fix $n > N_{\mathcal{S}}$, and let $T \in B^n_{\mathcal{S}}$ be given. Set $f = (\breve{\sigma}^n)^{-1}(T)$. Note that f is well-defined, and that $\|f\|_\infty \le 2\|T\|$ by Lemma 11.3. Then

$$\|T - \breve{\sigma}^n(\sigma^n_T)\| = \|\breve{\sigma}^n(f) - \breve{\sigma}^n(\sigma^n(\breve{\sigma}^n_f))\| \le \|f - \sigma^n(\breve{\sigma}^n_f)\| \le \delta_n^A L_A(f),$$

where we have used Inequality 11.2 for the last inequality just above. Because $(\breve{\sigma}^n)^{-1}$ is α-λ-equivariant and $\|(\breve{\sigma}^n)^{-1}\| \le 2$, we have $L_A(f) \le 2L_n^B(T)$. We have thus obtained:

Lemma 11.4. *(See proposition 4.19 of* [**35**]*.) For any $n > N_{\mathcal{S}}$ and any $T \in B^n_{\mathcal{S}}$ we have*

$$\|T - \breve{\sigma}^n(\sigma^n_T)\| \le 2\delta_n^A L_n^B(T).$$

Choose a faithful finite-dimensional unitary representation, π_0, of G that contains the trivial representation, and let $\pi = \pi_0 \otimes \bar{\pi}_0$, where $\bar{\pi}_0$ is the contragredient representation for π_0. Let χ be the character of π. Then χ is a non-negative real-valued function on G. Since π is faithful, we have the strict inequality $\chi(x) < \chi(e)$

for any $x \in G$ with $x \neq e$. Let χ^m denote the character of $\pi^{\otimes m}$, so that equally well it is the m^{th} pointwise power of χ. Set $\varphi_m = \chi^m/\|\chi^m\|_1$. Then the sequence $\{\varphi_m\}$ is a norm-1 approximate identity for the convolution algebra $L^1(G)$, as seen in the proof of theorem 8.2 of [**29**]. Furthermore, each φ_m is central in $L^1(G)$. Let β be an isometric strongly continuous action of G on a Banach space D, and let L^D be the corresponding seminorm for ℓ. Let β_{φ_n} denote the corresponding "integrated form" operator. As in the proof of lemma 8.3 of [**29**], for each $d \in D$ we have

$$\begin{aligned} \|d - \beta_{\varphi_m}(d)\| &= \|d \int \varphi_m(x)\ dx \ - \ \int \varphi_m(x)\beta_x(d)\ dx\| \\ &\leq \int \varphi_m \|d - \beta_x(d)\| dx \leq \left(\int \varphi_m(x)\ell(x) dx \right) L^D(d), \end{aligned}$$

and the sequence $\left\{ \int \varphi_m(x)\ell(x)dx \right\}$ converges to 0.

We can now argue exactly as in the rest of the proof of theorem 6.1 of [**30**] to obtain:

THEOREM 11.5. *(Theorem 6.1 of* [**30**]*) For each $n \geq 1$ let δ_n^B be as defined in Notation 8.4 but for B^n, so that it is the smallest constant such that*

$$\|T - \breve{\sigma}^n(\sigma_T^n)\| \leq \delta_n^B L_n^B(T)$$

for all $T \in B^n$. Then the sequence $\{\delta_n^B\}$ converges to 0.

PROOF OF THEOREM 11.5. Let $\varepsilon > 0$ be given. We can choose $\varphi = \varphi_m$ as just above such that for any ergodic action β of G on any unital C^*-algebra C we have $\|c - \beta_\varphi(c)\| \leq (\varepsilon/3)L(c)$ for all $c \in C$. Now φ is a positive function, and is a linear combination of the characters of a finite subset $\mathcal{S}$ of $\hat{G}$, and so the integrated operator β_φ is a completely positive unital equivariant map of C onto its S-isotypic component.

Then for every n and every $T \in B^n$ we have $\alpha_\varphi(T) \in B^n_{\mathcal{S}}$ and

$$\|T - \breve{\sigma}^n(\sigma_T^n)\| \leq (\varepsilon/3)L_n^B(T) + \|\alpha_\varphi(T) - \breve{\sigma}^n(\sigma^n_{\alpha_\varphi(T)})\| + (\varepsilon/3)L_n^B(T).$$

From Lemma 11.4 there is an integer N_ε such that for any $n > N_\varepsilon$ and any $T' \in B^n_{\mathcal{S}}$ we have

$$\|T' - \breve{\sigma}^n(\sigma^n(T'))\| \leq (\varepsilon/3)L_n^B(T').$$

Since $\alpha_\varphi(T) \in B^n_{\mathcal{S}}$, we can apply this to $T' = \alpha_\varphi(T)$. When we use the fact that $L_n^B(\alpha_\varphi(T)) \leq L_n^B(T)$, we see that for any $n > N_\varepsilon$ and any $T \in B^n$ we have

$$\|T - \breve{\sigma}^n(\sigma_T^n)\| \leq \varepsilon L_n^B(T).$$

This immediately implies the statement about the sequence $\{\delta_n^B\}$.

∎

12. The proof that $\gamma_n^B \to 0$

Consistent with the notation of Proposition 8.2, we have set

$$\gamma_n^B = \sup\{\|P^n \operatorname{tr}(P^nT) - P^nT\| : T \in B^n \text{ and } L_n^B(T) \leq 1\}.$$

PROPOSITION 12.1. *The sequence $\{\gamma_n^B\}$ converges to* 0.

PROOF. Let $\varepsilon > 0$ be given. With the notation that we used just before Theorem 11.5, choose m_0 such that for $\varphi = \varphi_{m_0}$ we have $\int \varphi(x)\ell(x)dx \le \varepsilon/4$. Then by the calculation done there we have

$$\|T - \alpha_\varphi(T)\| \le (\varepsilon/4)L_n^B(T)$$

for all n and for all $T \in B^n$. Then for any n and any $T \in B^n$

$$\begin{aligned}
&\|(P^n \operatorname{tr}(P^n T) - P^n T) - (P^n \operatorname{tr}(P^n \alpha_\varphi(T)) - P^n \alpha_\varphi(T))\| \\
&\le |\operatorname{tr}(P^n(T - \alpha_\varphi(T)))| + \|T - \alpha_\varphi(T)\| \\
&\le 2\|T - \alpha_\varphi(T)\| \le (\varepsilon/2)L_n^B(T),
\end{aligned}$$

where for the next-to-last inequality we have used the fact that $P^n(T - \alpha_\varphi(T))$ is of rank 1.

Now as discussed in the proof of Theorem 11.5, φ is a linear combination of the characters of a finite subset $\mathcal{S}$ of $\hat{G}$. Thus $\alpha_\varphi(T) \in B^n_{\mathcal{S}}$ and $L_n^B(\alpha_\varphi(T)) \le L_n^B(T)$, and so we now see that it suffices to prove:

MAIN LEMMA 12.2. *Let $\mathcal{S}$ be given. For any $\varepsilon > 0$ there is an integer N_ε such that for any $n \ge N_\varepsilon$ and any $T \in B^n_{\mathcal{S}}$ we have*

$$\|P^n \operatorname{tr}(P^n T) - P^n T\| \le (\varepsilon/2)L_n^B(T).$$

PROOF. Let $f \in A$, and let n be given. Because A is commutative and $\breve{\sigma}^n$ is positive, it follows from Kadison's generalized Schwarz inequality (e.g. 10.5.9 of [**12**]) that we have

$$\breve{\sigma}^n_f(\breve{\sigma}^n_f)^* \;\le\; \breve{\sigma}^n_{f\bar{f}}$$

for the usual order on positive operators. When we combine this with Proposition 8.3 we obtain

$$\begin{aligned}
\|P^n(\operatorname{tr}(P^n\breve{\sigma}^n_f)I - \breve{\sigma}^n_f)\|^2 &= \operatorname{tr}(P^n\breve{\sigma}^n_f(\breve{\sigma}^n_f)^*P^n) - |\operatorname{tr}(P^n\breve{\sigma}^n_f)|^2 \\
&\le \operatorname{tr}(P^n\breve{\sigma}^n_{f\bar{f}}) - |\operatorname{tr}(P^n\breve{\sigma}^n_f)|^2 = (\sigma^n(\breve{\sigma}^n_{f\bar{f}}))(e) - |\sigma^n(\breve{\sigma}^n_f)(e)|^2,
\end{aligned}$$

which by Inequality 11.2 above and theorem 3.4 of [**30**] converges to

$$(f\bar{f})(e) - |f(e)|^2 = 0$$

as n increases.

For each n define an operator, J^n, on B^n by

$$J^n(T) = P^n(\operatorname{tr}(P^n T)I \;-\; T).$$

The calculation above shows that the sequence $J^n(\breve{\sigma}^n_f)$ converges to 0 for any $f \in A$ with $L^A(f) < \infty$. For $\mathcal{S}$ as above it follows that the sequence of restrictions of $J^n \circ \breve{\sigma}^n$ to $A_{\mathcal{S}}$ converges to 0 in operator norm. Let $N_{\mathcal{S}}$ be as in Lemma 11.3, so that $\|(\breve{\sigma}^n)^{-1}\| \le 2$ for $n > N_{\mathcal{S}}$. It follows that for $n > N_{\mathcal{S}}$ we have $\|J_n\| \le 2\|J^n \circ \breve{\sigma}^n\|$, so that the sequence of restrictions of J^n to $B^n_{\mathcal{S}}$ converges to 0 in norm. Thus for any $\varepsilon' > 0$ we can find an $n_{\varepsilon'}$ such that for $n > n_{\varepsilon'}$ and all $T \in B^n_{\mathcal{S}}$ we have

$$\|J^n(T)\| \le \varepsilon'\|T\|.$$

Now $J^n(I) = 0$, and so it follows that

$$\|J^n(T)\| \le \varepsilon'\|\tilde{T}\|^\sim,$$

where much as before $\|\cdot\|^\sim$ denotes the quotient norm on $\tilde{B}^n = B^n/\mathbb{C}I$. But by lemma 2.4 of [**25**] the radius of each of the algebras B^n is no larger than $r =$

$\int \ell(x)dx$, in the sense that $\|\tilde{T}\|^{\sim} \leq rL_n^B(T)$ for all $T \in B^n$. We include a slightly simpler proof here. For $T \in B^n$ let $\eta(T) = \int \alpha_x(T)\ dx$, so that $\eta(T) \in \mathbb{C}I$ since U^n is irreducible. Then

$$\|\tilde{T}\|^{\sim} \leq \|T - \eta(T)\| = \|\int (T - \alpha_x(T))dx\| \leq L_n^B(T)\int \ell(x)dx.$$

It follows that

$$J^n(T) \leq r\varepsilon' L_n^B(T).$$

Consequently, if we choose $\varepsilon' = \varepsilon/(2r)$, and set $N_\varepsilon = n_{\varepsilon'} \vee N_{\mathcal{S}}$, we find that for $n \geq N_\varepsilon$ we have

$$\|P^n \operatorname{tr}(P^n T) - P^n T\| \leq (\varepsilon/2)L_n^B(T)$$

for all $T \in B^n_{\mathcal{S}}$, as needed. □

□

13. The proof of the main theorem

We now use the results of the previous sections to prove Theorem 9.1. For any n set $\gamma_n = \max(\gamma_n^A, \gamma_n^B)$, and define L_n on $A \oplus B^n$ by

$$L_n(f, T) = L_A(f) \vee L_n^B(T) \vee \gamma_n^{-1}(N_{\sigma^n}(f, T) \vee N_{\sigma^n}(\bar{f}, T^*)).$$

Then for each n we have $\gamma_n \geq \gamma_n^A$ so that the quotient of L_n on A is L_A by Proposition 8.1, and we have $\gamma_n \geq \gamma_n^B$ so that the quotient of L_n on B^n is L_n^B by Proposition 8.2. Thus L_n is in $\mathcal{M}_C(L_A, L_n^B)$ as defined in Notation 5.5.

Then according to Proposition 7.1 (with notation as in Proposition 8.1 and in the sentence before Proposition 10.1), $S(A)$ is in the γ_n-neighborhood of $S(B^n)$ for ρ_{L_n}. Furthermore, according to Proposition 8.5 (with notation as in Theorem 11.5) $S(B^n)$ is in the $(\gamma_n^A + \delta_n^B)$-neighborhood of $S(A)$. It follows that

$$\operatorname{dist}_H^{\rho_{L_n}}(S(A), S(B^n)) \leq \max\{\gamma_n^A + \delta_n^B, \gamma^n\} = \max\{\gamma_n^A + \delta_n^B, \gamma_n^B\},$$

and so

$$\operatorname{prox}(A, B^n) \leq \max\{\gamma_n^A + \delta_n^B, \gamma_n^B\}.$$

But γ_n^A, δ_n^B and γ_n^B all converge to 0 as n goes to ∞, according to Proposition 10.1, Theorem 11.5 (theorem 6.1 of [**30**]), and Proposition 12.1 respectively. Consequently $\operatorname{prox}(A, B^n)$ converges to 0 as n goes to ∞, as desired.

14. Matricial seminorms

In this section we will briefly describe the relations between the previous sections of this paper and several variants of quantum Gromov–Hausdorff distance.

The first variant is the matricial quantum Gromov–Hausdorff distance introduced by Kerr [**13**]. It has the advantage that if two C^*-algebras with Lip-norms are at distance 0 for his distance then the C^*-algebras are isomorphic. We will not repeat here Kerr's definitions and results for general operator systems; rather we will only indicate, somewhat sketchily, what Kerr's variant says in the context of the present paper. For any unital C^*-algebra A and each $q \in \mathbb{Z}_{>0}$ the $*$-algebra $M_q(A)$ of $q \times q$ matrices with entries in A has a unique C^*-norm. The collection of these C^*-norms forms a "matricial norm" for A. Given unital C^*-algebras A and B, a linear map $\varphi : A \to B$ determines for each q a linear map, φ^q, from $M_q(A)$ to $M_q(B)$, by entry-wise application. One says that φ is "completely positive" if each φ^q is positive as a map between C^*-algebras. For each q let $UCP_q(A)$ denote the collection of all unital completely positive maps from A into $M_q(\mathbb{C})$. The $UCP_q(A)$'s are called the "matricial state-spaces" of A. All these considerations apply equally well to unital C^*-normed algebras, where "positive" is with respect to the completions.

Let a Lip*-norm, L, on A be specified. Then Kerr defines a metric, ρ_L^q, on $UCP_q(A)$ by

$$\rho_L^q(\varphi, \psi) = \sup\{\|\varphi(a) - \psi(a)\| : L(a) \leq 1\},$$

and he shows that the topology on $UCP_q(A)$ determined by ρ_L^q agrees with the point-norm topology (and so is compact). Now let (A, L_A) and (B, L_B) be unital C^*-algebras with Lip*-norms. Essentially as in definition 4.2 of [**29**] let $\mathcal{M}(L_A, L_B)$ denote the set of Lip*-norms on $A \oplus B$ whose quotients on the self-adjoint part agree with L_A and L_B. Note that $UCP_n(A)$ and $UCP_n(B)$ can be viewed as subsets of $UCP_n(A \oplus B)$ in an evident way. Then for each q Kerr defines the q-distance, dist_s^q, between A and B by

$$\mathrm{dist}_s^q(A, B) = \inf\{\mathrm{dist}_H^{\rho_L^q}(UCP_q(A), UCP_q(B)) : L \in \mathcal{M}(A, B)\},$$

and he defines the complete distance, dist_s, by

$$\mathrm{dist}_s(A, B) = \sup_q\{\mathrm{dist}_s^q(A, B)\}.$$

Finally (for our purposes), he shows that for our setting of coadjoint orbits with $A = C(G/H)$ and $B^n = \mathcal{L}(\mathcal{H}_n)$ with their Lip*-norms from a length function ℓ, one has

$$\lim_{n\to\infty} \mathrm{dist}_s(A, B^n) = 0.$$

We can quickly adapt Kerr's arguments to our Leibniz setting. For C^*-algebras A and B equipped with C^*-metrics, we define $\mathcal{M}_C(L_A, L_B)$ exactly as in Notation 5.5. Any L in $\mathcal{M}_C(L_A, L_B)$ is, in particular, a Lip*-norm, and so defines for each q the metric ρ_L^q on $UCP_q(A \oplus B)$. We can then define, for each q,

$$\mathrm{prox}^q(A, B) = \inf\{\mathrm{dist}_H^{\rho_L^q}(UCP_q(A), UCP_q(B)) : L \in \mathcal{M}_C(A \oplus B)\}.$$

Then we can define "complete proximity" by

$$\mathrm{prox}_s(A, B) = \sup_q\{\mathrm{prox}^q(A, B)\}.$$

Of course, we have

$$\mathrm{dist}_s(A, B) \leq \mathrm{prox}_s(A, B).$$

THEOREM 14.1. *For $A = C(G/H)$ and $B^n = \mathcal{L}(\mathcal{H}_n)$ with their C^*-metrics L_A and L_B^n as defined earlier in terms of a length function on G, we have*

$$\lim_{n\to\infty} \mathrm{prox}_s(A, B^n) = 0.$$

PROOF. We follow the outline of Kerr's example 3.13 of [**13**], but for a given n we set, as earlier,

$$L_n = L_A \vee L_n^B \vee N_n \vee N_n^*$$

with $N_n = \gamma_n^{-1} N_{\sigma^n}$ and with γ_n chosen exactly as in the proof of Theorem 9.1 that is completed in Section 13. Thus $L_n \in \mathcal{M}_C(L_A, L_n^B)$. The key observation, for Kerr and for us, is that σ^n and $\breve{\sigma}^n$ are (unital) completely positive maps, so that if $\varphi \in UCP_q(A)$ then $\varphi \circ \sigma^n$ is in $UCP_q(B^n)$, and similarly for $\breve{\sigma}^n$. Given

$\varphi \in UCP_q(A)$, set $\psi = \varphi \circ \sigma^n$. Then exactly as in the proof of Proposition 7.1 we see that if $L_n(f, T) \leq 1$ then

$$\|\varphi(f) - \psi(T)\| \leq \|f - \sigma_T\| \leq \gamma_n,$$

so that $UCP_q(A)$ is in the γ_n neighborhood of $UCP_q(B^n)$. On the other hand, for any $\psi \in UCP_q(B^n)$ set $\varphi = \psi \circ \breve{\sigma}$. Then in the somewhat more complicated way given in Section 13 we find that $UCP_q(B^n)$ is in as small a neighborhood of $UCP_q(A)$ as desired if n is sufficiently large. □

We remark that in section 5 of [**13**] Kerr considers a weak form of the Leibniz property which he calls "f-Leibniz" (for which he comments that the corresponding distance may not satisfy the triangle inequality).

In [**18**] Hanfeng Li introduced a quite flexible variant of quantum Gromov–Hausdorff distance that in a suitable way uses the Hausdorff distance between the unit L-balls of two quantum metric spaces. Li called this "order-unit quantum Gromov–Hausdorff distance". In [**14**] Kerr and Li developed a matricial version of Li's variant, which they called "operator Gromov–Hausdorff distance". They then show (theorem 3.7) that this version coincides with Kerr's matricial quantum Gromov–Hausdorff distance. It would be interesting to have a version of our complete proximity above that is defined in terms of the unit L-balls, since it might well have certain technical advantages similar to those possessed by Li's order-unit Gromov–Hausdorff distance.

For the specific case of C^*-algebras, Li introduced [**19**] yet another variant of quantum Gromov–Hausdorff distance that explicitly uses the algebra multiplication. He calls this "C^*-algebraic quantum Gromov–Hausdorff distance". It would be interesting to know how this version relates to Leibniz seminorms and proximity. We should mention that in several places the later papers of Kerr and of Li discussed in this section again consider the f-Leibniz property that Kerr introduced in [**13**].

Hanfeng Li has pointed out to me that much the same arguments as given in the last part of Section 6, showing that prox is dominated by his dist_{nu}, also show that our "complete proximity" prox_s is dominated by dist_{nu}; and since, as mentioned in Section 6, the examples that have been studied so far for convergence for quantum Gromov-Hausdorff distance all involve nuclear C^*-algebras, and convergence for them holds for dist_{nu}, this gives for them a proof of convergence for prox_s.

The papers discussed above all begin just with a Lip-norm. In a different direction Wei Wu has defined and studied matricial Lipschitz seminorms [**41, 42, 43**]. Again, we will not repeat here his general definitions and results; rather we will only indicate somewhat sketchily how they can be adapted to the context of the present paper, I thank Wei Wu for answering several questions that I had about his papers.

Let G be a compact group equipped with a length function ℓ, and let α be an action of G on a unital C^*-algebra A. Then G has an evident entry-wise action on $M_q(A)$ for each $q \in \mathbb{Z}_{>0}$, and we can then use ℓ to define a seminorm, L^q, on each $M_q(A)$ as in Example 2.5. This family of seminorms satisfies Ruan-type axioms [**10**], in particular, $L(T_{ij}) \leq L^q(T)$ for $T = \{T_{ij}\} \in M_q(A)$. Wu presents this family as one example of what he calls a "matrix Lipschitz seminorm" on A. It is a very natural example, and it indicates how natural it is to consider matrix Lipschitz seminorms quite generally. However Wu does not make use of the fact that each of the seminorms L^q above is Leibniz (in fact, strongly Leibniz), and he uses the bridge from [**30**], which is not Leibniz.

For $A = C(G/H)$ and B^n as earlier we denote the seminorms by L_A^q and $L_B^{n,q}$. As Wu notes, the Berezin symbol map σ^n gives, by entry-wise application, a completely positive map from $M_q(B^n)$ to $M_q(A)$ for each $q \in \mathbb{Z}_{>0}$. We denote these maps still by σ^n. Much as in Section 7 we can then define a seminorm on $M_q(A \oplus B^n)$ by

$$\|M_f \circ \sigma^n - \sigma^n \circ \Lambda_T\|.$$

But the analogue of the alternative description in terms of seminorms N_x given in Proposition 7.2 is now more complicated, and so I have found it best just to work directly with the analogs of the N_x's. Specifically, we write $\mathrm{diag}(\alpha_x(P^n))$ for the matrix in $M_q(B^n)$ each of whose diagonal entries is $\alpha_x(P^n)$, with all other entries being 0. For each $x \in G$ (or G/H) we set

$$N_x^{n,q}(f,T) = \|\mathrm{diag}(\alpha_x(P))(f(x) \otimes I_n - T)\|$$

for any $(f,T) \in M_q(A \oplus B^n)$. It is easily seen that $N_x^{n,q}$ is strongly Leibniz. We then set

$$N_\sigma^{n,q}(f,T) = \sup\{N_x^{n,q}(f,T) : x \in G\}.$$

Then we set

$$N_{n,q}(f,T) = \gamma^{-1} N_\sigma^{n,q}(f,T),$$

where γ remains to be chosen for each n. Finally we set

$$L_{n,q}(f,T) = L_A^q(f) \vee L_B^{n,q}(T) \vee N_{n,q}(f,T) \vee N_{n,q}^*(f,T).$$

It is easily verified that the family $\{L_{n,q}\}$ is a "matrix Lipschitz seminorm" as defined in definition 3.1 of [**43**]. We would like to choose γ in such a way that the quotients of $L_{n,q}$ on $M_q(A)$ and $M_q(B^n)$ are L_A^q and $L_B^{n,q}$.

We consider the quotient on $M_q(A)$ first. We note, as does Wu, that $\breve{\sigma}^n$ gives, by entry-wise application, a unital completely positive map from $M_q(A)$ to $M_q(B^n)$. Given $f \in M_q(A)$, we set $T = \breve{\sigma}_f^n$. Then, much as in Section 8,

$$N_x^{n,q}(f,T) = \left\| \left\{ \alpha_x(P^n)(f_{ij}(x)I_n - d\int f_{ij}(y)\alpha_y(P^n)dy \right\} \right\|,$$

where $\{\cdot\}$ denotes a matrix. As in Section 8, the translation-invariance of L_A^q and the arbitrariness of f permit us to consider just the case in which $x = e$. Then, with manipulations as in Section 8, we see that

$$\begin{aligned} N_e^{n,q}(f,T) &\leq d\int \|\{f_{ij}(e) - f_{ij}(y)\}\| \|\mathrm{diag}(P^n\alpha_y(P^n))\| dy \\ &\leq L_A^q(f)\int \rho(e,y) d\|P^n\alpha_y(P^n)\| dy \\ &= L_A^q(f)\gamma_n^A, \end{aligned}$$

where γ_n^A is defined at the beginning of Section 9. Thus if $\gamma \geq \gamma_n^A$ then the quotient of $L_{n,q}$ on $M_q(A)$ will be L_A^q, which is exactly the same condition as for the case of $q = 1$ treated in Section 8.

We now consider the quotient on $M_q(B^n)$. Given $T \in M_q(B^n)$, we set $f = \sigma_T^n$. Then

$$N_x^{n,q}(f,T) = \|\{\alpha_x(P)(\mathrm{tr}(\alpha_x(P)T_{ij})I_n - T_{ij})\}\|.$$

I don't see a good way to estimate this except by the entry-wise estimate

$$\begin{aligned}&\le q \sup_{i,j} \|\alpha_x(P)(\mathrm{tr}(\alpha_x(P)T_{ij})I_n - T_{ij})\| \\ &\le q\gamma_n^B \sup_{i,j} L_n^B(T_{ij}) \le q\gamma_n^B L_B^{n,q}(T),\end{aligned}$$

where γ_n^B is defined at the beginning of Section 12, and where we have used the α-invariance of L_n^B, and the fact that for any $R \in M_q(B^n)$ we have $\|R\| \le q\sup_{i,j}\{\|R_{ij}\|\}$. (To see this latter, express R as the sum of the q matrices whose only non-zero entries are the entries R_{ij} of R for which $i-j$ is constant modulo q.) Thus if $\gamma \ge q\gamma_n^B$ then the quotient of $L_{n,q}$ on $M_q(B^n)$ will be $L_B^{n,q}$. The factor of q in this estimate has the quite undesirable effect that we seem not to be able to say that for a sufficiently large γ it is true that for all q simultaneously the quotient of $L_{n,q}$ on $M_q(B^n)$ is $L_B^{n,q}$. Thus the family $\{L_{n,q}\}$ cannot be used to estimate the "quantized Gromov–Hausdorff distance" defined by Wu in definition 4.5 of [**43**]. But for fixed q we will still have that $q\gamma_n^B$ converges to 0 as $n \to \infty$, and this may still be useful, for instance in dealing with vector bundles along the lines discussed in [**32**].

According to Wu's definition of "quantized Gromov–Hausdorff distance" we must now show that $UCP_q(A)$ and $UCP_q(B_n)$ are within suitable neighborhoods of each other in $UCP_q(A\oplus B)$ (once we have chosen $\gamma \ge \gamma_n^A \vee q\gamma_n^B$). Given $f \in M_q(A)$ and $\varphi \in UCP_q(A)$ (which Wu denotes by $CS_q(A)$), let $\langle\langle\varphi, f\rangle\rangle$ denote the element of $M_{q^2}(\mathbb{C})$ whose entries are the $\varphi_{ij}(f_{kl})$'s. (See 1.1.27 of [**10**].) Equivalently, view f as in $M_q \otimes A$, and let $\tilde{\varphi} = I_q \otimes \varphi$ so that $\tilde{\varphi} : M_q \otimes A \to M_q \otimes M_q$. Then $\langle\langle\varphi, f\rangle\rangle = \tilde{\varphi}(f)$. We can thus use L_A^q to define a metric, $D_{L_A^q}$, on $UCP_q(A)$, defined by

$$D_{L_A^q}(\varphi_1, \varphi_2) = \sup\{\|\langle\langle\varphi_1, f\rangle\rangle - \langle\langle\varphi_2, f\rangle\rangle\| : f \in M_q(A), L_A^q(f) \le 1\}.$$

(See proposition 3.1 of [**42**].) Wu shows that the topology on $UCP_q(A)$ from the metric $D_{L_A^q}$ coincides with the point-norm topology. In the same way $L_B^{n,q}$ defines a metric on $UCP_q(B^n)$, and $L_{n,q}$ defines a metric on $UCP_q(A\oplus B^n)$. Furthermore, when we view $UCP_q(A)$ and $UCP_q(B^n)$ as subsets of $UCP_q(A\oplus B^n)$, the restriction of $D_{L_{n,q}}$ to them will agree with $D_{L_A^q}$ and $D_{L_B^{n,q}}$ if the quotients of $L_{n,q}$ on $M_q(A)$ and $M_q(B^n)$ agree with L_A^q and $L_B^{n,q}$. (See proposition 3.6 of [**43**].)

We now show that $UCP_q(A)$ is in a suitably small neighborhood of $UCP_q(B^n)$ for $D_{L_n^q}$.

Lemma 14.2. *For any $(f,T) \in M_q(A \oplus B^n)$ we have*

$$\|f - \sigma_T^n\| \le qN_\sigma^{n,q}(f,T).$$

Proof.

$$\begin{aligned}\|f - \sigma_T^n\| &= \sup_x \|\{f_{ij}(x) - \mathrm{tr}(\alpha_x(P)T_{ij})\}\| \\ &\le q \sup_{x,i,j} |\,\mathrm{tr}(\alpha_x(P)(f_{ij}(x)I_n - T_{ij}))| \\ &\le q \sup_{x,i,j} \|\alpha_x(P)(f_{ij}(x)I_n - T_{ij})\| \\ &\le q \sup_x \|\{\alpha_x(P)(f_{ij}(x)I_n - T_{ij})\}\| = qN_\sigma^{n,q}(f,T).\end{aligned}$$

□

We can now proceed much as in the first half of Wu's proof of theorem 8.6 of [**43**]. Let q be fixed, and now set $\gamma_n = \gamma_n^A \vee q\gamma_n^B$ in the definition of $L_{n,q}$, so that $L_{n,q}$ has the right quotients. Let $\varphi \in UCP_q(A)$ be given. Set $\psi = \varphi \circ \sigma^n$, so that $\psi \in UCP_q(B^n)$. Suppose that $(f, T) \in M_q(A \oplus B^n)$ and that $L_n^q(f, T) \leq 1$, so that $N_\sigma^{n,q}(f, T) \leq \gamma_n$. Then by Lemma 14.2,

$$\begin{aligned}\|\langle\langle\varphi, f\rangle\rangle - \langle\langle\psi, T\rangle\rangle\| &= \|\langle\langle\varphi, f - \sigma_T^n\rangle\rangle\| \\ &\leq \|f - \sigma_T^n\| \leq qN_\sigma^{n,q}(f, T) \leq q\gamma_n.\end{aligned}$$

Thus $UCP_q(A)$ is in the $q\gamma_n$-neighborhood of $UCP_q(B_n)$. Since $\gamma_n^A \vee q\gamma_n^B$ converges to 0 as $n \to \infty$ we can make $q\gamma_n$ as small as desired by choosing n large enough.

We now show that $UCP_q(B^n)$ is in a suitably small neighborhood of $UCP_q(A)$. We can proceed as in the second half of Wu's proof of his theorem 8.6 of [**43**]. Let $\psi \in UCP_q(B^n)$ be given. Set $\varphi = \psi \circ \breve{\sigma}^n$, so that $\varphi \in UCP_q(A)$. For $L(f, T) \leq 1$ as above we have, much as in the proof of Proposition 8.5,

$$\begin{aligned}\|\langle\langle\varphi, f\rangle\rangle - \langle\langle\psi, T\rangle\rangle\| &= \|\langle\langle\psi, \breve{\sigma}_f^n - T\rangle\rangle\| \\ &\leq \|\breve{\sigma}_f^n - T\| \leq \|\breve{\sigma}_f^n - \breve{\sigma}^n(\sigma_T^n)\| + \|\breve{\sigma}^n(\sigma_T^n) - T\| \\ &\leq \|f - \sigma_T^n\| + \|\breve{\sigma}^n(\sigma_T^n) - T\| \\ &\leq q\gamma_n + \|\breve{\sigma}^n(\sigma_T^n) - T\|.\end{aligned}$$

We can deal with the second of these terms much as we do in Section 13, just as Wu does. One then sees that for a given $\varepsilon > 0$ one can (for a fixed q) choose N large enough that $UCP_q(A)$ and $UCP_q(B^n)$ are in each other's ε-neighborhood for $n \geq N$.

It will be interesting to see whether the von Neumann algebra approach to quantum metrics developed in [**15, 40**] eventually leads to a useful notion of quantum Gromov-Hausdorff distance, perhaps even in matricial form.

References

[1] Cristina Antonescu and Erik Christensen, *Metrics on group C^*-algebras and a noncommutative Arzelà-Ascoli theorem*, J. Funct. Anal. **214** (2004), no. 2, 247–259, arXiv:math.OA/0211312. MR 2083302 (2005d:46150)

[2] Scott Armstrong, Ken Dykema, Ruy Exel, and Hanfeng Li, *On embeddings of full amalgamated free product C^*-algebras*, Proc. Amer. Math. Soc. **132** (2004), no. 7, 2019–2030, arXiv:math.OA/0210448. MR 2053974 (2005b:46132)

[3] Bruce Blackadar, *Weak expectations and nuclear C^*-algebras*, Indiana U. Math. J. **27** (1978), no. 6, 1021–1026. MR 0467329 (57 #7188)

[4] Bruce Blackadar and Joachim Cuntz, *Differential Banach algebra norms and smooth subalgebras of C^*-algebras*, J. Operator Theory **26** (1991), no. 2, 255–282. MR 1225517 (94f:46094)

[5] Alain Connes, *Compact metric spaces, Fredholm modules, and hyperfiniteness*, Ergodic Theory Dynamical Systems **9** (1989), no. 2, 207–220. MR 90i:46124

[6] ______, *Noncommutative Geometry*, Academic Press Inc., San Diego, CA, 1994. MR 95j:46063

[7] Alain Connes and Giovanni Landi, *Noncommutative manifolds, the instanton algebra and isospectral deformations*, Comm. Math. Phys. **221** (2001), no. 1, 141–159, arXiv:math.QA/0011194. MR 1846904 (2003i:58014)

[8] Alain Connes and Henri Moscovici, *Type III and spectral triples*, Traces in number theory, geometry and quantum fields, Aspects Math., E38, 57–71, Friedr. Vieweg, Wiesbaden, 2008, arXiv:math.OA/0609703. MR 2427588

[9] Joachim Cuntz, *Bivariante K-Theorie für localkonvexe Algebren und der Chern-Connes-Charakter*, Doc. Math. **2** (1997), 139–182. MR 98h:19006

[10] E. G. Effros and Z.-J. Ruan, *Operator spaces*, The Clarendon Press Oxford University Press, New York, 2000. MR 1 793 753

[11] J. M. Gracia-Bondia, J. C. Várilly, and H. Figueroa, *Elements of noncommutative geometry*, Birkhäuser Boston Inc., Boston, MA, 2001. MR 1 789 831

[12] R. V. Kadison and J. R. Ringrose, *Fundamentals of the theory of operator algebras. Vol. I*, American Mathematical Society, Providence, RI, 1997, Elementary theory, Reprint of the 1983 original. MR 98f:46001a

[13] David Kerr, *Matricial quantum Gromov-Hausdorff distance*, J. Funct. Anal. **205** (2003), no. 1, 132–167. MR 2020211 (2004m:46153)

[14] David Kerr and Hanfeng Li, *On Gromov-Hausdorff convergence for operator metric spaces*, J. Operator Theory **62** (2009), no. 1, 83–109, arXiv:math.OA/0411157.

[15] Greg Kuperberg and Nik Weaver, *A von Neuman algebra approach to quantum metrics*, arXiv:1005.0353.

[16] Frédéric Latrémolière *Bounded-Lipschitz distances on the state space of a C^*-algebra*, Taiwanese J. Math. **11** (2007), no. 2, 447–469, arXiv:math/0510340. MR 2333358 (2008g:46121)

[17] Hanfeng Li, *θ-deformations as compact quantum metric spaces*, Comm. Math. Phys. **256** (2005), no. 1, 213–238, arXiv:math.OA/0311500. MR 2134342 (2006e:46079)

[18] ———, *Order-unit quantum Gromov-Hausdorff distance*, J. Funct. Anal. **231** (2006), no. 2, 312–360, arXiv:math.OA/0312001. MR 2195335 (2006k:46119)

[19] ———, *C*-algebraic quantum Gromov-Hausdorff distance*, arXiv:math.OA/0312003.

[20] ———, *Metric aspects of noncommutative homogeneous spaces*, J. Funct. Anal. 257 (2009), no. 7, 2325–2350, arXiv:0810.4694.

[21] Terry A. Loring, *Lifting solutions to perturbing problems in C^*-algebras*, Fields Institute Monographs **8**, American Mathematical Society, Providence, RI, 1997. MR 1420863 (98a:46090)

[22] Narutaka Ozawa and Marc A. Rieffel, *Hyperbolic group C^*-algebras and free-product C^*-algebras as compact quantum metric spaces*, Canad. J. Math. **57** (2005), no. 5, 1056–1079, arXiv:math.OA/0302310. MR 2164594 (2006e:46080)

[23] Vern I. Paulsen and Mark Tomforde, *Vector spaces with an order unit*, Indiana Univ. Math. J., **58** (2009), 1319-1359, arXiv:0712.2613. MR 2542089 (2010f:46008)

[24] Marc A. Rieffel, *Deformation quantization for actions of $\mathbb{R}^d$*, Mem. Amer. Math. Soc. **106** (1993), no. 506, x+93. MR 94d:46072

[25] ———, *Metrics on states from actions of compact groups*, Doc. Math. **3** (1998), 215–229, arXiv:math.OA/9807084. MR 99k:46126

[26] ———, *Metrics on state spaces*, Doc. Math. **4** (1999), 559–600, arXiv:math.OA/9906151. MR 1 727 499

[27] ———, *Group C^*-algebras as compact quantum metric spaces*, Doc. Math. **7** (2002), 605–651, arXiv:math.OA/0205195. MR 2015055 (2004k:22009)

[28] ———, *Compact quantum metric spaces*, Operator algebras, quantization, and noncommutative geometry, Contemp. Math., vol. 365, Amer. Math. Soc., Providence, RI, 2004, arXiv:math.OA/0308207, pp. 315–330. MR 2106826 (2005h:46099)

[29] ———, *Gromov-Hausdorff distance for quantum metric spaces*, Mem. Amer. Math. Soc. **168** (2004), no. 796, 1–65, arXiv:math.OA/0011063. MR 2055927

[30] ———, *Matrix algebras converge to the sphere for quantum Gromov-Hausdorff distance*, Mem. Amer. Math. Soc. **168** (2004), no. 796, 67–91, arXiv:math.OA/0108005. MR 2055928

[31] ———, *Lipschitz extension constants equal projection constants*, Operator theory, operator algebras, and applications, Contemp. Math., vol. 414, Amer. Math. Soc., Providence, RI, 2006, arXiv:math.FA/0508097, pp. 147–162. MR 2277209

[32] ———, *Vector bundles and Gromov-Hausdorff distance*, J. K-Theory **5** (2010), 39-103. arXiv:math.MG/0608266. MR 2600284

[33] ———, *Distances between matrix algebras that converge to coadjoint orbits*, arXiv:0910.1968.

[34] Walter Rudin, *Functional analysis*, second ed., International Series in Pure and Applied Mathematics, McGraw-Hill Inc., New York, 1991. MR 1157815 (92k:46001)

[35] Jeremy Sain, *Berezin Quantization from Ergodic Actions of Compact Quantum Groups, and Quantum Gromov-Hausdorff Distance*, Doctoral dissertation. University of California, Berkeley 2009. arXiv:0906.1829.

[36] Larry B. Schweitzer, *Spectral invariance of dense subalgebras of operator algebras*, International J. Math. **4** (1993), no. 2, 289–317. MR 1217384 (94i:46097)

[37] Walter Thirring, *Quantum mechanics of atoms and molecules*, A Course in Mathematical Physics. Vol. 3, Springer-Verlag, New York, 1981. MR 625662 (84m:81006)

[38] V. S. Varadarajan, *Lie groups, Lie algebras, and their representations*, Prentice-Hall Inc., Englewood Cliffs, N.J., 1974. MR 0376938 (51 #13113)

[39] Nik Weaver, *Lipschitz Algebras*, World Scientific, Singapore, 1999. MR 1832645 (2002g:46002)

[40] ———, *Quantum relations*, arXiv:1005.0354.

[41] Wei Wu, *Non-commutative metrics on matrix state spaces*, J. Ramanujan Math. Soc. **20** (2005), no. 3, 215–254, arXiv:math.OA/0411475. MR 2181130

[42] ———, *Non-commutative metric topology on matrix state space*, Proc. Amer. Math. Soc. **134** (2006), no. 2, 443–453, arXiv:math.OA/0410587. MR 2176013 (2006f:46072)

[43] ———, *Quantized Gromov-Hausdorff distance*, J. Funct. Anal. **238** (2006), no. 1, 58–98, arXiv:math.OA/0503344. MR 2234123

Department of Mathematics, University of California, Berkeley, CA 94720-3840
E-mail address: `rieffel@math.berkeley.edu`

Clay Mathematics Proceedings
Volume 11, 2010

Structured Vector Bundles Define Differential K-Theory

James Simons and Dennis Sullivan

To Alain who perceived and was inspired by the deeper meaning of noncommutativity

ABSTRACT. An equivalence relation, preserving the Chern-Weil form, is defined between connections on a complex vector bundle. Bundles equipped with such an equivalence class are called Structured Bundles, and their isomorphism classes form an abelian semiring. By applying the Grothedieck construction one obtains the ring $\hat{K}$, elements of which, modulo a complex torus of dimension the sum of the odd Betti numbers of the base, are uniquely determined by the corresponding element of ordinary K and the Chern-Weil form. This construction provides a simple model of differential K-theory, cf. [5], as well as a useful codification of vector bundles with connection.

Introduction

This paper grew out of the effort to construct a simple geometric model for differential K-theory, roughly speaking, the fibre product of usual K-theory with closed differential forms, [4],[5],[6]. More specifically, by differential K-theory we will mean any functor $\hat{K}$ on the smooth category which satisfies the diagram displayed below. The model which finally emerged also fulfilled our long-standing wish for a simple and straightforward codification of complex vector bundles with connection.

The set of pairs of connections whose Chern-Simons difference form is exact defines an equivalence relation in the space of all connections on a given bundle. We call a pair, $\mathcal{V} = (V, \{\nabla\})$, consisting of a vector bundle together with a particular such equivalence class, a **structured bundle**. As is true for vector bundles, structured bundles have additive inverses up to trivial structured bundles: given $\mathcal{V}$ there is a $\mathcal{W}$ such that their direct sum is equivalent to a bundle with trivial holonomy (Theorem 1.8).

By defining Struct to be the commutative semiring of isomorphism classes of structured bundles, and using the standard Grothedieck device to turn Struct into a commutative ring, we obtain $\hat{K}$, a functor from smooth compact manifolds with corners into commutative rings. As in ordinary K, every element of $\hat{K}$ may be

2010 *Mathematics Subject Classification.* Primary 19L50.

written as $\mathcal{V} - [n]$, where $\mathcal{V}$ is a structured bundle and $[n]$ is the trivial structured bundle of dim n. $\hat{K}$ achieves the above desired codification of connections and serves as the sought-after geometric model of differential K-theory.

Defining four natural transformations into and out of $\hat{K}$ we develop in the first four sections the diagram with exact diagonals and boundaries,

$$\begin{tikzcd}
0 \arrow[dr] & & & & 0 \\
& K^{odd}(\mathbb{C}/\mathbb{Z}) \arrow[rr, "\text{Bockstein}"] \arrow[dr, "j"] & & K^{even}(\mathbb{Z}) \arrow[ur] \arrow[dr, "c"] & \\
H^{odd}(\mathbb{C}) \arrow[ur, "\text{reduction mod } Z"] \arrow[dr, "\text{deR}"] & & \hat{K} \arrow[ur, "\delta"] \arrow[dr, "ch"] & & H^{even}(\mathbb{C}) \\
& \Lambda^{odd}/\Lambda_{GL} \arrow[ur, "i"] \arrow[rr, "d"] & & \Lambda_{BGL} \arrow[ur, "\text{deR}"] \arrow[dr] & \\
0 \arrow[ur] & & & & 0
\end{tikzcd}$$

where the sequence along the upper boundary may be identified (via $ch \otimes \mathbb{C}$) with the Bockstein sequence for complex K-theory (the long exact sequence associated to the short exact sequence of coefficients $0 \to \mathbb{Z} \to \mathbb{C} \to \mathbb{C}/\mathbb{Z} \to 0$), and that along the lower boundary comes from de Rham theory. Here, Λ_{BGL} means all even closed forms cohomologous to Chern characters of complex vector bundles, and Λ_{GL} means all odd closed forms cohomologous to pull-backs by maps into $GL =$ union of the $GL(n, \mathbb{C})$ of the transgression of the Chern character form. δ is the map which simply drops the connection, and ch is the Chern-Weil map applied to the Chern character polynomial. The fibre product statement above is related to the commutative square on the right half of the diagram.

The work's main technical innovation is embodied in Proposition 2.6, where it is shown that all odd forms modulo Λ_{GL} arise as the Chern-Simons difference forms between the trivial connection and arbitrary connections on trivial bundles. A corollary, as implied by the diagram above, is that every even total closed form cohomologous to the Chern character of a vector bundle may be realized as the Chern character form of a connection on a stabilized version of that bundle. In particular, if a bundle has zero characteristic classes over $\mathbb{C}$, then there is a connection on that bundle, stabilized by adding in a trivial bundle, with vanishing Chern-Weil forms.

By considering the simultaneous kernel of ch and δ, the diagram also shows that the ambiguity in determining a structured bundle (up to stabilization) solely by its characteristic forms and underlying element of K is measured by a complex torus, the dimension of which is the sum of the odd Betti numbers of the base manifold.

In showing that the kernel of ch is $K(\mathbb{C}/\mathbb{Z})$ we use the work of Karoubi [2] and Lott [1], which gave a related description of $K(\mathbb{C}/\mathbb{Z})$ involving a bundle with connection and an extra total odd form whose d is the Chern character form.

We also point out that the existence of a differential K-theory associated to K-theory, and indeed a differential theory associated to any exotic cohomology theory, was constructed in the paper of Hopkins and Singer [5]. Following their approach, Freed, as well as Hopkins and Singer and perhaps others like ourselves, were aware that a model for differential K-theory could be constructed based on pairs (E, O), where E is a bundle with connection and O is a total odd form with an equivalence relation generalizing that in [1]. One point of the present work is that this total odd form may be taken to be zero in the equivalent description of differential K-theory presented here.

There is a word for word variant of the above concerning complex vector bundles with Hermitian connection. Now there is a functor $\hat{K}_{\mathbb{R}}$, four natural transformations and the diagram

$$\begin{tikzcd}
0 \arrow[dr] & & & & 0 \\
& K^{odd}(\mathbb{R}/\mathbb{Z}) \arrow[rr, "\text{Bockstein}"] \arrow[dr, "j"] & & K^{even}(\mathbb{Z}) \arrow[ur] \arrow[dr, "c"] & \\
H^{odd}(\mathbb{R}) \arrow[ur, "\text{reduction mod } \mathbb{Z}"] \arrow[dr, "\text{deR}"] & & \hat{K}_{\mathbb{R}} \arrow[ur, "\delta"] \arrow[dr, "ch"] & & H^{even}(\mathbb{R}) \\
& \Lambda^{odd}/\Lambda_U \arrow[ur, "i"] \arrow[rr, "d"] & & \Lambda_{BU} \arrow[ur, "\text{deR}"] \arrow[dr] & \\
0 \arrow[ur] & & & & 0
\end{tikzcd}$$

This is discussed briefly in Section 5. As a corollary, for any bundle over a closed Riemannian manifold, after stabilizing there is a unitary connection on the bundle whose Chern-Weil form is the harmonic representative of the Chern character of the bundle. Moreover, when the odd Betti numbers vanish, this structured bundle is unique up to adding factors with trivial holonomy.

Our model of $\hat{K}$ or $\hat{K}_{\mathbb{R}}$ relates to two questions:

1. Up to a natural transformation, are $\hat{K}$ or $\hat{K}_{\mathbb{R}}$ uniquely determined by the diagram, as shown in [7] for ordinary differential cohomology.
2. Can one enrich the families index theorem by passing from K to $\hat{K}$ or $\hat{K}_{\mathbb{R}}$? cf. [3], [4], [6], and the Remark below.

Both of these questions are now answered in the affirmative. There are short proofs in a manuscript we are revising. The main point is to define a differential characters version of differential K-theory. Bunke and Schick have a general result applicable to the first question, and Freed and Lott have a general result applicable to the second question. Both have appeared in arxiv.

REMARK 0.1 (I.M. Singer). *Differential K-theory arose from anomaly cancellation problems in string theory and M-theory [6(b), 5]. In [6(d)], Freed and Hopkins showed how an anomaly could be cancelled if one had a refined index theorem for $\hat{K}$.*

Edward Witten described global anomalies in terms of connections on the determinant line bundle [8]. He gave a formula for the holonomy of the connection in terms of the eta invariant for an appropriate operator. Bismut and Freed gave a rigorous proof in [3(c)], as did Cheeger in [12].See also Freed's account of determinant line bundles in [6(a)].

This model of $\hat{K}$ or $\hat{K}_{\mathbb{R}}$ might be helpful for certain quantum theories and M-theory, in which it has already been observed that actions can be written more appropriately in the language of differential K-theory than in that of differential forms [6(c)]. In this respect we note Theorem 3.5, showing that $\hat{K}$ and $\hat{K}_{\mathbb{R}}$ satisfy the Mayer-Vietoris property, which relates to locality.

1. Structured Bundles

Let $[V, \nabla]$ be a complex vector bundle with connection over a smooth compact manifold with corners, X, and let $R \in \Lambda^2(X, \mathrm{End}(V))$ denote its curvature tensor.

Using the Chern-Weil homomorphism, the Chern character of V, $ch(V)$, may be represented by the total complex-valued closed form on X, $ch(\nabla)$, defined by

$$ch(\nabla) = \sum_{j=0} \frac{1}{j!}\left(\frac{1}{2\pi i}\right)^j \mathrm{tr}(\overbrace{R \wedge \cdots \wedge R}^{j}) \quad \in \Lambda^{even}(X, \mathbb{C}). \tag{1.1}$$

For $t \in [0,1]$ and $\gamma(t) = \nabla^t$ a smooth curve of connections, $(\nabla^t)' = A^t \in \Lambda^1(X, \mathrm{End}(V))$, and we set

(1.2)

$$cs(\gamma) = \int_0^1 \quad \sum_{j=1} \frac{1}{(j-1)!}\left(\frac{1}{2\pi i}\right)^j \mathrm{tr}(A^t \wedge \overbrace{R^t \wedge \cdots \wedge R^t}^{j-1}) \quad \in \Lambda^{odd}(X, \mathbb{C}).$$

It is a standard fact that

$$dcs(\gamma) = ch(\nabla^1) - ch(\nabla^0). \tag{1.3}$$

There is a second formulation of (1.2) which will be useful in what follows.

Let $\Pi : X \times [0,1] \to X$ be the standard projection, and set $W = \Pi^*(V)$. We may construct a connection, $\bar{\nabla}$, on W by defining $\bar{\nabla}_s = \nabla^t_{\Pi_*(s)}$ when s is tangent to the slice through t, and by making $\bar{\nabla}_{\partial/\partial t}(\Pi^*(f)) = 0$ for f any cross-section of V.

Let $\bar{R}$ be the curvature tensor of $\bar{\nabla}$. Then, if r, s are tangent to the slice through t,

$$\begin{aligned} \bar{R}_{r,s} &= R^t_{\Pi_*(r),\Pi_*(s)} \\ \bar{R}_{\partial/\partial t,\, s} &= A^t_{\Pi_*(s)}. \end{aligned} \tag{1.4}$$

The first is straightforward. To show the second, let $w \in W_{(x,t)}$, and extend it to be of the form $\Pi^*(f)$, where f is a cross-section of V. Also extend s to be the lift of a vector field on X. Clearly $[s, \partial/\partial t] = 0$. Thus

$$\bar{R}_{\partial/\partial t,\, s}\, w = \bar{\nabla}_{\partial/\partial t}\bar{\nabla}_s w - \bar{\nabla}_s \bar{\nabla}_{\partial/\partial t} w = \bar{\nabla}_{\partial/\partial t}\bar{\nabla}_s w = \frac{d}{dt}\nabla^t_{\Pi_*(s)} w = A^t_{\Pi_*(s)} w.$$

Now, let $\psi_t : X \to X \times [0,1]$ be the slice map, $\psi_t(x) = (x,t)$. Then by (1.4)

$$\begin{aligned} \mathrm{tr}(A^t \wedge \overbrace{R^t \wedge \cdots \wedge R^t}^{j-1}) &= \psi_t^*(\,\mathrm{tr}(i_{\partial/\partial t}\bar{R} \wedge \overbrace{\bar{R} \wedge \cdots \wedge \bar{R}}^{j-1})\,) \\ &= \psi_t^*(\, i_{\partial/\partial t}(\,\tfrac{1}{j}\,\mathrm{tr}(\overbrace{\bar{R} \wedge \cdots \wedge \bar{R}}^{j}))\,). \end{aligned}$$

From this we conclude

$$cs(\gamma) = \int_0^1 \psi_t^*\,(i_{\partial/\partial t}\, ch(\bar{\nabla})). \tag{1.5}$$

The following proposition is almost certainly well known, but we did not find a reference.

Proposition 1.1. *If α and γ are two paths connecting ∇^0 and ∇^1, then*

$$cs(\alpha) = cs(\gamma) + \textit{exact}.$$

Proof. It is sufficient to prove that if γ is a closed path of connections, then $cs(\gamma)$ is exact.
By (1.3) $cs(\gamma)$ is obviously closed. To show it is exact we show that $cs(\gamma)$ integrates to 0 on every cycle of X.
Let Z be such a cycle. Then by (1.5)

$$\begin{aligned} \int_Z cs(\gamma) = \int_{Z\times S^1} ch(\bar{\nabla}) &= ch(W)(Z \times S^1) \\ &= \Pi^*(ch(V))(Z \times S^1) = ch(V)(\Pi_*(Z \times S^1)) = 0 \end{aligned}$$

Thus $cs(\gamma)$ is exact. □

Since ∇^0 and ∇^1 may always be joined by a smooth path, using Proposition 1.1, we may set

$$CS(\nabla^0, \nabla^1) = cs(\gamma) \mod \text{exact}. \tag{1.6}$$

From Proposition 1.1 we also see

$$CS(\nabla^0, \nabla^1) + CS(\nabla^1, \nabla^2) = CS(\nabla^0, \nabla^2). \tag{1.7}$$

Definition 1.2. *∇^0 and ∇^1 will be called **equivalent**, and written $\nabla^0 \sim \nabla^1$, if $CS(\nabla^0, \nabla^1) = 0$. Equation (1.7) shows that $\sim$ is an equivalence relation.*

Definition 1.3. *A pair $\mathcal{V} = [V, \{\nabla\}]$, where $\{\nabla\}$ is an equivalence class of connections on V, will be called a **structured bundle**.*

If ∇^W is a connection on W and $\sigma : V \to W$ is a bundle isomorphism covering the identity map of X, σ induces $\sigma^*(\nabla^W)$, a connection on V, and it is easily seen that $\{\sigma^*(\nabla^W)\} = \sigma^*(\{\nabla^W\})$. $\mathcal{V} = [V, \{\nabla^V\}]$ and $\mathcal{W} = [W, \{\nabla^W\}]$ are called **isomorphic** if $\sigma^*(\{\nabla^W\}) = \{\nabla^V\}$.

If $\psi : X \to Y$ is C^∞, and V is a bundle over Y with connections ∇^0 and ∇^1, then, in the usual manner, $\psi^*(\nabla^0)$ and $\psi^*(\nabla^1)$ are connections on $\psi^*(V)$. Clearly

$$CS(\psi^*(\nabla^0), \psi^*(\nabla^1)) = \psi^*(CS(\nabla^0, \nabla^1)).$$

Thus, if $\mathcal{V} = [V, \{\nabla\}]$ is a structured bundle over Y then $\psi^*(\mathcal{V}) = [\psi^*(V), \{\psi^*(\nabla)\}]$ is well defined as a structured bundle over X.

Suppose $\psi_t : X \to Y$ is a smooth 1-parameter family of maps. If $\mathcal{V} = [V, \{\nabla\}]$ is a structured bundle over Y, then $\mathcal{V}^t = [\psi_t^*(V), \psi_t^*(\{\nabla\})]$ is a 1-parameter family of structured bundles over X. Assume $t \in [0,1]$ and let $\gamma_x : [0,1] \to Y$ be the curve $\gamma_x(t) = \psi_t(x)$. Let $\sigma_t : \psi_0^*(V) \to \psi_t^*(V)$ be parallel transport along the curves γ_t. Then, letting $W = \psi_0^*(V)$ and $\nabla^t = \sigma_t^*(\psi_t^*(\nabla))$, $\mathcal{W}^t = [W, \{\nabla^t\}]$ is a 1-parameter family of structured bundles over X, isomorphic to the family $\mathcal{V}^t$, having the same underlying vector bundle.

Letting $\gamma_x'(t)$ denote the tangent vector to γ_x at t, and using (1.5), we conclude

$$CS(\nabla^0, \nabla^1) = \int_0^1 \psi_t^*(i_{\gamma_x'(t)}\, ch(\nabla))\, dt. \tag{1.8}$$

If ∇^V and ∇^W are connections on V and W they determine connections on $V \oplus W$ and $V \otimes W$, denoted by $\nabla^V \oplus \nabla^W$ and $\nabla^W \otimes \nabla^W$. For f, g cross-sections in V and W and r a tangent vector to X,

$$\begin{aligned} (\nabla^V \oplus \nabla^W)_r(f,g) &= (\nabla_r^V f, \nabla_r^W g) \\ (\nabla^V \otimes \nabla^W)_r(f \otimes g) &= \nabla_r^V(f) \otimes g + f \otimes \nabla_r^W(g). \end{aligned}$$

It is well known that

$$ch(\nabla^V \oplus \nabla^W) = ch(\nabla^V) + ch(\nabla^W) \tag{1.9}$$

$$ch(\nabla^V \otimes \nabla^W) = ch(\nabla^V) \wedge ch(\nabla^W). \tag{1.10}$$

Lemma 1.4. *Let $\nabla^V, \bar\nabla^V, \nabla^W, \bar\nabla^W$ be connections on the indicated bundles. Then*

$$\begin{aligned} &a) \quad CS(\nabla^V \oplus \nabla^W, \bar\nabla^V \oplus \bar\nabla^W) &&= CS(\nabla^V, \bar\nabla^V) + CS(\nabla^W, \bar\nabla^W) \\ &b) \quad CS(\nabla^V \otimes \nabla^W, \bar\nabla^V \otimes \bar\nabla^W) &&= ch(\nabla^V) \wedge CS(\nabla^W, \bar\nabla^W) \\ &\qquad + ch(\bar\nabla^W) \wedge CS(\nabla^V, \bar\nabla^V) \end{aligned}$$

Proof. Using (1.7)

$$CS(\nabla^V \oplus \nabla^W, \bar\nabla^V \oplus \bar\nabla^W) = CS(\nabla^V \oplus \nabla^W, \nabla^V \oplus \bar\nabla^W) + CS(\nabla^V \oplus \bar\nabla^W, \bar\nabla^V \oplus \bar\nabla^W).$$

Direct calculation of each term using (1.2) shows a).

Again using (1.7)

$$CS(\nabla^V \otimes \nabla^W, \bar\nabla^V \otimes \bar\nabla^W) = CS(\nabla^V \otimes \nabla^W, \nabla^V \otimes \bar\nabla^W) + CS(\nabla^V \otimes \bar\nabla^W, \bar\nabla^V \otimes \bar\nabla^W)$$

and again from (1.2), direct calculation shows b). □

From Lemma 1.4 one immediately sees

PROPOSITION 1.5. *If $\mathcal{V} = [V, \{\nabla^V\}]$ and $\mathcal{W} = [W, \{\nabla^W\}]$ are structured bundles, then the equivalence classes $\{\nabla^V \oplus \nabla^W\}$ and $\{\nabla^V \otimes \nabla^W\}$ are independent of the choices of $\nabla^V \in \{\nabla^V\}$ and $\nabla^W \in \{\nabla^W\}$, and so*

$$\begin{aligned} \mathcal{V} \oplus \mathcal{W} &= [V \oplus W, \{\nabla^V \oplus \nabla^W\}] \qquad \text{and} \\ \mathcal{V} \otimes \mathcal{W} &= [V \otimes W, \{\nabla^V \otimes \nabla^W\}] \end{aligned}$$

are well defined structured bundles.

DEFINITION 1.6. *We define* Struct(X) *to be the set of isomorphism classes of structured bundles over X. By Proposition 1.5, the operations $\oplus$ and $\otimes$ make* Struct(X) *an abelian semigroup with commutative, distributive multiplication. A smooth map ψ from X to Y induces $\psi^* : \mathrm{Struct}(Y) \to \mathrm{Struct}(X)$ preserving these operations. Thus,* Struct *is a **functor** on the category of smooth compact manifolds with corners into that of commutative semirings.*

We conclude from (1.3) that $ch : \mathrm{Struct}(X) \to \Lambda^{even}(X)$ is a well defined natural transformation, and from (1.9) and (1.10)

$$\begin{aligned} ch(\mathcal{V} \oplus \mathcal{W}) &= ch(\mathcal{V}) + ch(\mathcal{W}) \\ ch(\mathcal{V} \otimes \mathcal{W}) &= ch(\mathcal{V}) \wedge ch(\mathcal{W}). \end{aligned} \tag{1.11}$$

DEFINITION 1.7. *A connection ∇ on V will be called **Flat** if its holonomy around every closed path is the identity. This implies the curvature $R \equiv 0$ and that V is naturally isomorphic to the product bundle with the trivial product connection. $\mathcal{V} = [V, \{\nabla\}]$ will be called Flat if some $\nabla \in \{\nabla\}$ is Flat. Since any two such of dim n are isomorphic, we shall denote this isomorphism class by $[n] \in \mathrm{Struct}(X)$.*

The following theorem is based on a related result in [11], stated without giving the proof. We employ that proof here in Lemma 1.9 below.

THEOREM 1.8. *Given any $\mathcal{V} \in \mathrm{Struct}(X)$ there is a $\mathcal{W} \in \mathrm{Struct}(X)$ such that $\mathcal{V} \oplus \mathcal{W} = [n]$ for some n. Any such $\mathcal{W}$ will be called an **inverse** of $\mathcal{V}$.*

To prove the Theorem we need

LEMMA 1.9. *Let ∇ be a connection on $V \oplus W$ with curvature R. Let ∇^V and ∇^W be the connections on V and W induced by ∇. E.g. if $\Pi^V : V \oplus W \to V$ is the projection, and f is a cross-section in V then $\nabla^V_r f = \Pi^V(\nabla_r f)$. Suppose $R_{r,s}(V) \subseteq V$ and $R_{r,s}(W) \subseteq W$ for all tangent vectors r, s at any point of X. Then,*

$$\nabla^V \oplus \nabla^W \sim \nabla.$$

PROOF. We may write

$$\nabla = \nabla^V \oplus \nabla^W + A$$

where $A \in \Lambda^1(X, \mathrm{End}(V \oplus W))$. For f a cross-section in V we see

$$A_r f = \nabla_r f - \Pi^V(\nabla_r f) = \Pi^W(\nabla_r f) \in W.$$

As the same holds for W, we see

$$A_r(V) \subseteq W \qquad \text{and} \qquad A_r(W) \subseteq V. \tag{1.12}$$

Setting $\bar{\nabla} = \nabla^V \oplus \nabla^W$, let $\bar{R}$ denote its curvature and $\bar{d}$ denote its exterior differentiation operator. Since $\bar{\nabla}$ preserves V and W, (1.12) implies

$$\bar{d}A_{r,s}(V) \subseteq W \qquad \textit{and} \qquad \bar{d}A_{r,s}(W) \subseteq V. \tag{1.13}$$

The usual formula for computing the curvature of one connection from that of another shows

$$R = \bar{R} + A \wedge A + \bar{d}A.$$

By hypothesis, R preserves V and W. So does $\bar{R}$, being the curvature of a direct sum connection, and so does $A \wedge A$ by (1.12). This implies that $\bar{d}A$ preserves them as well, but (1.13) shows the opposite. Thus $\bar{d}A = 0$ and

$$R = \bar{R} + A \wedge A. \tag{1.14}$$

Let $\nabla^t = \bar{\nabla} + tA$, a curve of connections joining $\nabla^V \oplus \nabla^W$ to ∇. Letting R^t denote the associated curvature, we see from (1.14)

$$R^t = \bar{R} + t^2 A \wedge A. \tag{1.15}$$

In the notation of (1.2), $A^t = (\nabla^t)' = A$, and so the CS integrand consists of terms of the form

$$\operatorname{tr}(A \wedge \overbrace{R^t \wedge \cdots \wedge R^t}^{j-1}).$$

But, by (1.15) R^t preserves both V and W, and, since A reverses them, all such trace terms must vanish. Thus $CS(\nabla, \bar{\nabla}) = 0$. □

Proof of Theorem 1.8

Proof. The classifying spaces $B_k GL(n, \mathbb{C}) = GL(n+k, \mathbb{C})/GL(n, \mathbb{C}) \times GL(k, \mathbb{C})$ come with natural bundles, V^n and W^k, of dimension n and k, and connections ∇^n and ∇^k induced by the standard Flat connection on $V^n \oplus W^k$. Lemma (1.9) shows that $\mathcal{V}^n = [V^n, \{\nabla^n\}]$ and $\mathcal{W}^k = [W^k, \{\nabla^k\}]$ are inverses of each other.

The theorem of Narasimhan-Ramanan [9] shows that for sufficiently large k, an n-dim $\mathcal{V} \in \operatorname{Struct}(X)$ may be obtained as the pull-back of $\mathcal{V}^n$ via a C^∞ map of $X \to B_k GL(n, \mathbb{C})$. The pull-back of $\mathcal{W}^k$ will then be an inverse of $\mathcal{V}$ in the sense of Theorem 1.8. □

2. The Stably Trivial Case

Let $GL = \lim\limits_{n} GL(n, \mathbb{C})$, the stabilized complex general linear group and $\mathcal{G}$ its Lie algebra. $\mathcal{G}$ consists of complex-valued matrices, all but a finite number of whose entries are 0. Let $\theta \in \Lambda^1(GL, \mathcal{G})$ denote the canonical left invariant $\mathcal{G}$-valued form on GL. Set

$$\Theta = \sum_{j=1} b_j \operatorname{tr}(\overbrace{\theta \wedge \theta \wedge \cdots \wedge \theta}^{2j-1}) \in \wedge^{odd}(GL) \tag{2.1}$$

where

$$b_j = \frac{1}{(j-1)!} \left(\frac{1}{2\pi i} \right)^j \int_0^1 (t^2 - t)^{j-1} \, dt.$$

It is well known that Θ is a bi-invariant closed odd form, and the free abelian group generated by all distinct products of its components represent the entire complex cohomology ring of GL. We define $\Lambda_{GL} \subseteq \Lambda^{odd}$ by

$$\Lambda_{GL}(X) = \{g^*(\Theta)\} + \Lambda^{odd}_{exact} \tag{2.2}$$

where $g : X \to GL$ runs through all C^∞ maps.

Note that if g, h map X into GL, then $g \oplus h : X \to GL$ may be defined, and $(g \oplus h)^*(\Theta) = g^*(\Theta) + h^*(\Theta)$. Moreover, $(g^{-1})^*(\Theta) = -g^*(\Theta)$. Thus $\Lambda_{GL}(X)$ is an abelian group.

Lemma 2.1. *Let V be a trivial bundle with the two Flat connections ∇ and $\bar{\nabla}$. Then*

$$CS(\nabla, \bar{\nabla}) \in \Lambda_{GL}/\Lambda^{odd}_{exact}.$$

Proof. Since ∇ and $\bar{\nabla}$ each have trivial holonomy, we can find a cross-section $g \in \mathrm{Aut}(V)$ such that

$$\bar{\nabla} = g^{-1} \circ \nabla \circ g.$$

Expressing g as a matrix with respect to a ∇-parallel framing of V, we see

$$\bar{\nabla} = \nabla + g^{-1}\, dg.$$

Now, regarding $g : X \to GL$, one easily sees that $g^{-1}dg = g^*(\theta)$. Thus

$$\bar{\nabla} = \nabla + g^*(\theta).$$

Setting $\bar{\nabla}^t = \nabla + tg^*(\theta)$, we see that

$$\bar{R}^t = R + t\, dg^*(\theta) + t^2\, g^*(\theta) \wedge g^*(\theta).$$

But, either calculating on GL, or directly with $g^{-1}dg$, we see that $dg^*(\theta) = -g^*(\theta) \wedge g^*(\theta)$. Moreover, since ∇ has trivial holonomy, $R \equiv 0$. Thus

$$\bar{R}^t = (t^2 - t) g^*(\theta) \wedge g^*(\theta).$$

It then follows from (1.2) that $CS(\nabla, \bar{\nabla}) = g^*(\Theta)$. $\square$

Definition 2.2.

$$\mathrm{Struct}_{\mathrm{ST}}(X) = \{\, [V, \{\nabla\}] \in \mathrm{Struct}(X) \mid V \textit{ is stably trivial}\}.$$

For $\mathcal{V} \in \mathrm{Struct}_{\mathrm{ST}}(X)$, let F and H be trivial bundles such that $V \oplus F = H$ and let ∇^F, ∇^H be Flat connections on F and H. We define

$$\widehat{CS} : \mathrm{Struct}_{ST}(X) \to \Lambda^{odd}/\Lambda_{GL}$$

by

$$\widehat{CS}(\mathcal{V}) = CS(\nabla^H, \nabla \oplus \nabla^F) \mod \Lambda_{GL}/\Lambda^{odd}_{exact}.$$

Proposition 2.3. *$\widehat{CS}$ is a well defined homomorphism.*

Proof. Suppose $\bar{F}, \bar{H}, \nabla^{\bar{F}}, \nabla^{\bar{H}}$ are another pair of trivial bundles with Flat connections with $V \oplus \bar{F} = \bar{H}$. Using (1.6), Lemma 1.4 and Lemma 2.1, and working mod Λ_{GL}, we see

$$\begin{aligned} CS(\nabla^{\bar{H}}, \nabla \oplus \nabla^{\bar{F}}) &= CS(\nabla^{\bar{H}} \oplus \nabla^F, \nabla \oplus \nabla^{\bar{F}} \oplus \nabla^F) \\ &= CS(\nabla^H \oplus \nabla^{\bar{F}}, \nabla \oplus \nabla^{\bar{F}} \oplus \nabla^F) \\ &= CS(\nabla^H, \nabla \oplus \nabla^F). \end{aligned}$$

Thus $\widehat{CS}$ is well defined. That $\widehat{CS}$ is a homomorphism follows immediately from Lemma 1.4. $\square$

DEFINITION 2.4. *$\mathcal{V} \in \mathrm{Struct}(X)$ is called **stably Flat** if there exists Flat $\mathcal{F}$ and $\mathcal{H}$ such that $\mathcal{V} \oplus \mathcal{F} = \mathcal{H}$. The set of these objects will be denoted by* $\mathrm{Struct}_{\mathrm{SF}}(X)$. *Clearly* $\mathrm{Struct}_{\mathrm{SF}}(X) \subseteq \mathrm{Struct}_{\mathrm{ST}}(X)$ *and is a sub-semigroup.*

PROPOSITION 2.5. $\ker(\widehat{CS}) = \mathrm{Struct}_{\mathrm{SF}}(X)$.

PROOF. Obviously $\mathrm{Struct}_{\mathrm{SF}} \subseteq \ker(\widehat{CS})$. Now suppose $\widehat{CS}(\mathcal{V}) = 0$. Let F, H, ∇^F and ∇^H be as in the definition of $\widehat{CS}$. Now, $\widehat{CS}(\mathcal{V}) = 0$ implies

$$CS(\nabla^H, \nabla \oplus \nabla^F) = g^*(\Theta) \bmod \Lambda^{odd}_{exact}$$

for some $g : X \to GL$. Again as in the proof of Lemma 2.1, choosing a ∇^H-parallel framing of H, we may regard $g \in \mathrm{Aut}(H)$ and set

$$\bar{\nabla}^H = g^{-1}(\nabla^H \circ g).$$

As in the Lemma we see $CS(\nabla^H, \bar{\nabla}^H) = g^*(\Theta)$ and thus $CS(\bar{\nabla}^H, \nabla^H) = -g^*(\Theta)$. Therefore

$$\begin{aligned} CS(\bar{\nabla}^H, \nabla \oplus \nabla^F) &= CS(\bar{\nabla}^H, \nabla^H) + CS(\nabla^H, \nabla \oplus \nabla^F) \\ &= -g^*(\Theta) + g^*(\Theta) = 0 \ \text{mod exact}. \end{aligned}$$

Setting $\bar{\mathcal{H}} = [H, \{\bar{\nabla}^H\}]$ and $\mathcal{F} = [F, \{\nabla^F\}]$ we see $\mathcal{V} \oplus \mathcal{F} = \bar{\mathcal{H}}$ and thus $\mathcal{V} \in \mathrm{Struct}_{\mathrm{SF}}(X)$. $\square$

PROPOSITION 2.6. $\mathrm{Im}(\widehat{CS}) = \Lambda^{odd}(X)/\Lambda_{GL}(X)$.

PROOF. If L is a trivialized line bundle over X then any connection on L is simply a complex-valued 1-form, w. Since $w \wedge w = 0$, the associated curvature, R^w, is dw, and $\{w\} = \{w + df \mid f \in C^\infty(X, \mathbb{C})\}$.

Let $\mathcal{L}_w = [L, \{w\}]$. Using tw as a curve of connections joining w to the trivial connection, noting that $R^{tw} = tR^w$, and working mod Λ_{GL}, (1.2) shows

$$\widehat{CS}(\mathcal{L}_w) = \sum_{j=1} \frac{1}{j!}\left(\frac{1}{2\pi i}\right)^j w \wedge (dw)^{j-1}. \tag{2.3}$$

We first suppose $X = R^n$. If $w = f\,dx$ then $w \wedge dw = 0$ and thus $\widehat{CS}(\mathcal{L}_{f\,dx}) = f\,dx$. Moreover, since $\widehat{CS}$ is a homomorphism

$$\widehat{CS}\left(\sum_i \oplus \mathcal{L}_{f_i dx_i}\right) = \sum f_i dx_i.$$

Thus $\Lambda^1(R^n)/\Lambda_G(R^n) \subseteq \mathrm{Im}(\widehat{CS})$.

Proceeding by induction on k, suppose

$$\left(\sum_{j=1}^{k} \Lambda^{2j-1}(R^n)\right)/\Lambda_{GL}(R^n) \subseteq \mathrm{Im}(\widehat{CS}). \tag{2.4}$$

Let $w = x_1 dx_2 + x_3 dx_4 + \cdots + x_{2k-1} dx_{2k} + f dx_{2k+1}$.

Claim: $w \wedge (dw)^k = (k+1)! f dx_1 \wedge \cdots \wedge dx_{2k+1} + \text{exact}$.

To show this, let $\gamma = dx_1 \wedge dx_2 + \cdots + dx_{2k-1} \wedge dx_{2k}$, and note

$$dw = \gamma + df \wedge dx_{2k+1} \qquad \Rightarrow \qquad (dw)^k = (\gamma + df \wedge dx_{2k+1})^k.$$

Since all powers of $df \wedge dx_{2k}$ vanish,

$$(dw)^k = \gamma^k + k\gamma^{k-1} \wedge df \wedge dx_{2k+1} = k! dx_1 \wedge \cdots \wedge dx_{2k} +$$
$$k! \left[\sum_{j=1}^{k} dx_1 \wedge dx_2 \wedge \cdots \wedge dx_{\widehat{2j-1} \wedge dx_{2j}} \wedge \cdots \wedge dx_{2k-1} \wedge dx_{2k} \right] \wedge df \wedge dx_{2k+1}$$

Thus,

$$\begin{aligned} & w \wedge (dw)^k = k! f dx_1 \wedge \cdots \wedge dx_{2k+1} + \\ & k! \left[\sum_{j=1}^{k} dx_1 \wedge \cdots \wedge dx_{2j-2} \wedge x_{2j-1} \wedge dx_{2j} \wedge \cdots \wedge dx_{2k} \right] \wedge df \wedge dx_{2k+1} \\ = \quad & k!\, f dx_1 \wedge \cdots \wedge dx_{2k+1} \\ & \quad -k! \sum_{j=1}^{k} dx_1 \wedge \cdots \wedge dx_{2j-1} \wedge x_{2j-1} df \wedge dx_{2j} \wedge \cdots \wedge dx_{2k+1} \\ = \quad & (k+1)!\, f dx_1 \wedge \cdots \wedge dx_{2k+1} + \text{exact}. \end{aligned}$$

Thus, working $\bmod \Lambda_{GL}(R^n)$,

$$\widehat{CS}(\mathcal{L}_{(2\pi i)^{k+1} w}) = f dx_1 \wedge \cdots \wedge dx_{2k+1} + \theta,$$

where

$$\theta \in \sum_{j=1}^{k} \Lambda^{2j-1}(R^n).$$

By induction, $\theta = \widehat{CS}(\mathcal{V})$ for some $\mathcal{V} \in \mathrm{Struct}_{\mathrm{ST}}(R^n)$. Theorem 1.8 shows $\mathcal{V}$ has an inverse $\mathcal{V}^{-1}$. Clearly $\mathcal{V}^{-1} \in \mathrm{Struct}_{\mathrm{ST}}(R^n)$ and by Proposition 2.3, $\widehat{CS}(\mathcal{V}^{-1}) = -\theta$. Thus

$$\widehat{CS}(\mathcal{L}_{(2\pi i)^{k+1} w} \oplus \mathcal{V}^{-1}) = f dx_1 \wedge \cdots \wedge dx_{2k+1}.$$

The general element of $\Lambda^{2k+1}(R^n)$ is the sum of such terms, and thus is the image under $\widehat{CS}$ of the direct sum of the inverse images of each of these terms.

For the general case let $\psi : X \to R^n$ be an imbedding. Since $\psi^* : \Lambda^{odd}(R^n) \to \Lambda^{odd}(X)$ is onto, and $\psi^*(\Lambda_{GL}(R^n)) \subseteq \Lambda_{GL}(X)$, $\psi^* : \Lambda^{odd}(R^n)/\Lambda_{GL}(R_n) \to \Lambda^{odd}(X)/\Lambda_{GL}(X)$ is onto. Moreover, $\psi^*(\mathrm{Struct}_{\mathrm{ST}}(R^n)) \subseteq \mathrm{Struct}_{\mathrm{ST}}(X)$, and finally $\widehat{CS} \circ \psi^* = \psi^* \circ \widehat{CS}$. Thus if $\rho \in \Lambda^{odd}(X)/\Lambda_{GL}(X)$, we can find $\bar{\rho} \in \Lambda^{odd}(R_n)/\Lambda_{GL}(R^n)$ with $\psi^*(\bar{\rho}) = \rho$. By the special case, $\bar{\rho} = \widehat{CS}(\mathcal{V})$ for some $\mathcal{V} \in \mathrm{Struct}_{\mathrm{ST}}(R^n)$. Then

$$\rho = \psi^*(\widehat{CS}(\mathcal{V})) = \widehat{CS}(\psi^*(\mathcal{V})).$$

□

From Propositions 2.3, 2.5, 2.6 we see

THEOREM 2.7.

$$\widehat{CS} : \mathrm{Struct}_{\mathrm{ST}}(X)/\mathrm{Struct}_{\mathrm{SF}}(X) \quad \underset{\to}{\cong} \quad \Lambda^{odd}(X)/\Lambda_{GL}(X).$$

3. $\hat{K}(X)$

Using the standard construction of K, which transforms an abelian semi-group into a group, we define

$$\hat{K} = K(\text{Struct}(X)).$$

$\hat{K}(X)$ is the free abelian group generated by isomorphism classes of structured bundles, modulo the relation $\mathcal{V} + \mathcal{W} - (\mathcal{V} \oplus \mathcal{W})$. Equivalently defined, $\hat{K}(X)$ is the quotient of the semigroup under $\oplus$ consisting of all pairs $(\mathcal{V}, \mathcal{W})$ modulo the sub-semigroup consisting of pairs $(\mathcal{V}, \mathcal{V})$. Since $(0, \mathcal{V})$ is obviously the additive inverse of $(\mathcal{V}, 0)$, we write $(\mathcal{V}, \mathcal{W})$ as $\mathcal{V} - \mathcal{W}$.

By Theorem 1.8, using the pairs definition above it is straightforward to show

(3.1) Every element of $\hat{K}(X)$ is of the form $\mathcal{V} - [n]$.

(3.2) $\mathcal{V} - [n] = 0 \Leftrightarrow \mathcal{V}$ is stably Flat and $n = \dim(\mathcal{V})$.

Again using the pairs definition, one sees that $\otimes$ is well defined in $\hat{K}(X)$, and thus $\hat{K}(X)$ becomes a commutative ring. (Defining $(\mathcal{V}, \mathcal{W}) \otimes (\mathcal{V}', \mathcal{W}')$ to be $(\mathcal{V} \otimes \mathcal{V}' \oplus \mathcal{W} \otimes \mathcal{W}', \mathcal{W} \otimes \mathcal{V}' \oplus \mathcal{V} \otimes \mathcal{W}')$ one sees that $\{(\mathcal{W}, \mathcal{W})\}$ is an ideal.)

We define $\wedge_{BGL} \subseteq \Lambda^{even}$ by

$$\Lambda_{BGL}(X) = \{ch(\mathcal{V})\} + \Lambda^{even}_{exact}$$

where $\mathcal{V}$ ranges over all elements of Struct(X). From (1.9) and (1.10) and Theorem 1.8 we see that $\Lambda_{BGL}(X)$ is a commutative ring.

By analogy with the definition of Λ_{GL}, and using the theorem of Narasimhan-Ramanan [9], we could alternatively have defined

$$\Lambda_{BGL}(X) = \{\phi^*(\Omega)\} + \Lambda^{even}_{exact}$$

where $\phi : X \to BGL$ ranges over all C^∞ maps, and Ω is the Chern character form of the standard connection on the classifying bundle over BGL.

Clearly, ch extends to $\hat{K}(X)$, and maps it to $\Lambda_{BGL}(\mathbb{C})$. We also define

$$\delta : \hat{K}(X) \to K(X)$$

by

$$\delta(\,[V, \{\nabla\}] - [W, \{\bar{\nabla}\}]\,) = V - W.$$

Letting $c : K(X) \to H^{even}(X, \mathbb{C})$ be the natural transformation defined by the Chern character, and deR : $\Lambda_{BGL}(X) \to H^{even}(X, \mathbb{C})$ be that defined by the de Rham Theorem, we see

(3.3)

$$\begin{tikzcd} & K(X) \arrow[dr, "c"] & \\ \hat{K}(X) \arrow[ur, "\delta"] \arrow[dr, "ch"'] \arrow[rr, phantom, "\otimes"] & & H^{even}(X, \mathbb{C}) \\ & \Lambda_{BGL}(X) \arrow[ur, "\text{deR}"'] & \end{tikzcd}$$

is a commutative diagram.

PROPOSITION 3.1. $\ker(\delta) \cong \Lambda^{odd}(X)/\Lambda_{GL}(X)$.

PROOF. Define $\Gamma : \mathrm{Struct}_{ST}/\mathrm{Struct}_{SF} \to \hat{K}$ by

$$\Gamma(\{\mathcal{V}\}) = \mathcal{V} - [n],$$

where $n = \dim(\mathcal{V})$. By (3.2), Γ is well defined and is an injection. Moreover, it is clear that $\mathrm{Im}(\Gamma) = \ker(\delta)$. Thus from Theorem 2.7,

$$\Gamma \circ \widehat{CS}^{-1} : \Lambda^{odd}(X)/\Lambda_{GL}(X) \xrightarrow{\cong} \ker(\delta).$$

□

Let $i = \Gamma \circ \widehat{CS}^{-1}$. Since δ is clearly onto,

$$0 \longrightarrow \Lambda^{odd}(X)/\Lambda_{GL}(X) \xrightarrow{i} \hat{K}(X) \xrightarrow{\delta} K(X) \longrightarrow 0 \tag{3.4}$$

is an exact sequence.

PROPOSITION 3.2. *$ch \circ i = d$, and ch is onto.*

PROOF. To show the first, note that from the definition of $\widehat{CS}$,

$$d\widehat{CS}(\mathcal{V}) = ch(\mathcal{V}) - \dim(\mathcal{V}) = ch(\Gamma(\mathcal{V}))$$

for any $\mathcal{V} \in \mathrm{Struct}_{ST}(X)$. Thus, for $\theta \in \Lambda^{odd}(X)$ and $\{\theta\}$ its equivalence class mod $\Lambda_{GL}(X)$,

$$ch(i(\{\theta\})) = ch(\Gamma(\widehat{CS}^{-1}(\{\theta\}))) = d(\{\theta\}) = d\theta.$$

To show the second, let $\mu \in \Lambda_{BGL}(X)$. By definition, $\exists\, \mathcal{V} \in \mathrm{Struct}(X)$ and $\theta \in \Lambda^{odd}$ so that $\mu = ch(\mathcal{V}) + d\theta$. By the above, $\mu = ch(\mathcal{V} + i(\{\theta\}))$.

□

Let $\mathrm{deR} : H^{odd}(X, \mathbb{C}) \to \Lambda^{odd}(X)/\Lambda_{GL}(X)$ be the obvious map induced by the de Rham Theorem. Since the image of deR consists of closed forms, $d \circ \mathrm{deR} = 0$, which by Proposition 3.2, implies $ch \circ i \circ \mathrm{deR} = 0$. Thus, $i \circ \mathrm{deR}(H^{odd}(X, \mathbb{C})) \subseteq \ker(ch)$. We have now established

PROPOSITION 3.3. *The following diagram of functors and natural transformations is commutative, and its diagonals are exact.*

$$\begin{tikzcd}
0 \arrow[dr] & & & & 0 \\
& \ker(ch) \arrow[rr, "\delta\,|"] \arrow[dr, "\subseteq"] & & K \arrow[ur] \arrow[dr, "c"] & \\
H^{odd}(\mathbb{C}) \arrow[ur, "i \circ \mathrm{deR}"] \arrow[dr, "\mathrm{deR}"] & & \hat{K} \arrow[ur, "\delta"] \arrow[dr, "ch"] & & H^{even}(\mathbb{C}) \\
& \Lambda^{odd}/\Lambda_{GL} \arrow[ur, "i"] \arrow[rr, "d"] & & \Lambda_{BGL} \arrow[ur, "\mathrm{deR}"] \arrow[dr] & \\
0 \arrow[ur] & & & & 0
\end{tikzcd}$$

Corollary 3.4. *The outside sequences*

$$H^{odd}(\mathbb{C}) \xrightarrow{i\circ\mathrm{deR}} \ker(ch) \xrightarrow{\delta\,|} K \xrightarrow{c} H^{even}(\mathbb{C})$$

$$H^{odd}(\mathbb{C}) \xrightarrow{\mathrm{deR}} \Lambda^{odd}/\Lambda_{GL} \xrightarrow{d} \Lambda_{BGL} \xrightarrow{\mathrm{deR}} H^{even}(\mathbb{C})$$

are exact.

Proof. Exactness of the first follows from diagram chasing, and that of the second from the de Rham Theorem. □

In the decomposition below we assume that D is a codimension zero submanifold with collar neighborhoods in each of A and B. Thus a smooth form on D can be extended to a smooth form on either A or B.

Theorem 3.5 (Mayer-Vietoris). *Let $A, B \subseteq X$ with $A \cap B = D$ and $A \cup B = X$. If $\mu_A \in \hat{K}(A)$ and $\mu_B \in \hat{K}(B)$ with $\mu_A|D = \mu_B|D$, then there exists $\mu \in \hat{K}(X)$ with $\mu|A = \mu_A$ and $\mu|B = \mu_B$.*

Proof. Following the diagram in Proposition 3.3, since $\delta(\mu_A)\,|\,D = \delta(\mu_B)\,|\,D$, the Mayer-Vietoris property for K produces $V - [n] \in K(X)$ with $(V - [n])\,|\,A = \delta(\mu_A)$ and $(V - [n])\,|\,B = \delta(\mu_B)$. Choose $\bar{\mu} \in \hat{K}(X)$ with $\delta(\bar{\mu}) = V - [n]$.

Now, $\delta(\bar{\mu}\,|\,A) = \delta(\bar{\mu})\,|\,A = \delta(\mu_A)$, and similarly for B. Thus, by the diagram

$$\begin{aligned} \bar{\mu}\,|\,A &= \mu_A + i(\{\alpha_A\}) \\ \bar{\mu}\,|\,B &= \mu_B + i(\{\alpha_B\}) \end{aligned} \tag{3.5}$$

where α_A, $\alpha_B \in \Lambda^{odd}(A)$, $\Lambda^{odd}(B)$ and $\{\alpha_A\}$, $\{\alpha_B\}$ represent their equivalence classes mod $\Lambda_{GL}(A)$, $\Lambda_{GL}(B)$.

By the above,

$$\begin{aligned}i(\{\alpha_A \,|\, D\}) - i(\{\alpha_B \,|\, D\}) &= i(\{\alpha_A\}) \,|\, D - i(\{\alpha_B\}) \,|\, D \\ &= (\bar{\mu} \,|\, A) \,|\, D - (\bar{\mu} \,|\, B) \,|\, D - \mu_A \,|\, D + \mu_B \,|\, D.\end{aligned}$$

The first pair vanishes since each term is $\bar{\mu} \,|\, D$, and the second pair vanishes by hypothesis. Since i is an injection,

$$\alpha_A \,|\, D = \alpha_B \,|\, D + w$$

where $w \in \Lambda_{GL}(D)$.

Case I: $w = d\rho$

Extend ρ to all of A, and set $\tilde{\alpha}_A = \alpha_A + d\rho$. Thus $\{\tilde{\alpha}_A\} = \{\alpha_A\}$, and $\tilde{\alpha}_A \,|\, D = \alpha_B \,|\, D$. The latter equation implies there is a unique $\alpha \in \Lambda^{odd}(X)$ with $\alpha \,|\, A = \tilde{\alpha}_A$ and $\alpha \,|\, B = \alpha_B$. Thus by (3.5)

$$\begin{aligned}\bar{\mu} \,|\, A &= \mu_A + i(\{\alpha\}) \,|\, A \\ \bar{\mu} \,|\, B &= \mu_B + i(\{\alpha\}) \,|\, B\end{aligned}$$

which implies that $\mu = \bar{\mu} - i(\{\alpha\})$ satisfies the conditions of the theorem.

Case II: $w = g^*(\Theta) + d\rho$, where $g : D \to GL$, and $g^*(\Theta)$ is not exact.

Using the clutching construction, we may construct a vector bundle V over X with the properties that $V \,|\, A$ and $V \,|\, B$ are each trivialized by cross-sections $\{E_i^A\}$ and $\{E_i^B\}$, and

$$E_j^B \,|\, D = \sum_i g_{ij} E_i^A \,|\, D. \tag{3.6}$$

Choose a connection, ∇', on V, set $\mathcal{V} = [V, \{\nabla'\}]$ and $\mu' = \mathcal{V} - [\dim(\mathcal{V})] \in \hat{K}(X)$.

By construction, $\delta(\mu' \,|\, A) = 0 = \delta(\mu' \,|\, B)$, and thus

$$\begin{aligned}\mu' \,|\, A &= i(\{\alpha'_A\}) \\ \mu' \,|\, B &= i(\{\alpha'_B\})\end{aligned}$$

where α'_A, $\alpha'_B \in \Lambda^{odd}(A)$, $\Lambda^{odd}(B)$ and $\{\alpha'_A\}$, $\{\alpha'_B\}$ represent their equivalence classes modulo $\Lambda_{GL}(A)$, $\Lambda_{GL}(B)$.

Let ∇^{AF} and ∇^{BF} be the Flat connections on $V \,|\, A$ and $V \,|\, B$ defined by making $\{E_i^A\}$, $\{E_i^B\}$ parallel. By the definition of i, and working mod exact, we may take

$$\begin{aligned}\alpha'_A &= CS(\nabla^{AF}, \nabla' \,|\, A) \\ \alpha'_B &= CS(\nabla^{BF}, \nabla' \,|\, B).\end{aligned}$$

Now, continuing to work mod exact,

$$\begin{aligned}\alpha'_A \,|\, D - \alpha'_B \,|\, D &= CS(\nabla^{AF} \,|\, D, \nabla' \,|\, D) - CS(\nabla^{BF} \,|\, D, \nabla' \,|\, D) \\ &= CS(\nabla^{AF} \,|\, D, \nabla^{BF} \,|\, D) = g^*(\Theta)\end{aligned}$$

by (3.6) and the argument of Lemma 2.1.

Thus, by taking $\bar{\bar{\mu}} = \bar{\mu} - \mu'$ and referring to (3.5) we see

$$\begin{aligned}\bar{\bar{\mu}} \,|\, A &= \mu_A + i(\{\alpha_A - \alpha'_A\}) \\ \bar{\bar{\mu}} \,|\, B &= \mu_B + i(\{\alpha_B - \alpha'_B\})\end{aligned}$$

and

$$(\alpha_A - \alpha'_A) \,|\, D \quad = \quad (\alpha_B - \alpha'_B) \,|\, D + \text{exact}.$$

The problem is now reduced to Case I. ☐

COROLLARY 3.6. $\ker(ch)$ *also satisfies the Mayer-Vietoris property.*

PROOF. In the theorem above, if $ch(\mu_A) = 0 = ch(\mu_B)$, then $ch(\mu) \,|\, A = 0 = ch(\mu) \,|\, B$. Since $ch(\mu)$ is a differential form, this implies $ch(\mu) = 0$. ☐

COROLLARY 3.7. $\ker(ch)$ *is a homotopy functor.*

PROOF. Any element of $\ker(ch)$ is of the form $\mathcal{V} - [\dim(\mathcal{V})]$, where $ch(\mathcal{V}) = \dim(\mathcal{V})$. By (1.8) the pull-backs of $\mathcal{V}$ under two smoothly homotopic C^∞ maps would be isomorphic, and so of course would pull-backs of $[\dim(\mathcal{V})]$. ☐

REMARK 3.8. *Now we can see using [10] that* $\ker(ch)$ *is a homotopy functor represented by homotopy classes of maps into some classifying space. Rather than determining this space now, we will shift gears in the next section and use the work of Lott and Karoubi in [1] and [2].*

4. $\ker(ch)$ is equivalent to K-theory with coefficients in $\mathbb{C}/\mathbb{Z}$

The K-groups are the even part of an exotic $\mathbb{Z}$-graded cohomology theory which is 2-periodic. For every $\mathbb{Z}$-graded cohomology theory, h, there is a long exact sequence, analogous to the Bockstein sequence in ordinary cohomology theory, associated to the coefficient sequence
$0 \to \mathbb{Z} \to \mathbb{C} \to \mathbb{C}/\mathbb{Z} \to 0,$

$$\cdots \to h^{i-1}(\mathbb{C}/\mathbb{Z}) \to h^i(\mathbb{Z}) \to h^i(\mathbb{C}) \to h^i(\mathbb{C}/\mathbb{Z}) \to \cdots$$

where $h^i(\mathbb{Z}) = h^i$, $h^i(\mathbb{C}) = h^i \oplus \mathbb{C}$, and the $h^{i-1}(\mathbb{C}/\mathbb{Z})$ may be defined by homotopy classes of mappings into a classifying space, in this case the homotopy fibre of the map of classifying spaces corresponding to the map $\otimes\mathbb{C}$ in degree i. cf. [10], [13]. For the convenience of the reader, we discuss and characterize homotopy fibres in the Appendix.

In the case of K-theory, Karoubi [2] and Lott [1] studied a smooth model of $K(\mathbb{C}/\mathbb{Z})$. In this model one considers the semigroup of triples $(\mathcal{V}, \nabla, \eta)$, where $[\mathcal{V}, \nabla]$ is a complex vector bundle with connection, and $\eta \in \Lambda^{odd}/\Lambda^{exact}$ with $d\eta = ch(\nabla) - \dim \mathcal{V}$. $(\mathcal{V}, \nabla, \eta)$ and $(\mathcal{V}, \bar{\nabla}, \bar{\eta})$ are called equivalent iff

$$\eta - \bar{\eta} = CS(\bar{\nabla}, \nabla).$$

Trivial elements are triples where $[\mathcal{V}, \nabla]$ is Flat and $\eta = 0$. Stabilizing this equivalence by the addition of trivial elements, passing to isomorphism classes, and dividing out by trivial elements, yields the model of $K(\mathbb{C}/\mathbb{Z})$.

In this model, the "Bockstein" sequence contains

$$H^{odd}(\mathbb{C}) \xrightarrow{\alpha} K^{odd}(\mathbb{C}/\mathbb{Z}) \xrightarrow{\beta} K^{even}(\mathbb{Z}) \tag{4.1}$$

where α is defined by choosing any closed form η representing a given cohomology class μ and setting $\alpha(\mu) = (F, \nabla^F, -\eta)$ where $[F, \nabla^F]$ is Flat. β is defined by taking $(\mathcal{V}, \nabla, \eta)$ into $\mathcal{V} - \dim\mathcal{V} \in K(\mathbb{Z})$.

There is a natural map, Φ, sending $\ker(ch)$ into this model, in which

$$\Phi([\mathcal{V}, \{\nabla\}] - [\dim \mathcal{V}]) = (\mathcal{V}, \nabla, 0).$$

Proposition 4.1. *$\Phi : \ker(ch) \longrightarrow K^{odd}(\mathbb{C}/\mathbb{Z})$ is a bijection. Moreover, using the notation of the diagram in Prop. 3.3, the diagram below is commutative.*

$$\begin{tikzcd}
& \ker(ch) \arrow[dr, "\delta\,|"] \arrow[dd, "\Phi"] & \\
H^{odd}(\mathbb{C}) \arrow[ur, "i \circ \mathrm{deR}"] \arrow[dr, "\alpha"'] & & K^{even}(\mathbb{Z}) \\
& K^{odd}(\mathbb{C}/\mathbb{Z}) \arrow[ur, "\beta"'] &
\end{tikzcd}$$

Proof. Injectivity of Φ follows from the definition of structured bundles, and surjectivity follows from the surjectivity of $\widehat{CS}$ in Prop. 2.6. Moreover, it is immediate from (4.1) that $\beta \circ \Phi = \delta\,|$.

To see that the left side commutes, let $\mu \in H^{odd}(\mathbb{C})$ be represented by a closed form η. From the definition of i in Prop. 3.3 we see

$$i \circ \mathrm{deR}(\mu) = [F, \{\nabla\}] - [\dim F]$$

where F may be taken to be trivial and $\widehat{CS}([F, \{\nabla\}]) = \eta \bmod \Lambda_{GL}$. Using Lemma 2.1 and enlarging F if necessary, we may find a suitable Flat connection ∇^F so that $CS(\nabla^F, \nabla) = \eta$ mod exact. Then

$$\Phi \circ i \circ \mathrm{deR}(\mu) = (F, \nabla, 0) \sim (F, \nabla^F, \eta) = \alpha(\mu).$$

$\square$

Thus, via Φ, we may substitute $K^{odd}(\mathbb{C}/\mathbb{Z})$ for $\ker(ch)$ in Prop 3.3, and so prove

Theorem 4.2. *The following diagram has exact diagonals and exact upper and lower boundaries.*

$$
\begin{tikzcd}
0 \arrow[dr] & & & & 0 \\
& K^{odd}(\mathbb{C}/\mathbb{Z}) \arrow[rr, "\mathit{Bockstein}"] \arrow[dr, "j"] & & K^{even}(\mathbb{Z}) \arrow[ur] \arrow[dr, "c"] & \\
H^{odd}(\mathbb{C}) \arrow[ur, "\mathit{reduction\ mod}\ \mathbb{Z}"] \arrow[dr, "\mathrm{deR}"] & & \hat{K} \arrow[ur, "\delta"] \arrow[dr, "ch"] & & H^{even}(\mathbb{C}) \\
& \Lambda^{odd}/\Lambda_{GL} \arrow[ur, "i"] \arrow[rr, "d"] & & \Lambda_{BGL} \arrow[ur, "\mathrm{deR}"] \arrow[dr] & \\
0 \arrow[ur] & & & & 0
\end{tikzcd}
$$

Corollary 4.3. *If we set*

$$K^{even}(\mathbb{Z}) \dagger \Lambda_{BGL} = \{(V, w) \in K^{even}(\mathbb{Z}) \times \Lambda_{BGL} \,|\, c(V) = \mathrm{deR}(w)\}$$

and

$$T = H^{odd}(\mathbb{C}) \,/\, \mathrm{deR}(\Lambda_{GL})$$

then the diagram shows that

$$0 \to T \to \hat{K} \to K^{even}(\mathbb{Z}) \dagger \Lambda_{BGL} \to 0$$

is an exact sequence, the kernel of which is a complex torus of dimension the sum of the odd Betti numbers of the underlying manifold.

5. Hermitian Vector Bundles

In all that preceded, the basic objects were complex vector bundles with connection. The entire approach immediately applies to Hermitian bundles with inner product preserving connection. The same definition of equivalence goes through and gives rise to a Hermitian version of Struct. Analogs of all results remain true, with proofs following identical lines.

Letting $\hat{K}_R = K$ (Hermitian Struct), we obtain the following commutative diagram,

$$\begin{tikzcd}
0 \arrow[dr] & & & & 0 \\
& K^{odd}(\mathbb{R}/\mathbb{Z}) \arrow[rr, "\text{Bockstein}"] \arrow[dr, "j"] & & K^{even}(\mathbb{Z}) \arrow[ur] \arrow[dr, "ch"] & \\
H^{odd}(\mathbb{R}) \arrow[ur, "\text{reduction mod } \mathbb{Z}"] \arrow[dr, "\text{deR}"] & & \hat{K}_{\mathbb{R}} \arrow[ur, "\delta"] \arrow[dr, "ch"] & & H^{even}(\mathbb{R}) \\
& \Lambda^{odd}/\Lambda_U \arrow[rr, "d"] \arrow[ur, "i"] & & \Lambda_{BU} \arrow[ur, "\text{deR}"] \arrow[dr] & \\
0 \arrow[ur] & & & & 0
\end{tikzcd}$$

where Λ_U and Λ_{BU} are real-valued forms, defined analogously to Λ_{GL} and Λ_{BGL}.

By analogy to Corollary 4.3, we see

Corollary 5.1. *For any bundle over a closed Riemannian manifold after stabilizing, there is a unitary connection on the bundle whose Chern-Weil form is the harmonic representative of the Chern character of the bundle. Moreover, when the odd Betti numbers vanish, this structured bundle is unique up to adding factors with trivial holonomy.*

Appendix

Recall in the homotopy theory of spaces homotopy equivalent to CW complexes that a map $X \to Y$ is homotopy equivalent to the projection map of a Serre fibration. To see this let us assume X and Y are connected. First replace $X \xrightarrow{p} Y$ by $X \xrightarrow{\tilde{p}} \tilde{Y}$ where $\tilde{p}$ is an inclusion by replacing Y by (the mapping cylinder of p) $X \times I \cup_{\sim} Y$ where $X \times 1$ is collapsed by p onto its image in Y.

Then replace X by $\tilde{X}$ where $\tilde{X}$ is all the paths in $\tilde{Y}$ that start in X. Then $\tilde{X}$ maps into $\tilde{Y}$ (continue to call it $\tilde{p}$) with the Serre path lifting property by evaluating a path at its endpoint in $\tilde{Y}$. Clearly, $\tilde{Y} \sim Y$, $\tilde{X} \sim X$, and $\tilde{p} \sim p$.

The fibre $F \to X$ of $X \xrightarrow{p} Y$ is defined up to homotopy to be the inclusion into $\tilde{X}$ of the paths in $\tilde{Y}$ starting in X and ending at a specific point $y \in Y$ (or $\tilde{Y}$).

Question: What properties characterize the homotopy fibre $F \xrightarrow{i} X$ of a map $X \xrightarrow{p} Y$?

In the following, we assume X, Y, F, and F' are connected.

Proposition 5.2. *Suppose we have a map $F' \xrightarrow{i'} X$ and further suppose the composition $F' \xrightarrow{i'} X \xrightarrow{p} Y$ is provided with a null homotopy so that the induced map* *of homotopy sets*

$$\pi_i(X, F') \to \pi_i(Y, \text{base point})$$

are bijections $i = 1, 2, \ldots$. *Then* $F' \xrightarrow{i'} X$ *is homotopy equivalent to the homotopy fibre* $F \to X$ *of* $X \xrightarrow{p} Y$.

Proof. By our hypothesis, F' is connected. By the path lifting property of Serre fibrations, the null homotopy of the composition $F' \xrightarrow{i} X \xrightarrow{p} Y$ defines a canonical homotopy class of maps $F' \to F$ so that

$$\begin{array}{ccccc} F' & \xrightarrow{i'} & X & \xrightarrow{p} & Y \\ \downarrow & & \| & & \| \\ F & \xrightarrow{i} & X & \xrightarrow{p} & Y \end{array}$$

is homotopy commutative.

Now we look at the exact sequence of homotopy groups and sets

$$\begin{array}{ccccccccccccc} \cdots & \to & \pi_2 X & \to & \pi_2(X, F') & \to & \pi_1 F' & \to & \pi_1 X & \to & \pi_1(X, F') & \to & \pi_0 F' & \cong & \text{pt} \\ & & \| & & \downarrow & & \downarrow & & \| & & \downarrow & & \cong\downarrow & & \\ \cdots & \to & \pi_2 X & \to & \pi_2(X, F) & \to & \pi_1 F & \to & \pi_1 X & \to & \pi_1(X, F) & \to & \pi_0 F & \cong & \text{pt} \end{array}$$

In a Serre fibration the path lifting property implies that the homotopy sets $\pi_i(X, F)$ are isomorphic to $\pi_i(Y, \text{base point})$ and thus become groups. By the above commutative diagram the maps $\pi_i(X, F') \to \pi_i(X, F)$ are bijections. Thus the proposition follows from the 5-lemma. □

References

[1] Lott, John. "$\mathbb{R}/\mathbb{Z}$ Index Theory". Comm Anal Geom 2. 1994. pp. 279-311.

[2] Karoubi, M.
- (a) "Homologie Cyclique et K-Théorie". Asterisque 149. 1987.
- (b) "Théorie Générale des Classes Caractéristiques Secondaires". K-Theory 4. 1987. pp. 55-87.

[3] Bismut, J.M.
- (a) "The Index Theorem for Families of Dirac Operators: Two Heat Equation Proofs". Inv. Math. 83. pp. 91-151. 1986.
- (b) with Jeff Cheeger. "η-Invariants and Their Adiabatic Limits". J. Amer. Math. Soc. 2. 1989. pp. 33-70.
- (c) with D Freed. "The Analysis of Elliptic Familes II". Comm. Math. Phys. 107. 1986. pp. 103-163.

[4] Bunke, U. and Schick, T. "Smooth K-Theory". arXiv : 0707.0046 (math.KT). 2007. 57 pp.

[5] Hopkins, M.J. and Singer, I.M. "Quadratic functions in Geometry, Topology, and M-theory". J. Diff. Geom 70. 2005. pp. 329-452.

[6] Freed, Daniel S.
- (a) "On Determinant Line Bundles". Mathematical Aspects of String Theory. World Scientific. 1987. pp. 189-238.
- (b) "Dirac Charge Quantization and Generalized Differential Cohomology". arXiv:hep-th/0011220. (hep-th,math.DG). 2000. 62 pp.
- (c) with Gregory W. Moore. "Setting the Quantum Integrand of M-Theory". arXiv:hep-th/0409135. 2004. 52 pp.
  Comm. Math. Physics 263. 2006. pp. 89-132.
- (d) with M.J. Hopkins. "On Ramond-Ramond fields and K-Theory". arXiv': hep-th/0002027v3. 2000. 12 pp.

[7] Simons, James and Sullivan, Dennis. "Axiomatic Characterization of Ordinary Differential Cohomology".
arXiv : math/0701077v1. (math.AT). 2007. 15 pp.
Journal of Topology 1 (1). 2008. pp 45-56.

[8] Witten, Edward. "Global Gravitational Anomalies". Commun. Math. Phys. 100. 1985. p. 197.

[9] Narasimhan, M.S. and Ramanan, S. "Existence of Universal Connections". Amer. J. Math. 83. 1961. pp. 563-572, 85. 1963. pp. 223-231.

[10] Adams, J.F. "A Variant of E.H. Brown's Representability Theorem". Topology 10. 1971. pp. 185-198. MR 44:1018.

[11] Cheeger, Jeff and Simons, James. "Differential Characters and Geometric Invariants". Notes of Stanford Conference 1973, Lecture Notes in Math. No. 1167. Springer-Verlag, New York. 1985. pp. 50-90.

[12] Cheeger, Jeff. "Eta-invariants, the adiabatic approximation and conical singularities". Journal of Differential Geometry 26. 1987. pp. 175-221.

[13] Anderson, Donald. "A New Cohomology Theory". UC-Berkeley Thesis, 1964. Unpublished.

Euclidean Capital, 160 5th Ave., New York, NY, 10010

Department of Mathematics, Stony Brook University, Stony Brook NY, 11794-3651, USA

E-mail address: dennis@math.sunysb.edu

Clay Mathematics Proceedings
Volume 11, 2010

Entropy in Operator Algebras

Erling Størmer

Dedicated to Alain Connes on the occasion of his 60th birthday

Introduction

Noncommutative entropy is one of the many mathematical topics in which Alain Connes has excelled. In these notes I shall give a quick survey of the subject with the main emphasis on Alain's contributions and influence.

After Ornstein in 1970 [**19**] published his result on the classification of Bernoulli shifts by entropy, entropy related studies in ergodic theory became a very active discipline in the early 1970's. At roughly the same time it became clear that operator algebras and ergodic theory had more in common than was previously recognized. It was therefore a natural problem to try to extend the concept of entropy to operator algebras. It started in the C^*-algebraic formalism of quantum statistical mechanics, where spin lattice systems with automorphisms arising from translations behaved very much like abelian systems, so that entropy could be generalized by replacing partitions by local algebras. The resulting mean entropy yielded most of the results one wanted. But in more general C^*-dynamical systems mean entropy was useless.

Alain and I met often at that time, and we both thought about possible noncommutative generalizations of entropy. One of our motivations was that it might yield useful invariants for operator algebras. This later on turned out to be true by Brown's characterization [**2**] of type I C^*-algebras as those C^*-algebras for which inner automorphisms have zero entropy with respect to invariant states. Alain's and my common interest developed into joint work on the problem, in which Alain's deep understanding and technical fluency made it possible to write our paper [**7**], which appeared in 1975, in which we gave a definition which has proved useful to this day. Twelve years later Alain jointly with Narnhofer and Thirring [**6**] extended the definition to C^*-algebras and obtained what I consider the best definition to date.

Some years later, in 1995, Voiculescu [**30**] defined topological entropy, which together with an extension by Brown to exact C^*-algebras [**3**] became the basic generalization of topological entropy as it is defined in ergodic theory.

2010 *Mathematics Subject Classification.* Primary: 46L53. Secondary: 46L05, 46L55.

There have been many other attempts at defining entropy of automorphisms in operator algebras. I shall indicate some of them in Section 8. Most of them suffer from a drawback in that they do not satisfy all the criteria I want for a suitable definition; namely

(i) They should restrict to the usual entropy in the abelian case.

(ii) They should be computable in the simpler noncommutative cases, e.g. give $\log n$ for the entropy with respect to the tracial state of the shift on the infinite tensor product of the $n \times n$ matrices.

(iii) There should be a kind of Kolmogorov-Sinai theorem in the theory.

Since the technical proofs of the theory are covered in the book [**18**] by Neshveyev and myself and also the survey article of mine [**27**], I shall mainly state results and definitions and try to present the main ideas involved.

1. Definition of noncommutative entropy

Let us first fix notation and recall the definition of entropy as it is given in the "classical" abelian case.

Let $(X, \mathcal{B}, \mu)$ be a probability space with X a set, $\mathcal{B}$ a $\sigma-$algebra, and μ a probability measure. Let $\mathcal{P} = \{P_1, ..., P_k\}$ be a partition of X with $P_i \in \mathcal{B}$, and let η denote the real function on [0,1] defined by $\eta(0) = 0, \eta(t) = -t \log t$ for $t > 0$. The *entropy* of $\mathcal{P}$ is

$$H(\mathcal{P}) = \sum_{i=1}^{k} \eta(\mu(P_i)).$$

We identify P_i with its characteristic function χ_{P_i} and $\mathcal{P}$ with the linear span of the χ_{P_i}, $i = 1, ...k$. Then $\mathcal{P}$ is a finite dimensional subalgebra of $L^\infty(X, \mathcal{B}, \mu)$. If $\mathcal{P}_1, ..., \mathcal{P}_k$ are finite dimensional subalgebras we define their joint entropy by

$$H(\mathcal{P}_1, ..., \mathcal{P}_k) = H(\bigvee_{i=1}^{k} \mathcal{P}_i), \tag{1}$$

where $\bigvee_{i=1}^{k} \mathcal{P}_i$ is the finite dimensional algebra they generate. If T is a nonsingular measure preserving transformation of X, define an automorphism α of $L^\infty(X, \mathcal{B}, \mu)$ by

$$\alpha(f)(x) = f(T^{-1}x)$$

for $f \in L^\infty(X, \mathcal{B}, \mu), x \in X$. Then we put

$$H(\alpha, \mathcal{P}) = \lim_n \frac{1}{n} H(\mathcal{P}, \alpha(\mathcal{P}), ..., \alpha^{n-1}(\mathcal{P})), \tag{2}$$

and the entropy of α as

$$H(\alpha) = \sup_{\mathcal{P}} H(\alpha, \mathcal{P}), \tag{3}$$

where the sup is taken over all finite dimensional subalgebras of $L^\infty(X, \mathcal{B}, \mu)$. In order to compute $H(\alpha)$ it is usually necessary to invoke the Kolmogorov-Sinai Theorem. There are several versions of this theorem. The one which is most useful for generalization to the noncommutative case states that if $\mathcal{P}_1 \subset \mathcal{P}_2 \subset \cdots$ are finite dimensional subalgebras with union weakly dense in $L^\infty(X, \mathcal{B}, \mu)$, then $H(\alpha) = \lim_n H(\alpha, \mathcal{P}_n)$. If we want to define entropy in the noncommutative case, the natural first approach would be to replace $L^\infty(X, \mathcal{B}, \mu)$ by a C^*-algebra or a von Neumann algebra and the measure μ by a state ϕ and a partition by a finite dimensional subalgebra. If M is a finite dimensional C^*-algebra with a tracial state τ, then a natural definition of the entropy of M is

$$H(M) = \sum_i \eta(\tau(e_i)),$$

where $e_1, ..., e_k$ is an orthogonal family of minimal projections in M with sum 1. But if N is another finite dimensional C^*-subalgebra then the algebra generated by M and N is often of infinite dimension, so equation (1) does not make sense. To overcome this problem we must rewrite (1) in a form which considers the $\mathcal{P}_i$ separately.

Consider first one finite dimensional subalgebra $\mathcal{P}$. Let

$$F = \{f_i \in \mathcal{P} : i = 1, \ldots, n, f_i \geq 0, \sum_{i=1}^{n} f_i = 1\}$$

be a finite partition of 1 in $\mathcal{P}$. Then we can show

$$H(\mathcal{P}) = \sup_F \sum_i (\eta(\mu(f_i)) - \mu(\eta(f_i))),$$

where the sup is taken over all finite partitions F. Let now $\mathcal{P}_1, ..., \mathcal{P}_n$ be finite dimensional subalgebras of $L^\infty(X, \mathcal{B}, \mu)$, $E_{\mathcal{P}_k}$ the $\mu-$ invariant conditional expectation of $L^\infty(X, \mathcal{B}, \mu)$ onto $\mathcal{P}_k$, and $F = \{f_{i_1,\ldots,i_n}\}$ a finite partition of 1 in $L^\infty(X, \mathcal{B}, \mu)$. Put

$$f_{i_k}^{(k)} = \sum_{\check{i_k}} f_{i_1,\ldots,i_n},$$

i.e. the sum holding the index i_k fixed.

Theorem 1. *With the above notation*

$$H(\mathcal{P}_1, ..., \mathcal{P}_n) = \sup_F \sum_{i_1,\ldots,i_n} \eta(\mu(f_{i_1,\ldots,i_n})) - \sum_{k=1}^{n} \sum_{i_k} \mu(\eta(E_{\mathcal{P}_k} f_{i_k}^{(k)})),$$

where the sup is taken over all finite partitions of $\bigvee_i \mathcal{P}_i$.

This formula for $H(\mathcal{P}_1, ..., \mathcal{P}_n)$ makes sense in the noncommutative case, and was behind the definition in [**7**]. Let M be a von Neumann algebra with a faithful normal tracial state τ. Let $N_1, ..., N_n \subset M$ be finite dimensional von Neumann subalgebras. Let E_N denote the $\tau-$invariant conditional expectation on a von Neumann subalgebra N of M, and let $F = \{f_{i_1,\ldots,i_n}\}$ be as before with $f_{i_1,\ldots,i_n} \in M^+$. Then with μ replaced by τ and the $\mathcal{P}_i$ by N_i, we define the *mutual entropy* $H(N_1, ..., N_n)$ as in the theorem. If α is an automorphism of M such that $\tau \circ \alpha = \tau$ define $H(\alpha, N)$ as in the classical case (2), and the entropy of α as in (3), i.e.

$$H(\alpha) = \sup_N H(\alpha, N).$$

$H(\alpha)$ has many features in common with classical entropy; for example if M is hyperfinite then $H(\alpha^p) = |p| H(\alpha)$ for an integer p, but it does not behave so well with respect to tensor products. If (M_i, α_i, τ_i), $i = 1, 2$, are two W^*-dynamical systems as above, then we can only conclude that

$$H(\alpha_1 \otimes \alpha_2) \geq H(\alpha_1) + H(\alpha_2),$$

because there are not enough finite dimensional subalgebras of $M_1 \otimes M_2$ of the form $N_1 \otimes N_2$ to conclude equality.

The Kolmogorov-Sinai Theorem assumes the form alluded to above, namely

THEOREM 2. *If $N_1 \subset N_2 \subset \cdots$ is an increasing sequence of finite dimensional von Neumann subalgebras of M with union weakly dense in M, then*

$$H(\alpha) = \lim_n H(\alpha, N_n).$$

A technical tool in the proof is that of *relative (or conditional) entropy* of two finite dimensional subalgebras N and P, defined by

$$H(N|P) = \sup \sum_i (\tau(\eta(E_P(x_i))) - \tau(\eta(E_N(x_i)))),$$

where the sup is taken over all finite partitions x_i of 1. $H(N|P)$ has very nice continuity properties, which are used in the proof, i.e. if N is approximately contained in P in the $\|\cdot\|_2$-norm defined by τ, then $H(N|P)$ is small. We shall encounter relative entropy again in Section 7 in a different situation.

The mutual entropy $H(N_1, ..., N_n)$ satisfies most of the properties that can be expected from the abelian case, like symmetry in the N_i, subadditivity, monotonicity etc. One property has been very useful for computation, namely, if P_i is a C^*-subalgebra of N_i for each i such that the P_i pairwise commute, and $\bigvee_i P_i = \bigvee_i N_i$, then $H(N_1, ..., N_n) = H(\bigvee_i N_i)$. A more abelian result was shown by Haagerup and myself [**10**]. We use the short hand "masa" for maximal abelian subalgebra.

THEOREM 3. *With notation as above suppose $N = \bigvee_{i=1}^n N_i$ is finite dimensional. Then $H(N_1, ..., N_n) = H(N)$ if and only if there exists a masa $A \subset N$ such that $A = \bigvee_{i=1}^n A \cap N_i$.*

The easy part of the proof is a direct consequence of the properties of H alluded to above, namely

$$\begin{aligned}
H(N) &\geq H(N_1, ..., N_n), \quad \textit{since} \;\; N_i \subset N \\
&\geq H(A \cap N_1, ..., A \cap N_n), \quad \textit{since} \;\; A \cap N_i \subset N_i \\
&= H(\bigvee_{i=1}^n A \cap N_i), \quad \textit{since the abelian algebras} \;\; A \cap N_i \;\; \textit{commute} \\
&= H(A) \\
&= H(N), \quad \textit{since } A \textit{ is a masa in } N.
\end{aligned}$$

The converse implication is much more involved as one then has to construct the masa A.

Example. Let $M = \bigotimes_{-\infty}^{\infty}(M_i, \tau_i)$ be the infinite tensor product of the $n \times n$ matrices M_i with itself with respect to the tracial state τ_i on M_i. Let α be the shift on M, and $D = \bigotimes_{-\infty}^{\infty} D_i$, where $D_i = \alpha(D_0), D_0$ a masa in M_0. Then

$$H(\alpha) = H(\alpha|_D) = \log n.$$

This follows immediately from the easy part of Theorem 3 together with the Kolmogorov-Sinai Theorem (Thm.2) applied to the sequence $N_k = \bigotimes_{-k}^{k} M_i \subset M$. The example can easily be generalized to general Bernoulli shifts by replacing τ_i by a state $\phi_i = \phi_0$ for all i. Then M must be replaced by the centralizer of $\phi = \otimes\phi_i$.

After the appearance of Alain's and my paper [**7**] not much happened to the theory for ten years until Alain realized that the formula for mutual entropy could be rewritten in terms of relative entropy of states, and thus the definition of entropy could be extended to nontracial states (for details of the following see Chapter 3 in [**18**]). Recall that if ϕ and ψ are positive linear functionals on a finite dimensional C^*-algebra with density operators Q_ϕ and Q_ψ with respect to a trace Tr, then their *relative entropy* is

$$S(\phi,\psi)=\mathrm{Tr}(Q_\phi(\log Q_\phi-\log Q_\psi)), \text{if} \quad \phi\leq\lambda\psi,\lambda>0,$$

and we put $S(\phi,\psi)=+\infty$ if no such λ exists. Note that $\tau\eta(x)=-S(\tau(x\cdot)|_N,\tau|_N)$ for N a finite dimensional algebra, and that the map $x\mapsto\tau(x\cdot)$ establishes a 1-1 correspondence between positive operators $0\leq x\leq 1$ and positive linear functionals $\phi\leq\tau$ on N. We can therefore rewrite the definition of mutual entropy as

$$H(N_1,...,N_n)=\sup_{\phi_{i_1,\dots,i_n}}\sum_{i_1,\dots,i_n}\eta(\phi_{i_1,\dots,i_n}(1))+\sum_k\sum_{i_k}S(\phi^{(k)}_{i_k}|_{N_k},\tau|_{N_k}),$$

where $\phi_{i_1,\dots,i_n}$ and $\phi^{(k)}_{i_k}$ are defined in analogy with the $f_{i_1,\dots,i_n}$ and $f^{(k)}_{i_k}$ in the abelian case. Alain [**5**] then defined the mutual entropy with respect to a normal state analogously, namely

$$H_\phi(N_1,...,N_n)=\sup_{\phi_{i_1,\dots,i_n}}\sum_{i_1,\dots,i_n}\eta(\phi_{i_1,\dots,i_n}(1))+\sum_k\sum_{i_k}S(\phi^{(k)}_{i_k}|_{N_k},\phi|N_k),$$

where $\sum\phi_{i_1,\dots,i_n}=\phi$, and S is a more general definition of relative entropy than the above. Thus he could define $H_\phi(\alpha)$ of a $\phi-$invariant automorphism α of a von Neumann algebra with respect to a normal state ϕ.

A couple of years later Alain together with Narnhofer and Thirring extended the definition to automorphisms of C^*-algebras [**6**]. Then one encounters a new problem, namely the lack of sufficiently many finite dimensional C^*-subalgebras of a C^*-algebra A. They solved this by restating the definition as

$$H_\phi(\gamma_1,...,\gamma_n)=\sup_{\phi_{i_1,\dots,i_n}}\sum_{i_1,\dots,i_n}\eta(\phi_{i_1,\dots,i_n}(1))+\sum_k\sum_{i_k}S(\phi^{(k)}_{i_k}\circ\gamma_k,\phi\circ\gamma_k), \tag{4}$$

where $\gamma_1,...,\gamma_n$ are completely positive maps from finite dimensional C^*-algebras into A.

A clever trick was to represent the definition as an abelian problem. They noticed that if C is a finite dimensional abelian C^*-algebra, $C_1,...,C_n\subset C$ subalgebras, μ a state on C, and $P\colon A\to C$ a unital positive map such that $\phi=\mu\circ P$, then a decomposition of ϕ is given by

$$\phi_{i_1,\dots,i_n}(a)=\mu(P(a)p^{(1)}_{i_1}\cdots p^{(n)}_{i_n}),$$

where $p^{(k)}_{i_k}$ is the set of atoms in C_k. Conversely each decomposition of ϕ gives rise to such an abelian model, denoted by $(C,\mu,(C_k),P)$. Thus (4) can be rewritten as

$$H_\phi(\gamma_1,...,\gamma_n)=\sup H_\mu(\bigvee_k C_k)+\sum_k\sum_{i_k}S(\mu(P\circ\gamma_k)p^{(k)}_{i_k},\phi\circ\gamma_k), \tag{5}$$

where the sup is taken over all abelian models. As before $H_\mu(\bigvee_k C_k)$ is the entropy of $\bigvee_k C_k)$ with respect to μ. Now, the entropy of a $\phi-$invariant automorphism α is defined as before, namely we put

$$H_\phi(\alpha,\gamma)=\lim\frac{1}{n}H_\phi(\gamma,\alpha\circ\gamma,...,\alpha^{n-1}\circ\gamma),$$

and

$$h_\phi(\alpha) = \sup_\gamma H_\phi(\alpha, \gamma).$$

This entropy is usually referred to as the *CNT-entropy* after Connes, Narnhofer and Thirring, and it has all the desired properties, including a Kolmogorov-Sinai theorem resembling that of Theorem 2. It is tied together with the von Neumann algebra definition via the GNS-representation π_ϕ of ϕ. Namely, let $\bar{\alpha}$ be the extension of α to $\pi_\alpha(A)$". Then

$$h_\phi(\alpha) = H_\phi(\bar{\alpha}).$$

Thus, after placing the theory of dynamical entropy on a firm basis, Alain stopped working actively on entropy; there were so many other things he wanted to do. But he kept his interest in the subject and had a strong influence on its further development. This I shall describe in the next sections.

2. Bogoliubov automorphisms

One of my favorite C^*-algebras is the CAR-algebra $A(H)$, which is isomorphic to the infinite tensor product of the 2×2 matrices. It is generated by operators $a(f)$ where $f \in H$, a complex Hilbert space, and $f \to a(f)$ is a linear map of H into a C^*-algebra satisfying the canonical anticommutation relations

$$a(f)a(g)^* + a(g)^*a(f) = (f, g)1,$$

$$a(f)a(g) + a(g)a(f) = 0,$$

for all $f, g \in H$. If U is a unitary operator on H, the Bogoliubov automorphism of $A(H)$ corresponding to U is defined by

$$\alpha_U(a(f)) = a(Uf).$$

Let $m(U)$ be the multiplicity function of the absolutely continuous part of U, and let τ be the unique tracial state on $A(H)$.

In a letter written in 1987, or perhaps it was in 1988, Alain asked me to prove the following formula:

$$h_\tau(\alpha_U) = \frac{\log 2}{2\pi} \int_0^{2\pi} m(U)(\theta) d\theta,$$

a formula which requires deep insight in the theory to guess. I started on the problem, but hit a wall after a while. Fortunately I met Voiculescu at a conference in Kansas and told him about the problem. It turned out that he could do what I could not do, so we could prove the above formula, even in the more general setting of certain quasi-free states. The ideas are explained in [**27**], Chapter 4, and details can be found in [**18**], Chapter 13. The main idea is the following. A unitary U can be written as a direct sum $U_s \oplus U_a$, where U_s has singular and U_a absolutely continuous spectral measure with respect to the Lebesgue measure $d\theta$ on the circle. The first step is to show $h_\tau(\alpha_{U_s}) = 0$. This is also true for the entropy with respect to a quasi-free state ω_A which is defined by an operator $0 \le A \le 1$ by the formula

$$\omega_A(a(g_1)^* \cdots a(g_m)^* a(f_1) \cdots a(f_n)) = \delta_{mn} \det(Af_i, g_j).$$

ω is α_U-invariant if and only if $AU = UA$. Note that $\tau = \omega_{\frac{1}{2}}$.

If A has pure point spectrum, ω_A is an infinite product state. The part U_a is then essentially a direct sum of powers of the bilateral shift. Elaborating on these vague ideas, Voiculescu and I [**28**] could prove the following theorem under the assumption of pure point spectrum.

Theorem 4.

$$h_{\omega_A}(\alpha_U) = \frac{\log 2}{2\pi} \int_0^{2\pi} \mathrm{Tr}(\eta(A(\theta)) + \eta(1 - A(\theta)))d\theta,$$

where $H = \int^{\oplus} H_\theta d\theta, A = \int^{\oplus} A(\theta)d\theta, U_a = \int^{\oplus} U_a(\theta)d\theta$, *are direct integral representations of* H, A, U_a *respectively.*

The assumption that A had pure point spectrum inspired others to improve the theorem. First Narnhofer and Thirring [**14**] and Park and Shin [**20**] proved generalizations, and then Neshveyev [**16**] proved it in general. His paper was the starting point of a fruitful collaboration with me, later resulting in a permanent position for Neshveyev at the University of Oslo. Thus Alain's influence on the entropy formula for Bogoliubov automorphisms was not only a proof of the, I think, most spectacular formula in the theory of dynamical entropy, but also it changed Neshveyev's life considerably.

3. Topological and a related entropy

When we defined entropy we encountered the problem of considering the algebra generated by several finite dimensional algebras. We overcame the problem by considering them separately and could thus define their mutual entropy; see Theorem 1. Alain and I [**8**] proposed another definition in which we considered finite subsets of operators of the von Neumann algebra instead of subalgebras, and thus reduced the problem with generated algebras to that of taking unions of sets. The definition was as follows. Let $M \subset B(H)$ be a von Neumann algebra with a cyclic and separating trace vector. Let τ be the corresponding trace. Let $U(M)$ denote the unitary group in M. For $\delta > O$ and $F \subset U(M)$ finite put

$$(6) \qquad h_\delta^\tau(F) = \inf\{S(\phi) : \phi \in B(H)_*^+, \phi\,|_M = \tau, S(\phi, \phi \circ \mathrm{Ad}\, u) < \delta \ \ \forall u \in F\}$$

The analogue of $H(\alpha, N)$ for a $\tau-$invariant automorphism α is now

$$h^\tau(\alpha, F) = \sup_{\delta>0} \limsup_n \frac{1}{n} \ \ h_\delta^\tau(\bigcup_{i=0}^{n-1} \alpha^i(F)),$$

and the entropy is

$$h^\tau(\alpha) = \sup_F h^\tau(\alpha, F),$$

where the sup is taken over all finite subsets F of $U(M)$. We could then show that the definition coincided with the classical definition in the abelian case, but did nothing else with it.

The above procedure reappeared in Voiculescu's definition of topological entropy [**30**]. Instead of taking the entropy $S(\phi)$ in (6) he used the idea of the McMillan Theorem. Let $(X, \mathcal{B}, \mu)$ be a probability space with a nonsingular measure preserving transformation T of X. Let α be the corresponding automorphism of $L^\infty(X, \mathcal{B}, \mu)$. If $\mathcal{P}$ is a finite partition, the entropy $H(\mathcal{P}, \alpha)$ is, up to approximation depending on μ, roughly determined by the number of sets in $\mathcal{P}$. The corresponding value for a finite dimensional C^*-algebra B is $\mathrm{rank}\, B$, the dimension of a masa in B. Voiculescu's definition now takes the form:

Let A be a C^*-algebra, F a finite subset of A and B a finite dimensional C^*-algebra. Let ρ and ψ be two unital completely positive maps, $\rho\colon A \to B$ and $\psi\colon B \to A$. Put

$$\operatorname{rcp}(F, \delta) = \inf\{\operatorname{rank} B : \|\psi \circ \rho(x) - x\| < \delta\}, \tag{7}$$

where the inf is over all triples (B, ρ, ψ). Let $\alpha \in \operatorname{Aut}(A)$ be an automorphism. Then the analogue of $H(\alpha, N)$ is now

$$\operatorname{ht}(\alpha, F) = \sup_{\delta>0} \limsup_{n} \frac{1}{n} \log \operatorname{rcp}\Big(\bigcup_{i=0}^{n-1} \alpha^i(F)\Big),$$

and the *topological entropy* of α is

$$\operatorname{ht}(\alpha) = \sup_F \operatorname{ht}(\alpha, F),$$

where the sup is taken over all finite subsets of A. This works well when A is a nuclear algebra, because we always find B, ρ, ψ such that $\|\psi \circ \rho(x) - x\| < \delta$ for all $x \in F$. More generally, as shown by Brown [**3**], it works when A is exact. $\operatorname{ht}(\alpha)$ satisfies most of the properties one expects from topological entropy. We have a Kolmogorov-Sinai Theorem, and $\operatorname{ht}(\alpha^p) = |p|\operatorname{ht}(\alpha)$ for an integer p. Moreover $h_\phi(\alpha) \leq \operatorname{ht}(\alpha)$ for all $\alpha-$invariant states ϕ. But the topological entropy of a tensor product of two automorphism satisfies only the inequality

$$\operatorname{ht}(\alpha \otimes \beta) \leq \operatorname{ht}(\alpha) + \operatorname{ht}(\beta),$$

which is the opposite inequality of what we have for dynamical entropy. In particular when there exist ϕ and ψ such that $h_\phi(\alpha) = \operatorname{ht}(\alpha)$ and similarly for β, then we have equalities in the above inequality and the corresponding for dynamical entropy.

This leads us to a problem which was left open in the first papers on noncommutative entropy; namely, is

$$h_{\phi\otimes\psi}(\alpha \otimes \beta) = h_\phi(\alpha) + h_\psi(\beta)?$$

Even when ϕ and ψ are traces the answer turned out to be negative. Narnhofer, Thirring and I [**15**] found an example of a binary shift α on the CAR-algebra for which $h_\tau(\alpha) = 0$, while $h_{\tau\otimes\tau}(\alpha \otimes \alpha) = \log 2$. In this example $\operatorname{ht}(\alpha) = \log 2$. Almost at the same time Sauvageot [**23**] exhibited another example where equality fails.

4. The variational principle

The variational principle has over the years attracted much attention both in classical ergodic theory and in the C^*-algebra setting for quantum statistical mechanics. Using mean entropy on spin lattice systems mathematical physicists, in particular Araki, Lanford, Robinson, and Ruelle showed that the variational principle holds with respect to translations and that the KMS condition characterizes equilibrium states. It was natural to expect that one might prove similar results for the CNT-entropy h_ϕ. Early work on this extension was done by Narnhofer [**13**], studying KMS states and then by Moriya [**12**], who showed that for lattice systems one could replace mean entropy by dynamical entropy in the formalism of the variational principle. The formulation of the simplest case is that for an automorphism α of a C^*-algebra one should have $\operatorname{ht}(\alpha) = \sup h_\phi(\alpha)$, where the sup is taken over all $\alpha-$invariant states. This is false for almost all binary shifts of the CAR-algebra, see Theorem 12.2.3 and Proposition 12.2.1 in [**18**]. Thus some assumptions are necessary. Again Alain's deep insight was necessary.

Alain is a very good friend of the Norwegian operator algebraists, and once he went with some of us to an old mining town, Røros, in Norway to go skiing. At dinner one day Alain suggested to me that I show the variational principle for asymptotically abelian C^*-algebras, because he was convinced that for some of the many definitions of asymptotic abelianness the variational principle should hold. I thought that it was a very nice problem and invited Sergey Neshveyev to come to Oslo and work on it with me. Later on I noticed that Alain had stated the problem in his book on noncommutative geometry.

Let me first start with the finite dimensional case. Let B be a finite dimensional C^*-algebra with a trace Tr_B satisfying $\mathrm{Tr}_B(e) = 1$ for all minimal projections in B. Let $\phi = \mathrm{Tr}_B(Q_\phi \cdot)$ be a state on B with density operator Q_ϕ. in B^+. Let

$$S(\phi) = \mathrm{Tr}_B(\eta(Q_\phi))$$

be the entropy of ϕ. The variational principle is formulated as follows.

Lemma 5. *With the above notation*

$$S(\phi) - \phi(h) \leq \log \mathrm{Tr}_B(e^{-h})$$

for $h \in B$ self-adjoint. Equality holds if and only if

$$Q_\phi = \frac{e^{-h}}{\mathrm{Tr}_B(e^{-h})}$$

is the Gibbs state. In the latter case ϕ is a KMS state with respect to the one-parameter automorphism group

$$\sigma_t(x) = e^{-ith} x e^{ith}.$$

That ϕ is KMS means that $\phi(a\sigma^h_{it}(b)) = \phi(ba)$ for all $a, b \in B$. This lemma gives a model for defining the analogue, called pressure, of $\log \mathrm{Tr}_B(e^{-h})$. We imitate the definition of topological entropy. So, let A be a C^*-algebra, $F \subset A$ a finite set, and B a finite dimensional C^*-algebra with completely positive maps $\rho \colon A \to B, \psi \colon B \to A$. Let $h \in A$ be self-adjoint and α an automorphism of A. If $\delta > 0$ let

$$P(h, F, \delta) = \inf_{(B,\rho,\psi)} \{\log \mathrm{Tr}_B(e^{-\rho(h)}) : \|\psi \circ \rho(x) - x\| < \delta, \quad \forall x \in F\}.$$

Note that if $h = 0$ then $\log \mathrm{Tr}_B(e^{-\rho(h)}) = \mathrm{rank}\, B$, so we have the situation of the definition of topological entropy. Now continue as in the definition of topological entropy and obtain the *pressure* $P(\alpha, h)$ of α with respect to h. In particular $P(\alpha, 0) = ht(\alpha)$. One can show that $P(\alpha, h)$ satisfies the main properties expected from the abelian case, see Section 9.1 in [**18**]. The choice of asymptotic abelianness which is required for the variational principle, is given by

Definition 6. *A C^*-dynamical system (A, α) is called asymptotically abelian with locality if there exists an $\alpha-$invariant dense *-subalgebra $\mathcal{A} \subset A$, called the local algebra, such that for each pair $a, b \in \mathcal{A}$ the C^*-algebra $C^*(a, b)$ they generate is finite dimensional, and there exists $p = p(a, b) \in \mathbb{N}$ such that $[\alpha^j(a), b] = 0$ whenever $|p| \leq j$, and $[\ ,\ ]$ denotes the commutator.*

A good example of the above situation is the infinite tensor product $\bigotimes_{-\infty}^{\infty} B_i$, with $B_i = B_0$ finite dimensional and α the shift. Then $\mathcal{A}$ consists of operators which are finite sums of finite tensors. Note that if A is separable, we can take a dense sequence in $\mathcal{A}$ and inductively define an increasing sequence of finite dimensional C^*-subalgebras of $\mathcal{A}$ with union dense in $\mathcal{A}$, hence in A. Thus A is an AF-algebra. We can now formulate the variational principle [**17**].

THEOREM 7. *Let (A, α) be a unital separable C^*-dynamical system which is asymptotically abelian with locality. Then*

$$P(\alpha, h) = \sup_{\phi}(h_{\phi}(\alpha) - \phi(h)),$$

where the sup is taken over all $\alpha-$invariant states. In particular the topological entropy satisfies

$$\mathrm{ht}(\alpha) = \sup_{\phi} h_{\phi}(\alpha).$$

The proof consists essentially of two lemmas. The first proves the theorem under the assumption that there exists a finite dimensional C^*-subalgebra N of $\mathcal{A}$ such that the images $\alpha^j(N)$ all commute for $j \in \mathbb{Z}$ and that they generate A as a C^*-algebra. In general there is no such N, but for each N there is $p \in \mathbb{N}$ such that $\alpha^i(N)$ and $\alpha^j(N)$ commute whenever $|i - j| \geq p$, and some estimation is necessary to reduce to the first lemma.

In order to obtain the KMS part for the analogue of Gibbs states one must be more careful in defining the corresponding 1-parameter automorphism group. So assume (A, α) is asymptotically abelian with locality and h a self-adjoint operator in the local algebra $\mathcal{A}$. Put

$$\delta_h(x) = \sum_{i \in \mathcal{Z}} [\alpha^j(h), x]$$

for $x \in \mathcal{A}$,. For each $x \in \mathcal{A}$ the sum is finite, and δ_h is a derivation of $\mathcal{A}$ which defines a 1-parameter group $\sigma_t^h = \exp(it\delta_h)$ of automorphisms of A. Let $\beta \geq 0$. We say an $\alpha-$invariant state ϕ is an *equilibrium state at h at inverse temperature β* if

$$P(\alpha, \beta h) = h_{\phi}(\alpha) - \beta\phi(h)(= \sup_{\psi}(h_{\psi}(\alpha) - \beta\psi(h)).$$

Then the KMS condition for equilibrium states (or Gibbs states) can be formulated.

THEOREM 8. *With (A, α) and notation as above assume* $\mathrm{ht}(\alpha) < \infty$, *and let $h \in \mathcal{A}$ be self-adjoint. Suppose ϕ is an equilibrium state at h at inverse temperature β. Then ϕ is a KMS state for σ^h at β. In particular, if* $\mathrm{ht}(\alpha) = h_{\phi}(\alpha)$, *then ϕ is a trace.*

5. Free products

In this and the following two sections I want to present a glimpse of three major developments in the theory which have not been under as direct influence of Alain as the previous sections. As we have seen, entropy was best behaved and took values expected from the abelian theory when the C^*-dynamical system (A, α) had a great deal of abelianness built into it. To understand entropy better it is therefore necessary to consider the opposite case when the system is highly noncommutative. The first study was of the free shift, defined as follows.

Let $\mathcal{F}_\infty$ be the free group on an infinite number of generators $(g_i)_{i \in \mathcal{Z}}$. Let $L(\mathcal{F}_\infty)$ be the II_1- factor obtained from the left regular representation of $\mathcal{F}_\infty$, and let α be the free shift on $L(\mathcal{F}_\infty)$ defined by $\alpha(g_i) = g_{i+1}$. α is extremely ergodic; for example if $N \subset L(\mathcal{F}_\infty)$ is a finite dimensional C^*-subalgebra, or abelian, or even injective, and $\alpha(N) = N$, then $N = \mathbb{C}1$, [**22**]. Our experience from the abelian case would make us guess that such an automorphism must have infinite entropy, but from the results in Section 1 the extreme noncommutativity of $L(\mathcal{F}_\infty)$ indicates that this may not be so. Furthermore $\mathcal{F}_\infty$ has some features in common with compact groups in that powers of the free shift α do not move elements "very far away". The answer is given in the next theorem, [**26**].

THEOREM 9. *The free shift of* $L(\mathcal{F}_\infty)$ *has entropy* $H(\alpha) = 0$ *taken with respect to the trace.*

The result was subsequently extended to free products of C^*−algebras by Choda, Brown, Dykema, and Shlyakhtenko, see [**17**], Chapter 14.

THEOREM 10. *Let* A_0 *be an exact* C^*−*algebra, and* $A_n = A_0, n \in \mathbb{Z}$. *Let* ϕ_0 *be a state of* A_0, *and* $\phi_n = \phi_0$ *the same state on* A_n. *Let* α *be the free shift on the free product* $(A, \phi) = (*A_n, *\phi_n)$. *Then* $\mathrm{ht}(\alpha) = 0$.

A similar theorem was proved by me for dynamical entropy of general C^*−algebras. I mentioned in Section 4 that almost all binary shifts on the CAR-algebra have entropy zero. In those cases too, the reason is the highly noncommutative situation of the example. We can summarize the discussion with the following general principle:

If (A, α) is a highly noncommutative C^*−dynamical system, then $\mathrm{ht}(\alpha)$ and $h_\phi(\alpha)$ tend to be zero.

6. Crossed products

If (A, α) is a C^*−dynamical system and ϕ an α−invariant state, the crossed product $A \ltimes_\alpha \mathbb{Z}$ is generated by an isomorphic image of A and a unitary operator u such that $\alpha(x) = uxu^*$ for $x \in A$. This and more general constructions yield many of the main examples of C^*− and von Neumann algebras in operator algebra theory. Define an automorphism $\bar{\alpha}$ on $A \ltimes_\alpha \mathbb{Z}$ by $\bar{\alpha}(x) = \alpha(x)$ if $x \in A$, and $\bar{\alpha}(u) = u$. There is a natural conditional expectation $E\colon A \ltimes_\alpha \mathbb{Z} \to A$ defined by $E(xu^n) = 0$ if $n \neq 0$, and $E(x) = x$ for $x \in A$. Then $\bar{\phi} = \phi \circ E$ is an $\bar{\alpha}$−invariant state. From the early days of the theory it was an open problem to find the entropy of $\bar{\alpha}$ with respect to $\bar{\phi}$. This was solved by Voiculescu [**30**] in 1995.

THEOREM 11. *With the above notation* $h_{\bar{\phi}}(\bar{\alpha}) = h_\phi(\alpha)$.

The result has later been extended to more general situations. In the special case when $A = L^\infty(X, \mathcal{B}, \mu)$ with X nonatomic, and α is a freely acting ergodic μ−invariant automorphism, then the von Neumann algebra closure $(A \ltimes_\alpha \mathbb{Z})$" of $A \ltimes_\alpha \mathbb{Z}$ is a II_1−factor with A as a masa. Then the entropy of $\bar{\alpha}$ is the entropy of the restriction α to the invariant masa A.

7. Subfactors and relative entropy

In classical ergodic theory we have relative entropy $H(\mathcal{P}_1|\mathcal{P}_2)$ for two partitions. If T is a measure preserving transformation then

$$H(T) = \lim_n H\Big(\bigvee_0^n T^{-i}\mathcal{P} \Big| \bigvee_1^n T^{-i}\mathcal{P}\Big), \tag{8}$$

when $\mathcal{P}$ is a generator. In Section 1 we saw how to define relative entropy with respect to a trace between two finite dimensional C^*-subalgebras of a finite von Neumann algebra and mentioned the extension by Pimsner and Popa [**21**] to infinite dimensions when one algebra was contained in the other. Pimsner and Popa noticed that relative entropy is closely related to the Jones index $[M : N]$ of a von Neumann algebra M with respect to a subfactor N. Their most striking result is

THEOREM 12. *Let $N \subset M$ be II_1-factors with $[M : N] < \infty$ and $N' \cap M = \mathbb{C}$. Then*

$$H(M|N) = \log[M : N].$$

If $N \subset M$ are II_1-factors with finite Jones index, let $E_N : M \to N$ be the trace invariant conditional expectation of M onto N, and let e_N denote the projection in $B(L^2(M), \tau)$ such that

$$e_N x e_N = E_N(x)e_N, \forall x \in M.$$

We then get an increasing sequence of II_1- factors

$$M_{-1} = N \subset M_0 = M \subset M_1 = (M, e_N) \subset M_2 = (M_1, e_M) \subset ...,$$

where (M, e_N) denotes the von Neumann algebra generated by M and e_N, etc. The relative commutants $M_k' \cap M_l, k < l$, are finite dimensional, and we obtain an automorphism, denoted by Γ, called the *canonical shift* on

$$A_\infty = \lim_{k\to-\infty, l\to+\infty} M_k' \cap M_l,$$

where the limit is the inductive limit as C^*-algebras, such that $\Gamma(M_k' \cap M_l) = M_{k+2}' \cap M_{l+2}$. The canonical shift was introduced by Pimsner and Popa and Ocneanu, and its entropy has been studied by Choda and Hiai, see e.g. [**4**] and a little by myself. The result is a striking relationship between relative entropy, dynamical entropy and index theory reminiscent of equation (8).

THEOREM 13. *With the above notation let $R = (\bigcup_{n=1}^\infty M' \cap M_{2n})$". Then*

$$\frac{1}{2}H(R|\Gamma(R)) \le H(\Gamma) \le \mathrm{ht}(\Gamma) \le \log[M : N].$$

If furthermore $N \subset M$ has finite depth, then all inequalities above are equalities.

8. Other definitions of entropy

I want to conclude these notes with a few words on some of the many other attempts at defining noncommutative entropy. For more details see the Notes in Chapters 3,6,7, and also Chapter 5 in [**18**].

Shortly before Alain and I wrote our paper on entropy G. Emch introduced a definition [**9**]. It is too involved to describe here, but it did not catch on because it lacked a Kolmogorov-Sinai Theorem, and it had the deficiency that his definition of $H(\alpha, N)$ is not increasing in N.

When Alain and I worked on our paper, we had another possible candidate, namely the abelian entropy

$$H_{\mathrm{ab}}(\alpha) = \sup_A H(\alpha|_A),$$

where the sup is taken over all abelian subalgebras of the finite von Neumann algebra M such that $\alpha(A) = A$. Note that by monotonicity $H_{\mathrm{ab}}(\alpha) \le H(\alpha)$. We dropped it because it is not satisfactory for computations, and it is not necessarily true that $\alpha^k(A) = A$ for some $k \in \mathbb{N}$ implies that $\alpha(A) = A$. However, if we look at the examples discussed in these notes, we see that it coincides with the entropy $H(\alpha)$ for shifts on infinite tensor products whenever $H(\alpha) = 0$, in particular the free shift on $L(\mathcal{F}_\infty)$, and crossed products with abelian algebras. As far as I know it is still an open question whether $H_{\mathrm{ab}}(\alpha) = H(\alpha)$, or even if they are equal for the entropy $h_\phi(\alpha)$ defined with respect to an invariant state.

Alain and I had another candidate for entropy, namely the one presented in Section 3 just before I introduced Voiculescu's definition of topological entropy. Recall equation (7) in Section 3,

$$\mathrm{rcp}(F) = \inf\{\operatorname{rank} B : \|\psi \circ \rho(x) - x\| < \delta \quad \forall x \in F\}.$$

Voiculescu also studied the analogous expression with different norms. An important case is when ϕ is a faithful $\alpha-$invariant state, and we replace the norm above by the norm $\|x\|_\phi = \phi(x^*x)^{\frac{1}{2}}$. Then we get a dynamical entropy $h\,\mathrm{cpa}_\phi(\alpha)$, called the *completely positive approximation entropy.* It satisfies roughly the same properties as topological entropy and has been useful in several contexts.

Sauvageot and Thouvenot [**25**] introduced a definition of entropy, which coincides with the CNT-entropy h_ϕ when we are in the situation when the Kolmogorov-Sinai Theorem holds, see Theorem 5.1.5 in [**18**]. Let (A, ϕ, α) be a C^*-dynamical system and (X, μ, T) a probability space with a nonsingular measure preserving transformation T of X. Let $\beta(f) = f \circ T^{-1}$ be the corresponding automorphism of $L^\infty(X, \mu)$. By a *stationary coupling* also called *joining*, of (A, ϕ, α) with (X, μ, T) we mean an $\alpha \otimes \beta-$invariant state λ on $A \otimes L^\infty(X, \mu)$ such that $\lambda|_A = \phi$, and $\lambda|_{L^\infty(X,\mu)} = \mu$. Via the same formula as defined equation (5) we can give a definition of an entropy $h_\phi^{ST}(\alpha)$ which makes the action of T on X more explicit than it did in the definition of CNT-entropy. Just as we did before, we take the sup over all couplings to define the entropy, see Definition 5.1.1 in [**18**].

Alicki and Fannes [**1**] suggested a promising definition of entropy. Let A_0 be an $\alpha-$invariant *-subalgebra of A. They considered finite subsets $X = \{x_1, ..., x_n\} \subset A_0$ such that $\sum_1^n x_k x_k^* = 1$, called a partition of unity. They defined a completely positive map $\theta_X \colon M_n(\mathbb{C}) \to A$ by $\theta_X(e_{ij}) = x_i^* x_j$ and considered the entropy $H[\phi, X]$ of the state $\phi \circ \theta_X$. If $Y = \{y_1, ..., y_m\}$ is another partition of unity we obtain a new partition $X \circ Y$ consisting of the elements $x_k y_l$. Then put

$$h[\phi, \alpha, X] = \limsup_n \frac{1}{n} H[\phi, \alpha^{n-1} \circ \alpha^{n-2} \circ \cdots \circ X].$$

This construction defines an entropy $h[\phi, \alpha, A_0]$ with respect to A_0 by taking the sup over all X. In some cases there exist good natural choices for A_0 to give useful answers for the entropy. But the theory lacks a Kolmogorov-Sinai Theorem with the drawback that implies.

A tempting approach to noncommutative entropy is to consider the representation of the self-adjoint part A_{sa} of A as the continuous real functions on the state space $S(A)$ of A by $a \mapsto \hat{a}$ defined by $\hat{a}(\phi) = \phi(a)$, and then use the classical abelian entropy for the transformation T on $S(A)$ defined by $T(\phi) = \phi \circ \alpha^{-1}$. This does not work well, because there are too few continuous affine functions to give enough information. But some information on $h_\phi(\alpha)$ has been obtained [**24**]; for example if $h_\mu(T) = 0$ for all $T-$invariant probability measures μ on $S(A)$ with barycenter ϕ, then $h_\phi(\alpha) = 0$. In particular, if the topological entropy of T is zero, then $h_\phi(\alpha) = 0$ for all $\alpha-$invariant states ϕ. For more information on this approach consider Chapter 7 in [**18**].

Concerning other definitions of topological entropy I should mention that Hudetz, see e.g. [**11**] and Thomsen [**29**] have given definitions based on the same approach as that taken by Alicki and Fannes.

I conclude this section with a wish. Find a definition of entropy which gives useful nonzero values for entropy of highly noncommutative C^*-dynamical systems.

References

[1] Alicki, R. and Fannes, M. *Defining quantum dynamical entropy*, Lett.Math.Phys.,32, (1994),75-82.

[2] Brown, N.P. *Characterizing type I C^*-algebras via entropy*, C.R. Math.Acad.Sci. Paris (2004), 827-829.

[3] Brown, N.P. *Topological entropy in exact C^*-algebras*, Math.Ann.,314, (1999),347-367.

[4] Choda, M. and Hiai, F. *Entropy for canonical shifts. II*, Publ. Res.Inst.Math.Sci.,27,(1991), 461-489.

[5] Connes, A. *Entropie de Kolmogoroff-Sinai et mecanique statistique quantique. C.R.Acad.Sci.Paris Sr. I Math., 301 (1985),1-6.*

[6] Connes. A., Narnhofer, H., and Thirring, W. *Dynamical entropy of C^*-algebras and von Neumann algebras*, Comm. Math. Phys.,112,(1987),691-719.

[7] Connes, A. and Størmer, E. *Entropy of automorphisms in II_1 von Neumann algebras*, Acta Math.,134,(1975), 289-306.

[8] Connes, A. and Størmer, E. *A connection between the classical and the quantum mechanical entropies*, in: Operator algebras and group representations, Vol.1 (Neptun 1980), 113-123, Monogr.Stud.Math.,17,Pitman, Boston,MA (1984)

[9] Emch, G.G. *Positivity of the K-entropy on non-abelian K-flows*, Z.Wahrscheinlichkeitstheorie und Verw.Gebiete, 29,(1974),241-252.

[10] Haagerup, U. and Størmer, E. *Maximality of entropy in finite von Neumann algebras*, Invent. Math., 132, (1998),433-455.

[11] Hudetz, T. *Topological entropy for approximately approximated C^*-algebras*, J.Math.Phys.,35, (1994),4303-4333.

[12] Moriya, H., *Variational principle and the dynamical entropy of space translations*, Rev.Math.Phys.,11 (1999),1315-1328.

[13] Narnhofer, H. *Free energy and dynamical entropy of space translations*, Rep.Math.Phys.,25(1988), 345-356.

[14] Narnhofer, H. and Thirring, W. Dynamical entropy of quasifree automorphisms, Lett.Math.Phys., 14 (1987),89-96.

[15] Narnhofer, H.,Størmer, E., and Thirring, W. *C*-dynamical systems for which the tensor product formula for entropy fails*, Ergod.Th.Dynam.Sys., 15 (1995), 961-968.

[16] Neshveyev, S.,*Entropy of Bogoliubov automorphisms of CAR and CCR algebras with respect to quasi-free states*, Rev.Math.Phys.,13 (2001), 29-50.

[17] Neshveyev, S. and Størmer, E. *The variational principle for a class of asymptotically abelian C^*-algebras*, Comm.Math.Phys.,215 (2000),177-196.

[18] Neshveyev, S. and Størmer, E. *Dynamical entropy in operator algebras*, Ergebnisse Math., 50. Springer-Verlag, Berlin 2006.

[19] Ornstein, D. *Bernoulli shifts with the same entropy are isomorphic*, Advances in Math., 4 (1970),337-352.

[20] Park, Y.M. and Shin, H.H.*Dynamical entropy of space translations of CAR and CCR algebras with respect to quasi-free states*, Comm.Math.Phys., 152 (1993),497-537.

[21] Pimsner, M. and Popa, S. *Entropy and index for subfactors*, Ann.Sci.Ecole Norm.Sup.,19 (1986), 57-106.

[22] Popa, S.*Maximal injective subalgebras in factors associated with free groups*, Adv.in Math., 50 (1983),27-48.

[23] Sauvageot, J.-L.*Ergodic properties of the action of a matrix in $SL_2(2,\mathbb{Z})$ on a noncommutative 2-torus*,Preprint (1995).

[24] Sauvageot, J.-L. *The canonical channel for dynamical quantum entropy*, In: Quantum probability communications,QP-PQ, X. World Sci.Publishing, River Edge NJ (1998), 321-328.

[25] Sauvageot, J.-L. and Thouvenot, J.-P. *Une nouvelle definition de l'entropie dynamique des systems non-commutatifs*, Comm.Math.Phys., 145, (1992), 411-423.

[26] Størmer, E. *Entropy of some automorphisms of the II_1-factor of the free group in infinite number of generators*, Invent.Math.,110, (1992), 63-73.

[27] Størmer, E. *A survey of noncommutative dynamical entropy, In: Classification of nuclear C^*-algebras. Entropy in operator algebras*, Encyclopaedia Math. Sci., 126,147-198. Springer, Berlin, 2002.

[28] Størmer, E. and Voiculescu, D. *Entropy of Bogoliubov automorphisms of the canonical anticommutation relations*, Comm.Math.Phys., 133, (1990), 521-542.

[29] Thomsen, K. *Topological entropy for endomorphisms of local C^*-algebras*, Comm.Math.Phys.,164,(1994),181-193.

[30] Voiculescu, D. *Dynamical approximation entropies and topological entropy in operator algebras*, Comm.Math.Phys.,170, (1995), 249-281.

Department of Mathematics, University of Oslo, 0316 Oslo, Norway.
E-mail address: erlings@math.uio.no

Clay Mathematics Proceedings
Volume 11, 2010

Non-Compact Spectral Triples with Finite Volume

Raimar Wulkenhaar

Dedicated to Alain Connes on the occasion of his 60th birthday

Abstract. In order to extend the spectral action principle to non-compact spaces, we propose a framework for spectral triples where the algebra may be non-unital but the resolvent of the Dirac operator remains compact. We show that an example is given by the supersymmetric harmonic oscillator which, interestingly, provides two different Dirac operators. This leads to two different representations of the volume form on the Hilbert space, and only their product is the grading operator. The index of the even-to-odd part of each of these Dirac operators is 1.

We also compute the spectral action for the corresponding Connes-Lott two-point model. There is an additional harmonic oscillator potential for the Higgs field, whereas the Yang-Mills action is unchanged. The total Higgs potential shows a two-phase structure with smooth transition between them: In the spontaneously broken phase below a critical radius, all fields are massive, with the Higgs field mass slightly smaller than the NCG prediction. In the unbroken phase above the critical radius, gauge fields and fermions are massless, whereas the Higgs field remains massive.

1. Introduction

One of the greatest achievements of noncommutative geometry [**1**] is the conceptual understanding of the Standard Model of particle physics. This was not reached in one step. It took more than 15 years

- from the first appearance of the Higgs potential in noncommutative models [**2, 3**]
- via the two-sheeted universe of Connes-Lott [**4**] with its bimodule structure [**1**],
- the discovery of the real structure [**5**] (which eliminated one redundant $U(1)$ group),
- the understanding of gauge fields as inner fluctuations in an axiomatic setting [**6**] and the move from the Dixmier trace based action functional to the spectral action principle [**7**], which unifies the Standard Model with gravity,

2010 *Mathematics Subject Classification.* 58J42, 81T75, 81R60.

- the superseding of the unimodularity condition [**8**] (which eliminated the second redundant $U(1)$ group),
- to the spectacular rebirth [**9**] with the explanation [**10**] of the $\mathbb{C} \oplus \mathbb{H} \oplus M_3(\mathbb{C})$ Standard Model matrix algebra as the distinguished maximal subalgebra of $M_2(\mathbb{H}) \oplus M_4(\mathbb{C})$ compatible with a non-trivial first order condition (i.e. Majorana masses) and a six-dimensional real structure (i.e. charge conjugation).

There is one important message of this evolution: One should never be completely satisfied with one's achievements! The description given in Alain Connes' book [**1**] definitely has its beauty. The little annoyance with the redundant $U(1)$ found its solution in the real structure [**5**] which soon was realised as a key to unlocking the secrets of spin manifolds [**6**] in noncommutative geometry. This axiomatic setting initiated many examples of noncommutative manifolds and culminated in the recent spectral characterisation of manifolds [**11**].

Let me give a wish list for further improvements—not as a criticism of the model, but rather as a possible source of insight.

(1) Quantisation. The outcome of the spectral action principle is a classical action functional valid at a distinguished (grand unification) scale. It is connected to the scale realised in a particle accelerator by the renormalisation group flow. This flow can be computed by rules from perturbative quantum field theory. The input is not directly the spectral action, but a gauge-fixed version of it which involves Faddeev-Popov ghosts. It is highly desirable to include these ghosts in the spectral action, because in this way unitary invariance is realised as cohomology of the BRS complex. We may speculate that the BRS cohomology of the spectral action is deeply connected to the wealth of noncommutative cohomology theories. As a starting point one might use results of Perrot [**12**], who identifies the BRS coboundary as the de Rham differential in the loop space $C^\infty(S^1,\mathcal{U}(\mathcal{A}))$ and connects the chiral anomaly with the local index formula [**13**].

(2) Big desert. The present form of the spectral action is based on the big desert hypothesis which asserts that, apart from the Higgs boson, all particles relevant at the grand unification scale are already discovered. The minor mismatch between observed and predicted $U(1)$ coupling constant (see Figure 1 in [**9**]) might suggest some new physics in the desert. Candidates include supersymmetry and dark matter, but also noncommutativity of space itself could alter the slope of the running $U(1)$ coupling.

   The latter question concerning the renormalisation group flow of field theories on noncommutative geometries was intensely studied in the last decade. After unexpected difficulties with UV/IR-mixing, we established perturbative renormalisability of scalar field theories on Moyal-deformed Euclidean space [**14, 15**]. The key is a deformation also of the differential calculus, namely from the Laplace operator to the harmonic oscillator Schrödinger operator. It turned out indeed that the combined Moyal-harmonic oscillator deformation removes the Landau ghost of the commutative scalar model [**16**] by altering the slope of the running coupling constant [**17**]. Since the $U(1)$-part of the Standard Model has the same Landau ghost problem, we might expect that, once the Standard Model has been grounded in an appropriate noncommutative geometry, the three running couplings of Figure 1 in [**9**] will eventually intersect in a single point.

The first step in this programme is to construct a spectral triple with its canonically associated spectral action for the combined Moyal-harmonic oscillator deformation. The present paper achieves an intermediate goal: We construct and investigate a *commutative* harmonic oscillator spectral triple. Its Moyal isospectral deformation will be treated in [**18**], building on ideas developed in [**19**]. The main obstacle was to identify a Dirac operator whose square is the harmonic oscillator Hamiltonian of [**14**]. The solution which we give in this paper is deeply connected to supersymmetric quantum mechanics [**20**], in particular to Witten's approach to Morse theory [**21**]. It would be interesting to reformulate Witten's results in noncommutative index theory using the spectral triple we suggest.

(3) Time. The spectral action relies on compact Euclidean geometry. For the Standard Model one typically chooses the manifold $S^3 \times S^1$, where S^3 is for "space" and S^1 for "temperature", not "time". Although the universe is filled with thermal background radiation, it is desirable to allow for a genuine time evolution of the spectral geometry. In fact, noncommutative von Neumann algebras carry their own time evolution through the modular automorphism group, and it has been argued [**22**] that this is the source of the physical time flow. So far the modular automorphisms seem disconnected from the spectral action. The most ambitious project to reconcile time development and spectral geometry within generally covariant quantum field theory was initiated by Paschke and Verch [**23**].

(4) Compactness. As mentioned above, the spectral action presumes compactness, namely, compactness of the resolvent of the Dirac operator. The example we study in this paper shows that compactness of the resolvent does not imply spatial compactness. It is eventually a matter of experiment to determine the type of compactness of the universe.

The paper is organised as follows: We propose in Section 2 a definition of non-unital spectral triples, but with compactness of the resolvent of the Dirac operator. We show in Section 3 that the supersymmetric harmonic oscillator is an example of such a spectral triple: In Section 3.1 we introduce the supercharges in a slightly generalised framework and briefly discuss their cohomology. The supercharges give rise to two distinct Dirac operators. In Section 3.2 we identify for the harmonic oscillator the algebra and the smooth part of the Hilbert space. In Section 3.3 and Appendix A we compute the dimension spectrum. The novel orientability structure is studied in Section 3.4, and Section 3.5 discusses the index formula for the Dirac operators. The spectral action is computed in Section 4 and Appendix B. In the final Section 5 we study the solution of the equations of motion.

2. Non-compact spectral triples

Motivated by the spectral characterisation of manifolds [**11**], we propose here a definition of spectral triples which does not require the algebra to be unital. There are several proposals in the literature for a non-compact generalisation of spectral triples; see [**24**] and references therein. To include $\mathbb{R}^d$ with its standard Dirac operator, these proposals relax the compactness of the resolvent of $\mathcal{D}$ to the requirement that $\pi(a)(\mathcal{D}+\mathrm{i})^{-1}$ is compact for all $a \in \mathcal{A}$. However, compactness of the resolvent (or similar regularisation [**25**]) is essential for a well-defined spectral action. Moreover, the usual Dirac operator on $\mathbb{R}^d$ is not suited for an index formula [**26**]. We therefore keep compactness of the resolvent (and thus exclude standard $\mathbb{R}^d$), but to achieve this in the non-compact situation we are forced to give up (at least in our example)

(1) the universality of dimensions,
(2) the connection between volume form and $\mathbb{Z}_2$-grading.

We give some comments after the definition. To simplify the presentation we require the algebra to be commutative; the noncommutative generalisation involves the real structure J.

Definition 1. *A (possibly non-compact) commutative spectral triple with finite volume $(\mathcal{A},\mathcal{H},\mathcal{D})$ is given by a (possibly non-unital) commutative and involutive algebra $\mathcal{A}$ represented on a Hilbert space $\mathcal{H}$ and a selfadjoint unbounded operator $\mathcal{D}$ in $\mathcal{H}$ with compact resolvent fulfilling the conditions 1-5 below.*

(1) `Regularity and dimension spectrum.` *For any $a \in \mathcal{A}$, both a and $[\mathcal{D},a]$ belong to $\bigcap_{n=1}^{\infty}\mathrm{dom}(\delta^n)$, where $\delta T := [\langle\mathcal{D}\rangle,T]$ and $\langle\mathcal{D}\rangle := (\mathcal{D}^2+1)^{\frac{1}{2}}$.*
For any element ϕ of the algebra $\Psi_0(\mathcal{A})$ generated by $\delta^m a$ and $\delta^m[\mathcal{D},a]$, with $a \in \mathcal{A}$, the function $\zeta_\phi(z) := \mathrm{Tr}(\phi\langle\mathcal{D}\rangle^{-z})$ extends holomorphically to $\mathbb{C}\setminus\mathrm{Sd}$ for some discrete set $\mathrm{Sd}\subset\mathbb{C}$ (the dimension spectrum), and all poles of ζ_ϕ at $z\in\mathrm{Sd}$ are simple.

(2) `Metric dimension.` *The maximum $d := \max\{r \in \mathbb{R}\cap\mathrm{Sd}\}$ belongs to $\mathbb{N}$. The noncommutative integral $⨍ a\langle\mathcal{D}\rangle^{-d}$ is finite for any $a\in\mathcal{A}$ and positive for positive elements of $\mathcal{A}$.*

(3) `Orientability.` *For the preferred unitisation*
$$\mathcal{B} := \{b\in\mathcal{A}'' \;:\; b,[\mathcal{D},b]\in\bigcap_{n\in\mathbb{N}}\mathrm{dom}(\delta^m)\}\;,$$
there is a Hochschild d-cycle $\boldsymbol{c}\in Z_d(\mathcal{B},\mathcal{B})$, i.e. a finite sum of terms $b_0\otimes b_1\otimes\cdots\otimes b_d$. Its representation $\boldsymbol{\gamma} := \pi_{\mathcal{D}}(\boldsymbol{c})$, with $\pi_{\mathcal{D}}(b_0\otimes b_1\otimes\cdots\otimes b_d) := b_0[\mathcal{D},b_1]\cdots[\mathcal{D},b_d]$, satisfies $\boldsymbol{\gamma}^2=1$ and $\boldsymbol{\gamma}^*=\boldsymbol{\gamma}$. Additionally, $\boldsymbol{\gamma}$ defines the volume form on $\mathcal{A}$, i.e.
$$\phi_{\boldsymbol{\gamma}}(a_0,\dots,a_d) := ⨍\big(\boldsymbol{\gamma}a_0[\mathcal{D},a_1]\cdots[\mathcal{D},a_d]\langle\mathcal{D}\rangle^{-d}\big)$$
provides a non-vanishing Hochschild d-cocycle $\phi_{\boldsymbol{\gamma}}$ on $\mathcal{A}$.

(4) `First order.` *$[[\mathcal{D},b],b'] = 0$ for all $b,b'\in\mathcal{B}$.*

(5) `Finiteness.` *The subspace $\mathcal{H}_\infty := \bigcap_{k=0}^{\infty}\mathrm{dom}(\mathcal{D}^k)\subset\mathcal{H}$ is a finitely generated projective $\mathcal{A}$-module $e\mathcal{A}^n$, for some $n\in\mathbb{N}$ and some projector $e=e^2=e^*\in M_n(\mathcal{B})$. The composition of the noncommutative integral with the induced Hermitian structure $(\;|\;):\mathcal{H}_\infty\times\mathcal{H}_\infty\to\mathcal{A}$ coincides with the scalar product $\langle\;,\;\rangle$ on $\mathcal{H}_\infty$,*
$$\langle\xi,\eta\rangle = ⨍\Big((\xi|\eta)\,\langle\mathcal{D}\rangle^{-d}\Big)\;,\qquad \xi,\eta\in\mathcal{H}_\infty\;.$$

The dimension spectrum was introduced by Connes and Moscovici [**13**] precisely to describe by a local formula the lower-dimensional pieces in the Chern character that are ignored by the top-dimensional Hochschild cohomology class. The local index formula was generalised in [**27**] to a larger class of examples. We are interested in a similar situation. For non-unital algebras we may have the characteristic values of the resolvent of $\mathcal{D}$ run as $\mathcal{O}(n^{-\frac{1}{p}})$ for p greater than the metric dimension d. The dimension spectrum is the right tool to deal with this case.

It would be interesting to know whether Definition 1, despite its differences from Connes' original definition [**11**], allows reconstruction of a manifold structure on the spectrum $X = \mathrm{Spec}(A)$ of the norm closure A of $\mathcal{A}$. At first sight, the construction of candidates for local charts only uses the measure λ on X defined by the noncommutative integral $\lambda(f) = ⨍ f\langle\mathcal{D}\rangle^{-d}$ for $f \in A = C(X)$ and the fact that the Hilbert space $\mathcal{H}$ is precisely the L^2-closure of $\mathcal{H}_\infty$ with respect to λ. The details of how $⨍ f\langle\mathcal{D}\rangle^{-d}$ is constructed, whether as a state-independent Dixmier trace or as a residue in the dimension spectrum, do not seem to enter. In particular, Lemma 2.1 of [**11**] holds: if $1 \in \mathcal{A}$, then $\mathcal{B} = \mathcal{A}$ (in the notation of Definition 1), so that conditions (3),(4),(5) are the same as in [**11**], with the sole exception that $\boldsymbol{\gamma}$ is not necessarily the $\mathbb{Z}_2$-grading for even d or $\boldsymbol{\gamma} = 1$ for odd d. However, this was only used for uniqueness of the noncommutative integral, which we achieve alternatively from the dimension spectrum. But [**11**, §9] makes heavy use of the asymptotics of the eigenvalues of $\langle\mathcal{D}\rangle^{-1}$ to prove injectivity of the local charts; we do not know how to achieve this from the dimension spectrum.

3. A spectral triple for the harmonic oscillator

3.1. Supersymmetric quantum mechanics. Supersymmetric quantum mechanics provides an elegant approach to exactly solvable quantum-mechanical models [**20**] and is also a powerful tool in mathematics [**21**]. Our notation is a compromise between [**20**] and [**21**].

Let X be a d-dimensional smooth manifold with trivial cotangent bundle and ∂_μ, for $\mu = 1, \dots, d$, be the basis of the tangent space T_xX induced by the coordinate functions. On the Hilbert space $L^2(X)$ we consider the unbounded operators

$$a_\mu = e^{-\omega h}\partial_\mu e^{\omega h} = \partial_\mu + W_\mu\ , \qquad a^\dagger_\mu = -e^{\omega h}\partial_\mu e^{-\omega h} = -\partial_\mu + W_\mu\ , \tag{1}$$

where h is some real-valued function on X, the Morse function [**21**], and $W_\mu(x) = \omega(\partial_\mu h)(x)$. It is convenient to keep the frequency ω separate from h. The resulting commutation relations are

$$[a_\mu, a_\nu] = [a^\dagger_\mu, a^\dagger_\nu] = 0\ , \qquad\qquad [a_\mu, a^\dagger_\nu] = 2\omega\partial_\mu\partial_\nu h\ . \tag{2}$$

We define fermionic ladder operators $b^\mu, b^{\dagger\mu}$ which satisfy the anticommutation relations

$$\{b^\mu, b^\nu\} = 0\ , \qquad \{b^{\dagger\mu}, b^{\dagger\nu}\} = 0\ , \qquad \{b^\mu, b^{\dagger\nu}\} = \delta^{\mu\nu}\ . \tag{3}$$

We also let all mixed commutators vanish, $[a^{(\dagger)}_\mu, b^{(\dagger)\nu}] = 0$. We introduce the supercharges $\mathfrak{Q}, \mathfrak{Q}^\dagger$ by

$$\mathfrak{Q} := a_\mu \otimes b^{\dagger\mu}\ , \qquad \mathfrak{Q}^\dagger := a^\dagger_\mu \otimes b^\mu\ . \tag{4}$$

Unless otherwise stated, we use Einstein's summation convention, i.e. summation over a pair of upper/lower Greek indices from 1 to d is understood. The supercharges satisfy

$$\{\mathfrak{Q},\mathfrak{Q}\} = \{\mathfrak{Q}^\dagger,\mathfrak{Q}^\dagger\} = 0\ , \quad \{\mathfrak{Q},\mathfrak{Q}^\dagger\} =: \mathfrak{H}\ , \qquad [\mathfrak{Q},\mathfrak{H}] = [\mathfrak{Q}^\dagger,\mathfrak{H}] = 0\ . \tag{5}$$

The Hamiltonian $\mathfrak{H}$ introduced by the anticommutator reads explicitly (index raising by $\delta^{\mu\nu}$)

$$\begin{aligned}\mathfrak{H} &= \frac{1}{2}\delta^{\mu\nu}\{a_\mu, a_\nu^\dagger\}\otimes 1 + \frac{1}{2}[a_\mu,a_\nu^\dagger]\otimes[b^{\dagger\mu},b^\nu] \\ &= \big(-\partial_\mu\partial^\mu + \omega^2(\partial_\mu h)(\partial^\mu h)\big)\otimes 1 + \omega(\partial_\mu\partial_\nu h)\otimes[b^{\dagger\mu},b^\nu]\ .\end{aligned} \tag{6}$$

The supercharges give rise to *two* anticommuting Dirac operators

$$\mathcal{D}_1 = \mathfrak{Q}+\mathfrak{Q}^\dagger\ , \qquad\qquad \mathcal{D}_2 = \mathrm{i}\mathfrak{Q}-\mathrm{i}\mathfrak{Q}^\dagger\ , \tag{7}$$

$$\mathcal{D}_i^2 = \mathfrak{H} \quad \text{for } i=1,2\ , \qquad\qquad \mathcal{D}_1\mathcal{D}_2+\mathcal{D}_2\mathcal{D}_1 = 0\ . \tag{8}$$

We let $|0\rangle_f$ be the fermionic vacuum with $b^\mu|0\rangle_f = 0$. By repeated application of $b^{\dagger\mu}$ one constructs out of $|0\rangle_f$ the 2^d-dimensional fermionic Hilbert space $\bigwedge(\mathbb{C}^d)$ in which we label the standard orthonormal basis as follows:

$$|s_1,\dots,s_d\rangle_f = (b^{\dagger 1})^{s_1}\dots(b^{\dagger d})^{s_d}|0\rangle_f\ , \qquad s_\mu\in\{0,1\}\ . \tag{9}$$

The fermionic number operator is $N_f = b_\mu^\dagger b^\mu$, with

$$N_f|s_1,\dots,s_d\rangle_f = (s_1+\dots+s_d)|s_1,\dots,s_d\rangle_f\ .$$

The fermionic Hilbert space is $\mathbb{N}$-graded by $\bigwedge(\mathbb{C}^d) = \bigoplus_{p=0}^d \Lambda^p(\mathbb{C}^d)$ with $\dim(\Lambda^p(\mathbb{C}^d)) = \binom{d}{p}$. Accordingly, the total Hilbert space $\mathcal{H} = L^2(X)\otimes\bigwedge(\mathbb{C}^d)$ is graded by the fermion number, $\mathcal{H} = \bigoplus_{p=0}^d \mathcal{H}_p$. Note that $\mathfrak{Q}:\mathcal{H}_p\to\mathcal{H}_{p+1}$ and $\mathfrak{Q}^\dagger:\mathcal{H}_p\to\mathcal{H}_{p-1}$. The induced $\mathbb{Z}_2$-grading operator is

$$\Gamma = (-1)^{N_f}\ , \qquad \Gamma^2 = 1\ ,\ \Gamma=\Gamma^*\ , \qquad \Gamma\mathcal{D}_i+\mathcal{D}_i\Gamma = 0\ . \tag{10}$$

Let $B_p(\omega)$ be the dimension of the p-th cohomology group of $\mathfrak{Q}$, i.e. the number of linearly independent $\psi_p\in\ker\mathfrak{Q}\cap\mathcal{H}_p$ that cannot be written as $\psi_p = \mathfrak{Q}\eta_{p-1}$ for some $\eta\in\mathcal{H}_{p-1}$. According to Witten [**21**], $B_p(\omega)$ coincides with the Betti number B_p and is deeply connected with the Morse index M_p for the function h: Let x_α be a critical point of h, i.e. $(\partial_\mu h)(x) = 0$. If $\partial_\mu\partial_\nu h$ is regular at each of these critical points, then M_p is the number of critical points at which $\partial_\mu\partial_\nu h$ has p negative eigenvalues. The weak Morse inequalities $M_p\geq B_p$ follow from the eigenvalue problem for $\mathfrak{H}$ in the limit of large ω.

By Hodge theory, which relies on the Hilbert space structure, every generator of the p-th cohomology group of $\mathfrak{Q}$ has a unique representative ψ which is also $\mathfrak{Q}^\dagger$-exact (and thus belongs to $\ker\mathfrak{H}$). Since the $b^\mu, b^{\dagger\mu}$ generate linearly independent subspaces, this means (no summation over $\bar\mu,\bar\nu$)

$$(a_{\bar\mu}\otimes b^{\dagger\bar\mu})\psi = 0 \quad\text{and}\quad (a_{\bar\nu}^\dagger\otimes b^{\bar\nu})\psi = 0 \quad \text{for all } \bar\mu,\bar\nu = 1,\dots,d\ . \tag{11}$$

The only candidates are (up to a multiplicative constant)

$$\psi_0 = e^{-\omega h}|0\rangle_f \quad\text{and}\quad \psi_d = e^{\omega h}b^{\dagger 1}\cdots b^{\dagger d}|0\rangle_f\ . \tag{12}$$

For compact manifolds, where both $e^{\pm\omega h}$ are integrable, this yields $B_0 = 1$ and $B_d = 1$ as the only non-vanishing Betti numbers. In the non-compact case one should choose $e^{-\omega h}$ integrable, so that $e^{\omega h}$ is not integrable, and hence $B_p = \delta_{p0}$.

Of course, this behaviour is due to the assumption of a trivial cotangent bundle. For more interesting topology one should define the smooth subspace of the Hilbert space as a finitely generated projective module.

3.2. The harmonic oscillator. In the following we propose a spectral triple in the sense of Definition 1 with objects related to the harmonic oscillator. We will check the axioms, but no attempt will be made to reconstruct a manifold.

The harmonic oscillator is obtained from the Morse function $h = \frac{\|x\|^2}{2} = \frac{1}{2}\delta^{\mu\nu}x_\mu x_\nu$ on the manifold $\mathbb{R}^d$. This leads to the relation

$$[a_\mu, a_\nu^\dagger] = 2\omega\delta_{\mu\nu}\ , \tag{13}$$

which in turn permits a complete reconstruction of the eigenfunctions by repeated application of $a_\mu^\dagger, b^{\dagger\nu}$ to the ground state $\psi_0 = |0\rangle_b \otimes |0\rangle_f \in \ker \mathfrak{H}$, with $|0\rangle_b = (\frac{\omega}{\pi})^{\frac{d}{4}} e^{-\frac{\omega}{2}\|x\|^2}$. Defining

$$|n_1, \dots, n_d\rangle_b = \frac{1}{\sqrt{n_1! \dots n_d! (2\omega)^{n_1+\dots+n_d}}} (a_1^\dagger)^{n_1} \cdots (a_d^\dagger)^{n_d} |0\rangle_b\ , \qquad n_\mu \in \mathbb{N}\ , \tag{14}$$

the tensor products $|n_1, \dots, n_d\rangle_b \otimes |s_1, \dots, s_d\rangle_f$ of (14) with (9) form an orthonormal basis of the Hilbert space $\mathcal{H} = \ell^2(\mathbb{N}^d) \otimes \mathbb{C}^{2^d} \simeq L^2(\mathbb{R}^d) \otimes \bigwedge(\mathbb{C}^d)$.

There are two ways of viewing the Hamiltonian (6). In the $L^2(\mathbb{R}^d)$-representation, we have

$$\mathfrak{H} = H \otimes 1 + \omega \otimes \Sigma\ , \qquad H = -\partial_\mu \partial^\mu + \omega^2 x_\mu x^\mu\ , \qquad \Sigma = [b_\mu^\dagger, b_\mu]\ , \tag{15}$$

i.e. the total Hamiltonian is the sum of the harmonic oscillator Hamiltonian and ω times the spin matrix Σ. This representation will be useful when considering the algebra $\mathcal{A}$ later on which is also realised in the $L^2(\mathbb{R}^d)$-representation. In the $\ell^2(\mathbb{N}^d)$-representation, we have

$$\mathcal{D}_1^2 = \mathcal{D}_2^2 = \mathfrak{H} = a_\mu^\dagger a^\mu \otimes 1 + 2\omega \otimes b_\mu^\dagger b^\mu = 2\omega(N_b + N_f)\ , \tag{16}$$

which is up to a factor of 2ω the supersymmetric number operator:

$$\begin{aligned} &\mathcal{D}_i^2(|n_1, \dots, n_d\rangle_b \otimes |s_1, \dots, s_d\rangle_f) \\ &\quad = \Big(2\omega \sum_{\mu=1}^{d} (n_\mu + s_\mu)\Big)(|n_1, \dots, n_d\rangle_b \otimes |s_1, \dots, s_d\rangle_f)\ . \end{aligned} \tag{17}$$

In particular, the kernel of $\mathcal{D}_i$ is one-dimensional, and the resolvent of $\mathcal{D}_i$ is compact. To deal with the kernel, we introduce

$$\langle \mathcal{D} \rangle := (\mathcal{D}_1^2 + 1)^{\frac{1}{2}} = (\mathcal{D}_2^2 + 1)^{\frac{1}{2}}\ , \qquad \delta T := [\langle \mathcal{D} \rangle, T] \quad \text{for } T \in \mathcal{B}(\mathcal{H})\ . \tag{18}$$

Counting the number of eigenvalues $\leq N$ one finds that $\langle \mathcal{D} \rangle^{-1}$ is a noncommutative infinitesimal of order $2d$, and $\langle \mathcal{D} \rangle^{-p}$ is trace-class for $p > 2d$. Formula (17) also shows that

$$\mathcal{H}_\infty := \bigcap_{m \geq 0} \mathrm{dom}(\mathcal{D}^n) = \mathcal{S}(\mathbb{N}^d) \otimes \bigwedge(\mathbb{C}^d) \simeq \mathcal{S}(\mathbb{R}^d) \otimes \bigwedge(\mathbb{C}^d) \simeq \big(\mathcal{S}(\mathbb{R}^d)\big)^{2^d}\ , \tag{19}$$

which is required to be a finitely generated projective module over the algebra of the spectral triple. We are interested here in the commutative case, so that we are led to consider the algebra

$$\mathcal{A} = \mathcal{S}(\mathbb{R}^d) \tag{20}$$

of Schwartz class functions with standard commutative product. The Hermitian structure is pointwise the scalar product in $\bigwedge(\mathbb{C}^d)$, i.e. $(\xi|\eta) = \sum_{i=1}^{2^d} \xi_i^* \eta_i$ for $\xi = (\xi_1, \dots, \xi_{2^d}), \eta = (\eta_1, \dots, \eta_{2^d}) \in \mathcal{H}_\infty = \big(\mathcal{S}(\mathbb{R}^d)\big)^{2^d}$.

As usual, we represent the algebra $\mathcal{A}$ on $\mathcal{H}$ by pointwise multiplication in $L^2(\mathbb{R}^d)$:

$$f(\psi \otimes \rho) := (f\psi) \otimes \rho \qquad \text{for } f \in \mathcal{A}\,,\ \psi \in L^2(\mathbb{R}^d)\,,\ \rho \in \bigwedge(\mathbb{C}^d)\,. \tag{21}$$

The action of $\mathcal{A}$ commutes with $b^\mu, b^{\dagger\mu}$ so that we obtain

$$[\mathcal{D}_1, f] = \partial_\mu f \otimes (b^{\dagger\mu} - b^\mu)\,, \qquad\qquad [\mathcal{D}_2, f] = \partial_\mu f \otimes (\mathrm{i}b^{\dagger\mu} + \mathrm{i}b^\mu)\,. \tag{22}$$

In particular, the first-order condition is satisfied. For $f \in \mathcal{A}$, the expansion coefficients $\langle n_1, \dots, n_d | f | n_1', \dots, n_d' \rangle$ are Schwartz sequences in n_μ, n_μ'. Therefore, f and $[\mathcal{D}_i, f]$ belong for any $m \in \mathbb{N}$ to the domain of δ^m.

We show in joint work with V. Gayral [**18**] (which supersedes [**19**]), that the Moyal-deformation of $\mathcal{S}(\mathbb{R}^d)$ together with the same Dirac operator and Hilbert space forms a noncommutative spectral triple in the sense of Definition 1, i.e. an isospectral deformation.

3.3. Dimension spectrum. In this subsection we take for $\mathcal{D}$ either of $\mathcal{D}_1$ or $\mathcal{D}_2$. We consider the algebra $\Psi_0(\mathcal{A})$ generated by $\delta^m f$ and $\delta^m[\mathcal{D}, f]$. As $\langle \mathcal{D} \rangle^{-z}$ is trace-class for $\mathrm{Re}(z) > 2d$, the ζ-function $\zeta_\phi(z) := \mathrm{Tr}(\phi \langle \mathcal{D} \rangle^{-z})$ exists for such $z \in \mathbb{C}$ and $\phi \in \Psi_0(\mathcal{A})$ and can possibly be extended to a meromorphic function on $\mathbb{C}$. The following theorem identifies the poles and the structure of the residues:

THEOREM 2. *The spectral triple $(\mathcal{A}, \mathcal{H}, \mathcal{D})$ has dimension spectrum* $\mathrm{Sd} = d - \mathbb{N}$ *and hence metric dimension d. All poles of ζ_ϕ at $z \in \mathrm{Sd}$ are simple with local residues, i.e. for $\phi = \delta^{n_1} f_1 \cdots \delta^{n_v} f_v$, any residue $\mathrm{res}_{z \in \mathrm{Sd}} \zeta_\phi(z)$ is a finite sum of $\displaystyle\int_{\mathbb{R}^d} dx\ x^{\alpha_0} (\partial^{\alpha_1} f_1) \cdots (\partial^{\alpha_v} f_v)$, where α_i are multi-indices. The analogous result holds when f_i in ϕ is replaced by $[\mathcal{D}, f_i]$.*

This theorem is the central result of this paper. We give the rather long proof in Appendix A.

A special case of the proof of Theorem 2 is the computation of the Dixmier trace:

PROPOSITION 3. $\displaystyle ⨍ f \langle \mathcal{D} \rangle^{-d} = \frac{1}{(4\pi)^{\frac{d}{2}} \Gamma(\frac{d+2}{2})} \int_{\mathbb{R}^d} dx\ f(x)$ *for any $f \in \mathcal{A}$.*

Proof. As the dimension spectrum is simple, the Dixmier trace can be computed as a residue [**28**], is independent of the state ω, and defines unambiguously the noncommutative integral:

$$⨍ f \langle \mathcal{D} \rangle^{-d} = \mathrm{res}_{s=1} \mathrm{Tr}(f \langle \mathcal{D} \rangle^{-sd})\,. \tag{23}$$

Taking $v = 1$ and $n_1 = 0$ in (82) and inserting $\det Q$ and Q^{-1} from (85) and (86) as well as (80), we have

$$⨍ f \langle \mathcal{D} \rangle^{-d} = \mathrm{res}_{s=1} \bigg(\frac{1}{\Gamma(\frac{sd}{2})} \int_0^\infty dt_0\ t_0^{\frac{sd}{2}-1}\, e^{-t_0} \int_{\mathbb{R}^d} \frac{dp}{(2\pi)^d}\ \hat{f}(p) \frac{e^{-\frac{p^2}{\omega \tanh(\omega t_0)}}}{\tanh^d(\omega t_0)} \bigg)\,. \tag{24}$$

We write $\hat{f}(p) = \hat{f}(0) + p_\mu \frac{\partial \hat{f}}{\partial p_\mu}(0) + p_\mu p_\nu \int_0^1 d\lambda \, (1-\lambda) \frac{\partial^2 \hat{f}}{\partial p_\mu \partial p_\nu}(\lambda p_\mu)$ and get

$$
\begin{aligned}
(25) \qquad & \frac{1}{\Gamma(\frac{sd}{2})} \int_0^\infty dt_0 \, t_0^{\frac{sd}{2}-1} e^{-t_0} \int_{\mathbb{R}^d} \frac{dp}{(2\pi)^d} \, \hat{f}(0) \frac{e^{-\frac{p^2}{\omega \tanh(\omega t_0)}}}{\tanh^d(\omega t_0)} \\
&= \frac{\hat{f}(0)}{(4\pi)^{\frac{d}{2}} \Gamma(\frac{sd}{2})} \int_0^\infty dt_0 \, t_0^{\frac{(s-1)d}{2}-1} \, e^{-t_0} \underbrace{\Big(\frac{\omega t_0}{\tanh(\omega t_0)}\Big)^{\frac{d}{2}}}_{g(t_0)} \\
&= \frac{\hat{f}(0)}{(4\pi)^{\frac{d}{2}}} \frac{\Gamma(\frac{(s-1)d}{2})}{\Gamma(\frac{sd}{2})} + \frac{\hat{f}(0)}{(4\pi)^{\frac{d}{2}} \Gamma(\frac{sd}{2})} \int_0^\infty dt_0 \, t_0^{\frac{(s-1)d}{2}} \, e^{-t_0} \int_0^1 d\lambda \, g'(\lambda t_0) \,.
\end{aligned}
$$

As $|g'(y)| \leq \frac{d}{2} y^{\frac{d}{2}-1}$ for all $y \in \mathbb{R}_+$, we have

$$
(26) \qquad \Big| \int_0^\infty dt_0 \, t_0^{\frac{(s-1)d}{2}} \, e^{-t_0} \int_0^1 d\lambda \, g'(\lambda t_0) \Big| \leq \int_0^\infty dt_0 \, t_0^{\frac{s}{2}-1} \, e^{-t_0} = \Gamma(\tfrac{s}{2}) \,,
$$

which is regular for $s = 1$. The first-order term $p_\mu \frac{\partial \hat{f}}{\partial p_\mu}(0)$ does not contribute as an odd function in p. In the remainder, $\int_0^1 d\lambda (1-\lambda) \frac{\partial^2 \hat{f}}{\partial p_\mu \partial_\nu}(\lambda p_\mu)$ is bounded, and

$$
(27) \qquad \int \frac{dp}{(2\pi)^d} \, p_\mu p_\nu \frac{e^{-\frac{p^2}{\omega \tanh(\omega t_0)}}}{\tanh^d(\omega t_0)} = \frac{\omega^2}{2} \frac{\delta_{\mu\nu}}{(4\pi)^{\frac{d}{2}}} \Big(\frac{\omega}{\tanh(\omega t_0)}\Big)^{\frac{d}{2}-1}
$$

provides another factor of t_0 so that the remainder does not contribute to the residue at $s = 1$. The assertion follows from $\hat{f}(0) = \int_{\mathbb{R}^d} dx \, f(x)$. □

Therefore, with the normalisation $\langle \xi, \eta \rangle = \frac{1}{(4\pi)^{\frac{d}{2}} \Gamma(\frac{d+2}{2})} \int_{\mathbb{R}^d} dx \, (\xi | \eta)$ of the scalar product in $\mathcal{H}$, the finiteness condition is satisfied.

It remains to discuss the orientability, for which we need the algebra

$$
(28) \qquad \mathcal{B} := \{ b \in \mathcal{A}'' \; : \quad b, [\mathcal{D}, b] \in \bigcap_{m \in \mathbb{N}} \mathrm{dom}(\delta^m) \} \,.
$$

Clearly, $\mathcal{B}$ is unital and commutative; we now show that it contains the plane waves $u_\mu = e^{ix_\mu}$.

LEMMA 4. $u_\mu = e^{ix_\mu} \in \mathcal{B}$.

Proof. From (73), which applies without change to $T = u_\mu$, we get (no summation over μ)

$$
(29) \qquad \delta^n u_\mu = \frac{(-\mathrm{i})^n}{\pi^n} \int_0^\infty \prod_{i=1}^n \frac{d\lambda_i \sqrt{\lambda_i}}{\langle \mathcal{D} \rangle^2 + \lambda_i} \{ \underbrace{\partial_\mu, \ldots, \{\partial_\mu}_{n \text{ derivatives}}, e^{\mathrm{i}x^\mu} \} \ldots \} \prod_{j=1}^n \frac{1}{\langle \mathcal{D} \rangle^2 + \lambda_j} \,.
$$

We have

$$
\begin{aligned}
(30) \qquad & \Big(\prod_{i=1}^n \frac{1}{A + \lambda_i} \Big) B \\
&= \Big(\sum_{S \in \{1,2,\ldots,n\}} (-1)^{|S|} \Big(\prod_{i \in S} \frac{1}{A + \lambda_i} \Big) (\mathrm{ad}(A))^{|S|}(B) \Big) \Big(\prod_{j=1}^n \frac{1}{A + \lambda_j} \Big) \,,
\end{aligned}
$$

where the sum runs over all subsets $S \subset \{1, 2, \dots, n\}$ including the empty set. After relabelling of the $|S|$ elements of S, which gives a factor $\binom{n}{|S|}$, we have

$$\delta^n(u_\mu) = \frac{(-\mathrm{i})^n}{\pi^n} \sum_{k=0}^{n} \binom{n}{k} \mathrm{i}^k \times \int_0^\infty \prod_{i=1}^{k} \frac{1}{\langle\mathcal{D}\rangle^2 + \lambda_i} \{ \underbrace{\partial_\mu, \dots, \{\partial_\mu}_{n+k \text{ derivatives}}, e^{\mathrm{i}x^\mu}\} \dots \} \prod_{j=1}^{n} \frac{d\lambda_j \sqrt{\lambda_j}}{(\langle\mathcal{D}\rangle^2 + \lambda_j)^2} . \tag{31}$$

The anticommutators can be arranged as a finite sum with $r \le n$ derivatives on the right and $l \le k$ derivatives on the left of $e^{\mathrm{i}x^\mu}$. Each such term is estimated by

$$\begin{aligned} &\Big\| \int_0^\infty \prod_{i=1}^{k} \frac{1}{\langle\mathcal{D}\rangle^2 + \lambda_i} (\partial_\mu)^l e^{\mathrm{i}x^\mu} (\partial_\mu)^r \langle\mathcal{D}\rangle^{-n} \prod_{j=1}^{n} \frac{d\lambda_j \sqrt{\lambda_j}\langle\mathcal{D}\rangle}{(\langle\mathcal{D}\rangle^2 + \lambda_j)^2} \Big\| \\ &\le \big\| \langle\mathcal{D}\rangle^{-2k} (\partial_\mu)^l \big\| \, \big\| (\partial_\mu)^r \langle\mathcal{D}\rangle^{-n} \big\| \Big\| \int_0^\infty \frac{d\lambda \sqrt{\lambda}\langle\mathcal{D}\rangle}{(\langle\mathcal{D}\rangle^2 + \lambda)^2} \Big\|^n , \end{aligned} \tag{32}$$

which is bounded because the integral in the second line evaluates to $\frac{\pi}{2}$. ☐

By the same arguments one shows that the algebra $C_b^\infty(\mathbb{R}^d)$ of smooth bounded functions with all derivatives bounded is contained in $\mathcal{B}$, and it is plausible that actually $\mathcal{B} = C_b^\infty(\mathbb{R}^d)$.

3.4. Orientability. Here the distinction between $\mathcal{D}_1$ and $\mathcal{D}_2$ is crucial again. It follows from the standard example of the compact case that

$$\boldsymbol{c} = \sum_{\sigma \in S_d} \epsilon(\sigma) \frac{\mathrm{i}^{\frac{d(d-1)}{2}}}{d!} (u_1 \cdots u_d)^{-1} \otimes u_{\sigma(1)} \otimes \cdots u_{\sigma(d)} \in Z_d(\mathcal{B}, \mathcal{B}) \tag{33}$$

is a Hochschild d-cycle, $b\boldsymbol{c} = 0$. From (22) and (3) we obtain

$$\begin{aligned} \boldsymbol{\gamma}_1 &:= \pi_{\mathcal{D}_1}(\boldsymbol{c}) = \mathrm{i}^{\frac{d(d+1)}{2}} (b^{\dagger 1} - b^1) \cdots (b^{\dagger d} - b^d) , \\ \boldsymbol{\gamma}_2 &:= \pi_{\mathcal{D}_2}(\boldsymbol{c}) = \mathrm{i}^{\frac{d(d+3)}{2}} (b^{\dagger 1} + b^1) \cdots (b^{\dagger d} + b^d) . \end{aligned} \tag{34}$$

Both $\boldsymbol{\gamma}_i$ commute with every element of $\mathcal{A}$ or $\mathcal{B}$. Using the anticommutation relations (3) and $(b^\mu)^* \equiv b^{\dagger\mu}$, we have

$$\boldsymbol{\gamma}_1^2 = 1 = \boldsymbol{\gamma}_2^2 , \qquad \boldsymbol{\gamma}_1^* = \boldsymbol{\gamma}_1 , \quad \boldsymbol{\gamma}_2^* = \boldsymbol{\gamma}_2 . \tag{35}$$

Decomposing the fermionic part of the Dirac operators $\mathcal{D}_i$ in $b^{\dagger\mu} \pm b^\mu$, we have

$$(b^{\dagger\mu} \pm b^\mu)\boldsymbol{\gamma}_1 = \pm(-1)^d \boldsymbol{\gamma}_1 (b^{\dagger\mu} \pm b^\mu) , \quad (b^{\dagger\mu} \pm b^\mu)\boldsymbol{\gamma}_2 = \mp(-1)^d \boldsymbol{\gamma}_2 (b^{\dagger\mu} \pm b^\mu) . \tag{36}$$

Therefore, $b^{\dagger\mu} \pm b^\mu$ and hence the $\mathcal{D}_i$ *always* (d even or odd) anticommute with the product $\boldsymbol{\gamma}_1\boldsymbol{\gamma}_2$, which turns out to be (up to a factor) the $\mathbb{Z}_2$-grading $(-1)^{N_f}$ of the Hilbert space:

$$(-\mathrm{i})^d \boldsymbol{\gamma}_1\boldsymbol{\gamma}_2 = \mathrm{i}^d \boldsymbol{\gamma}_2\boldsymbol{\gamma}_1 = (b^1 b^{\dagger 1} - b^{\dagger 1} b^1) \cdots (b^d b^{\dagger d} - b^{\dagger d} b^d) = (-1)^{N_f} . \tag{37}$$

This is quite different from conventional spectral triples [**11**] with a single operator $\mathcal{D}$.

3.5. The index formula. We let $\mathcal{H} = \mathcal{H}_{ev} \oplus \mathcal{H}_{odd}$ be the decomposition into even and odd subspaces with respect to the grading $(-1)^{N_f}$ induced by the fermion number operator N_f. The $\mathcal{D}_i$ are off-diagonal in this decomposition, $\mathcal{D}_i = \mathcal{D}_i^+ + \mathcal{D}_i^-$, with $\mathcal{D}_i^+ = \mathcal{D}_i\big|_{\mathcal{H}_{ev}} : \mathcal{H}_{ev} \to \mathcal{H}_{odd}$ and $\mathcal{D}_i^- = (\mathcal{D}_i^+)^* = \mathcal{D}_i\big|_{\mathcal{H}_{odd}} : \mathcal{H}_{odd} \to \mathcal{H}_{ev}$.

There is a well-defined index problem for $\mathcal{D}_i^+$ due to Elliott, Natsume and Nest [**26**]. The $\mathcal{D}_i^+$ are elliptic pseudodifferential operators in the sense of Shubin [**29**] with symbol $\mathfrak{a}_i$. Then, the analytic index

$$\operatorname{index}(\mathcal{D}_i^+) = \dim\ker\mathcal{D}_i^+ - \dim\ker\mathcal{D}_i^- \tag{38}$$

can be computed by an index formula for the symbol $\mathfrak{a}_i$ as described below.

Following [**26**], we associate to (appropriate) operators $\mathcal{P}_\mathfrak{a} : \mathcal{S}(\mathbb{R}^n;\mathbb{C}^k) \to \mathcal{S}(\mathbb{R}^n;\mathbb{C}^k)$ the symbol $\mathfrak{a} \in M_k(C^\infty(T^*\mathbb{R}^n))$ by

$$(\mathcal{P}_\mathfrak{a}\eta)(x) = \frac{1}{(2\pi)^n}\int_{\mathbb{R}^n\times\mathbb{R}^n} d\xi\, dy\; e^{\mathrm{i}\langle x-y,\xi\rangle}\,\mathfrak{a}_i(x,\xi)\,\eta(y)\,, \qquad \eta \in \mathcal{S}(\mathbb{R}^n;\mathbb{C}^k)\,. \tag{39}$$

The symbol $\mathfrak{a}$ is said to be *elliptic of order* m if there exist $C, R > 0$ such that $\mathfrak{a}(x,\xi)^*\mathfrak{a}(x,\xi) \geq C(\|x\|^2 + \|\xi\|^2)^m 1_k$ for $\|x\|^2 + \|\xi\|^2 \geq R$.

For $m > 0$ one defines the graph projector

$$e_\mathfrak{a} = \begin{pmatrix} (1+\mathfrak{a}^*\mathfrak{a})^{-1} & (1+\mathfrak{a}^*\mathfrak{a})^{-1}\mathfrak{a} \\ \mathfrak{a}^*(1+\mathfrak{a}^*\mathfrak{a})^{-1} & \mathfrak{a}^*(1+\mathfrak{a}^*\mathfrak{a})^{-1}\mathfrak{a} \end{pmatrix} \in M_{2k}(C(T^*\mathbb{R}^n)) \tag{40}$$

and the matrix $\hat{e}_\mathfrak{a} = e_\mathfrak{a} - \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1\end{smallmatrix}\right) \in M_{2k}(C_0(T^*\mathbb{R}^n))$, i.e. $\hat{e}_\mathfrak{a}$ vanishes at infinity for $m > 0$ (the entries of $\hat{e}_\mathfrak{a}$ are of order $-m$). Using continuous fields of C^*-algebras, the following index theorem is proven in [**26**]:

THEOREM 5. *If $\mathcal{P}_\mathfrak{a}$ is an elliptic pseudodifferential operator of positive order, then*

$$\operatorname{index}(\mathcal{P}_\mathfrak{a}) = \frac{1}{(2\pi\mathrm{i})^n n!}\int_{T^*\mathbb{R}^n} \operatorname{tr}\big(\hat{e}_\mathfrak{a}(d\hat{e}_\mathfrak{a})^{2n}\big)\,, \tag{41}$$

where $T^*\mathbb{R}^n$ is oriented by $dx_1 \wedge d\xi_1 \wedge \cdots \wedge dx_n \wedge d\xi_n > 0$.

Let us return to our example. Restricting $\mathcal{D}_i^+$ to the even part of $\mathcal{H}_\infty$, we regard $\mathcal{D}_i^+$ as an operator $\mathcal{D}_i^+ : \mathcal{S}(\mathbb{R}^d;\mathbb{C}^{2^{d-1}}) \to \mathcal{S}(\mathbb{R}^d;\mathbb{C}^{2^{d-1}})$. The symbol $\mathfrak{a}_i \in M_{2^{d-1}}(C^\infty(T^*\mathbb{R}^d))$ of $\mathcal{D}_i^+$ is obtained from the action of $\mathfrak{Q}, \mathfrak{Q}^\dagger$ on the basis $e^{\mathrm{i}\langle\xi,x\rangle}|s_1,\dots,s_d\rangle_f$. For example, we have for $d = 2$ in the matrix bases $\left(\begin{smallmatrix}|0,0\rangle_f \\ |1,1\rangle_f\end{smallmatrix}\right)$ of $\big(\bigwedge(\mathbb{C}^d)\big)_{ev}$ and $\left(\begin{smallmatrix}|1,0\rangle_f \\ |0,1\rangle_f\end{smallmatrix}\right)$ of $\big(\bigwedge(\mathbb{C}^d)\big)_{odd}$ the representation

$$\mathfrak{a}_1(x_1,x_2,\xi_1,\xi_2) = \begin{pmatrix} \mathrm{i}\xi_1 + \omega x_1 & -(-\mathrm{i}\xi_2 + \omega x_2) \\ \mathrm{i}\xi_2 + \omega x_2 & -\mathrm{i}\xi_1 + \omega x_1 \end{pmatrix}\,. \tag{42}$$

The product $\mathfrak{a}_i(x,\xi)^*\mathfrak{a}_i(x,\xi)$ is the restriction of the symbol of H to the even subspace. This implies

$$\mathfrak{a}_i(x,\xi)^*\mathfrak{a}_i(x,\xi) = (\omega^2\|x\|^2 + \|\xi\|^2)1_{2^{d-1}}\,, \tag{43}$$

i.e. ellipticity of order 1 if $\omega > 0$. Note that the usual Dirac operator $\mathrm{i}\gamma^\mu\partial_\mu$ on $\mathbb{R}^d$ is not elliptic in this sense.

For $d = 2$ an already lengthy computation shows

$$\operatorname{tr}\big(\hat{e}_{\mathfrak{a}_1}\, d\hat{e}_{\mathfrak{a}_1} \wedge d\hat{e}_{\mathfrak{a}_1} \wedge d\hat{e}_{\mathfrak{a}_1} \wedge d\hat{e}_{\mathfrak{a}_1}\big) = -\frac{96\omega^2\, dx_1 \wedge d\xi_1 \wedge dx_2 \wedge d\xi_2}{(1+\omega^2x_1^2+\omega^2x_2^2+\xi_1^2+\xi_2^2)^5}\,, \tag{44}$$

which yields

$$\text{index}\,(\mathcal{D}_1^+) = \frac{1}{(2\pi\mathrm{i})^2\cdot 2}\int_0^\infty 2\pi x\,dx \int_0^\infty 2\pi\xi\,d\xi\,\frac{(-96\omega^2)}{(1+\omega^2x^2+\xi^2)^5} = 1\;. \tag{45}$$

This is of course expected in any dimension d: the (one-dimensional) kernel of $\mathcal{D}_i^+$ is spanned by the Gaußian $e^{-\frac{\omega}{2}\|x\|^2}|0,\dots,0\rangle_f$, and the cokernel is trivial.

4. The spectral action for the $U(1)$-Higgs model

In the Connes-Lott spirit [**4**] we take the tensor product of the $(d=4)$-dimensional spectral triple $(\mathcal{A},\mathcal{H},\mathcal{D}_1)$ with the finite Higgs spectral triple $(\mathbb{C}\oplus\mathbb{C},\mathbb{C}^2,M\sigma_1,\sigma_3)$, which is even with $\mathbb{Z}_2$-grading σ_3. Here, M is a real number, and σ_k are the Pauli matrices. For the bosonic sector considered here only the spectrum of $\mathcal{D}_i$ matters, so that $\mathcal{D}_1$ and $\mathcal{D}_2$ give identical results. The total Dirac operator $\mathcal{D}=\mathcal{D}_1\otimes\sigma_3+1\otimes M\sigma_1$ of the product triple becomes

$$\mathcal{D} = \begin{pmatrix} \mathcal{D}_1 & M \\ M & -\mathcal{D}_1\end{pmatrix}\;. \tag{46}$$

In this representation, the algebra is $\mathcal{A}\oplus\mathcal{A}\ni(f,g)$ with diagonal action by pointwise multiplication on $\mathcal{H}_{tot}=\mathcal{H}\oplus\mathcal{H}$. The commutator of $\mathcal{D}$ with (f,g) is

$$[\mathcal{D},(f,g)] = \begin{pmatrix} \partial_\mu f\otimes(b^{\dagger\mu}-b^\mu) & M(g-f) \\ M(f-g) & -\partial_\mu g\otimes(b^{\dagger\mu}-b^\mu)\end{pmatrix}\;. \tag{47}$$

This shows that selfadjoint fluctuated Dirac operators $\mathcal{D}_A=\mathcal{D}+\sum_i a_i[\mathcal{D},b_i]$ are of the form

$$\mathcal{D}_A = \begin{pmatrix} \mathcal{D}_1+\mathrm{i}A_\mu\otimes(b^{\dagger\mu}-b^\mu) & \phi\otimes 1 \\ \bar{\phi}\otimes 1 & -\mathcal{D}_1-\mathrm{i}B_\mu\otimes(b^{\dagger\mu}-b^\mu)\end{pmatrix}\;, \tag{48}$$

for real fields $A_\mu=\overline{A_\mu}$, $B_\mu=\overline{B_\mu}\in\mathcal{A}$ and a complex field $\phi\in\mathcal{A}$. The square of $\mathcal{D}_A$ is

$$\mathcal{D}_A^2 = \begin{pmatrix} H\otimes 1+\omega\otimes\Sigma+\mathrm{i}F_A+|\phi|^2\otimes 1 & D_\mu\phi\otimes(b^{\dagger\mu}-b^\mu) \\ -\overline{D_\mu\phi}\otimes(b^{\dagger\mu}-b^\mu) & H\otimes 1+\omega\otimes\Sigma+\mathrm{i}F_B+|\phi|^2\otimes 1\end{pmatrix}\;, \tag{49}$$

where

$$D_\mu\phi := \partial_\mu\phi+\mathrm{i}(A_\mu-B_\mu)\phi\;, \tag{50}$$

$$\begin{aligned} F_A &:= \{\mathcal{D}_1,A_\mu\otimes(b^{\dagger\mu}-b^\mu)\}+\mathrm{i}A_\mu A_\nu\otimes(b^{\dagger\mu}-b^\mu)(b^{\dagger\nu}-b^\nu) \\ &= (-\{\partial_\mu,A^\mu\}-\mathrm{i}A_\mu A^\mu)\otimes 1+\frac{1}{4}F^A_{\mu\nu}\otimes[b^{\dagger\mu}-b^\mu,b^{\dagger\nu}-b^\nu] \end{aligned} \tag{51}$$

and similarly for F_B. Here, $F^A_{\mu\nu}=\partial_\mu A_\nu-\partial_\nu A_\mu$ is the $U(1)$-curvature (field strength), and the explicit appearance of x has dropped out in F_A because of $\{b^{\dagger\mu}+b^\mu,b^{\dagger\nu}-b^\nu\}=0$.

According to the spectral action principle [**6, 7**], the bosonic action depends only on the spectrum of the Dirac operator. Thus, by functional calculus, the most general form of the bosonic action is

$$S(\mathcal{D}_A) = \mathrm{Tr}\big(\chi(\mathcal{D}_A^2)\big) = \int_0^\infty dt\;\mathrm{Tr}(e^{-t\mathcal{D}_A^2})\hat{\chi}(t)\;, \tag{52}$$

for some function $\chi : \mathbb{R}_+ \to \mathbb{R}_+$ for which the operator trace exists. The second equality is obtained by Laplace transformation, which produces the inverse Laplace transform $\hat{\chi}$ of $\chi(s) = \int_0^\infty dt\, e^{-st}\hat{\chi}(t)$. One has

$$\chi_z := \int_0^\infty dt\, t^z \hat{\chi}(t) = \begin{cases} \dfrac{1}{\Gamma(-z)} \displaystyle\int_0^\infty ds\, s^{-z-1}\chi(s) & \text{for } z \notin \mathbb{N}\,, \\ (-1)^k \chi^{(k)}(0) & \text{for } z = k \in \mathbb{N}\,. \end{cases} \tag{53}$$

To compute the traces $\mathrm{Tr}(e^{-t\mathcal{D}_A^2})$ we write $\mathcal{D}_A^2 = \mathrm{H}_0 - V$, with $\mathrm{H}_0 := H + \omega\Sigma$, and consider the Duhamel expansion

(54)

$$\begin{aligned}
&e^{-t_0(\mathrm{H}_0 - V)} \\
&= e^{-t_0 \mathrm{H}_0} - \int_0^{t_0} dt_1\, \frac{d}{dt_1}\big(e^{-(t_0-t_1)(\mathrm{H}_0-V)} e^{-t_1\mathrm{H}_0}\big) \\
&= e^{-t_0 \mathrm{H}_0} + \int_0^{t_0} dt_1\, \big(e^{-(t_0-t_1)(\mathrm{H}_0-V)} V e^{-t_1\mathrm{H}_0}\big) \\
&= e^{-t_0 \mathrm{H}_0} + \int_0^{t_0} dt_1\, \big(e^{-(t_0-t_1)\mathrm{H}_0} V e^{-t_1\mathrm{H}_0}\big) \\
&+ \int_0^{t_0} dt_1 \int_0^{t_0-t_1} dt_2\, \big(e^{-(t_0-t_1-t_2)\mathrm{H}_0} V e^{-t_2\mathrm{H}_0} V e^{-t_1\mathrm{H}_0}\big) + \dots \\
&+ \int_0^{t_0} dt_1 \dots \int_0^{t_0-t_1-\dots-t_{n-1}} dt_n\, \big(e^{-(t_0-t_1-\dots-t_n)\mathrm{H}_0} (V e^{-t_n\mathrm{H}_0}) \cdots (V e^{-t_1\mathrm{H}_0})\big) + \dots \\
&= e^{-t_0 \mathrm{H}_0} + \sum_{n=1}^\infty t_0^n \int_{\Delta^n} d^n\alpha \Big(e^{-t_0(1-|\alpha|)\mathrm{H}_0} \prod_{j=1}^n (V e^{-t_0\alpha_j \mathrm{H}_0})\Big)\,,
\end{aligned}$$

where the integration is performed over the standard n-simplex $\Delta^n := \{\alpha := (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n\,,\ \alpha_i \geq 0\,,\ |\alpha| := \alpha_1 + \dots + \alpha_n \leq 1\}$.

Using $\mathrm{tr}(e^{\omega\Sigma t}) = (2\cosh(\omega t))^4$ and the Mehler kernel (76), the vacuum contribution without V is

$$\begin{aligned}
\mathrm{Tr}(e^{-t(H+\omega\Sigma)\otimes 1_2}) &= \big(2\,\mathrm{tr}(e^{\omega\Sigma t})\big) \int_{\mathbb{R}^4} dx\, e^{-tH}(x,x) \\
&= 2(2\cosh(\omega t))^4 \cdot \Big(\frac{\omega}{2\pi\sinh(2\omega t)}\Big)^2 \int_{\mathbb{R}^4} dx\, e^{-\omega\tanh(\omega t)\|x\|^2} \\
&= \frac{2}{\tanh^4(\omega t)}\,.
\end{aligned} \tag{55}$$

With $\coth^4(\omega t) = \frac{1}{(\omega t)^4} + \frac{4}{3(\omega t)^2} + \frac{26}{45} + \mathcal{O}(t^2)$ we get under the usual assumption $\chi^{(k)}(0) = 0$ for $k = 1, 2, 3, \dots$ the asymptotic expansion[1]

$$S_0(\mathcal{D}_A) = \frac{2\chi_{-4}}{\omega^4} + \frac{8\chi_{-2}}{3\omega^2} + \frac{52\chi_0}{45}\,. \tag{56}$$

[1] The Laplace transformation for the vacuum contribution can be performed exactly. For powers of $\coth x = \frac{1+e^{-2x}}{1-e^{-2x}}$ we have

$$\Big(\frac{1+y}{1-y}\Big)^n = 1 + \sum_{k=1}^\infty \underbrace{\frac{(k+n-1)!}{k!}\, {}_2F_1\Big(\begin{matrix} -k\,,\ -n \\ 1-k-n \end{matrix}\Big| -1\Big)}_{=F_n(k)} y^k\,.$$

For the further computation we distinguish the vertices (see (49), (50) and (51))

$$V_1 := \mathrm{diag}\big(\mathrm{i}\{\partial^\mu, A_\mu\}\otimes 1, \mathrm{i}\{\partial^\mu, B_\mu\}\otimes 1\big)\;, \tag{57}$$

$$V_2 := \mathrm{diag}\big(-A_\mu A^\mu \otimes 1 - |\phi^2|\otimes 1, -B_\mu B^\mu\otimes 1 - |\phi^2|\otimes 1\big)\;,$$

$$V_3 := \mathrm{diag}\big(-\mathrm{i}F^A_{\mu\nu}\otimes \tfrac{1}{4}[b^{\dagger\mu}-b^\mu, b^{\dagger\nu}-b^\nu],\; -\mathrm{i}F^B_{\mu\nu}\otimes \tfrac{1}{4}[b^{\dagger\mu}-b^\mu, b^{\dagger\nu}-b^\nu]\big)\;,$$

$$V_4 = \begin{pmatrix} 0 & -D_\mu\phi\otimes(b^{\dagger\mu}-b^\mu) \\ \overline{D_\mu\phi}\otimes(b^{\dagger\mu}-b^\mu) & 0\end{pmatrix}\;.$$

We compute the traces of the spectral action in the same way as the residues of the ζ-function in Appendix A. The main step consists in computing the following trace:

$$S_{t_1,\dots,t_v}(\tilde{V}_1,\dots,\tilde{V}_v) := \mathrm{Tr}\Big(\tilde{V}_1 e^{-t_1H}\tilde{V}_2 e^{-t_2H}\cdots \tilde{V}_v e^{-t_vH}\Big)\;, \tag{58}$$

either with $\tilde{V}_i = f_i$ or $\tilde{V}_i = -\mathrm{i}\{\partial_\mu, f_i^\mu\} = -\mathrm{i}(\partial_\mu f_i^\mu) - 2\mathrm{i}f_i^\mu\partial_\mu$. We realise this alternative as $\tilde{V}_i = f_i^{1-n_i}\{-\mathrm{i}\partial_\mu, f^\mu\}^{n_i}$ with $n_i\in\{0,1\}$:

(59)

$$\begin{aligned}
&S^{n_1\dots n_v}_{t_1,\dots t_v}(f_1,\dots,f_v)\\
&=\sum_{k_1=0}^{n_1}\cdots\sum_{k_v=0}^{n_v}\omega^{k_1+\dots+k_v}\int_{(\mathbb{R}^4\times\mathbb{R}^4)^v}\Big(\prod_{i=1}^v\frac{dx_i dp_i}{(2\pi)^4}\Big)\\
&\times\Big(\prod_{i=1}^v \hat{f}_i^{1-n_i}(p_i)\Big(\hat{f}_i^{\mu_1}(p_i)p_{i,\mu_i}^{1-k_i}P_{\mu_i}^{k_i}\Big(2\omega t_i,\frac{\partial}{\partial p_i},\frac{\partial}{\partial p_{i+1}}\Big)\Big)^{n_i}\Big)\Big(\prod_{i=1}^v e^{-t_iH}(x_i,x_{i+1})e^{\mathrm{i}p_ix_i}\Big)\\
&=\sum_{k_1=0}^{n_1}\cdots\sum_{k_v=0}^{n_v}\int_{(\mathbb{R}^4)^v}\Big(\prod_{i=1}^v\frac{dp_i}{(2\pi)^4}\Big)\frac{\omega^{k_1+\dots+k_v}}{(2\sinh(\omega(t_1+\dots+t_v)))^4}\\
&\times\Big(\prod_{i=1}^v \hat{f}_i^{1-n_i}(p_i)\Big(\hat{f}_i^{\mu_1}(p_i)p_{i,\mu_i}^{1-k_i}P_{\mu_i}^{k_i}\Big(2\omega t_i,\frac{\partial}{\partial p_i},\frac{\partial}{\partial p_{i+1}}\Big)\Big)^{n_i}\Big)e^{-\frac{1}{4}pQ^{-1}p}\;,
\end{aligned}$$

where P_μ and Q^{-1} are given in (83) and (86). From the formulae analogous to (88) and (90) we thus obtain

(60)

$$\begin{aligned}
&S^{n_1\dots n_v}_{t_1,\dots t_v}(f_1,\dots,f_v)\\
&=\sum_{\substack{k_1+r_{11}+\dots+r_{1v}=n_1,\dots,\\ k_1+r_{v1}+\dots+r_{vv}=n_v,\\ r_{ii}=0\;,\; r_{ij}=r_{ji}}}\int_{(\mathbb{R}^4)^v}\Big(\prod_{i=1}^v\frac{dp_i}{(2\pi)^4}\Big)\frac{1}{(2\sinh(\omega t))^4}\Big(\prod_{i=1}^v \hat{f}_i^{1-n_i}(p_i)\big(\hat{f}_i^{\mu_i}(p_i)\big)^{n_i}\Big)\\
&\times\Big(\prod_{i=1}^v\Big(\sum_{j\neq i}\frac{\sinh(\omega t_{ji})}{\sinh(\omega t)}p_{j,\mu_i}\Big)^{k_i}\Big)\Big(\prod_{i\leq j}\Big(2\omega\delta_{\mu_i\mu_j}\frac{\cosh(\omega t_{ji})}{\sinh(\omega t)}\Big)^{r_{ij}}\Big)e^{-\frac{1}{4}pQ^{-1}p}\;,
\end{aligned}$$

Particular values are $F_1(k)=2$, $F_2(k)=4k$, $F_3(k)=8k^2+4$, $F_4(k)=16k^3+32k$ and $F_5(k)=32k^4+160k^2+48$. Inserted into (52) we obtain after Laplace transformation

$$S_0(\mathcal{D}_A) = 2\chi(0)+\sum_{k=1}^\infty(32k^3+64k)\chi(2\omega k)\;.$$

where $t_{ji} := t_j + \cdots + t_{i-1} - t_i - \cdots - t_{j-1}$ and $t := t_1 + \cdots + t_v$

For the spectral action we are interested in the small-t behaviour. From (86) we know that the singularity in $\sinh^{-4-\sum_{i<j} r_{ij}}(\omega t)$ is protected by $\exp(-\frac{(p_1+\cdots+p_v)^2}{4\omega \tanh(\omega t)})$ unless the total momentum is conserved. Thus, Taylor-expanding the prefactor about $p_v = -(p_1 + \cdots + p_{v-1})$ up to order ρ and Gaußian integration in p_v yields

$$S^{n_1,\dots n_v}_{t_1,\dots,t_v} = \mathcal{O}(t^{-2-\lfloor \frac{n_1+\cdots+n_v}{2} \rfloor + \lceil \frac{\rho}{2} \rceil}) .$$

To obtain the spectral action, there are apart from the (at most) t-neutral matrix trace the v integrations over $t_1, \dots, t_v$ which contribute another power of t^v. If there are v_i vertices of type V_i present, with $v_1 + \cdots + v_4 = v$, then $n_1 + \cdots + n_v = v_1$, and we have for such a contribution

$$S_t(V_1^{v_1}, \dots, V_4^{v_4}) = \mathcal{O}(t^{-2+v_2+v_3+v_4+\lceil \frac{v_1}{2} \rceil + \lceil \frac{\rho}{2} \rceil}) .$$

Only the non-positive exponents contribute to the asymptotic expansion so that it suffices to compute the following traces of vertex combinations:

(1) V_2 with Taylor expansion up to order $\rho = 2$ (V_3 and V_4 are traceless, and in V_1 alone there is necessarily $k_1 = n_1 = 1$ and then no sum over $i \neq j$),
(2) $V_1 V_1$ with Taylor expansion up to order $\rho = 2$,
(3) $V_1 V_2$, $V_2 V_1$ and $V_1 V_1 V_2$, $V_1 V_2 V_1$, $V_2 V_1 V_1$ with Taylor expansion up to order $\rho = 0$,
(4) $V_2 V_2$, $V_3 V_3$ and $V_4 V_4$ with Taylor expansion up to order $\rho = 0$ (mixed products are traceless),
(5) $V_1 V_1 V_1$ and $V_1 V_1 V_1 V_1$ with Taylor expansion up to order $\rho = 0$.

We compute these vertex combinations in Appendix B. The spectral action is the sum of (100), (104), (107), (109), (111), (113) and (115). Altogether, the spectral action of the Abelian Higgs model reads

$$\begin{aligned} S(\mathcal{D}_A) = {} & \frac{2\chi_{-4}}{\omega^4} + \frac{8\chi_{-2}}{3\omega^2} + \frac{52\chi_0}{45} \\ & + \frac{\chi_0}{\pi^2} \int_{\mathbb{R}^4} dx \, \Big\{ \frac{5}{12} (F_A^{\mu\nu} F_{A\mu\nu} + F_B^{\mu\nu} F_{B\mu\nu}) + \overline{D_\mu \phi} (D^\mu \phi) \\ & \qquad - \frac{2\chi_{-1}}{\chi_0} |\phi|^2 + |\phi|^4 + 2\omega^2 \|x\|^2 |\phi|^2 \Big\}(x) . \end{aligned} \tag{61}$$

The scalar sector (putting $A = B = 0$ and ignoring the constant) is almost identical to the commutative version of the renormalisable ϕ^4-action [**14**],

$$S(\mathcal{D}_A)|_{A=b=0} = \frac{\chi_0}{\pi^2} \int_{\mathbb{R}^4} dx \, \Big\{ \partial_\mu \bar{\phi} (\partial^\mu \phi) + 2\omega^2 \|x\|^2 |\phi|^2 - \frac{2\chi_{-1}}{\chi_0} |\phi|^2 + |\phi|^4 \Big\}(x) . \tag{62}$$

The crucial difference is the negative mass squared term, which leads to a drastically different vacuum structure, as shown in the next section.

5. Field equations

We can assume the solution of the corresponding equation of motion to be given by $A = B = 0$ and ϕ a real function. Then, the Euler-Lagrange equation reads

$$-\Delta \phi + 2\omega^2 \|x\|^2 \phi + 2\phi^3 - 2\tfrac{\chi_{-1}}{\chi_0} \phi = 0 . \tag{63}$$

In terms of the rescaled radius $r = 2^{\frac{1}{4}}\sqrt{\omega}\|x\|$ and the rescaled field $\phi = \frac{\pi}{\sqrt{2}\chi_0}\varphi$ we have the rotationally invariant equation

$$
\begin{aligned}
(64)\qquad & -\varphi''(r) - \frac{3}{r}\varphi'(r) + (r^2 - 4\mu^2)\varphi(r) = -\lambda\varphi^3(r)\;, \\
& \mu^2 = \frac{\chi_{-1}}{\sqrt{8}\omega\chi_0}\;, \qquad \lambda = \frac{\pi^2}{\sqrt{2}\omega\chi_0}\;.
\end{aligned}
$$

We expand φ in terms of eigenfunctions of the four-dimensional harmonic oscillator,

$$
\begin{aligned}
(65)\qquad & \varphi = \frac{2}{\sqrt{\lambda}}\sum_{n=0}^{\infty} c_n\varphi_n\;, \\
& \varphi_n := e^{-\frac{r^2}{2}}L_n^1(r^2)\;, \qquad \Big(-\frac{d^2}{dr^2} - \frac{3}{r}\frac{d}{dr} + r^2\Big)\varphi_n = 4(n+1)\varphi_n\;.
\end{aligned}
$$

We are thus left with the equation

$$
(66)\qquad \sum_{n=0}^{\infty} c_n(\mu^2 - n - 1)\varphi_n = \sum_{k,l,m=0}^{\infty} c_k c_l c_m \varphi_k\varphi_l\varphi_m
$$

or, using the orthogonality relation,

$$
(67)\qquad c_n(\mu^2 - n - 1) = \frac{1}{(n+1)}\sum_{k,l,m=0}^{\infty} c_k c_l c_m \int_0^{\infty} dt\; e^{-2t}\, t\, L_k^1(t)\, L_l^1(t)\, L_m^1(t)\, L_n^1(t)\;.
$$

The generating function $(1-z)^{-\alpha-1}\exp(-\frac{xz}{1-z}) = \sum_{k=0}^{\infty} L_k^{\alpha}(t)z^k$ is used to obtain

$$
\begin{aligned}
(68)\qquad & c_n(\mu^2 - n - 1) \\
& = \sum_{k,l,m=0}^{\infty} \frac{c_k c_l c_m}{k!l!m!}\Big(\frac{d^k}{dw^k}\frac{d^l}{dy^l}\frac{d^m}{dz^m}\frac{(1 - yz - yw - wz + 2wyz)^n}{(2 - y - z - w + yzw)^{n+2}}\Big)_{w=y=z=0}\;.
\end{aligned}
$$

With a cut-off N for the matrix indices, this equation can be solved numerically. It turns out that except for a region about $r = 4\mu^2$ the convergence is quite good. Figure 1 contains plots of the vacuum solution $\varphi_{vac}(r)$ for $4\mu^2 = 9$ and $4\mu^2 = 13$ compared with the ellipse $\varphi^2 + \frac{1}{4}r^2 = \mu^2$. We learn that $\varphi_{vac}(r) < \sqrt{\mu^2 - \frac{1}{4}r^2}$ due to the negative curvature $\frac{1}{\varphi}(\varphi'' + \frac{3}{r}\varphi') < 0$ which effectively reduces μ^2. For $r > 2\mu$ we should have $\varphi_{vac}(r) = 0$ as the only solution[2]. We also expect that for $\mu \to \infty$, where the ellipse becomes flat, the vacuum solution approaches its limiting ellipse. This limit is connected to the limit $\omega \to 0$, i.e $r = 2^{\frac{1}{4}}\sqrt{\omega}\|x\| \to 0$. In this limit the usual constant Higgs vacuum is recovered:

$$
(69)\qquad \lim_{\omega\to 0}\phi^2 = \frac{\pi^2}{2\chi_0^2}\frac{4\mu^2}{\lambda} = \frac{\chi_{-1}}{\chi_0^2} = \text{const}\;.
$$

For finite ω the cut-off for φ_{vac} at $r = 2\mu$ implies that φ_{vac} is an integrable function.

The vacuum solution

$$
(70)\qquad \frac{2}{\sqrt{\lambda}}\varphi_{vac} = \sqrt{\frac{4\mu^2}{\lambda}}\frac{\varphi_{vac}}{\mu} = \sqrt{\frac{2\chi_{-1}}{\pi^2}}\frac{\varphi_{vac}}{\mu}
$$

[2]The numerical convergence in the figure is bad for $r \approx 2\mu$.

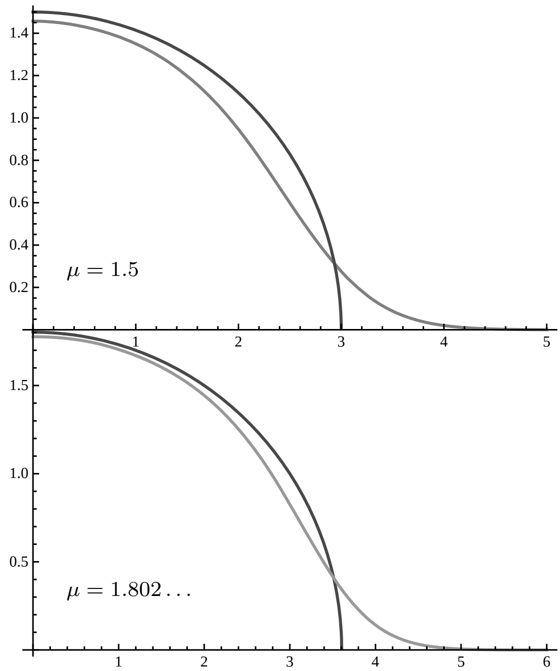


FIGURE 1. The lower curve at $r = 0$ shows $\varphi_{vac}(r)$ in units of $\frac{2}{\sqrt{\lambda}}$, with cut-off at $N = 10$. The upper curve at $r = 0$ is the ellipse $\varphi^2 + \frac{1}{4}r^2 = \mu^2$. The error is below 1% for $r < 1.8\mu$. The true curve $\varphi_{vac}(r)$ is expected to stay always below the ellipse and to connect smoothly (at least C^2) to $\varphi_{vac} = 0$ for $r > 2\mu$.

sets the scale for the bare masses of gauge fields and fermions. On the other hand, the bare mass of the Higgs field is obtained from the shift of the Higgs potential into its minimum and therefore reads

$$\sqrt{\sqrt{2}\omega((r^2 - 4\mu^2) + 12\mu^2\frac{\varphi_{vac}^2}{\mu^2})} = \sqrt{\frac{4\chi_{-1}}{\chi_0}}\frac{\sqrt{\frac{3}{2}\varphi_{vac}^2 - \frac{1}{2}\mu^2 + \frac{1}{8}r^2}}{\mu}\,. \tag{71}$$

We compare in Figure 2 the scale $\frac{\varphi_{vac}}{\mu}$ of gauge field mass with the scale $\frac{1}{\mu}\sqrt{\frac{3}{2}\varphi_{vac}^2 - \frac{1}{2}\mu^2 + \frac{1}{8}r^2}$ of the bare Higgs mass. Reinserting ω we obtain the following *two-phase structure*:

- *A spontaneously broken phase for* $\omega^2\|x\|^2 < \frac{\chi_{-1}}{\chi_0}$.
  Fermions, gauge fields and Higgs field are all massive, with the Higgs mass slightly smaller than the prediction from noncommutative geometry [**9**]. In particular, this phase is the only existing one in the limit $\omega \to 0$, and in this limit the NCG prediction is recovered.
- *An unbroken phase for* $\omega^2\|x\|^2 > \frac{\chi_{-1}}{\chi_0}$.
  Fermions and gauge fields are massless, whereas the Higgs field remains massive.

The model we have studied is a toy model. But, as it is a noncommutative geometry like that of the NCG-formulation of the Standard Model [**9**], it is ultimately an experimental question to set limits on the frequency parameter ω. To

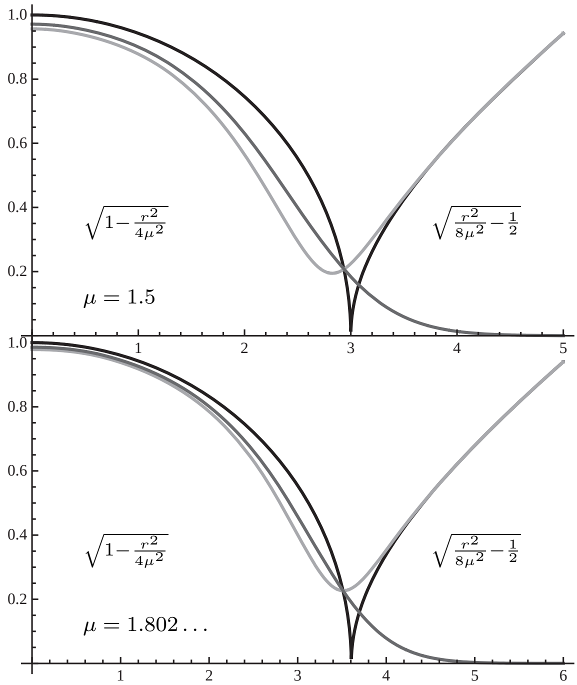


FIGURE 2. The scale $\frac{\varphi_{vac}}{\mu}(r)$ (middle curve at $r=0$) of the gauge field mass compared with the scale $\frac{1}{\mu}\sqrt{\frac{3}{2}\varphi^2_{vac}(r)-\frac{1}{2}\mu^2+\frac{1}{8}r^2}$ of the Higgs field mass (lowest curve at $r=0$) and the limiting ellipse $s^2+\frac{r^2}{4\mu^2}=1$ and hyperbola $\frac{r^2}{8\mu^2}-s^2=\frac{1}{2}$. Cutoff again at $N=10$. The true curve $\frac{\varphi_{vac}}{\mu}(r)$ should always stay below the ellipse and connect smoothly to $\frac{\varphi_{vac}}{\mu}=0$ for $r>2\mu$. The true curve $\frac{1}{\mu}\sqrt{\frac{3}{2}\varphi^2_{vac}(r)-\frac{1}{2}\mu^2+\frac{1}{8}r^2}$ should stay below $\frac{\varphi_{vac}}{\mu}$ for $r<2\mu$, whereas for $r>2\mu$ one should exactly have $\frac{1}{\mu}\sqrt{\frac{3}{2}\varphi^2_{vac}-\frac{1}{2}\mu^2+\frac{1}{8}r^2}(r)=\sqrt{\frac{r^2}{8\mu^2}-\frac{1}{2}}$.

be compatible with both high energy and cosmological data, ω has to be extremely small. We definitely live in the spontaneously broken phase $\omega^2\|x\|^2<\frac{\chi_{-1}}{\chi_0}$, and the observable universe is very close to $\omega^2\|x\|^2=0$. Nevertheless, a regulating $\omega\neq 0$ has some nice consequences such as integrability of the Higgs vacuum and integrability of the cosmological constant.

One may speculate how an $\omega\neq 0$ can be detected. We mentioned the reduction of the ratio between Higgs mass and Z mass compared with the NCG prediction. However, in the presence of $\omega\neq 0$ the β-functions must be recomputed so that at the moment no prediction is possible. In cosmology, limits for ω could be obtained from precision measurements of the ratio between the proton mass and the electron mass at far distance. The electron mass which governs the atomic spectra via the Rydberg frequency should vary in the same way as the Higgs scale $\frac{\varphi_{vac}}{\mu}$. On the other hand, the proton mass arises mainly from broken scale invariance in QCD and therefore can be regarded as constant. This means that the gravitational energy of a standard star is constant whereas its transition into radiation energy might vary with the position of the star in the universe. Observational limits on such a variation would limit the value of ω.

Another observable consequence could be a variation of the cosmological constant. The Higgs potential at the vacuum solution is negative and hence reduces the volume term of the cosmological constant. Thus, the effective cosmological constant would increase with the radius (the masses of gauge fields and fermions dissolve into the cosmological constant).

6. Conclusion and perspectives

We have proposed a definition for non-compact spectral triples $(\mathcal{A}, \mathcal{H}, \mathcal{D})$ where the algebra is allowed to be non-unital but the resolvent of the operator $\mathcal{D}$ remains compact. The metric dimension is defined via the dimension spectrum; it is (in general) different from the noncommutative dimension given by the decay rate of the characteristic values of the resolvent.

Our definition excludes non-compact manifolds with the standard Dirac operator, but this is necessary for a well-defined index problem and a well-defined spectral action in the non-compact case. An example for our definition is given by operators $\mathcal{D}$ which are square roots of the d-dimensional harmonic oscillator Hamiltonian $-\Delta + \omega^2 x^2$. These square roots are constructed by conjugation of the partial derivatives with $e^{\pm\omega h}$, where h is the Morse function. This relates to supersymmetric quantum mechanics, in particular to a special case of Witten's work [**21**] on Morse theory.

The most involved piece of work was the computation of the dimension spectrum which showed that the metric dimension is the oscillator dimension and that all residues of the operator zeta function are local. Due to its relation to supersymmetry, there are in fact two Dirac operators $\mathcal{D}_1$ and $\mathcal{D}_2$, which define two distinct images $\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$ of the d-dimensional volume form, and only the product $\boldsymbol{\gamma}_1\boldsymbol{\gamma}_2$ defines the $\mathbb{Z}_2$-grading.

We have computed the spectral action for the corresponding Connes-Lott two-point model. In contrast to standard $\mathbb{R}^d$, the spectral action is finite also in the cosmological constant part. The result is an Abelian Higgs model with additional harmonic oscillator potential for the Higgs field. The resulting field equations show a phase transition phenomenon: There is a spontaneously broken phase below a critical radius determined by the oscillator frequency ω, which for small enough ω is qualitatively identical to standard Higgs models. Possible observable consequences are discussed at the end of the previous section. Above the critical radius we have an unbroken phase with massless gauge fields. This phase is necessary to have an integrable vacuum solution for the Higgs field.

The class of spectral triples we proposed deserves further investigation. We show with V. Gayral [**18**] that there is an isospectral Moyal deformation of the harmonic oscillator spectral triple. Some ideas appeared already in our preprint [**19**] with H. Grosse, but the mathematical structure was unclear at that point. The field equations of the preprint [**19**] are correct, but their "solution" is wrong. It misses the phase transitions which we first observed for the commutative model in the present paper. We expect that the phase structure is much richer in the Moyal-deformed model. A hint can already be found in the pure gauge field sector, which leads in terms of "covariant coordinates" to the field equation $[X^\mu, [X_\mu, X_\nu]] = 0$.

This equation has the Moyal deformation $[X_\mu, X_\nu] = \mathrm{i}\Theta_{\mu\nu} = \text{const}$ as a solution, but also commutative coordinates $[X_\mu, X_\nu] = 0$; the preferred solution arises from a subtle interplay with the boundary conditions. One may speculate that these boundary conditions change with the temperature of the universe, so that the (non)commutative geometry could emerge through a cascade of phase transitions when the universe cools down. The Moyal-deformed harmonic oscillator spectral triple could serve as an excellent toy model to study these transitions.

On the mathematical side, the relation to supersymmetric quantum mechanics needs further study. In particular, a real structure (or better several real structures) must be identified to reduce the multiplicity of the action of the algebra from its present value 2^d to $2^{\frac{d}{2}}$ in order to support a Spin^c structure. One should also allow for a non-trivial projection e to define the smooth subspace $\mathcal{H}_\infty = e\mathcal{A}^n$ of the Hilbert space. The corresponding action of $\mathcal{D}_i$ or its components $\mathfrak{Q}, \mathfrak{Q}^\dagger$ would then permit a complete reformulation of Witten's approach [**21**] to Morse theory in the framework of spectral triples and noncommutative index theory.

Acknowledgements

I would like to thank the Clay Mathematics Institute for inviting me to write this contribution to the birthday volume for Alain Connes.

It is a pleasure to thank Harald Grosse for the long-term collaboration in which we found the need for ω in noncommutative scalar models and the possibility of a spectral triple for these models. The relation to supersymmetric quantum mechanics and Morse theory appeared in discussions with Christian Voigt. Several hints by Alain Connes on the $\mathbb{Z}_2$-grading were important in identifying the structure observed in section 3.4. The index formula was inspired by lectures of Toshikazu Natsume at the Tehran conference on noncommutative geometry.

The scientific exchange was supported by the grants SFB 478 of the Deutsche Forschungsgemeinschaft and MRTN-CT-2006-031962 of the European Union.

Appendix A. Proof of Theorem 2

Let $\mathcal{D}$ denote $\mathcal{D}_1$ or $\mathcal{D}_2$. The spectral identity $A = \dfrac{1}{\pi}\displaystyle\int_0^\infty \frac{d\lambda}{\sqrt{\lambda}} \frac{A^2}{A^2+\lambda}$ for a positive selfadjoint operator A leads to

$$\delta T = \frac{1}{\pi}\int_0^\infty d\lambda\, \sqrt{\lambda}\frac{1}{\langle\mathcal{D}\rangle^2+\lambda}[\mathcal{D}^2, T]\frac{1}{\langle\mathcal{D}\rangle^2+\lambda}\;. \tag{72}$$

From (15) we recall that $\mathcal{D}^2 = H + \omega\Sigma$, where $H = -\partial^\mu\partial_\mu + \omega^2 x_\mu x^\mu$ and $\Sigma = [b^\dagger_\mu, b^\mu]$ satisfy $[H, \Sigma] = 0$. This implies

$$\delta^n T = \sum_{k=0}^m \binom{n}{k}\big(\omega\,\mathrm{ad}(\Sigma)\big)^{n-k}\Big(\frac{1}{\pi^n}\int_0^\infty \prod_{i=1}^n \frac{d\lambda_i\,\sqrt{\lambda_i}}{\langle\mathcal{D}\rangle^2+\lambda_i}(\mathrm{ad}(H))^k(T)\prod_{j=1}^n\frac{1}{\langle\mathcal{D}\rangle^2{+}\lambda_j}\Big). \tag{73}$$

The case $T = [\mathcal{D}_1, f] = \partial_\mu f \otimes (b^{\dagger\mu} - b^\mu)$ or $T = [\mathcal{D}_2, f] = \partial_\mu f \otimes (\mathrm{i}b^{\dagger\mu} + \mathrm{i}b^\mu)$ is also reduced to $T = f$; only $\mathrm{ad}(\Sigma)$ distinguishes them, and each application of $\mathrm{ad}(\Sigma)$ makes δT more regular. It is therefore sufficient to study $T = f$ and $k = n$. Using

$[H,f] = -(\Delta f) - 2(\partial^\mu f)\partial_\mu = -\{\partial_\mu, \partial^\mu f\}$, we have

$$\delta^n f = \sum_{k=0}^{n} \binom{n}{k} 2^k \frac{(-1)^n}{\pi^n} \int_0^\infty \prod_{i=1}^{n} \frac{d\lambda_i \sqrt{\lambda_i}}{\langle\mathcal{D}\rangle^2 + \lambda_i} \times (\Delta^{n-k}\partial^{\mu_1}\cdots\partial^{\mu_k} f)\partial_{\mu_1}\cdots\partial_{\mu_k} \prod_{j=1}^{n} \frac{1}{\langle\mathcal{D}\rangle^2 + \lambda_j} \,. \tag{74}$$

By linearity, it suffices to consider $\phi = (\delta^{n_1} f_1)\cdots(\delta^{n_v} f_v)$. The most convenient way is to compute $\zeta_\phi(z)$ as a trace over position space kernels,

$$\begin{aligned} \zeta_\phi(z) &:= \mathrm{Tr}\big((\delta^{n_1} f_1)\cdots(\delta^{n_v} f_v)\langle\mathcal{D}\rangle^{-z}\big) \\ &= \mathrm{tr}\bigg(\int_0^\infty dt_0\, \frac{t_0^{\frac{z}{2}-1}}{\Gamma(\frac{z}{2})} \int_{(\mathbb{R}^d)^v} \Big(\prod_{i=1}^{v} dy_i\Big)(\delta^{n_1} f_1)(y_1,y_2)\cdots(\delta^{n_{v-1}} f_{v-1})(y_{v-1},y_v) \\ &\qquad \times (\delta^{n_v} f_v)(y_v,y_0)(e^{-t_0\langle\mathcal{D}\rangle^2})(y_0,y_1)\bigg) \,. \end{aligned} \tag{75}$$

The remaining trace tr is taken in $\bigwedge(\mathbb{C}^d)$. Further evaluation is possible thanks to the d-dimensional Mehler kernel

$$e^{-tH}(x,y) = \Big(\frac{\omega}{2\pi\sinh(2\omega t)}\Big)^{\frac{d}{2}} e^{-\frac{\omega}{4}\coth(\omega t)\|x-y\|^2 - \frac{\omega}{4}\tanh(\omega t)\|x+y\|^2} \,, \tag{76}$$

for $x,y\in\mathbb{R}^d$, which solves the differential equation $(\frac{d}{dt} + H_x)e^{-tH}(x,y) = 0$ with initial condition $\lim_{t\to 0} e^{-tH}(x,y) = \delta(x-y)$. Uniqueness of the solution implies

$$\int_{\mathbb{R}^d} dy\; e^{-t_1 H}(x,y)e^{-t_2 H}(y,z) = e^{-(t_1+t_2)H}(x,z) \,. \tag{77}$$

We can therefore recombine left and right Mehler kernels

$$\frac{1}{\langle\mathcal{D}\rangle^2 + \lambda_{i,j_i}} = \int_0^\infty dt_{i,j_i}\; e^{-t_{i,j_i}(H+\omega\Sigma+1+\lambda_{i,j_i})} \tag{78}$$

in (74) and integrate over λ_{i,j_i}:

$$\begin{aligned} &(\delta^{n_i} f_i)(y_i,y_{i+1}) \\ &= \sum_{k_i=0}^{n_i} \binom{n_i}{k_i} 2^{k_i} \frac{(-1)^{n_i}}{(2\sqrt{\pi})^{n_i}} \int_0^\infty \prod_{j_i=1}^{n_i} \frac{dt_{i,j_i} ds_{i,j_i}}{(t_{i,j_i}+s_{i,j_i})^{\frac{3}{2}}} e^{-(1+\omega\Sigma)(S_i+T_i)} \\ &\times \int_{\mathbb{R}^d} dx_i\, e^{-S_i H}(y_i,x_i)(\Delta^{n_i-k_i}\partial^{\mu_1^i}\cdots\partial^{\mu_{k_i}^i} f_i)(x_i) \frac{\partial^{k_1}}{\partial x_i^{\mu_1^i}\cdots\partial x_i^{\mu_{k_i}^i}} e^{-T_i H}(x_i,y_{i+1}) \,, \end{aligned} \tag{79}$$

where $S_i := \sum_{j_i=1}^{n_i} s_{j_i}$ and $T_i := \sum_{j_i=1}^{n_i} t_{j_i}$. We insert this into (75), move $e^{-S_1 H}$ under the trace to the end, and perform the y_i-integrations which combine the Mehler kernels into $e^{-\frac{\tau_i}{2\omega}H}(x_i,x_{i+1})$, with $\tau_i = 2\omega(T_i + S_{i+1} + \delta_{iv}t_0)$ and the convention $v+1\equiv 1$. The remaining trace in $\bigwedge(\mathbb{C}^d)$ is

$$\mathrm{tr}(e^{-\Sigma y}) = \mathrm{tr}(e^{-y[b^{\dagger 1},b^1]}\cdots e^{-y[b^{\dagger d},b^d]}) = (2\cosh y)^d \,. \tag{80}$$

Now the k_i partial derivatives of the Mehler kernel read

(81)
$$\begin{aligned}&\sum_{k_i=0}^{n_i}\binom{n_i}{k_i}2^{k_i}(-1)^{n_i}(\Delta^{n_i-k_i}\partial^{\mu^i_1}\cdots\partial^{\mu^i_{k_i}}f_i)(x_i)\frac{\partial^{k_1}}{\partial x_i^{\mu^i_1}\cdots\partial x_i^{\mu^i_{k_i}}}e^{-\frac{\tau_i H}{2\omega}}(x_i,x_{i+1})\\&=\sum_{k_i+2l_i+r_i=n_i}\frac{n_i!}{l_i!k_i!r_i!}\omega^{n_i-k_i-l_i}(-1)^{k_i+l_i}2^{l_i}\coth^{l_i}(\tau_i)\big(\Delta^{k_i+l_i}\partial^{\mu^i_1}\cdots\partial^{\mu^i_{r_i}}f_i\big)(x_i)\\&\qquad\times\Big(\prod_{j=1}^{r_i}\big((x_i-x_{i+1})\coth\tfrac{\tau_i}{2}+(x_i+x_{i+1})\tanh\tfrac{\tau_i}{2}\big)_{\mu^i_j}\Big)e^{-\frac{\tau_i H}{2\omega}}(x_i,x_{i+1})\,.\end{aligned}$$

We represent the f_i by their Fourier transforms $f_i(x)=\int_{\mathbb{R}^d}\frac{dp_i}{(2\pi)^d}\;\hat{f}_i(p_i)\,e^{\mathrm{i}p_ix_i}$, write the x_i,x_{i+1} in (81) as derivatives with respect to p_i,p_{i+1}, respectively, and obtain after Gaußian integration of the x_i

$$\begin{aligned}\zeta_\phi(z)=&\sum_{\substack{k_1+2l_1+r_1=n_1,\dots,\\k_v+2l_v+r_v=n_v}}\Big(\prod_{i=1}^v\frac{n_i!}{l_i!k_i!r_i!}\omega^{n_i-k_i}\Big)\frac{1}{\Gamma(\frac{z}{2})(2\sqrt{\pi})^{n_1+\dots+n_v}}\\&\times\int_0^\infty dt_0\;t_0^{\frac{z}{2}-1}\int_0^\infty\prod_{i=1}^v\prod_{j_i=1}^{n_i}\frac{dt_{i,j_i}ds_{i,j_i}}{(t_{i,j_i}+s_{i,j_i})^{\frac{3}{2}}}e^{-(t_0+\sum_{i=1}^n(S_i+T_i))}\\&\times\big(2\cosh\tfrac{\tau_1+\dots+\tau_v}{2}\big)^d\Big(\prod_{i=1}^v\Big(\frac{2}{\omega}\coth\tau_i\Big)^{l_i}\Big(\frac{\omega}{2\pi\sinh\tau_i}\Big)^{\frac{d}{2}}\Big)\\&\times\int_{(\mathbb{R}^d)^v}\Big(\prod_{i=1}^v\frac{dp_i}{(2\pi)^d}\Big)\Big(\prod_{i=1}^v(p_i^2)^{k_i+l_i}p_i^{\mu^i_1}\cdots p_i^{\mu^i_{r_i}}\hat{f}_i(p_i)\Big)\\&\qquad\times\Big(\prod_{i=1}^v\prod_{j=1}^{r_i}P_{\mu^i_j}\Big(\tau_i;\frac{\partial}{\partial p_i},\frac{\partial}{\partial p_{i+1}}\Big)\Big)\Big(\frac{\sqrt{\pi}^{dv}\,e^{-\frac{1}{4}pQ^{-1}p}}{(\det Q)^{\frac{d}{2}}}\Big)\,,\end{aligned}\tag{82}$$

where

$$P_{\mu^i_j}\Big(\tau_i;\frac{\partial}{\partial p_i},\frac{\partial}{\partial p_{i+1}}\Big):=\coth\frac{\tau_i}{2}\Big(\frac{\partial}{\partial p_i^{\mu^i_j}}-\frac{\partial}{\partial p_{i+1}^{\mu^i_j}}\Big)+\tanh\frac{\tau_i}{2}\Big(\frac{\partial}{\partial p_i^{\mu^i_j}}+\frac{\partial}{\partial p_{i+1}^{\mu^i_j}}\Big)\tag{83}$$

and

(84)
$$Q=\frac{\omega}{2}\begin{pmatrix}\frac{\sinh(\tau_v+\tau_1)}{\sinh\tau_v\sinh\tau_1} & \frac{-1}{\sinh\tau_1} & 0 & \cdots & 0 & \frac{-1}{\sinh\tau_v}\\ \frac{-1}{\sinh\tau_1} & \frac{\sinh(\tau_1+\tau_2)}{\sinh\tau_1\sinh\tau_2} & \frac{-1}{\sinh\tau_2} & \ddots & \ddots & 0\\ 0 & \frac{-1}{\sinh\tau_2} & \frac{\sinh(\tau_2+\tau_3)}{\sinh\tau_2\sinh\tau_3} & \ddots & \ddots & 0\\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots\\ 0 & \ddots & \ddots & \ddots & \frac{\sinh(\tau_{v-2}+\tau_{v-1})}{\sinh\tau_{v-2}\sinh\tau_{v-1}} & \frac{-1}{\sinh\tau_{v-1}}\\ \frac{-1}{\sinh\tau_v} & 0 & 0 & \cdots & \frac{-1}{\sinh\tau_{v-1}} & \frac{\sinh(\tau_{v-1}+\tau_v)}{\sinh\tau_{v-1}\sinh\tau_v}\end{pmatrix}.$$

By Gauß-Jordan elimination and multiple use of the addition theorems for sinh it is straightforward to compute the determinant and the inverse of the symmetric matrix Q (the result also holds for $v=1$):

$$\det Q = \Big(\frac{\omega}{2}\Big)^v \frac{4\sinh^2(\frac{1}{2}(\tau_1+\dots+\tau_v))}{\prod_{i=1}^v \sinh\tau_i}\;, \tag{85}$$

$$(Q^{-1})_{ij} = \frac{1}{\omega\tanh(\frac{1}{2}(\tau_1+\dots+\tau_v))} + \tilde{Q}_{ij}\;, \tag{86}$$

$$\tilde{Q}_{ij} = -\frac{2\sinh(\frac{1}{2}(\tau_i+\dots+\tau_{j-1}))\sinh(\frac{1}{2}(\tau_j+\dots+\tau_{i-1}))}{\omega\sinh(\frac{1}{2}(\tau_1+\dots+\tau_v))}\;, \tag{87}$$

where in $\tilde{Q}_{ij}$ one of the chains $\tau_i+\dots+\tau_{j-1}$ or $\tau_j+\dots+\tau_{i-1}$ passes through the index $v\equiv 0$. The determinant can also be obtained from the fact that for $p=0$ we just have the trace over the concatenation of Mehler kernels (77).

The action of $(P_{\mu^i_j})$ on $e^{-\frac{1}{4}pQ^{-1}p}$ is partitioned into k'_i out of r_i single contractions, l'_i double contractions and r_{ij} halves of mixed contraction with another index $j\neq i$ such that $k'_i+l'_i+\sum_{j\neq i} r_{ij} = r_i$. Their number is $\frac{r_i!}{2^{l'_i}l'_i!k'_i!r_{i1}!\cdots r_{iv}!}$ if we put $r_{ii}=0$ and $r_{ij}=r_{ji}$. Together with the multiplying factor $p_i^{\mu^i_j}$, a single contraction gives a factor

$$p_i^{\mu^i}P_{\mu^i}(-\tfrac{1}{4}pQ^{-1}p) = -\frac{p_i^2}{\omega} - \sum_{j\neq i}\frac{\sinh(\frac{\tau_i+\dots+\tau_{j-1}-\tau_j-\dots-\tau_{i-1}}{2})}{\omega\sinh(\frac{\tau_1+\dots+\tau_v}{2})}p_ip_j\;. \tag{88}$$

A double contraction with respect to the same index i gives a factor

$$p_i^{\mu^i}p_i^{\nu^i}P_{\mu^i}P_{\nu^i}(-\tfrac{1}{4}pQ^{-1}p) = \Big(-\frac{4\coth\tau_i}{\omega}+\frac{2}{\omega}\coth(\tfrac{\tau_1+\dots+\tau_v}{2})\Big)p_i^2\;. \tag{89}$$

A mixed contraction with respect to different indices $i\neq j$ gives a factor

$$p_i^{\mu^i}p_j^{\mu^j}P_{\mu^i}P_{\mu^j}(-\tfrac{1}{4}pQ^{-1}p) = 2\frac{\cosh(\frac{\tau_j+\dots+\tau_{i-1}-\tau_i-\dots-\tau_{j-1}}{2})}{\omega\sinh(\frac{\tau_1+\dots+\tau_v}{2})}p_ip_j\;. \tag{90}$$

We insert these formulae into (82) and notice that the sum over l_i, l'_i combines to a joint sum (with new index l_i) involving only the factor $\frac{p_i^2}{\omega}\coth(\frac{1}{2}(\tau_1+\dots+\tau_v))$ from (89), whereas $\coth\tau_i$ cancels. In the same way, the sum over k_i,k'_i cancels the term $-p_i^2$ from (88) so that only the sum over $j\neq i$ remains:

(91)

$$\begin{aligned}\zeta_\phi(z) = &\sum_{\substack{k_1+2l_1+r_1=n_1,\dots,\\ k_v+2l_v+r_v=n_v\\ r_1+\dots+r_v=2m}}\;\sum_{\substack{r_{11}+\dots+r_{1v}=r_1,\dots,\\ r_{v1}+\dots+r_{vv}=r_v}}\Big(\prod_{i=1}^v\frac{n_i!}{l_i!k_i!}\Big)\frac{2^m\omega^{l_1+\dots+l_v+m}}{\Gamma(\frac{z}{2})(2\sqrt{\pi})^{n_1+\dots+n_v}}\\
&\times\int_0^\infty dt_0\,t_0^{\frac{z}{2}-1}\int_0^\infty\prod_{i=1}^v\prod_{j_i=1}^{n_i}\frac{dt_{i,j_i}ds_{i,j_i}}{(t_{i,j_i}+s_{i,j_i})^{\frac{3}{2}}}\frac{e^{-(t_0+\sum_{i=1}^n(S_i+T_i))}}{\big(\tanh\frac{\tau_1+\dots+\tau_v}{2}\big)^{d+l_1+\dots+l_v}}\\
&\times\int_{(\mathbb{R}^d)^v}\Big(\prod_{i=1}^v\frac{dp_i}{(2\pi)^d}\Big)\Big(\prod_{i<j}\frac{1}{r_{ij}!}\Big(\frac{\cosh\big(\frac{\tau_j+\dots+\tau_{i-1}}{2}-\frac{\tau_i+\dots+\tau_{j-1}}{2}\big)}{\sinh(\frac{\tau_1+\dots+\tau_v}{2})}p_ip_j\Big)^{r_{ij}}\Big)\\
&\times\Big(\prod_{i=1}^v\Big(\sum_{j\neq i}\frac{\sinh\big(\frac{\tau_j+\dots+\tau_{i-1}}{2}-\frac{\tau_i+\dots+\tau_{j-1}}{2}\big)}{\sinh(\frac{\tau_1+\dots+\tau_v}{2})}p_ip_j\Big)^{k_i}(p_i^2)^{l_i}\hat{f}_i(p_i)\Big)e^{-\frac{1}{4}pQ^{-1}p}.\end{aligned}$$

The zeta-function potentially has a singularity for $\tau = \tau_1 + \cdots + \tau_v \to 0$ of order $\tau^{\frac{z+k_1+\cdots+k_v}{2}-d}$. The contribution $\frac{z}{2}$ is from $dt_0\ t_0^{\frac{z}{2}-1}$, the measure $\prod \frac{dtds}{(t+s)^{\frac{3}{2}}}$ contributes $\frac{n_1+\cdots+n_v}{2}$, and $(\tanh\frac{\tau}{2})^{-d-l_1-\cdots-l_v}(\sinh\frac{\tau}{2})^{-r_{12}+\cdots+r_{v-1,v}}$ contribute $-(d+l_1+\cdots+l_v+\frac{r_1+\cdots+r_v}{2})$. However, the independence of the leading term in Q^{-1}_{ij} from i,j shows that this singularity is protected by $e^{-\frac{(p_1+\cdots+p_v)^2}{\omega\tanh\frac{\tau}{2}}}$ unless the total momentum is conserved, $p_1+\cdots+p_v=0$. The remaining singularity is identified by a Taylor expansion in p_v about $\bar{p}_v := -(p_1+\cdots+p_{v-1})$ up to order ρ to be determined later:

$$
\begin{aligned}
(92)\quad & F(p_1,\dots,p_v)\\
&= \sum_{|\alpha|\le\rho} \frac{(p_v-\bar{p}_v)^\alpha}{\alpha!}\frac{\partial^{|\alpha|}F}{\partial p_v^\alpha}(p_1,\dots,p_{v-1},\bar{p}_v)\\
&+ \sum_{|\alpha|=\rho+1}\frac{(p_v-\bar{p}_v)^\alpha}{\rho!}\int_0^1 d\lambda\,(1-\lambda)^\rho\frac{\partial^{|\alpha|}F}{\partial p_v^\alpha}(p_1,\dots,p_{v-1},\bar{p}_v+\lambda(p_v-\bar{p}_v))\,,
\end{aligned}
$$

where α is a multi-index. Together with the measure dp_v, the last line combines with $\tanh^{-d}(\frac{\tau}{2})$ to a factor $dP\ P^{\rho+1}\,e^{-P^2}\tanh^{\frac{\rho+1-d}{2}}(\frac{\tau}{2})$, where $P=\frac{p_v-\bar{p}_v}{\sqrt{\tanh\frac{\tau}{2}}}$. For sufficiently large but finite ρ we shall see in (96) that the potential singularity in $t_0^{\frac{z}{2}}$ is cancelled so that the last line of (92) is regular. The bilinear form in the exponent has the form

(93)

$$
e^{-\frac{1}{4}pQ^{-1}p} = e^{-\frac{(p_v-\bar{p}_v)^2}{4\omega\tanh\frac{\tau}{2}}-\frac{1}{2}(p_v-\bar{p}_v)q-\frac{1}{2}\bar{p}_v\sum_{j=1}^{v-1}\tilde{Q}_{vj}p_j-\frac{1}{4}\sum_{i,j=1}^{v-1}\tilde{Q}_{ij}p_ip_j}\,,\qquad q:=\sum_{j=1}^{v-1}\tilde{Q}_{vj}p_j\,.
$$

We can thus perform the Gaußian integration over p_v and obtain for the restricted zeta function ζ^r, where the second line of (92) is removed:

(94)

$$
\begin{aligned}
&\zeta^r_\phi(z)\\
&= \sum_{\substack{k_1+2l_1+r_{11}+\ldots+r_{1v}\,=\,n_1,\,\ldots,\\ k_1+2l_1+r_{v1}+\ldots+r_{vv}\,=\,n_v,\\ r_{ii}\,=\,0\,,\ r_{ij}\,=\,r_{ji}}} \frac{n_1!\cdots n_v!}{\Gamma(\frac{z}{2})\pi^{\frac{d}{2}}(2\sqrt{\pi})^{n_1+\cdots+n_v}}\\
&\times\int_0^\infty dt_0\ t_0^{\frac{z}{2}-1}\int_0^\infty\prod_{i=1}^v\prod_{j_i=1}^{n_i}\frac{dt_{i,j_i}ds_{i,j_i}}{(t_{i,j_i}+s_{i,j_i})^{\frac{3}{2}}}\ e^{-t}\Big(\frac{\omega}{\tanh(\omega t)}\Big)^{\frac{d}{2}+\sum_{i=1}^v l_i+\sum_{i<j}r_{ij}}\\
&\times\int_{(\mathbb{R}^d)^{v-1}}\Big(\prod_{i=1}^{v-1}\frac{dp_i}{(2\pi)^d}\Big)\sum_{|\alpha|\le\rho}\frac{(-2)^{|\alpha|}}{\alpha!}\frac{\partial^{|\alpha|}}{\partial q^\alpha}\Big(e^{\frac{\omega}{4}q^2\tanh(\omega t)-\frac{1}{2}\bar{p}_v\sum_{j=1}^{v-1}\tilde{Q}_{vj}p_j-\frac{1}{4}\sum_{i,j=1}^{v-1}\tilde{Q}_{ij}p_ip_j}\Big)\\
&\times\frac{\partial^{|\alpha|}}{\partial p_v^\alpha}\Bigg(\Bigg(\prod_{i<j}\frac{\Big(\frac{2\cosh(\omega t_{ij})}{\cosh(\omega t)}p_ip_j\Big)^{r_{ij}}}{r_{ij}!}\Bigg)\Bigg(\prod_{i=1}^v\frac{\Big(\sum_{j\neq i}\frac{\sinh(\omega t_{ij})}{\sinh(\omega t)}p_ip_j\Big)^{k_i}}{k_i!}\frac{(p_i^2)^{l_i}}{l_i!}\hat{f}_i(p_i)\Bigg)\Bigg)_{p_v\mapsto\bar{p}_v}\,,
\end{aligned}
$$

where $t = \frac{1}{2\omega}\tau = t_0 + \sum_{i=1}^{v}(T_i + S_i)$ and $t_{ij} = \frac{1}{2\omega}(\tau_j + \cdots + \tau_{i-1}) - \frac{1}{2\omega}(\tau_i + \cdots + \tau_{j-1})$. The q-derivatives and the quadratic form in the exponent become with $\tilde{Q}_{ij} = \frac{\cosh(\omega t_{ij}) - \cosh(\omega t)}{\omega \sinh(\omega t)}$

$$
\begin{aligned}
(95)\qquad & \sum_{|\alpha|\leq\rho} \frac{(-2)^{|\alpha|}}{\alpha!}\frac{\partial^{|\alpha|}}{\partial q^\alpha}\Big(e^{\frac{\omega}{4}q^2\tanh(\omega t) - \frac{1}{2}\bar{p}_v\sum_{j=1}^{v-1}\tilde{Q}_{vj}p_j - \frac{1}{4}\sum_{i,j=1}^{v-1}\tilde{Q}_{ij}p_ip_j}\Big)\frac{\partial^{|\alpha|}}{\partial p_v^\alpha} \\
&= \sum_{|\alpha|+2a\leq\rho} (\omega\tanh(\omega t))^{a+|\alpha|} e^{-\Big(\sum_{i,j=1}^{v-1}\frac{\sinh(\omega t_{ij}^-)\sinh(\omega t_{ij}^+)}{2\omega\sinh(2\omega t)}p_ip_j\Big)} \\
&\qquad \times \frac{1}{a!}\Big(\frac{\partial^2}{\partial p_v^\mu \partial p_{v\mu}}\Big)^a \frac{1}{\alpha!}\Big(\sum_{j=1}^{v-1}\frac{2\sinh(\omega\frac{t+t_{vj}}{2})\sinh(\omega\frac{t+t_{jv}}{2})}{\omega\sinh(\omega t)}p_j\Big)^\alpha \frac{\partial^{|\alpha|}}{\partial p_v^\alpha}\;,
\end{aligned}
$$

where $t_{ij}^- = t + t_{kv}\big|_{k=\min(i,j)}$ and $t_{ij}^+ = t + t_{vk}\big|_{k=\max(i,j)}$. Note that (95) is bounded for all t.

We insert (95) into (94). We change the integration variables to $t_0 = (1-u)t$, $\sum_{i=1}^{v}(S_i + T_i) = ut$ with integration over t from 0 to ∞, over u from 0 to 1 and over the surface Δ given by $\sum_{i=1}^{v}(S_i + T_i) = 1$. We write the denominators $\frac{1}{\sinh(\omega t)} = \frac{1}{\omega t}\cdot\frac{\omega t}{\sinh(\omega t)}$ and $\frac{1}{\tanh(\omega t)} = \frac{1}{\omega t}\cdot\frac{\omega t}{\tanh(\omega t)}$ and expand the bounded (at 0) fractions $\frac{\omega t}{\sinh(\omega t)}$ and $\frac{\omega t}{\tanh(\omega t)}$ into a Taylor series in (ωt). The numerators in hyperbolic functions of (ωt) and (ωt_{ij}) and $\frac{1}{\cosh(\omega t)}$ are expanded into a Taylor series in their arguments. Then, for each term in the sum, the u,t-integral is of the form

$$
\begin{aligned}
(96)\quad & \frac{1}{\Gamma(\frac{z}{2})}\int_0^\infty dt\; t^{(\frac{z}{2}-\frac{d}{2}+\frac{k_1+\cdots+k_v}{2}+a+2|\alpha|+b-1)}e^{-t}\int_0^1 du(1-u)^{\frac{z}{2}-1}u^{\frac{n_1+\cdots+n_v}{2}+c-1} \\
&= \frac{\Gamma(\frac{z}{2}-\frac{d}{2}+\frac{k_1+\cdots+k_v}{2}+a+2|\alpha|+b)\,\Gamma(\frac{n_1+\cdots+n_v}{2}+c)}{\Gamma(\frac{z}{2}+\frac{n_1+\cdots+n_v}{2}+c)}\;,
\end{aligned}
$$

where the integers $b \geq c \geq 0$ arise from the Taylor expansion. The remaining integration over the simplex Δ is regular because from the Taylor expansion only positive powers of the integration variables appear. From (96) we deduce the following information about the pole structure:

- For $z \notin \mathbb{Z}$ or for $z > d$ there is no pole.
- For $z = d - N$ with $N \in \mathbb{N}$, and $n_1, \ldots, n_v$ such that $z + n_1 + \cdots + n_v$ is even, there is a pole for a finite (and non-vanishing) number of index combinations and finite Taylor order $\rho = d + n_1 + \cdots + n_v - k_1 - \cdots - k_v$.

This concludes the proof that $\mathrm{Sd} = d - \mathbb{N}$.

It remains to characterise the nature of the residues. From (94) we conclude that the residues are given by the integral over $p_1, \ldots, p_{v-1}$ of an integrand which is a polynomial in $p_1, \ldots, p_{v-1}$ times $\prod_{i=1}^{v-1}\hat{f}_i(p_i)$ times possible derivatives of $\hat{f}_v(\bar{p}_v)$. Reconstructing the p_v-variable by a δ-function and integrating by parts the derivatives of $\hat{f}_v(\bar{p}_v)$, the residue becomes a finite sum of the form

$$(97)\qquad \mathrm{res}_{z=d-N}(\zeta(z))$$

$$= \sum_{\alpha_0,\dots,\alpha_v} c_{\alpha_0\dots\alpha_v} \int_{(\mathbb{R}^d)^v} \Big(\prod_{i=1}^v \frac{dp_i}{(2\pi)^d}\Big) \int_{\mathbb{R}^d} dx\; e^{\mathrm{i}(p_1+\dots+p_v)x} x^{\alpha_0} \prod_{i=1}^v p_i^{\alpha_i} \hat{f}_i(p_i)$$

$$= \sum_{\alpha_0,\dots,\alpha_v} \int_{\mathbb{R}^d} dx\; c_{\alpha_0\dots\alpha_v} (-\mathrm{i})^{|\alpha_1|+\dots+|\alpha_v|} x^{\alpha_0} \prod_{i=1}^v (\partial^{\alpha_i} f_i)(x_i)\;,$$

where the α_j are multi-indices which contract to a Lorentz scalar. The prefactor $c_{\alpha_0\dots\alpha_v}$ results from the integration over the t-variables. Thus, the residues are local. □

We would like to stress that it was important to keep track of the combinatorial factors which led to the cancellation of denominators $\frac{1}{\sinh\tau_i}$. Such denominators in the final formula (94) would be fatal because in that case the u-integral of (96) would produce a hypergeometric function instead of the beta function and therefore an infinite sum for the residue, which could be non-local.

Appendix B. Vertices contributing to the spectral action

We compute here the individual vertex contributions (54) to the spectral action. This is done by inserting the vertices (57) into (60) and then computing the t_i-integrals.

B.1. V_2**.** The contribution of a single V_2-vertex is

$$S_t(V_2) = \int_0^t dt_1\; \mathrm{tr}(e^{-\omega\Sigma t}) S_t^0(f)\;, \quad f = -2|\phi|^2 - A_\mu A^\mu - B_\mu B^\mu\;. \tag{98}$$

With $\mathrm{tr}(e^{-\omega\Sigma t}) = (2\cosh(\omega t))^4$ we have after second order Taylor expansion, ignoring the remainder and the odd first-order term,

$$\begin{aligned} S_t(V_2) &= \int_{\mathbb{R}^4} \frac{dp}{(2\pi)^4} \frac{t}{(\tanh(\omega t))^4} \Big(\hat{f}(0) + \frac{1}{2} p_\mu p_\nu \frac{\partial^2 \hat{f}}{\partial p_\mu \partial p_\nu}(0)\Big) e^{-\frac{p^2}{4\omega\tanh(\omega t)}} \\ &= \frac{\omega^2 t}{\pi^2 \tanh^2(\omega t)} \Big(\hat{f}(0) + \omega\tanh(\omega t)\delta_{\mu\nu} \frac{\partial^2 \hat{f}}{\partial p_\mu \partial p_\nu}(0)\Big) \\ &= \frac{\omega^2 t}{\pi^2 \tanh^2(\omega t)} \int_{\mathbb{R}^4} dx \Big(f(x) - \omega\|x\|^2 \tanh(\omega t)\, f(x)\Big)\;, \end{aligned} \tag{99}$$

after Fourier transformation $\hat{f}(p) = \int_{\mathbb{R}^4} dx\; e^{-ipx} f(x)$. Inserting f we obtain after Laplace transformation the leading terms of the asymptotic expansion to

$$\begin{aligned} S_2(\mathcal{D}_A) &= \frac{\chi_{-1}}{\pi^2} \int_{\mathbb{R}^4} dx\; \big(-2|\phi|^2 - A_\mu A^\mu - B_\mu B^\mu\big)(x) \\ &\quad + \frac{\chi_0}{\pi^2} \int_{\mathbb{R}^4} dx\; (\omega^2 |x|^2 \big(2|\phi|^2 + A_\mu A^\mu + B_\mu B^\mu\big)(x)\;. \end{aligned} \tag{100}$$

B.2. $V_1 V_1$**.** The contribution of two V_1-vertices is

$$S_t(V_1, V_1) = \int_0^t dt_1 \int_0^{t-t_1} dt_2\; \mathrm{tr}(e^{-\omega\Sigma t}) S^{1,1}_{t_2, t-t_2}(-A, -A) + (A \mapsto B)\;. \tag{101}$$

This is the most involved computation. To (60) there are the two contributions $k_1 = k_2 = 1$ up to order 0 and $r_{12} = r_{21} = 1$ with Taylor expansion about $p_2 = -p_1$ up to order 2:

(102)

$$
\begin{aligned}
&S_t(V_1,V_1)\\
&=\int_0^t dt_1\int_0^{t-t_1}dt_2\int_{(\mathbb{R}^4)^2}\frac{dp_1dp_2}{(2\pi)^8}\frac{1}{\tanh^4(\omega t)}\;e^{-\frac{(p_1+p_2)^2}{4\omega\tanh(\omega t)}+p_1p_2\frac{\sinh(\omega t_2)\sinh(\omega(t-t_2))}{\omega\sinh(\omega t)}}\\
&\times\Big\{\hat{A}^\mu(p_1)\hat{A}^\nu(-p_1)\Big(\frac{\sinh^2(\omega(t-2t_2))}{\sinh^2(\omega t)}p_{1\mu}p_{1\nu}+2\omega\delta_{\mu\nu}\frac{\cos(\omega(t-2t_2))}{\sinh(\omega t)}\Big)\\
&+(p_1+p_2)^\rho\hat{A}^\mu(p_1)\frac{\partial\hat{A}^\nu}{\partial p_2^\rho}(-p_1)\cdot 2\omega\delta_{\mu\nu}\frac{\cos(\omega(t-2t_2))}{\sinh(\omega t)}\\
&+\frac{1}{2}(p_1+p_2)^\rho(p_1+p_2)^\sigma\hat{A}^\mu(p_1)\frac{\partial^2\hat{A}^\nu}{\partial p_2^\rho\partial p_2^\sigma}(-p_1)\cdot 2\omega\delta_{\mu\nu}\frac{\cos(\omega(t-2t_2))}{\sinh(\omega t)}\Big\}+(A\mapsto B)\\
&=\int_0^t dt_1\int_0^{t-t_1}dt_2\int_{\mathbb{R}^4}\frac{dp_1}{(2\pi)^4}\frac{\omega^2}{\pi^2\tanh^2(\omega t)}\;e^{-\frac{\sinh(2\omega t_2)\sinh(2\omega(t-t_2))}{2\omega\sinh(2\omega t)}p_1^2}\\
&\times\Big\{\hat{A}^\mu(p_1)\hat{A}^\nu(-p_1)\Big(\frac{\sinh^2(\omega(t-2t_2))}{\sinh^2(\omega t)}p_{1\mu}p_{1\nu}+2\omega\delta_{\mu\nu}\frac{\cos(\omega(t-2t_2))}{\sinh(\omega t)}\Big)\\
&+4\omega p_1^\rho\hat{A}^\mu(p_1)\frac{\partial\hat{A}_\mu}{\partial p_2^\rho}(-p_1)\cdot\frac{\sinh(\omega t_2)\sinh(\omega(t-t_2))}{\cos(\omega t)}\frac{\cos(\omega(t-2t_2))}{\sinh(\omega t)}\\
&+\omega\Big(2\delta^{\rho\sigma}\omega\tanh(\omega t)+4p_1^\rho p_1^\sigma\frac{\sinh^2(\omega t_2)\sinh^2(\omega(t-t_2))}{\cos^2(\omega t)}\Big)\\
&\qquad\times\hat{A}^\mu(p_1)\frac{\partial^2\hat{A}_\mu}{\partial p_2^\rho\partial p_2^\sigma}(-p_1)\cdot\frac{\cos(\omega(t-2t_2))}{\sinh(\omega t)}\Big\}+(A\mapsto B)\;.
\end{aligned}
$$

Up to $\mathcal{O}(t)$ this reduces to

(103)

$$
\begin{aligned}
&S_t(V_1,V_1)\\
&=\int_0^t dt_1\int_0^{t-t_1}dt_2\int_{\mathbb{R}^4}\frac{dp_1}{(2\pi)^4}\frac{\omega^2}{\pi^2\tanh^2(\omega t)}\\
&\times\Big\{\hat{A}^\mu(p_1)\hat{A}^\nu(-p_1)\Big(\frac{\sinh^2(\omega(t-2t_2))}{\sinh^2(\omega t)}p_{1\mu}p_{1\nu}\\
&\qquad-\delta_{\mu\nu}\frac{\cosh(\omega(t-2t_2))}{\sinh(\omega t)}\frac{\sinh(2\omega t_2)\sinh(2\omega(t-t_2))}{\sinh(2\omega t)}p_1^2\Big)\\
&+2\omega\hat{A}^\mu(p_1)\hat{A}_\mu(-p_1)\frac{\cosh(\omega(t-2t_2))}{\sinh(\omega t)}+2\omega^2\hat{A}^\mu(p_1)\frac{\partial^2\hat{A}_\mu}{\partial p_2^\rho\partial p_{2\rho}}(-p_1)\frac{\cosh(\omega(t-2t_2))}{\cosh(\omega t)}\Big\}\\
&+(A\mapsto B)\\
&=\int_{\mathbb{R}^4}\frac{dp_1}{(2\pi)^4}\frac{\omega^2}{\pi^2\tanh^2(\omega t)}\\
&\times\Big\{\hat{A}^\mu(p_1)\hat{A}^\nu(-p_1)\Big(\Big(\frac{t}{4\omega\tanh(\omega t)}-\frac{t^2}{4\sinh^2(\omega t)}\Big)p_{1\mu}p_{1\nu}-\delta_{\mu\nu}\frac{t\tanh(\omega t)}{6\omega}p_1^2\Big)
\end{aligned}
$$

$$+ t\,\hat{A}^\mu(p_1)\hat{A}_\mu(-p_1) + (\omega t)\tanh(\omega t)\,\hat{A}^\mu(p_1)\frac{\partial^2 \hat{A}_\mu}{\partial p_2^\rho \partial p_{2\rho}}(-p_1)\Big\} + (A \mapsto B)\;.$$

After Fourier and Laplace transformation, the leading contribution to the spectral action becomes

$$\begin{aligned}(104)\qquad S_{11}(\mathcal{D}_A) &= \frac{\chi_{-1}}{\pi^2}\int_{\mathbb{R}^4} dx\,\big(A_\mu A^\mu + B_\mu B^\mu\big)(x)\\ &\quad - \frac{\chi_0}{\pi^2}\int_{\mathbb{R}^4} dx\,(\omega^2\|x\|^2)\big(A_\mu A^\mu + B_\mu B^\mu\big)(x)\\ &\quad - \frac{\chi_0}{12\pi^2}\int_{\mathbb{R}^4} dx\,\big(F^A_{\mu\nu}F^{A\mu\nu} + F^B_{\mu\nu}F^{B\mu\nu}\big)(x)\;.\end{aligned}$$

B.3. V_2V_2**,** V_3V_3**,** V_4V_4**.** We have

$$\begin{aligned}(105)\qquad &\sum_{i=2}^4 S_t(V_i,V_i)\\ &= \int_0^t dt_1\int_0^{t-t_1} dt_2\Big\{\mathrm{tr}(e^{-\omega\Sigma t})\Big(S^{0,0}_{t_2,t-t_2}(-|\phi|^2 - A_\mu A^\mu, -|\phi|^2 - A_\mu A^\mu)\\ &\qquad\qquad + S^{0,0}_{t_2,t-t_2}(-|\phi|^2 - B_\mu B^\mu, -|\phi|^2 - B_\mu B^\mu)\Big)\\ &\quad + \mathrm{tr}\Big(\frac{\mathrm{i}}{4}[b^{\dagger\mu} - b^\mu, b^{\dagger\nu} - b^\nu]e^{-\omega\Sigma t_2}\frac{\mathrm{i}}{4}[b^{\dagger\rho} - b^\rho, b^{\dagger\sigma} - b^\sigma]e^{-\omega\Sigma(t-t_2)}\Big)\\ &\qquad \times\Big(S^{0,0}_{t_2,t-t_2}(F^A_{\mu\nu},F^A_{\rho\sigma}) + S^{0,0}_{t_2,t-t_2}(F^B_{\mu\nu},F^B_{\rho\sigma})\Big)\\ &\quad + \mathrm{tr}\Big((b^{\dagger\mu} - b^\mu)e^{-\omega\Sigma t_2}(b^{\dagger\nu} - b^\nu)e^{-\omega\Sigma(t-t_2)}\Big)\\ &\qquad \times\Big(S^{0,0}_{t_2,t-t_2}(-D_\mu\phi,\overline{D_\nu\phi}) + S^{0,0}_{t_2,t-t_2}(\overline{D_\mu\phi},-D_\nu\phi)\Big)\Big\}\;.\end{aligned}$$

Since the $S^{0,0}_{t_2,t-t_2}$ are at least $\mathcal{O}(t^{-2})$, only the $\mathcal{O}(t^0)$-parts of $e^{-\omega\Sigma t_2}$ and $e^{-\omega\Sigma(t-t_2)}$ will contribute to the spectral action. Now the traces in $\bigwedge(\mathbb{C}^4)$ are easy to compute:

$$\begin{aligned}(106)\qquad \mathrm{tr}\big(e^{\omega\Sigma t}\big) &= (2\cosh(\omega t))^4\;,\\ \mathrm{tr}\big((b^{\dagger\mu} - b^\mu)e^{-\omega\Sigma t_2}(b^{\dagger\nu} - b^\nu)e^{-\omega\Sigma(t-t_2)}\big) &= -16\delta^{\mu\nu} + \mathcal{O}(t)\;,\\ \mathrm{tr}\big(\tfrac{\mathrm{i}}{4}[b^{\dagger\mu}-b^\mu, b^{\dagger\nu}-b^\nu]e^{-\omega\Sigma t_2}\tfrac{\mathrm{i}}{4}[b^{\dagger\rho}-b^\rho, b^{\dagger\sigma}-b^\sigma]e^{-\omega\Sigma(t-t_2)}\big) &= 8(\delta^{\mu\rho}\delta^{\nu\sigma}-\delta^{\mu\sigma}\delta^{\nu\rho})+\mathcal{O}(t).\end{aligned}$$

After Taylor expansion about $p_2 = -p_1$ up to order 0, integration over p_2, t_1, t_2 and Laplace transformation, we obtain

(107)

$$\begin{aligned}(S_{22}+S_{33}+S_{44})(\mathcal{D}_A) = \frac{\chi_0}{2\pi^2}\int_{\mathbb{R}^4} dx\Big\{&2\overline{D_\mu\phi}(D^\mu\phi) + (|\phi|^2 + A_\mu A^\mu)^2\\ &+ F^A_{\mu\nu}F^{A\mu\nu} + (|\phi|^2 + B_\mu B^\mu)^2 + F^B_{\mu\nu}F^{B\mu\nu}\Big\}(x)\;.\end{aligned}$$

B.4. V_1V_2**,** V_2V_1**.** With the abbreviation $f_{\phi A} := |\phi|^2 + A_\mu A^\mu$, we have

(108)

$$\begin{aligned}&S_t(V_1,V_2) + S_t(V_2,V_1)\\ &= \int_0^t dt_1\int_0^{t-t_1} dt_2\;\mathrm{tr}(e^{-\omega\Sigma t})\Big(S^{1,0}_{t_2,t-t_2}(-A,-f_{\phi A}) + S^{0,1}_{t_2,t-t_2}(-f_{\phi A},-A)\Big)\end{aligned}$$

$$
\begin{aligned}
&+(A\mapsto B)\\
&=\int_0^t dt_1\int_0^{t-t_1}dt_2\int_{\mathbb{R}^4\times\mathbb{R}^4}\frac{dp_1dp_2}{(2\pi)^8}\frac{1}{\tanh^4(\omega t)}\\
&\times\Big(p_{2,\mu}\hat{A}^\mu(p_1)\hat{f}_{\phi A}(p_2)-p_{1,\mu}\hat{A}^\mu(p_2)\hat{f}_{\phi A}(p_1)\Big)\frac{\sinh(\omega(t-2t_2))}{\sinh(\omega t)}e^{-\frac{1}{4}pQ^{-1}p}+(A\mapsto B)\\
&=\int_0^t dt_1\int_0^{t-t_1}dt_2\int_{\mathbb{R}^4}\frac{dp_1}{(2\pi)^4}\frac{\omega^2}{\pi^2\tanh^2(\omega t)}\frac{\sinh(\omega(t-2t_2))}{\sinh(\omega t)}\\
&\times\Big(-p_{1,\mu}\hat{A}^\mu(p_1)\hat{f}_{\phi A}(-p_1)-p_{1,\mu}\hat{A}^\mu(-p_1)\hat{f}_{\phi A}(p_1)\Big)+(A\mapsto B)+\mathcal{O}(t)\\
&=\mathcal{O}(t)\;.
\end{aligned}
$$

We thus have

$$S_{12}(\mathcal{D}_A)=0\;. \tag{109}$$

B.5. $V_1V_1V_2$**,** $V_1V_2V_1$**,** $V_2V_1V_1$**.** Only the $k_i=0$ terms in (60) contribute to the leading order. With the abbreviation $f_{\phi A}:=|\phi|^2+A_\mu A^\mu$, these give

(110)

$$
\begin{aligned}
&S_t(V_1,V_1,V_2)+S_t(V_1,V_2,V_1)+S_t(V_2,V_1,V_1)\\
&=\int_0^t dt_1\int_0^{t-t_1}dt_2\int_0^{t-t_1-t_2}dt_3\;\mathrm{tr}(e^{-\omega\Sigma t})\Big(S^{1,1,0}_{t_3,t_2,t-t_2-t_3}(-A,-A,-f_{\phi A})\\
&\qquad+S^{1,0,1}_{t_3,t_2,t-t_2-t_3}(-A,-f_{\phi A},-A)+S^{0,1,1}_{t_3,t_2,t-t_2-t_3}(-f_{\phi A},-A,-A)\Big)+(A\mapsto B)\\
&=\int_0^t dt_1\int_0^{t-t_1}dt_2\int_0^{t-t_1-t_2}dt_3\int_{(\mathbb{R}^4)^3}\frac{dp_1\,dp_2\,dp_3}{(2\pi)^{12}}\frac{-2\omega}{\tanh^4(\omega t)\sinh(\omega t)}\\
&\times\Big(\hat{A}_\mu(p_1)\hat{A}^\mu(p_2)\hat{f}_{\phi A}(p_3)\cosh(\omega(t-2t_3))\\
&\qquad+\hat{A}_\mu(p_1)\hat{f}_{\phi A}(p_2)\hat{A}^\mu(p_3)\cosh(\omega(t-2t_2-2t_3))\\
&\qquad+\hat{f}_{\phi A}(p_1)\hat{A}_\mu(p_2)\hat{A}^\mu(p_3)\cosh(\omega(t-2t_2))\Big)e^{-\frac{1}{4}pQ^{-1}p}+(A\mapsto B)+\mathcal{O}(t)\\
&=\int_0^t dt_1\int_0^{t-t_1}dt_2\int_0^{t-t_1-t_2}dt_3\int_{(\mathbb{R}^4)^2}\frac{dp_1\,dp_2}{(2\pi)^8}\frac{(-2\omega^3)}{\pi^2\tanh^2(\omega t)\sinh(\omega t)}\\
&\times\hat{A}_\mu(p_1)\hat{A}^\mu(p_2)\hat{f}_{\phi A}(-p_1-p_2)\Big(\cosh(\omega(t-2t_3))+\cosh(\omega(t-2t_2-2t_3))\\
&\qquad\qquad+\cosh(\omega(t-2t_2))\Big)+(A\mapsto B)+\mathcal{O}(t)\\
&=\int_{(\mathbb{R}^4)^2}\frac{dp_1\,dp_2}{(2\pi)^8}\frac{(-\omega^2t^2)}{\pi^2\tanh^2(\omega t)}\hat{A}_\mu(p_1)\hat{A}^\mu(p_2)\hat{f}_{\phi A}(-p_1-p_2)+(A\mapsto B)+\mathcal{O}(t)\;.
\end{aligned}
$$

After Fourier and Laplace transformation we obtain

$$S_{112}(\mathcal{D}_A)=-\frac{\chi_0}{\pi^2}\int_{\mathbb{R}^4}dx\Big\{A_\mu A^\mu(|\phi|^2+A_\nu A^\nu)+B_\mu B^\mu(|\phi|^2+B_\nu B^\nu)\Big\}(x)\;. \tag{111}$$

B.6. $V_1V_1V_1$**.** The leading order in (60) is given by the $(k_1=1, r_{23}=1)$ and the other two cyclic permutations:

(112)

$$
\begin{aligned}
&S_t(V_1,V_1,V_1)\\
&=\int_0^t dt_1\int_0^{t-t_1}dt_2\int_0^{t-t_1-t_2}dt_3\ \mathrm{tr}(e^{-\omega\Sigma t})S^{1,1,1}_{t_3,t_2,t-t_2-t_3}(-A,-A,-A)+(A\mapsto B)\\
&=\int_0^t dt_1\int_0^{t-t_1}dt_2\int_0^{t-t_1-t_2}dt_3\int_{(\mathbb{R}^4)^3}\frac{dp_1\,dp_2\,dp_3}{(2\pi)^{12}}\frac{-2\omega}{\tanh^4(\omega t)\sinh^2(\omega t)}\hat{A}_\mu(p_1)\hat{A}_\nu(p_2)\hat{A}_\rho(p_3)\\
&\times\Big((p_2^\mu\sinh(\omega(t-2t_3))+p_3^\mu\sinh(\omega(t-2t_2-2t_3)))\delta^{\nu\rho}\cosh(\omega(t-2t_2))\\
&\quad+(p_3^\nu\sinh(\omega(t-2t_2))+p_1^\nu\sinh(\omega(2t_3-t)))\delta^{\rho\mu}\cosh(\omega(t-2t_2-2t_3))\\
&\quad+(p_1^\rho\sinh(\omega(2t_2+2t_3-t))+p_2^\rho\sinh(\omega(2t_2-t)))\delta^{\mu\nu}\cosh(\omega(t-2t_3))\Big)\\
&\times e^{-\frac{1}{4}pQ^{-1}p}+(A\mapsto B)\\
&=\int_0^t dt_1\int_0^{t-t_1}dt_2\int_0^{t-t_1-t_2}dt_3\int_{(\mathbb{R}^4)^3}\frac{dp_1\,dp_2\,dp_3}{(2\pi)^{12}}\frac{-2\omega}{\tanh^4(\omega t)\sinh^2(\omega t)}\hat{A}_\mu(p_1)\hat{A}_\nu(p_2)\hat{A}_\rho(p_3)\\
&\times p_2^\mu\Big(\sinh(\omega(2t_2-2t_3))+\sinh(\omega(4t_3+2t_2-2t)+\sinh(\omega(2t-4t_2-2t_3))\Big)\\
&+(A\mapsto B)\\
&=\mathcal{O}(t)\;.
\end{aligned}
$$

(The integral without $e^{-\frac{1}{4}pQ^{-1}p}$ cancels exactly.) We thus have

$$S_{111}(\mathcal{D}_A)=0\;. \tag{113}$$

B.7. $V_1V_1V_1V_1$**.** The leading order in (60) is given by the three possibilities with $k_i=0$:

(114)

$$
\begin{aligned}
&S_t(V_1,V_1,V_1,V_1)\\
&=\int_0^t dt_1\int_0^{t-t_1}dt_2\int_0^{t-t_1-t_2}dt_3\int_0^{t-t_1-t_2-t_3}dt_4\ \mathrm{tr}(e^{-\omega\Sigma t})\\
&\qquad\times S^{1,1,1,1}_{t_4,t_3,t_2,t-t_2-t_3-t_4}(-A,-A,-A,-A)+(A\mapsto B)\\
&=\int_{(\mathbb{R}^4)^3}\frac{dp_1\,dp_2\,dp_3}{(2\pi)^{12}}\frac{(2\omega)^2}{\tanh^4(\omega t)\sinh^2(\omega t)}\hat{A}_\mu(p_1)\hat{A}_\nu(p_2)\hat{A}_\rho(p_3)\hat{A}_\sigma(p_4)\\
&\times\int_0^t dt_1\int_0^{t-t_1}dt_2\int_0^{t-t_1-t_2}dt_3\int_0^{t-t_1-t_2}dt_3\int_0^{t-t_1-t_2-t_3}dt_4\Big(\cosh(\omega t_{21})\cosh(\omega t_{43})\delta^{\mu\nu}\delta^{\rho\sigma}\\
&\qquad+\cosh(\omega t_{31})\cosh(\omega t_{42})\delta^{\mu\rho}\delta^{\nu\sigma}+\cosh(\omega t_{41})\cosh(\omega t_{32})\delta^{\mu\sigma}\delta^{\nu\rho}\Big)+(A\mapsto B)\;,
\end{aligned}
$$

with $t_{21}=t-2t_4$, $t_{43}=t-2t_2$, $t_{31}=t-2t_3-2t_4$, $t_{42}=t-2t_2-2t_3$, $t_{41}=t-2t_2-2t_3-2t_4$ and $t_{32}=t-2t_3$. Taylor expansion in p_4 and Gaußian integration over $\frac{dp_4}{(2\pi)^4}$ yield, as usual, a factor $\frac{\omega^2}{\pi^2}\tanh(\omega t)$ and an exponential function that can be ignored in leading order. The $t_1,\dots,t_4$ integrals evaluate to $\frac{t^2\sinh^2(\omega t)}{8\omega^2}$, so that we conclude

$$S_{1111}(\mathcal{D}_A) = \frac{\chi_0}{2\pi^2} \int_{\mathrm{R}^4} dx \left\{ A_\mu A^\mu A_\nu A^\nu + B_\mu B^\mu B_\nu B^\nu \right\}(x) . \tag{115}$$

References

[1] A. Connes, *Noncommutative geometry*, Academic Press Inc., San Diego (1994)
[2] M. Dubois-Violette, J. Madore and R. Kerner, "Classical bosons in a noncommutative geometry," Class. Quant. Grav. **6** (1989) 1709.
[3] A. Connes, "Géométrie non commutative," Inter éditions, Paris 1990.
[4] A. Connes and J. Lott, "Particle models and noncommutative geometry (expanded version)," Nucl. Phys. Proc. Suppl. **18B** (1991) 29.
[5] A. Connes, "Noncommutative geometry and reality," J. Math. Phys. **36**, 6194 (1995).
[6] A. Connes, "Gravity coupled with matter and the foundation of non-commutative geometry," Commun. Math. Phys. **182** (1996) 155 [arXiv:hep-th/9603053].
[7] A. H. Chamseddine and A. Connes, "The spectral action principle," Commun. Math. Phys. **186** (1997) 731 [arXiv:hep-th/9606001].
[8] S. Lazzarini and T. Schücker, "A farewell to unimodularity," Phys. Lett. B **510** (2001) 277 [arXiv:hep-th/0104038].
[9] A. H. Chamseddine, A. Connes and M. Marcolli, "Gravity and the standard model with neutrino mixing," Adv. Theor. Math. Phys. **11** (2007) 991 [arXiv:hep-th/0610241].
[10] A. H. Chamseddine and A. Connes, "Conceptual Explanation for the Algebra in the Noncommutative Approach to the Standard Model," Phys. Rev. Lett. **99** (2007) 191601 [arXiv:0706.3690 [hep-th]].
[11] A. Connes, "On the spectral characterization of manifolds," arXiv:0810.2088 [math.OA].
[12] D. Perrot, "Anomalies and noncommutative index theory," in: *Geometric and topological methods for quantum field theory*, Contemp. Math. **434** (2007) 125 [arXiv:hep-th/0603209].
[13] A. Connes and H. Moscovici, "The local index formula in noncommutative geometry," Geom. Func. Anal. **5** (1995) 174.
[14] H. Grosse and R. Wulkenhaar, "Renormalisation of ϕ^4-theory on noncommutative $\mathbb{R}^4$ in the matrix base," Commun. Math. Phys. **256** (2005) 305 [arXiv:hep-th/0401128].
[15] H. Grosse and R. Wulkenhaar, "Renormalisation of ϕ^4-theory on non-commutative $\mathbb{R}^4$ to all orders," Lett. Math. Phys. **71** (2005) 13 [arXiv:hep-th/0403232].
[16] M. Disertori, R. Gurau, J. Magnen and V. Rivasseau, "Vanishing of beta function of non-commutative ϕ^4_4 theory to all orders," Phys. Lett. B **649** (2007) 95 [arXiv:hep-th/0612251].
[17] H. Grosse and R. Wulkenhaar, "The beta-function in duality-covariant noncommutative ϕ^4-theory," Eur. Phys. J. C **35** (2004) 277 [arXiv:hep-th/0402093].
[18] V. Gayral and R. Wulkenhaar, "A spectral triple for harmonic oscillator Moyal space," in preparation.
[19] H. Grosse and R. Wulkenhaar, "8D-spectral triple on 4D-Moyal space and the vacuum of noncommutative gauge theory," arXiv:0709.0095 [hep-th].
[20] F. Cooper, A. Khare and U. Sukhatme, "Supersymmetry and quantum mechanics," Phys. Rept. **251** (1995) 267 [arXiv:hep-th/9405029].
[21] E. Witten, "Supersymmetry and Morse theory," J. Diff. Geom. **17** (1982) 661.
[22] A. Connes and C. Rovelli, "Von Neumann algebra automorphisms and time thermodynamics relation in general covariant quantum theories," Class. Quant. Grav. **11** (1994) 2899 [arXiv:gr-qc/9406019].
[23] M. Paschke and R. Verch, "Local covariant quantum field theory over spectral geometries," Class. Quant. Grav. **21** (2004) 5299 [arXiv:gr-qc/0405057].
[24] V. Gayral, J. M. Gracia-Bondía, B. Iochum, T. Schücker and J. C. Várilly, "Moyal planes are spectral triples," Commun. Math. Phys. **246** (2004) 569 [arXiv:hep-th/0307241].
[25] V. Gayral and B. Iochum, "The spectral action for Moyal planes," J. Math. Phys. **46** (2005) 043503 [arXiv:hep-th/0402147].
[26] G. A. Elliott, T. Natsume and R. Nest, "The Atiyah-Singer index theorem as passage to the classical limit in quantum mechanics," Commun. Math. Phys. **182** (1996) 505.
[27] A. Carey, J. Phillips, A. Rennie and F. A. Sukochev, "The Chern character of semifinite spectral triples," J. Noncommut. Geom. **2** (2008) 141 [arXiv:math.OA/0611227].

[28] A. Carey, J. Phillips and F. A. Sukochev, "Spectral flow and Dixmier traces," Adv. Math. **173** (2003) 68–113 [arXiv:math.OA/0205076].
[29] M. A. Shubin, "Pseudodifferential Operators and Spectral Theory," Springer-Verlag, Tokyo (1987).

Mathematisches Institut der Westfälischen Wilhelms-Universität, Einsteinstrasse 62, D-48149 Münster, Germany
E-mail address: raimar@math.uni-muenster.de

Clay Mathematics Proceedings
Volume 11, 2010

A Characterization of the Image of the Baum-Connes Map

Guoliang Yu

Dedicated to Alain Connes on the occassion of his sixtieth birthday

1. Introduction

A central problem in noncommutative geometry is the Baum-Connes conjecture. The Baum-Connes conjecture provides an algorithm for computing the K-theory of reduced group C^*-algebras and higher indices of elliptic differential operators on compact manifolds. The conjecture implies the Novikov conjecture on homotopy invariance of higher signatures, the Gromov-Lawson-Rosenberg conjecture on positive scalar curvature, and the Kadison-Kaplansky idempotent conjecture. Despite great progress, the Baum-Connes conjecture remains open. The purpose of this article is to give a characterization of the K-theory elements in the image of the Baum-Connes map. A very different characterization was given by Joachim Cuntz in [10].

This paper is organized as follows. In Section 2, we give an elementary definition of the Baum-Connes map. In Section 3, we state the main results. In Section 4, we introduce a certain local K-theory to indicate the proofs of the main results of this article.

I would like to thank Alain for his encouragement over the years and for many enlightening discussions. It has always been a great joy to hear him talk about mathematics.

The author wishes to express his gratitude to the referee for several suggestions which improved the exposition of this paper.

2. A reformulation of the Baum-Connes map

In this section, we give an elementary definition of the Baum-Connes map which will be useful for the rest of this article.

Let Γ be a countable discrete group. Let X be a locally compact metric space with a proper and cocompact isometric action of Γ. Let $C_0(X)$ be the algebra of all complex-valued continuous functions on X which vanish at infinity.

The following definition due to John Roe will play an important role in our discussions [15].

2010 *Mathematics Subject Classification.* Primary: 19K56, Secondary: 19K35, 46L55, 58J22.
The author is partially supported by NSF and CNSF..

Definition 2.1. *Let H be a Hilbert space and let ϕ be a $*$-homomorphism from $C_0(X)$ to $B(H)$, the C^*-algebra of all bounded operators on H.*

(1) *Let T be a bounded linear operator acting on H. The support of T is defined to be the complement (in $X\times X$) of the set of all points $(x,y)\in X\times X$ for which there exists $f\in C_0(X)$ and $g\in C_0(X)$ satisfying $\phi(f)T\phi(g)=0$ and $f(x)\neq 0$ and $g(y)\neq 0$;*
(2) *The propagation of T is defined to be: sup$\{d(x,y):(x,y)\in \mathrm{Supp}(T)\}$;*
(3) *T is said to be locally compact if $\phi(f)T$ and $T\phi(f)$ are compact for all $f\in C_0(X)$.*

Let H be a Hilbert space with a Γ-action and let ϕ be a $*$-homomorphism from $C_0(X)$ to $B(H)$ such that it is covariant in the sense that $\phi(\gamma f)h=(\gamma(\phi(f))\gamma^{-1})h$ for all $\gamma\in\Gamma$, $f\in C_0(X)$ and $h\in H$. Such a triple $(C_0(X),\Gamma,\phi)$ is called a covariant system.

Definition 2.2. *We define the covariant system $(C_0(X),\Gamma,\phi)$ to be admissible if*

(1) *the Γ-action on X is proper and cocompact;*
(2) *$\phi(f)$ is noncompact for any nonzero function $f\in C_0(X)$;*
(3) *for each $x\in X$, the action of the stabilizer group Γ_x on H is regular in the sense that it is isomorphic to the action of Γ_x on $l^2(\Gamma_x)\otimes W$ for some infinite dimensional Hilbert space W, where the Γ_x action on $l^2(\Gamma_x)$ is regular and the Γ_x action on W is trivial.*

We remark that condition (3) in the above definition can be dropped if Γ acts on X freely. In particular, if M is a compact manifold and $\Gamma=\pi_1(M)$, then $(C_0(\widetilde{M}),\Gamma,\phi)$ is an admissible covariant system, where $\widetilde{M}$ is the universal cover of M and $\phi(f)\xi=f\xi$ for each $f\in C_0(\widetilde{M})$ and all $\xi\in L^2(\widetilde{M})$. In general, for each locally compact metric space with a proper and cocompact isometric action of Γ, there exists an admissible covariant system $(C_0(X),\Gamma,\phi)$.

Definition 2.3. *Let $(C_0(X),\Gamma,\phi)$ be an admissible covariant system. We define $\mathbb{C}(\Gamma,X,H)$ to be the algebra of Γ-invariant locally compact operators acting on H with finite propagation. The C^*-algebra $C_r^*(\Gamma,X,H)$ is the operator norm closure of $\mathbb{C}(\Gamma,X,H)$.*

Proposition 2.4. *If $(C_0(X),\Gamma,\phi)$ is an admissible covariant system, then $C_r^*(\Gamma,X,H)$ is $*$-isomorphic to $C_r^*\Gamma\otimes K$, where $C_r^*\Gamma$ is the reduced group C^*-algebra and K is the algebra of all compact operators.*

When the group Γ is torsion-free, this result is due to John Roe [16].

Let H be a Hilbert space, let F be an operator acting on H, let ϕ be a $*$-homomorphism from $C_0(X)$ to $B(H)$ such that F is Γ-equivariant, i.e. $\gamma F\gamma^{-1}=F$ for all $\gamma\in\Gamma$, and $\phi(f)F-F\phi(f)$, $\phi(f)(FF^*-I)$ and $\phi(f)(F^*F-I)$ are compact operators for all $f\in C_0(X)$. Such an F should be viewed as a generalized Γ-equivariant elliptic operator on X.

(H,ϕ,F) gives a Kasparov cycle representing a K-homology class in $K_0^\Gamma(X)$. It is not difficult to prove that every K-homology class $K_0^\Gamma(X)$ is equivalent to (H,ϕ,F) such that $(C_0(X),\Gamma,\phi)$ is an admissible covariant system. This follows from the fact that there exists a $*$-homomorphism $\phi':C_0(X)\to B(H')$ for some Hilbert space H' with a Γ action such that $(C_0(X),\Gamma,\phi)$ is a covariant system and the covariant system $(C_0(X),\Gamma,\phi\oplus\phi')$ is admissible.

For any $\epsilon > 0$, let $\{U_i\}_{i\in I}$ be a locally finite and Γ-equivariant open cover of X satisfying diameter$(U_i) < \epsilon$ for all i. Let $\{\psi_i\}$ be a Γ-equivariant partition of unity subordinate to $\{U_i\}_{i\in I}$. We define

$$F_\epsilon = \sum_{i\in I} \phi(\sqrt{\psi_i})F\phi(\sqrt{\psi_i}),$$

where the infinite sum converges in the strong topology.

Note that F_ϵ has propagation ϵ and (H, ϕ, F_ϵ) is equivalent to (H, ϕ, F) in $K_0^\Gamma(X)$.

F_ϵ is a multiplier of $C_r^*(\Gamma, X, H)$. F_ϵ is invertible modulo $C_r^*(\Gamma, X, H)$, i.e. F_ϵ is a generalized Fredholm operator.

Let ∂ be the boundary map in K-theory:

$$K_1(M(C_r^*(\Gamma, X, H))/C_r^*(\Gamma, X, H)) \to K_0(C_r^*(\Gamma, X, H)),$$

where $M(C_r^*(\Gamma, X, H))$ is the multiplier algebra of $C_r^*(\Gamma, X, H)$. We can define the Baum-Connes map

$$\mu : K_0^\Gamma(X) \to K_0(C_r^*(\Gamma, X, H)) \cong K_0(C_r^*\Gamma)$$

by

$$\mu([(H, \phi, F)]) = \partial([F_\epsilon]).$$

More precisely, the Baum-Connes map can be implemented as follows.

Let p_ϵ be the idempotent:

$$\begin{pmatrix} F_\epsilon F_\epsilon^* + (I - F_\epsilon F_\epsilon^*)F_\epsilon F_\epsilon^* & F_\epsilon(I - F_\epsilon^* F_\epsilon) + (I - F_\epsilon F_\epsilon^*)F_\epsilon(I - F_\epsilon^* F_\epsilon) \\ (I - F_\epsilon^* F_\epsilon)F_\epsilon^* & (I - F_\epsilon^* F_\epsilon)^2 \end{pmatrix}.$$

Observe that the propagation of p_ϵ is at most 5ϵ.

Let

$$p_0 = \begin{pmatrix} I & 0 \\ 0 & 0 \end{pmatrix}.$$

We have

$$\mu([(H, \phi, F)]) = [p_\epsilon] - [p_0].$$

$\mu([(H, \phi, F)])$ should be interpreted as the index of the generalized Fredholm operator F_ϵ.

In the odd case, we can similarly define the Baum-Connes map μ as follows.

Any K-homology class in $K_1^\Gamma(X)$ can be represented by (H, ϕ, F) such that the covariant system $(C_0(X), \Gamma, \phi)$ is admissible, where H is a Hilbert space, F is an operator acting on H, ϕ is a $*$-homomorphism from $C_0(X)$ to $B(H)$ such that F is Γ-equivariant, i.e. $\gamma F \gamma^{-1} = F$ for all $\gamma \in \Gamma$, and $\phi(f)F - F\phi(f)$, $\phi(f)(F^2 - I)$ and $\phi(f)(F^* - F)$ are compact operators for all $f \in C_0(X)$.

For each $\epsilon > 0$, let F_ϵ be defined as in the even case. Let

$$q_\epsilon = \frac{1}{2}(F_\epsilon + I).$$

q_ϵ gives rise to an element in $K_0(M(C_r^*(\Gamma, X, H))/C_r^*(\Gamma, X, H))$.

Let ∂ be the boundary map in K-theory:

$$K_0(M(C_r^*(\Gamma, X, H))/C_r^*(\Gamma, X, H)) \to K_1(C_r^*(\Gamma, X, H)).$$

We can define the Baum-Connes map

$$\mu : K_1^{\Gamma}(X) \to K_1(C_r^*(\Gamma, X, H)) \cong K_1(C_r^*\Gamma)$$

by

$$\mu([(H, \phi, F)]) = \partial([q_\epsilon]).$$

More precisely, we have

$$\mu([(H, \phi, F)]) = \exp(2\pi i q_\epsilon).$$

The following is the well known Baum-Connes conjecture [2] [3].

CONJECTURE 2.1. *Let $\underline{E}\Gamma$ be the universal space for proper Γ actions. The Baum-Connes map $\mu : K_*^{\Gamma}(\underline{E}\Gamma) \to K_*(C_r^*\Gamma)$, is an isomorphism.*

$K_*^{\Gamma}(\underline{E}\Gamma)$ should be understood as the inductive limit of $K_*^{\Gamma}(X)$ for all closed Γ-invariant subspaces X such that the quotient X/Γ is compact.

3. The image of the Baum-Connes map

In this section, we give a characterization of the image of the Baum-Connes map.

For any $\delta > 0$, an operator q is said to be a δ-quasi-projection if $q^* = q$ and

$$||q^2 - q|| < \delta.$$

When $\delta = \frac{1}{100}$, a δ-quasi-projection will be called a quasi-projection.

THEOREM 3.1. *An element $[p]$ in $K_0(C_r^*\Gamma)$ is in the image of the Baum-Connes map if and only if there exists an admissible covariant system $(C_0(X), \Gamma, \phi)$ for some locally compact and finite dimensional simplicial polyhedron with the simplicial metric and dimension n such that $[p]$ is equivalent to $[q] - [p_0]$ and q is a quasi-projection in $M_{k_n}(C_r^*(\Gamma, X, H))$ for some natural number k_n with propagation at most ϵ_n, where k_n depends only on n and ϵ_n is a positive constant depending only on n, $p_0 = I \oplus 0$, and the propagation of an element in $M_k(C_r^*(\Gamma, X, H))$ is defined to be the maximal propagation of its entries.*

The "only if" part of Theorem 3.1 follows from the construction of the Baum-Connes map in the previous section. To prove the "if" part of Theorem 3.1, we will need certain localization techniques in the next section. The proof of Theorem 3.1 yields concrete estimations of k_n and ϵ_n (k_n grows exponentially in n and ϵ_n decays exponentially in n).

We remark that, as a consequence of the proof of Theorem 3.1, if an element $[p]$ in $K_0(C_r^*\Gamma)$ is in the image of the Baum-Connes map, then there exists an admissible covariant system $(C_0(X), \Gamma, \phi)$ for some locally compact and finite dimensional simplicial polyhedron with the simplicial metric such that there exists a natural number k_n depending on n for which, given any $\epsilon > 0$, $[p]$ is equivalent to $[q] - [p_0]$ and q is a quasi-projection in $M_{k_n}(C_r^*(\Gamma, X, H))$ with propagation at most ϵ. When Γ is a finitely generated torsion-free group, in the equivalence $C_r^*(\Gamma, X, H) \cong C_r^*\Gamma \otimes K$, small propagation in $C_r^*(\Gamma, X, H)$ corresponds to propagation at most 1 in $C_r^*\Gamma \otimes K$ with respect to the word metric of Γ. As a consequence, we obtain the following result.

Corollary 3.2. *Let Γ be a finitely generated torsion-free group with a finite generating set S. Every element in the image of the Baum-Connes map in $K_0(C_r^*\Gamma)$ is equivalent to $[q]-[p_0]$ such that q is a quasi-projection in $M_k(C_r^*\Gamma)$ such that each of its entries is a linear combination of elements in $S\cup\{e\}$, where e is the identity of Γ.*

We remark that if the classifying space for the torsion-free group Γ is finite dimensional, then the matrix size k in the above corollary depends only on the dimension of the classifying space.

For any $\delta>0$, an operator v is said to be a δ-quasi-unitary if

$$||v^*v-I||<\delta$$

and

$$||vv^*-I||<\delta.$$

When $\delta=\frac{1}{100}$, a δ-quasi-unitary will be called a quasi-unitary.

We have a similar result in the odd case.

Theorem 3.3. *An element $[u]$ in $K_1(C_r^*\Gamma)$ is in the image of the Baum-Connes map if and only if there exists an admissible covariant system $(C_0(X),\Gamma,\phi)$ for some locally compact and finite dimensional simplicial polyhedron with the simplicial metric and dimension n such that u is equivalent to a quasi-unitary in $M_{k_n}(C_r^*(\Gamma,X,H)^+)$ for some natural number k_n with propagation at most ϵ_n, where k_n depends only on n and ϵ_n is a positive constant depending only on n.*

The proof of Theorem 3.3 yields concrete estimations of k_n and ϵ_n (k_n grows exponentially in n and ϵ_n decays exponentially in n). It is an open question whether we can replace the quasi-unitary condition on v by simply an invertibility condition.

We remark that, as a consequence of the proof of Theorem 3.3, if an element $[u]$ in $K_1(C_r^*\Gamma)$ is in the image of the Baum-Connes map, then there exists an admissible covariant system $(C_0(X),\Gamma,\phi)$ for some locally compact and finite dimensional simplicial polyhedron with the simplicial metric such that there exists a natural number k_n depending on n for which, given for any $\epsilon>0$, $[u]$ is equivalent to a quasi-unitary in $M_{k_n}(C_r^*(\Gamma,X,H)^+)$ for some natural number k with propagation at most ϵ. When Γ is a finitely generated torsion-free group, in the equivalence $C_r^*(\Gamma,X,H)\cong C_r^*\Gamma\otimes K$, small propagation in $C_r^*(\Gamma,X,H)$ corresponds to propagation at most 1 in $C_r^*\Gamma\otimes K$ with respect to the word metric of Γ. As a consequence, we obtain the following result.

Corollary 3.4. *Let Γ be a finitely generated torsion-free group with a finite generating set S. Every element in the image of the Baum-Connes map in $K_1(C_r^*\Gamma)$ is equivalent to $[v]$ such that v is a quasi-unitary in $M_k(C_r^*\Gamma)$ and each of its entries is a linear combination of elements in $S\cup\{e\}$, where e is the identity of Γ.*

We remark that if the classifying space for the torsion-free group Γ is finite dimensional, then the matrix size k in the above corollary depends only on the dimension of the classifying space.

4. Localization techniques in K-theory

In this section, we develop several localization techniques necessary to prove Theorem 3.1 and 3.3.

Let X be a locally compact and finite dimensional simplicial polyhedron. We endow X with the simplicial metric. Let $(C_0(X), \Gamma, \phi)$ be an admissible covariant system introduced in the previous section, where ϕ is a $*$-homomorphism from $C_0(X)$ to $B(H)$ for some Hilbert space H.

DEFINITION 4.1. *The algebraic localization algebra $C^*_{L,\mathrm{alg}}(\Gamma, X, H)$ is defined to the algebra of all bounded and uniformly continuous functions $f : [0,\infty) \to \mathbb{C}(\Gamma, H)$ such that the propagation of $f(t)$ goes to 0 as $t \to \infty$. The localization algebra $C^*_L(\Gamma, X, H)$ is the norm closure of $C^*_{L,\mathrm{alg}}(\Gamma, X, H)$ with respect to the following norm:*

$$||f|| = sup_{t\in[0,\infty)}||f(t)||.$$

It is not difficult to prove that, up to a $*$-isomorphism, $C^*_{L,\mathrm{alg}}(\Gamma, X, H)$ and $C^*_L(\Gamma, X, H)$ are independent of the choice of the admissible covariant system $(C_0(X), \Gamma, \phi)$. The localization algebra is an equivariant analogue of the algebra introduced in [19].

Any K-homology class in $K_0^\Gamma(X)$ can be represented by (H, ϕ, F) such that the covariant system $(C_0(X), \Gamma, \phi)$ is admissible, where H is a Hilbert space, F is an operator acting on H, ϕ is a $*$-homomorphism from $C_0(X)$ to $B(H)$ such that F is Γ-equivariant, $\phi(f)F - F\phi(f)$, $\phi(f)(FF^* - I)$ and $\phi(f)(F^*F - I)$ are compact operators for all $f \in C_0(X)$.

For each natural number n, we let $F_{\frac{1}{n}}$ be defined as in the previous section. We define a operator valued function $F(t)$ on $[0,\infty)$ by

$$F(t) = (t - n + 1)F_{\frac{1}{n}} + (t - n)F_{\frac{1}{n+1}}$$

for all $t \in [n, n+1]$.

$F(t)$ is a multiplier of $C^*_L(\Gamma, X, H)$ and is invertible modulo $C^*_L(\Gamma, X, H)$. We define the local Baum-Connes map

$$\mu_L : K_0^\Gamma(X) \to K_0(C^*_L(\Gamma, X, H)),$$

by

$$\mu_L[H, \phi, F)] = \partial[F(t)],$$

where

$$\partial : K_1(M(C^*_L(\Gamma, X, H))/C^*_L(\Gamma, X, H)) \to K_0^*(C^*_L(\Gamma, X, H)),$$

is the boundary map in K-theory and $M(C^*_L(\Gamma, X, H))$ is the multiplier algebra of $C^*_L(\Gamma, X, H)$.

Similarly we can define the local Baum-Connes map

$$\mu_L : K_1^\Gamma(X) \to K_1(C^*_L(\Gamma, X, H)).$$

We remark that the local Baum-Connes map is very much in the spirit of the local index theory of elliptic differential operators. The local Baum-Connes map is an equivariant analogue of the local index map introduced in [19].

THEOREM 4.2. *The local Baum-Connes map μ_L is an isomorphism from $K_*^\Gamma(X)$ to $K_*(C^*_L(\Gamma, X, H))$ if X is a finite dimensional simplicial polyhedron.*

Let e be the evaluation map:

$$C^*_L(\Gamma, X, H) \to C^*_r(\Gamma, X, H) \cong C^*_r\Gamma \otimes K$$

defined by:

$$e(f) = f(0)$$

for all $f \in C_L^*(\Gamma, X, H)$.

We have

$$\mu = e_* \circ \mu_L.$$

Next we shall define a quantitative version of local K-theory.

DEFINITION 4.3. *Let $(C_0(X), \Gamma, \phi)$ be an admissible covariant system introduced as in the previoius section, where ϕ is a $*$-homomorphism from $C_0(X)$ to $B(H)$ for some Hilbert space H. For each $\epsilon > 0$, we define the quantitative K-group $K_0^\epsilon(C_r^*(\Gamma, X, H))$ to be the Grothendieck group of the semi-abelian group of the equivalence classes of all quasi-projections of propagation at most ϵ with two quasi-projections being equivalent if and only there exists a homotopy path of $\frac{1}{10}$-quasi-projections with propagation at most 10ϵ.*

DEFINITION 4.4. *Let $(C_0(X), \Gamma, \phi)$ be an admissible covariant system introduced as in the previoius section, where ϕ is a $*$-homomorphism from $C_0(X)$ to $B(H)$ for some Hilbert space H. For each $\epsilon > 0$, we define the quantitative K-group $K_1^\epsilon(C_r^*(\Gamma, X, H))$ to be the abelian group of the equivalence classes of all quasi-unitaries in $C_r^*(\Gamma, X, H)^+$ with propagation at most ϵ with two quasi-unitaries being equivalent if and only there exists a homotopy path of $\frac{1}{10}$-quasi-unitaries with propagation at most 10ϵ.*

The above quantitative K-groups are equivariant analogues of concepts introduced in [20].

Next we shall formulate a Mayer-Vietoris sequence for the quantitative local K-theory. Together with Lipschitz homotopy invariance, the Mayer-Vietoris sequence provides an algorithm to compute the quantitative local K-theory when ϵ is small.

If ϕ is a $*$-homomorphism from $C_0(X)$ to $B(H)$ for some Hilbert space H, ϕ can be extended to a $*$-homomorphism from the algebra of all bounded Borel functions on X to $B(H)$. If Y is a closed subset of X, by identifying $C_0(Y)$ with a subalgebra of the algebra of all bounded Borel functions on X, ϕ induces a $*$-homomorphism from $C_0(Y)$ to $B(H)$ (still denoted by ϕ). We let $\overset{\circ}{Y}$ be the set of points in Y which are interior points of X. If $\overset{\circ}{Y}$ is dense in Y and $\phi(f)$ is noncompact for any nonzero function $f \in C_0(X)$, then the induced $*$-homomorphism from $C_0(Y)$ to $B(H)$ satisfies the same condition.

For each $r \geq 0$, we define

$$N_r(Y) = \{x \in X : d(x, Y) \leq r\}.$$

THEOREM 4.5. *Let X be a locally compact and finite dimensional polyhedron with the simplicial metric and $X = Y \cup Z$, where Y and Z are closed subsets of X. Assume that $(C_0(X), \Gamma, \phi)$ is an admissible covariant system introduced as in the previoius section, where ϕ is a $*$-homomorphism from $C_0(X)$ to $B(H)$ for some Hilbert space H. If Y and Z are Γ-invariant, $\overset{\circ}{Y}$ and $\overset{\circ}{Z}$ are respectively dense in Y and Z, then there exists a univeral constant $c \geq 1$ such that the following sequence is asymptotically exact:*

$$K_1^\epsilon(C_r^*(\Gamma, Y\cap Z, H)) \xrightarrow{i} K_1^\epsilon(C_r^*(\Gamma, Y, H)) \oplus K_1^\epsilon(C_r^*(\Gamma, Z, H)) \xrightarrow{j} K_1^\epsilon(C_r^*(\Gamma, X, H))$$
$$\xrightarrow{\partial} K_0^{c\epsilon}(C_r^*(\Gamma, N_{c\epsilon}(Y) \cap N_{c\epsilon}(Z), H))$$
$$\xrightarrow{i} K_0^{c\epsilon}(C_r^*(\Gamma, N_{c\epsilon}(Y), H)) \oplus K_0^{c\epsilon}(C_r^*(\Gamma, N_{c\epsilon}(Z), H)) \xrightarrow{j} K_0^{c\epsilon}(C_r^*(\Gamma, X, H)),$$

in the sense that

(1) $j \circ i = 0$;
(2) *the kernel of* $j : K_1^{\epsilon}(C_r^*(\Gamma, Y, H)) \oplus K_1^{\epsilon}(C_r^*(\Gamma, Z, H)) \to K_1^{\epsilon}(C_r^*(\Gamma, X, H))$ *in* $K_1^{c^2\epsilon}(C_r^*(\Gamma, Y, H)) \oplus K_1^{c^2\epsilon}(C_r^*(\Gamma, Z, H))$ *is contained in the image of* $i : K_1^{c^2\epsilon}(C_r^*(\Gamma, Y \cap Z, H)) \to K_1^{c^2\epsilon}(C_r^*(\Gamma, Y, H)) \oplus K_1^{c^2\epsilon}(C_r^*(\Gamma, Z, H))$;
(3) $\partial \circ j = 0$;
(4) *the kernel of* ∂ *in* $K_1^{c^2\epsilon}(C_r^*(\Gamma, X, H))$ *is contained in the image of* $j : K_1^{c^2\epsilon}(C_r^*(\Gamma, N_{c\epsilon}(Y), H)) \oplus K_1^{c^2\epsilon}(C_r^*(\Gamma, N_{c\epsilon}(Z), H)) \to K_1^{c^2\epsilon}(C_r^*(\Gamma, X, H))$;
(5) $i \circ \partial = 0$;
(6) *the kernel of* $i : K_0^{\epsilon}(C_r^*(\Gamma, Y \cap Z, H)) \to K_0^{\epsilon}(C_r^*(\Gamma, Y, H)) \oplus K_0^{\epsilon}(C_r^*(\Gamma, Z, H))$ *in* $K_0^{c^2\epsilon}(C_r^*(\Gamma, N_{c\epsilon}(Y) \cap N_{c\epsilon}(Z), H))$ *is contained in the image of* $\partial : K_1^{c\epsilon}(C_r^*(\Gamma, X, H)) \to K_0^{c^2\epsilon}(C_r^*(\Gamma, N_{c\epsilon}(Y) \cap N_{c\epsilon}(Z), H))$.

Let X_1 and X_2 be two metric spaces with proper and cocompact Γ-actions. Assume that $(C_0(X_k), \Gamma, \phi_k)$ is an admissible covariant system for each $k = 1, 2$, where ϕ_k is a $*$-homomorphism from $C_0(X_k)$ to $B(H_k)$ for some Hilbert space H_k. A map $f : X_1 \to X_2$ is called a Lipschitz map if there exists a constant $C > 0$ satisfying $d(f(x), f(y)) \leq Cd(x, y)$ for all $x, y \in X_1$, where C is called the Lipschitz constant. A homotopy $F : X_1 \times [0, 1] \to X_2$, is called a Lipschitz homotopy if $F(\cdot, t)$ is Lipschitz with the same Lipschitz constant.

A Γ-equivariant Lipschitz map f induces a homomorphism

$$f_* : K_*^{\epsilon}(C_r^*(\Gamma, X_1, H_1)) \to K_*^{C'\epsilon}(C_r^*(\Gamma, X_2, H_2)),$$

where C' is any constant greater than the Lipschitz constant C of f.

The following result on Lipschitz homotopy invariance is a very useful tool in computing quantitative K-groups.

Theorem 4.6. *If F is a Γ-equivariant Lipschitz homotopy from X_1 to X_2 with Lipschitz constant C, then*

$$F(\cdot, 0)_* = F(\cdot, 1)_* : K_*^{\epsilon}(C_r^*(\Gamma, X_1, H_1)) \to K_*^{10C\epsilon}(C_r^*(\Gamma, X_2, H_2)).$$

For any $\epsilon > o$, we can define a quantitative Baum-Connes map:

$$\mu_\epsilon : K_*^{\Gamma}(X) \to K_*^{\epsilon}(C_r^*(\Gamma, X, H)).$$

Theorem 4.7. *Let X be a locally compact and finite dimensional polyhedron with the simplicial metric and dimension n and let $(C_0(X), \Gamma, \phi)$ be an admissible covariant system. There exists $\epsilon_n > 0$ such that the quantitative Baum-Connes map μ_ϵ is an isomorphism for all positive $\epsilon \leq \epsilon_n$.*

The proof of the above theorem follows from a standard five-lemma argument using Theorem 4.5, Theorem 4.6, and a quantitative version of Bott periodicity. Now Theorems 3.1 and 3.3 follow from Theorem 4.7.

References

[1] M. F. Atiyah, Global theory of elliptic operators. 1970 Proc. Internat. Conf. on Functional Analysis and Related Topics (Tokyo, 1969) pp. 21–30 Univ. of Tokyo Press, Tokyo
[2] P. Baum and A. Connes, K-theory for discrete groups, Operator Algebras and Applications, (D. Evans and M. Takesaki, editors), Cambridge University Press (1989), 1–20.
[3] P. Baum, A. Connes and N. Higson, Classifying space for proper actions and K-theory of group C^*-algebras. C^*-algebras: 1943–1993 (San Antonio, TX, 1993), 240–291, Contemp. Math., 167, Amer. Math. Soc., Providence, RI, 1994.

[4] A. Connes, Noncommutative differential geometry. Inst. Hautes Études Sci. Publ. Math. No. 62 (1985), 257–360.
[5] A. Connes, Noncommutative Geometry, Academic Press, 1994.
[6] A. Connes, Cyclic cohomology and the transverse fundamental class of a foliation. Geometric methods in operator algebras (Kyoto, 1983), 52–144, Pitman Res. Notes Math. Ser., 123, Longman Sci. Tech., Harlow, 1986.
[7] A. Connes and H. Moscovici, Cyclic cohomology, the Novikov conjecture and hyperbolic groups. Topology 29 (1990), no. 3, 345–388.
[8] A. Connes and H. Moscovici, The local index formula in noncommutative geometry. Geom. Funct. Anal. 5 (1995), no. 2, 174–243.
[9] A. Connes, H.Moscovici and M. Gromov, Group cohomology with Lipschitz control and higher signatures. Geom. Funct. Anal. 3 (1993), no. 1, 1–78.
[10] J. Cuntz, Noncommutative simplicial complexes and the Baum-Connes conjecture. Geom. Funct. Anal. 12 (2002), no. 2, 307–329.
[11] N. Higson and G. G. Kasparov, E-theory and KK-theory for groups which act properly and isometrically on Hilbert space. Invent. Math. 144 (2001), no. 1, 23–74.
[12] G. G. Kasparov, Equivariant KK-theory and the Novikov conjecture, Invent. Math. 91 (1988), 147–201.
[13] V. Lafforgue, K-théorie bivariante pour les algébres de Banach et conjecture de Baum-Connes. Invent. Math. 149 (2002), no. 1, 1–95.
[14] I. Mineyev and G.Yu, The Baum-Connes conjecture for hyperbolic groups, Invent. math. 149 (2002), 97–122.
[15] J. Roe, Coarse cohomology and index theory on complete Riemannian manifolds. Mem. Amer. Math. Soc. 104 (1993), no.497.
[16] J. Roe, Index theory, coarse geometry, and topology of manifolds. CBMS Regional Conference Series in Mathematics, the American Mathematical Society, Providence, RI, 1996.
[17] J. Rosenberg, C^*-algebras, positive scalar curvature and the Novikov Conjecture, Publ. I.H.E.S., No. 58 (1983), 197–212.
[18] M. Puschnigg, The Kadison-Kaplansky conjecture for word-hyperbolic groups. Invent. Math. 149 (2002), no. 1, 153–194.
[19] G. Yu, Localization algebras and the coarse Baum–Connes conjecture, K-Theory 11 (1997), 307–318.
[20] G. Yu, The Novikov Conjecture for groups with finite asymptotic dimension, Ann. of Math. Vol. 147, 2 (1998), 325–355.
[21] G. Yu, The coarse Baum-Connes conjecture for spaces which admit a uniform embedding into Hilbert space. Invent. Math. 139 (2000), no. 1, 201–240.

Department of Mathematics, 1326 Stevenson Center, Vanderbilt University, Nashville, TN 37240, USA

E-mail address: guoliang.yu@vanderbilt.edu

Clay Mathematics Proceedings
Volume 11, 2010

Quantum Modular Forms

Don Zagier

To Alain Connes on his 60th birthday, in friendship and admiration

A classical modular form is a holomorphic function f in the complex upper half-plane $\mathfrak{H}$ satisfying the transformation equation

$$(1) \qquad (f|_k\gamma)(z) \;:=\; (cz+d)^{-k}\, f\Bigl(\frac{az+b}{cz+d}\Bigr) \;=\; f(z)$$

for all $z \in \mathfrak{H}$ and all matrices $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}(2,\mathbb{Z})$, where k, the weight of the modular form, is a fixed integer. Of course, there are many variants: one can replace the group $\mathrm{SL}(2,\mathbb{Z})$ by a group commensurable with it or by a more general Fuchsian subgroup of $\mathrm{SL}(2,\mathbb{R})$; the automorphy factor $(cz+d)^{-k}$ may be multiplied by a character or replaced by a more general multiplier system; the weight k may be half-integral or even rational; the function f can be vector-valued rather than scalar-valued; there may be a further additive correction on the right-hand side of (1); one can allow non-holomorphic functions of specified type (e.g., Maass wave forms); etc. But in all of these generalizations, as well as the higher-dimensional generalizations of modular forms to Hilbert or Siegel modular forms or to automorphic forms of more general type, the functions considered are defined on a symmetric space $X = G/K$ associated to a Lie group G and transform suitably with respect to the action of a discrete subgroup $\Gamma \subset G$ on X.

In this note we want to discuss, in the simplest cases, another type of modular object which, because it has the "feel" of the objects occurring in perturbative quantum field theory and because several of the examples come from quantum invariants of knots and 3-manifolds, we call *quantum modular forms*. These are objects which live at the boundary of the space X, are defined only asymptotically, rather than exactly, and have a transformation behavior of a quite different type with respect to some modular group. We will consider only the case when G is $\mathrm{SL}(2,\mathbb{R})$, X is $\mathfrak{H}$, and Γ is $\mathrm{SL}(2,\mathbb{Z})$ or a group commensurable with it. Then, as is well-known, the natural boundary of X is $\mathbb{P}^1(\mathbb{Q}) = \mathbb{Q} \cup \{\infty\}$, the set of "cusps" of Γ.

A quantum modular form should therefore be a complex-valued function f on $\mathbb{Q}$, or possibly on $\mathbb{P}^1(\mathbb{Q}) \smallsetminus S$ for some finite subset $S \subset \mathbb{P}^1(\mathbb{Q})$, having a certain behavior under the action of Γ on $\mathbb{P}^1(\mathbb{Q})$. Here neither of the properties which are

2010 *Mathematics Subject Classification.* Primary 11F99; Secondary 33D99, 57M27.

required of classical modular forms—analyticity and Γ-covariance—are reasonable things to require: the former because $\mathbb{P}^1(\mathbb{Q})$, viewed as the set of cusps of the action on Γ on $\mathfrak{H}$, is naturally equipped only with the discrete topology, not with its induced topology as a subset of $\mathbb{P}^1(\mathbb{R})$, so that any requirement of continuity or analyticity is vacuous; and the latter because Γ acts on $\mathbb{P}^1(\mathbb{Q})$ transitively or with only finitely many orbits, so that any requirement of Γ-covariance of a function on this set would lead to a trivial definition. So we do not demand either continuity/analyticity or modularity, but require instead that the failure of one precisely offsets the failure of the other. In other words, our quantum modular form should be a function $f:\mathbb{Q}\to\mathbb{C}$ for which the function $h_\gamma:\mathbb{Q}\smallsetminus\{\gamma^{-1}(\infty)\}\to\mathbb{C}$ defined by

$$h_\gamma(x) \;=\; f(x) \,-\, (f|_k\gamma)(x) \tag{2}$$

has some property of continuity or analyticity (now with respect to the real topology) for every element $\gamma\in\Gamma$. This is purposely a little vague, since examples coming from different sources have somewhat different properties, and we want to consider all of them as being quantum modular forms. For the sake of definiteness we will take as our canonical definition of a quantum modular form a function $f:\mathbb{Q}\to\mathbb{C}$ for which the function h_γ defined by (2) extends to a real-analytic function on $\mathbb{P}^1(\mathbb{R})\smallsetminus S_\gamma$, where $S_\gamma\subset\mathbb{P}^1(R)$ is a finite set (typically just $\{\infty,\gamma^{-1}(\infty)\}$), for each $\gamma\in\Gamma$. Notice that this property need only be checked for a set of generators of Γ, and hence for only finitely many elements, because its validity for γ_1 and γ_2 automatically implies its validity for $\gamma_1\gamma_2$. In fact, the function $\gamma\mapsto h_\gamma$ is a cocycle on Γ (i.e., it satisfies $h_{\gamma_1\gamma_2}=h_{\gamma_1}|_k\gamma_2+h_{\gamma_2}$), so that any quantum modular form defines a cohomology class in the first cohomology group of Γ with coefficients in the space of piecewise analytic functions on $\mathbb{P}^1(\mathbb{R})$ with the action $h\mapsto h|_k\gamma$ of Γ.

The definition just given describes what one can call a *weak quantum modular form*. A *strong quantum modular form*—and most of our examples will belong to this category—is an object with a stronger (and more interesting) structure: it associates to each element of $\mathbb{Q}$ a formal power series over $\mathbb{C}$, rather than just a complex number, with a correspondingly stronger requirement on its behavior under the action of Γ. To describe this, we write the power series in $\mathbb{C}[[\varepsilon]]$ associated to $x\in\mathbb{Q}$ as $f(x+i\varepsilon)$ rather than, say, $f_x(\varepsilon)$, so that f is now defined in the union of (disjoint!) formal infinitesimal neighborhoods of all points $x\in\mathbb{Q}\subset\mathbb{C}$. Since the function h_γ in (2) was required to be real-analytic on the complement of a finite subset S_γ of $\mathbb{P}^1(\mathbb{R})$, it extends holomorphically to a neighborhood of $\mathbb{P}^1(\mathbb{R})\smallsetminus S_\gamma$ in $\mathbb{P}^1(\mathbb{C})$, and in particular has a power series expansion (convergent in some disk of positive radius) around each point $x\in\mathbb{Q}$. Our stronger requirement is now that the equation

$$f(z) \,-\, (f|_k\gamma)(z) \;=\; h_\gamma(z) \qquad (\gamma\in\Gamma,\quad z\to x\in\mathbb{Q}) \tag{3}$$

holds as an identity between countable collections of formal power series.

Finally, there is a further property which holds for all the examples of strong quantum modular forms that we know, namely, that the formal function $f(z)$ just described extends to an actual function $f:(\mathbb{C}\smallsetminus\mathbb{R})\cup\mathbb{Q}\to\mathbb{C}$ that is analytic on $\mathbb{C}\smallsetminus\mathbb{R}$ and whose asymptotic expansion as one approaches any rational point $x\in\mathbb{Q}$ vertically from above or below coincides to all orders with the formal power series f at x. (Here "analytic" can mean "holomorphic" or merely "real-analytic," depending on the example.) Of course such an extension, even if it exists, isn't canonical since it can be modified by adding an analytic function in $\mathfrak{H}^\pm$ which vanishes to infinite order as one approaches any rational point, but in our examples there will often be a natural choice. One then gets a peculiar kind of object: an analytic function in the upper half-plane which "leaks" into the lower half-plane through the infinitely many "holes" $\mathbb{Q} \subset \mathbb{R}$ in the real axis to another analytic function in $\mathfrak{H}^-$ in such a way that the combined function on $\mathfrak{H} \cup \mathbb{Q} \cup \mathfrak{H}^-$ is C^∞ on any vertical line passing through a rational point, or more generally on any smooth curve in $\mathbb{C}$ which intersects $\mathbb{R}$ only orthogonally and in rational points. The sheaf defined by functions of this type gives $(\mathbb{C} \smallsetminus \mathbb{R}) \cup \mathbb{Q}$ a bizarre "hybrid topology" in which it is a 1-dimensional complex manifold at all points outside of $\mathbb{Q}$ and a kind of 1-dimensional real C^∞-manifold at all points of $\mathbb{Q}$.

All of this sounds somewhat abstract. Let us turn for the rest of the paper to the examples, which are taken from a variety of fields: number theory, combinatorics (q-series) and, as already mentioned, quantum invariants of 3-manifolds and knots.

Example 0. We begin with a function which is is more of a prototype than a true example because it does not fit precisely into the scheme described above, but which is in the same spirit and is very familiar to number theorists. This is the classical Dedekind sum, defined on pairs of coprime integers (c, d) with $c > 0$ by the formula

$$s(d, c) \;=\; \sum_{0<k<c} \Big(\Big(\frac{k}{c}\Big)\Big)\, \Big(\Big(\frac{kd}{c}\Big)\Big)\,,$$

where $((x))$ denotes $x - [x] - \frac{1}{2}$ for $x \notin \mathbb{Z}$. It satisfies the well-known identities

$$s(d+c, c) = s(d, c), \quad s(-d, c) = -s(d, c), \quad s(d, c) + s(c, d) = \frac{c^2 + d^2 + 1 - 3cd}{12cd}\,,$$

which determine it completely. Hence the function $S : \mathbb{Q} \to \mathbb{Q}$ defined by $S(d/c) = 12s(d, c)$ satisfies the functional equations

$$S(x) - S(x{+}1) \;=\; 0, \qquad S(x) - S(-1/x) \;=\; x + \frac{1}{x} \pm 3 + \frac{1}{\mathrm{Num}(x)\mathrm{Den}(x)} \quad (x \lessgtr 0)\,.$$

If we ignore the last term, which is the reason why we said that this example does not quite fit in with our general scheme, then we see that we have precisely an example of the type of transformation property described above. (The reason for the anomaly is that this example is related to the Eisenstein series of weight 2 on $\mathrm{SL}(2, \mathbb{Z})$, which is a quasimodular rather than a modular form.)

We mention that a function with quantum modular properties very similar to those of the Dedekind sum occurs in a recent preprint of Brian Conrey [**5**].

Example 1. We consider the following two q-hypergeometric functions, the first of which was given in Ramanujan's "Lost" Notebook and the second, its partner, discovered later:

$$\begin{aligned} \sigma(q) &= \sum_{n=0}^{\infty} \frac{q^{n(n+1)/2}}{(1+q)(1+q^2)\cdots(1+q^n)} \\ &= 1 + q - q^2 + 2\,q^3 - 2\,q^4 + q^5 + q^7 - 2\,q^8 + \cdots\;, \\ \sigma^*(q) &= 2\,\sum_{n=1}^{\infty} \frac{(-1)^n q^{n^2}}{(1-q)(1-q^3)\cdots(1-q^{2n-1})} \\ &= -2\,q - 2\,q^2 - 2\,q^3 + 2\,q^7 + 2\,q^8 + 2\,q^{10} + \cdots\;. \end{aligned}$$

In a beautiful paper by George Andrews, Freeman Dyson and Dean Hickerson [**2**]—the story is told in more detail in the last section of Dyson's famous survey article [**8**]—several identities expressing these two q-series as theta series associated to indefinite quadratic forms were proved, thereby explaining in particular the otherwise amazing experimental fact that the coefficients of both are very small, even though the individual terms have huge coefficients. (For instance, no coefficient of q^n in $\sigma(q)$ for $n \le 1600$ is greater than 4 in absolute value, even though some coefficients of the individual terms in the sum in the same range exceed 10^{13}.) A typical identity they proved is

$$q\,\sigma\bigl(q^{24}\bigr) \;=\; \sum_{\substack{a,\,b\,\in\mathbb{Z}\\ a>6|b|}} \left(\frac{12}{a}\right)(-1)^b\,q^{a^2-24b^2}\,, \tag{4}$$

the right-hand side of which is very similar to that of the modular identity

$$\frac{\eta(24z)^3}{\eta(48z)} \;=\; \sum_{\substack{a,\,b\,\in\mathbb{Z}\\ a>6|b|}} \left(\frac{-12}{a}\right)(-1)^b\,q^{a^2-24b^2}\,, \tag{5}$$

where $\eta(z) = q^{1/24}\prod_{n=1}^{\infty}(1-q^n)$ denotes the classical Dedekind eta function.

In an equally beautiful paper [**4**] which appeared side-by-side with the Andrews-Dyson-Hickerson paper, Henri Cohen interpreted these identities in terms, first of algebraic number theory, and then of the theory of Maass wave forms. Define coefficients $\bigl\{T(n)\bigr\}_{n\in 24\mathbb{Z}+1}$ by

$$q\,\sigma\bigl(q^{24}\bigr) \;=\; \sum_{n\ge 0} T(n)\,q^n\,, \qquad q^{-1}\,\sigma^*\bigl(q^{24}\bigr) \;=\; \sum_{n<0} T(n)\,q^{|n|}\,. \tag{6}$$

Then the identities of [**2**] are equivalent to the fact that $T(n)$ is the coefficient of $|n|^{-s}$ in the Dirichlet series

$$L(s) \;=\; \prod_{\substack{p\equiv\pm 3\\ (\mathrm{mod}\ 8)}} \frac{1}{1-p^{-2s}} \prod_{\substack{p\equiv\pm 7\\ (\mathrm{mod}\ 24)}} \frac{1}{1+p^{-2s}} \prod_{\substack{p\equiv\pm 1\\ (\mathrm{mod}\ 24)}} \frac{1}{(1-\varepsilon(p)\,p^{-s})^2}\,,$$

where $\varepsilon(p)$ is defined for $p = |P|$ with $P \in 24\mathbb{Z}+1$ by $\varepsilon(p) = (-1)^b = \bigl(\frac{12}{c}\bigr) = \bigl(\frac{24}{f}\bigr)$ if P has the representations $P = a^2 - 72b^2 = c^2 - 96d^2 = e^2 - 192f^2$ as a norm in the three quadratic orders $\mathbb{Z}[6\sqrt{2}]$, $\mathbb{Z}[4\sqrt{6}]$ and $\mathbb{Z}[8\sqrt{3}]$, respectively. Cohen observed that this is an Artin L-function that can be expressed via the identities $L(s) = \zeta_{\mathbb{Q}(\sqrt{3+\sqrt{3}})}(s)/\zeta_{\mathbb{Q}(\sqrt{3})}(s) = \zeta_{\mathbb{Q}(\sqrt{3+\sqrt{6}})}(s)/\zeta_{\mathbb{Q}(\sqrt{3})}(s)$ as a quotient of Dedekind zeta functions. This implies the functional equation $\widehat{L}(s) = -\widehat{L}(1-s)$, where $\widehat{L}(s) = (24\sqrt{2}/\pi)^s\Gamma(s/2)^2L(s)$, and from this in turn one deduces that the function

$$u(z) \;=\; \sqrt{y}\sum_{n\in 24\mathbb{Z}+1} T(n)\,K_0\bigl(2\pi|n|y/24\bigr)\,e^{2\pi i n x/24} \qquad (z = x+iy \in \mathfrak{H}) \tag{7}$$

satisfies $u(-1/2z) = \overline{u(z)}$ as well as the more obvious functional equation $u(z+1) = e^{2\pi i/24}u(z)$, whence also $u(z/(2z+1)) = e^{2\pi i/24}u(z)$. Since $u(z)$ is also an eigenfunction of the hyperbolic Laplace operator $-y^2\bigl(\partial^2/\partial x^2 + \partial^2/\partial y^2\bigr)$ with eigenvalue $1/4$, this shows that $u(z)$ is a Maass wave form on the congruence subgroup $\Gamma_0(2)$ and thus that the identity (4) is just as modular in nature as the identity (5), but now using non-holomorphic rather than holomorphic modular forms.

All of this seems to have nothing to do with quantum modular forms. However, Cohen also observed a further phenomenon, and it is this which concerns us here. One has the two q-series identities (the first due to Andrews, the second derived in a similar way by Cohen)

$$(8)\qquad \begin{aligned} \sigma(q) &= 1 + \sum_{n=0}^{\infty}(-1)^n\, q^{n+1}\,(1-q)(1-q^2)\cdots(1-q^n)\,, \\ \sigma^*(q) &= -2\sum_{n=0}^{\infty} q^{n+1}\,(1-q^2)(1-q^4)\cdots(1-q^{2n})\,. \end{aligned}$$

Cohen observed that the right-hand side of each of these expressions, as well as being a convergent series in the disk $|q|<1$, also makes sense whenever q is a root of unity, because the series is then terminating in both cases. He then discovered the following surprising fact about these functions.

Lemma. *Define σ and σ^* at roots of unity by* (8). *Then $\sigma(q) = -\sigma^*(q^{-1})$ for every root of unity q.*

The first cases of this can be checked by hand: $\sigma(1) = -\sigma(1) = 2$, $\sigma(-1) = -\sigma^*(-1) = -2$, $\sigma(\omega) = -\sigma^*(\omega^2) = 2\omega + 6$ for $\omega^2+\omega+1 = 0$, and $\sigma(\pm i) = -\sigma^*(\mp i) = \mp 2i - 4$.

Proof. The Laurent series

$$S_k \;=\; \sum_{n=1}^{k} q^{-n(n-1)/2}(1+q)(1+q^2)\cdots(1+q^{k-n}) \quad\in\; \mathbb{Z}[q,q^{-1}]$$

satisfies the recursion $S_{k+1} - S_k = q^{k+1}\bigl(S_{k+1} - (1+q)\cdots(1+q^k)\bigr) - q^{-k(k+1)/2}$, so by induction

$$(9)\quad \sum_{n=0}^{k-1}\bigl(q^{-1}-1\bigr)\cdots\bigl(q^{-n}-1\bigr) - \sum_{n=0}^{k-1} q^{n+1}(1-q^2)\cdots(1-q^{2n}) \;=\; (1-q)\cdots(1-q^k)\,S_k$$

for every $k \ge 0$. If q is a root of unity and k is bigger than or equal to the order of q, then the right-hand side of (9) vanishes and the left-hand side is easily seen to be $\frac12\sigma(q^{-1}) + \frac12\sigma^*(q)$. ■

We can now define our quantum modular form. Define a function $f:\mathbb{Q}\to\mathbb{C}$ by

$$(10)\qquad f(x) \;=\; q^{1/24}\,\sigma(q) \;=\; -q^{1/24}\,\sigma^*(q^{-1}) \qquad (x\in\mathbb{Q},\;\; q = e^{2\pi i x})\,,$$

where the equality of the two formulas is precisely the content of the lemma. This function, whose values for x with denominator ≤ 4 were given (up to the factor $q^{1/24}$) before the proof of the lemma, jumps around erratically as x runs through the rational numbers, but the cocycle defined by (3) with $\Gamma = \Gamma_0(2)$ and $k = 1$ is almost everywhere analytic:

Proposition. *The function $f:\mathbb{Q}\to\mathbb{C}$ defined by* (10) *satisfies*

$$(11)\qquad f(x+1) \;=\; e^{2\pi i/24} f(x)\,, \qquad \frac{1}{2x+1}\, f\Bigl(\frac{x}{2x+1}\Bigr) \;=\; e^{2\pi i/24} f(x) \,+\, h(x)$$

where $h:\mathbb{R}\to\mathbb{C}$ is C^∞ on $\mathbb{R}$ and real-analytic except at $x = -1/2$.

We illustrate this behavior by plotting in Figure 1 the real part of $f(x)$ for all rational numbers $x \in [-1.7, 1.1]$ with denominator ≤ 100 (the imaginary part looks very similar), and in Figure 2 the values of the real and imaginary parts of $h(x)$ for the same values of x.

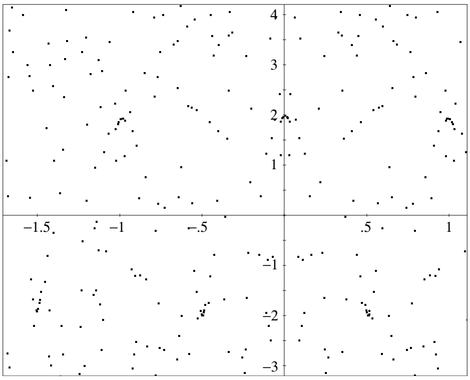


Figure 1. Graph of $\Re(f(x))$

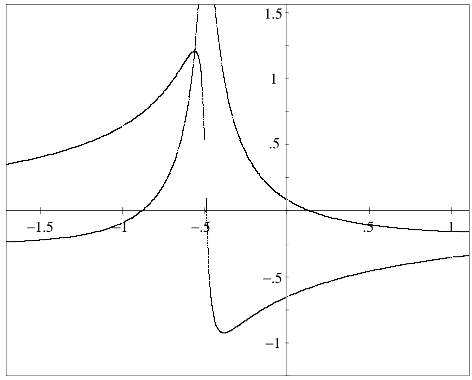


Figure 2. Graph of $\Re(h(x))$ and $\Im(h(x))$

This proposition, which we will prove in a moment, shows that f is a quantum modular form in the sense explained in the introduction, and the figures depict graphically what this means. In fact, f is a strong quantum modular form. Indeed, the two expressions in (8) are not only well-defined complex numbers when q is a root of unity, but well-defined power series in t, with coefficients in $\mathbb{Q}[\xi]$, when we take $q = \xi e^{-t}$ with ξ a root of unity. Furthermore, the identity $\sigma(q) = -\sigma^*(q^{-1})$ of the lemma remains true as an identity in $\mathbb{Q}[\xi][[t]]$, with the same proof, because the right-hand side of (7) is $\mathrm{O}(t^m)$ for k larger than m times the order of ξ. For instance, if we take $\xi = 1$ we find

$$\sigma\big(e^{-t}\big) \;=\; -\sigma^*\big(e^{t}\big) \;=\; 2 - 2\,t + 5\,t^2 - \frac{55}{3}\,t^3 + \frac{1073}{12}\,t^4 - \frac{32671}{60}\,t^5 + \frac{286333}{72}\,t^6 - \cdots .$$

If we extend the definition of f to infinitesimal neighborhoods of all rational points by interpreting (10) in the obvious way when x is replaced by $x + iy$ with $x \in \mathbb{Q}$ and y infinitesimal (so $q = \xi e^{-t}$ with $\xi = e^{2\pi i x}$ and $t = 2\pi y$), then (11) then still holds, where $h(x)$ is extended to a neighborhood of $\mathbb{R} \smallsetminus \{-1/2\}$ by analytic continuation. Here we can also clearly see the phenomenon of "leaking through the rational numbers" mentioned in the introduction, because we can extend the formally defined function f to a globally defined function $f : \mathfrak{H} \cup \mathfrak{H}^- \cup \mathbb{Q} \to \mathbb{C}$ by setting

$$f(z) \;=\; \begin{cases} q^{1/24}\,\sigma(q) & \text{if } z \in \mathfrak{H} \cup \mathbb{Q}\,, \\ -q^{1/24}\,\sigma^*(q^{-1}) & \text{if } z \in \mathfrak{H}^- \cup \mathbb{Q}, \end{cases} \tag{12}$$

where $q = e^{2\pi i z}$. Then the argument just given shows that f, which is obviously analytic in both $\mathfrak{H}$ and $\mathfrak{H}^-$, is C^∞ on any curve passing vertically through a rational point. In fact, the function $f(z)$ is the key to the proof of the proposition. Inserting the Fourier expansions (6) into (12) we can rewrite the definition of f in $\mathbb{C} \smallsetminus \mathbb{R}$ as

$$f(z) \;=\; \begin{cases} \sum_{n>0} T(n)\,q^n & \text{if } z \in \mathfrak{H}\,, \\ -\sum_{n<0} T(n)\,q^n & \text{if } z \in \mathfrak{H}^-, \end{cases}$$

which stands exactly is the same relation to the Maass wave form (7) as the functions denoted in the same way in the earlier work of J. Lewis and the author on Maass cusp forms on $\mathrm{SL}(2,\mathbb{Z})$ and their associated "period functions" [**12, 13**]. Making the needed minor changes in the results given there, we find that the holomorphic function f in $\mathbb{C}\smallsetminus\mathbb{R}$ can be expressed in terms of the Maass form u by the integral formulas

$$(13)\qquad f(z) \;=\; \begin{cases} \int_z^\infty [u(\tau), r_z(\tau)] & \text{if } z\in\mathfrak{H}\,, \\ -\int_{\bar z}^\infty [r_z(\tau), u(\tau)] & \text{if } z\in\mathfrak{H}^-, \end{cases}$$

where the function $r_z:\mathfrak{H}\to\mathbb{C}$ is defined by $r_z(\tau)=(\Im(\tau)/(\tau-z)(\bar\tau-z))^{1/2}$ and, like u, is an eigenfunction of the hyperbolic Laplace operator (with respect to τ) with eigenvalue $1/4$, and where $[\,\cdot\,,\,\cdot\,]$ denotes the *Green's form*

$$\big[u(\tau),v(\tau)\big] \;=\; \frac{\partial u(\tau)}{\partial\tau}\,v(\tau)\,d\tau \;+\; u(\tau)\,\frac{\partial v(\tau)}{\partial\bar\tau}\,d\bar\tau\,,$$

which is a closed 1-form whenever u and v are eigenfunctions of the hyperbolic Laplace operator with the same eigenvalue. From this and the modularity property $u(\gamma\tau)=\chi(\gamma)u(\tau)$ for $\gamma\in\Gamma$ of $u(\tau)$, where $\chi:\Gamma_0(2)\to\mathbb{C}^*$ is the character sending both generators $\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}1&0\\2&1\end{smallmatrix}\right)$ to $e^{2\pi i/24}$, together with the easy equivariance property $r_{\gamma z}(\gamma\tau)=\pm(cz+d)r_z(\tau)$ for $\gamma=\left(\begin{smallmatrix}\cdot&\cdot\\c&d\end{smallmatrix}\right)\in\mathrm{SL}(2,\mathbb{R})$, we deduce, apart from the obvious periodicity property $f(z+1)=e^{2\pi i/24}f(z)$, the formula

$$(14)\qquad (2z+1)\,f\Big(\frac{z}{2z+1}\Big)\;-\;e^{2\pi i/24}f(z)\;=\;-\int_{-1/2}^{\infty}[u(\tau),r_z(\tau)]$$

for z in either the upper or the lower half-plane, where the integral is taken along any path from $-1/2$ to ∞ passing to the left of z or $\bar z$. But the right-hand side now makes sense for any z lying to the right of both the chosen path and its reflection in the x-axis, so (if we push the path of integration far to the left) defines a holomorphic function on all of $\mathbb{C}\smallsetminus(-\infty,0]$. The function $h(x)$ occurring in (11) for $x>0$ is the restriction of this function to $\mathbb{R}_+$ and hence is real-analytic, and a similar argument works for $z\in\mathbb{C}\smallsetminus[0,\infty)$ and $x<0$ if we change the minus sign on the left-hand side of (14) to a plus sign and take a path of integration passing to the right of z and $\bar z$. This establishes the real-analyticity of h on $\mathbb{R}^*$. The fact that it is C^∞ also at $x=-1/2$ follows by looking more closely at the integral and using that u is a cusp form, as was done in [**13**] for the period functions of Maass forms on the full modular group. ∎

A similar discussion applies to other Maass wave forms on groups commensurable with $\mathrm{SL}(2,\mathbb{Z})$. We refer to the article [**3**] by R. Bruggeman for a treatment of this more general case.

Example 2. Our second example comes from [**14**], where the following elementary but rather surprising facts were proved.

1. Let $\mathcal{Q}_5$ denote the set of all quadratic functions $Q(x)=ax^2+bx+c$ with $a,\,b,\,c\in\mathbb{Z}$, $a<0$, and discriminant b^2-4ac equal to 5. Then for every rational number x we have

$$\sum_{Q\in\mathcal{Q}_5}\max(Q(x),0)\;=\;2\,,$$

the sum always being finite. (For example, the only $Q\in\mathcal{Q}_5$ with $Q(\frac13)>0$ are $-x^2+x+1$, $-x^2-x+1$, $-5x^2+5x-1$ and $-11x^2+7x-1$ and the corresponding values $Q(\frac{1}{3}) = \frac{11}{9}, \frac{5}{9}, \frac{1}{9}, \frac{1}{9}$ add up to 2.) More generally, if for every positive non-square integer D we define $\mathcal{Q}_D$ like $\mathcal{Q}_5$ but with the discriminant of Q now being the given number D, then we have

$$\sum_{Q\in\mathcal{Q}_D} \max(Q(x),0) \;=\; \alpha_D \tag{15}$$

for all $x \in \mathbb{Q}$, where α_D is a rational number that depends only on D and is equal to a simple multiple of the value of the Dedekind zeta function of $\mathbb{Q}(\sqrt{D})$ at $s=2$.
2. If one replaces the expression $\max(Q(x),0)$ by its cube, then the same thing happens: one has

$$\sum_{Q\in\mathcal{Q}_D} \max(Q(x),0)^3 \;=\; \beta_D \tag{16}$$

for all $x \in \mathbb{Q}$, where $\beta_D \in \mathbb{Q}$ is related to $\zeta_{\mathbb{Q}(\sqrt{D})}(4)$. But for the fifth power one has instead

$$\sum_{Q\in\mathcal{Q}_D} \max(Q(x),0)^5 \;=\; \gamma_D \,+\, \delta_D\,\Phi(x) \tag{17}$$

where γ_D (again related to $\zeta_{\mathbb{Q}(\sqrt{D})}(6)$) and δ_D are rational numbers depending only on D and $\Phi : \mathbb{Q} \to \mathbb{Q}$ is an even periodic function satisfying $q^{10}\,\Phi(\frac{p}{q}) \in \mathbb{Z}$ for all $\frac{p}{q} \in \mathbb{Q}$, the first values being

$p/q \pmod 1$	0	$1/2$	$\pm1/3$	$\pm1/4$	$\pm1/5$	$\pm2/5$	$\pm1/6$
$q^{10}\,\Phi(p/q)$	1	-1049	-29399	12076	3132025	-8012423	30839551

The function Φ satisfies—and, if one fixes one value, is uniquely characterized by—the two functional equations

$$\Phi(x+1) \;=\; \Phi(x)\,, \qquad x^{10}\,\Phi(-1/x) \;=\; \Phi(x) + x^{10} - \frac{691}{36}x^2(x^2-1)^3 - 1\,. \tag{18}$$

Therefore $\Phi(x)$ (and hence also $\sum_{Q\in\mathcal{Q}_D}\max(Q(x),0)^5$ for any D) is a quantum modular form. This example is unusual in that the cocycle $r_\gamma = \Phi - \Phi|_{-10}\gamma$ is analytic on all of $\mathbb{R}$ (it is a polynomial) and that Φ itself extends continuously (and even differentiably, though not C^∞) from $\mathbb{Q}$ to $\mathbb{R}$.

Here, again, the explanation is modular, but much simpler than in our first example because now only holomorphic modular forms on the full modular group are involved. The reason for the different behavior of the functions in (15) and (16) and in (17) is that there are no holomorphic modular forms except for Eisenstein series of weight 4 or 8 on $\mathrm{SL}(2,\mathbb{Z})$, while in weight 12 one has the cusp form

$$\Delta(z) \;=\; q\prod_{n=1}^{\infty}\bigl(1-q^n\bigr)^{24} \;=\; \sum_{n=1}^{\infty}\tau(n)\,q^n \qquad \bigl(z\in\mathfrak{H},\ q=e^{2\pi iz}\bigr),$$

as well as the Eisenstein series. The existence of the quantum modular form Φ follows directly from the existence of the cusp form Δ, as a consequence the classical Eichler-Shimura-Manin theory of periods of holomorphic modular forms. Specifically, we associate to $\Delta(z)$ its *Eichler integral*

$$\widetilde{\Delta}(z) \;=\; \frac{(2\pi/i)^{11}}{10!}\int_z^{\infty}\Delta(z')\,(z'-z)^{10}\,dz' \;=\; \sum_{n=1}^{\infty}\frac{\tau(n)}{n^{11}}\,q^n \qquad (z\in\mathfrak{H})\,. \tag{19}$$

(For an arbitrary cusp form f of weight k, $\widetilde{f}$ would be defined the same way with 10 and 11 replaced by $k-2$ and $k-1$.) Then $\dfrac{d^{10}\widetilde{\Delta}(z)}{dz^{10}} = \Delta(z)$ and from this and the modularity of Δ one deduces easily that

$$\widetilde{\Delta}(z) \;-\; (cz+d)^{10}\,\widetilde{\Delta}\big(\frac{az+b}{cz+d}\big) \;=\; P_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)}(z) \tag{20}$$

(or, more succinctly, $\widetilde{\Delta}\big|_{-10}(\gamma-1) = P_\gamma$) for all $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_1 = \mathrm{PSL}(2,\mathbb{Z})$, where $P_\gamma(z)$ is a polynomial of degree ≤ 10, given explicitly by

$$P_\gamma(z) \;=\; \frac{(2\pi/i)^{11}}{10!}\int_{\gamma^{-1}(\infty)}^{\infty} \Delta(z')\,(z-z')^{10}\,dz'\,. \tag{21}$$

These polynomials satisfy the cocycle relation $P_{\gamma\gamma'} = P_\gamma\big|_{-10}\gamma' + P_{\gamma'}$ and hence are determined by their values for the generators $T = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ and $S = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ of Γ_1, which are $P_T = 0$ (obviously) and

$$P_S(z) \;=\; -\Omega_1\left(z^{10} \;-\; \frac{691}{36}z^2(z^2-1)^3 \;-\; 1\right) \;+\; \Omega_2\left(z(z^2-1)^2(z^2-4)(4z^2-1)\right)$$

with $\Omega_1 = 0.98943291\cdots \in \mathbb{R}$, $\Omega_2 = 1.53908051\ldots i \in i\mathbb{R}$. From this and (18) we deduce that

$$\Phi(x) \;=\; \Re\big(\widetilde{\Delta}(x)/\Omega_1\big) \;=\; \frac{1}{\Omega_1}\sum_{n=1}^{\infty}\frac{\tau(n)}{n^{11}}\,\cos(2\pi n x) \tag{22}$$

for $x \in \mathbb{R}$, where $\widetilde{\Delta}(x)$ is defined by either of the formulas in (19), both of which remain convergent also when z lies on the real axis. The above-mentioned "continuous but not infinitely differentiable" properties of the function Φ follow from this: it is known that $\tau(n)$ is $\mathrm{O}\big(n^{11/2}\big)$ but not $\mathrm{o}\big(n^5\big)$ for n large, so the function $\Phi(x)$ on $\mathbb{R}$ is 4 times but not 6 times continuously differentiable.

In this example, too, we find a function that "leaks" from $\mathfrak{H}$ into $\mathfrak{H}^-$ through the rational holes in the real axis. To do this, we extend the definition (19) to the lower half-plane by

$$\begin{aligned} \widetilde{\Delta}(z) \;&=\; \frac{(2\pi/i)^{11}}{10!}\int_{\bar{z}}^{\infty}\Delta(z')\,(z'-z)^{10}\,dz' \\ &=\; \frac{1}{10!}\sum_{n=1}^{\infty}\frac{\tau(n)}{n^{11}}\,\gamma_{11}(4\pi n|y|)\,q^n \qquad (z\in\mathfrak{H}^-), \end{aligned} \tag{23}$$

where $z = x+iy \in \mathfrak{H}$ and $\gamma_{11}(t) = \int_t^\infty e^{-u}\,u^{10}\,du$, the incomplete gamma function (which is equal to e^{-t} times a polynomial in t). For $z = x \in \mathbb{R}$ the integrals in both (19) and (23) are convergent, because $\Delta(x+iy) = \mathrm{O}\big(y^{-6}\big)$ as $|y| \to 0$, so $\widetilde{\Delta}$ extends in this case to a continuous function in all of $\mathbb{C}$. This extended function still satisfies the functional equation (20), with the same polynomials P_γ as before, and because Δ is a cusp form and hence vanishes to infinite order as τ approaches any rational point, one sees easily that its restriction to any vertical line passing through a rational point is infinitely often differentiable. However, unlike the situation in our first example, here the function that "leaks" is only real-analytic, not holomorphic, in the lower half-plane.

Example 3. Our next example of a quantum modular form comes from the unusual series

$$F(q) \;=\; \sum_{m=0}^{\infty} (1-q)(1-q^2)\cdots(1-q^m)\,, \tag{24}$$

invented by Maxim Kontsevich, which has the peculiarity of not converging on any open subset of $\mathbb{C}$ but nevertheless makes sense as a function on the set of roots of unity because the series terminates after N terms if $q^N = 1$. We will be fairly brief in our treatment here, since this function was studied in detail in [**15**], and will only discuss the quantum modular aspect. Define $\varphi : \mathbb{Q} \to \mathbb{C}$ by $\varphi(x) = q^{1/24}F(q)$, where $q = e^{2\pi i x}$ as usual. Then $\varphi(1/n)$ has an asymptotic expansion of the form

$$\varphi(1/n) \;\sim\; n^{3/2}\, e^{2\pi i(3-n)/24} \;+\; \sum_{j=0}^{\infty} c_j\,(-2\pi i/n)^j \tag{25}$$

as $n \to \infty$, where $c_0 = 1$, $c_1 = \frac{23}{24}$, $c_2 = \frac{1681}{1152}, \ldots$ are certain rational coefficients. From the trivial functional equation $\varphi(x+1) = e^{2\pi i/24}\varphi(x)$ one sees that $e^{2\pi i(3-n)/24}$ equals $\sqrt{i}\,\varphi(-n)$, so (25) says that the function $g(x)$ defined by the second of the two equations

$$\varphi(x{+}1) = e^{2\pi i/24}\varphi(x)\,, \qquad \varphi(x) \mp i^{1/2}|x|^{3/2}\,\varphi(-1/x) \;=\; g(x) \qquad (x \in \mathbb{Q},\ \pm x > 0) \tag{26}$$

is smooth (i.e., has a well-defined Taylor expansion) at $x = 0$, and in fact it is real-analytic on the rest of the real axis, so that (26) presents $\varphi(x)$ as a quantum modular form.

The explanation is quite similar to that in the last example, except that the cusp form $\Delta(z)$ is replaced by its 24th root $\eta(z)$, which is a modular form of half-integral weight. Again we have a function $\widetilde{\eta}(z)$ in $\mathfrak{H} \cup \mathfrak{H}^-$, related to $\eta(z)$ in the same way as $\widetilde{\Delta}(z)$ in the previous example was related to $\Delta(z)$. (The direct analogues of the integrals in (19) and (23) diverge, because η has weight $1/2$, so that the exponent "10" in the integrand would have to be replaced by "$-3/2$," but they can be made sense of by integrating by parts once, or alternatively, we can use the definitions via sums rather than integrals.) In particular, since $\eta(z) = \sum_{n=0}^{\infty} n\left(\frac{12}{n}\right) q^{n^2/24}$ (Euler), this gives that $\widetilde{\eta}(z)$, appropriately normalized, is given by

$$\widetilde{\eta}(z) \;=\; \sum_{n=0}^{\infty} n\left(\frac{12}{n}\right) q^{n^2/24} \;=\; q^{1/24}\left(1 \,-\, 5q \,-\, 7q^2 \,+\, 11q^5 \,+\, \cdots\right), \tag{27}$$

and now the relation to Kontsevich's function follows from the formula

$$\sum_{n=0}^{\infty}\left(q^{1/24}(1-q)(1-q^2)\cdots(1-q^n) \,-\, \eta(z)\right) \;=\; -\frac{1}{2}\,\widetilde{\eta}(z) \,+\, \eta(z)\left(\frac{1}{2} - \sum_{n=1}^{\infty}\frac{q^n}{1-q^n}\right)$$

([**15**], Theorem 2), which implies that $-2\varphi(x)$ for $x \in \mathbb{Q}$ is the limiting value of $\widetilde{\eta}(z)$ as z approaches x from either the upper or the lower half-plane. We also deduce (26), with an explicit formula for the cocycle function $g(x)$ as an integral of the Dedekind eta-function along a path from 0 to ∞ in the upper half-plane.

We observe in passing that the function of this example, like those of Examples 4 and 5, belongs to the Habiro ring of "analytic functions of roots of unity" [**10**]. These functions, which are also related to the (now so very fashionable) $\mathbb{F}_1$-story, occur in many contexts connected with quantum topological invariants and quantum groups, and would be a natural setting to look for more examples of quantum modular forms.

Example 4. The next example, taken from [**11**], is similar in many ways to the last one, but is interesting because it comes from topology and more particularly from the theory of quantum invariants of 3-manifolds. Again we shall be brief and refer to the original paper for details. To any 3-manifold one can associate the so-called Witten-Reshetikhin-Turaev invariant, defined by the first of these authors by a path integral that can be made sense of only perturbatively or in the sense of topological quantum field theory, and by the second two in a rigorous, but less illuminating, algebraic way. The invariant makes sense at roots of unity of the form $\zeta_K = e^{2\pi i/K}$ with $K > 0$ integral. For manifolds of very special types (such as torus knots or Seifert fibrations) there are explicit formulas for it, and in particular for the Poincaré homology sphere $\Sigma(2,3,5)$ it is given by

$$W(q) \;=\; \frac{1}{2G} \sum_{\substack{\beta \pmod{60K} \\ \beta\not\equiv 0 \pmod{K}}} \frac{(1-\alpha^{24\beta})(1-\alpha^{40\beta})}{1+\alpha^{60\beta}}\,\alpha^{-(\beta+1)^2}\,, \tag{28}$$

if $q=\zeta_K$, where $\alpha=\zeta_{120K}$ and $G = \sum_{\beta \bmod 60K} \alpha^{-\beta^2} \;=\; (1-i)\sqrt{30K}$ (Gauss sum). We extend this to other roots of unity by Galois invariance $W(q)^\sigma = W(q^\sigma)$, or equivalently by formula (28) for q equal to any primitive Kth root of unity, with α being any primitive $(120K)$-th root of unity with $\alpha^{120}=q$. Let $\chi_+(n)$ be the odd periodic function of period 60 defined by the formula

$$\chi_+(n) \;=\; \begin{cases} (-1)^{[n/30]} & \text{if } (n,6)=1 \text{ and } n\equiv \pm 1 \pmod 5, \\ \quad 0 & \text{otherwise,} \end{cases}$$

and let $\Theta_+(z)$ be the theta series

$$\Theta_+(z) \;=\; \sum_{n=1}^{\infty} n\,\chi_+(n)\,q^{\frac{n^2}{120}} \;=\; q^{\frac{1}{120}}\bigl(1+11q+19q^3+29q^7-31q^8-\cdots\bigr) \quad (z\in\mathfrak{H}),$$

which is a modular form of weight $3/2$ on a certain congruence subgroup of $\mathrm{SL}(2,\mathbb{Z})$ (and in fact is the first component of a 2-component vector-valued modular form of weight $3/2$ on the full group $\mathrm{SL}(2,\mathbb{Z})$). Then for every $x\in\mathbb{Q}$ the number

$$f(x) \;=\; 2\,e^{\pi i x/60}\bigl(1\,-\,W\bigl(e^{2\pi i x}\bigr)\bigr) \tag{29}$$

is equal to the limit as $z\to x$ of the Eichler integral

$$\widetilde{\Theta}_+(z) \;=\; \sum_{n=1}^{\infty} \chi_+(n)\,q^{\frac{n^2}{120}} \;=\; q^{\frac{1}{120}}\bigl(1\,+\,q\,+\,q^3\,+\,q^7\,-\,q^8\,-\,\cdots\bigr) \tag{30}$$

([**11**], Theorem 1), and from this it follows that the function $f:\mathbb{Q}\to\mathbb{C}$ is a quantum modular form (in fact, a strong quantum modular form). The whole story is quite similar to that in Example 3 except that this time the modular form whose Eichler integral is involved has weight $3/2$ rather than $1/2$. There is also an expression

$$\sum_{n=0}^{\infty}\bigl[q^n(1-q)(1-q^6)\cdots(1-q^{5n-4})\;+\;q^{4n+3}(1-q^4)(1-q^9)\cdots(1-q^{5n-1})\bigr]$$

for $q^{-1/120}\widetilde{\Theta}_+(q)$ (due to Zwegers) of the same type as (24), which terminates and hence gives a closed formula for $W(q)$ whenever q is a root of unity of order not divisible by 5, as well as relations (also pointed out by Zwegers) to the mock theta functions of Ramanujan. See [**11**] for more details.

Example 5. The last example, which again comes from topology, is the most mysterious and in many ways the most interesting. The function from $\mathbb{Q}/\mathbb{Z}$ to $\mathbb{R}$ that we obtain in this case is not a quantum modular form in the strict sense of the definition we gave in the introduction, let alone a strong quantum modular form, because the associated cocycle is no longer analytic or even continuous, but it nevertheless will turn out to have a clearly defined modularity property.

To any knot and any integer $n \ge 2$ one can associate a Laurent polynomial $J_n(q) \in \mathbb{Z}[q, q^{-1}]$, called the n-colored Jones polynomial. The definition, which involves the theory of quantum groups, will not be reviewed here since we will only look at one example and here the Jones polynomials can simply be given by an explicit formula. We will consider the figure-eight knot, the simplest hyperbolic knot. The Jones polynomial of this knot is given by

$$J_n(q) \;=\; \sum_{m=0}^{n-1} q^{-mn} \prod_{j=1}^{m} (1-q^{n-j})(1-q^{n+j}) \;.$$

(Here the sum could also be taken from $m=0$ to ∞ since the mth summand is 0 for $m \ge n$.) If we fix a root of unity q, then the function $n \mapsto J_n(q)$ is periodic, of period N if $q^N = 1$, so we can extrapolate it backwards to define $J_n(q)$ also for $n \le 0$. Of particular interest to us is the $\overline{\mathbb{Q}}$-valued function on roots of unity defined by

$$(31)\quad J_0(q) \;:=\; J_N(q) \;=\; \sum_{m=0}^{\infty} \bigl|(1-q)(1-q^2)\cdots(1-q^m)\bigr|^2 \qquad (q \in \mathbb{C}^*,\ q^N = 1\,)$$

(compare the sum on the right-hand side to (24)), the first few values of which are as follows:

q	1	-1	$\zeta_3^{\pm1}$	$\pm i$	$\zeta_5^{\pm1}$	$\zeta_5^{\pm2}$	$\zeta_6^{\pm1}$
$J_0(q)$	1	5	13	27	$46+2\sqrt{5}$	$46-2\sqrt{5}$	89

The function J_0, which is related to perturbative $\mathrm{SL}(2,\mathbb{C})$ Chern-Simons theory (cf. [**7**]), is of a very different nature than the Jones polynomials themselves. For instance, the values of the Jones polynomials $J_n(q)$ when q is a root of unity are of only polynomial growth if $q^n \ne 1$, but the values of $J_0(\zeta_N) = J_N(\zeta_N)$ are exponentially big, as one can see in the following table:

N	100	200	300
$\max_{0<n<N} \lvert J_n(\zeta_N)\rvert$	12.07	18.62	24.99
$J_0(\zeta_N) = J_N(\zeta_N)$	8.20×10^{16}	2.48×10^{31}	4.89×10^{45}

Explicitly, $J_0(\zeta_N)$ is given by the the asymptotic formula [**1**]

$$J_0\bigl(e^{2\pi i/N}\bigr) \;\sim\; \frac{1}{\sqrt[4]{3}}\, N^{3/2}\, e^{CN} \qquad\quad (n \to \infty)\,,$$

where $C = 0.3230659\ldots$ is $1/2\pi$ times the hyperbolic volume of the complement of the knot, and in fact one has a complete asymptotic expansion [**6, 9, 16**]

$$J_0\bigl(e^{2\pi i/N}\bigr) \;=\; \frac{1}{3^{1/4}}\,N^{3/2}\,e^{CN}\left(1+\frac{11}{36\sqrt{3}}\,\frac{\pi}{N}+\frac{697}{7776}\,\frac{\pi^2}{N^2}+\frac{724351}{4199040\sqrt{3}}\,\frac{\pi^3}{N^3}+\cdots\right) \tag{32}$$

as $N\to\infty$, where the factor in parentheses is a power series in $\pi/N\sqrt{3}$ with rational coefficients. Conjecturally [**7**], the corresponding expansion for an arbitrary hyperbolic knot would be a power series in $\pi i/N$ with coefficients in the trace field of the knot, this trace field being $\mathbb{Q}(\sqrt{-3})$ for the figure 8 knot.

But since $J_0(q)$ is defined for all roots of unity, we can look at its expansion near some other point than 1, e.g., we can consider the values $q=-\zeta_N$ rather than $q=\zeta_N$. It is here that the phenomenon of most interest to us appears: these values are given (experimentally) by the asymptotic series

$$J_0\bigl(-e^{2\pi i/N}\bigr) \;=\; \kappa(N)\cdot\frac{3^{1/4}}{2^{3/2}}\,N^{3/2}\,e^{CN/4}\left(1+\frac{41}{36\sqrt{3}}\,\frac{\pi}{N}+\frac{12625}{7776}\,\frac{\pi^2}{N^2}+\cdots\right), \tag{33}$$

of the same general form as (32), but this time involving an extra factor

$$\kappa(N) \;=\; \begin{cases} 27 & \text{if } N\equiv 1 \pmod 2,\\ 1 & \text{if } N\equiv 2 \pmod 4,\\ 5 & \text{if } N\equiv 0 \pmod 4\end{cases}$$

that depends on the value of $N \pmod 4$. Comparing with the table of values of $J_0(q)$ given above, we find that this factor is given in all cases by $\kappa(N) = J_0\bigl(i^{N+2}\bigr)$. What's more, if we now try *rational* rather than integral values for N, but with bounded denominators, then we find that (33) still holds, with $\kappa(N) = J_0\bigl(e^{\pi i(N+2)/2}\bigr)$. Going back to (32), we find exactly the same behavior there: if N goes to infinity, not through integers, but through rational numbers, say with denominator 2, 3 or 4, then (32) remains true if we multiply the right-hand side multiplied by 5, 13, and 27, respectively (and in general by $J_0(e^{2\pi iN})$). More generally, the experimentally found asymptotic behavior of the function

$$\mathbf{J}\,:\,\mathbb{Q}/\mathbb{Z}\;\to\;\overline{\mathbb{Q}}\cap\mathbb{R}\,,\qquad \mathbf{J}(x)\;:=\;J_0(e^{2\pi ix}) \tag{34}$$

as x tends to a fixed rational number $\alpha = a/c$ ($a,\,c\in\mathbb{Z}$, $(a,c)=1$) from the right or left is given by the formula

$$\mathbf{J}(\alpha\pm\varepsilon)\;=\;\mathbf{J}(\alpha^*\pm\beta)\cdot\frac{\exp(C/c^2\varepsilon)}{\varepsilon^{3/2}}\bigl(A_\pm(\alpha)\,+\,B_\pm(\alpha)\varepsilon\,+\,C_\pm(\alpha)\varepsilon^2\,+\,\cdots\bigr) \tag{35}$$

as ε tends to 0 through positive rational values with $1/c^2\varepsilon\equiv\beta \pmod 1$ for some fixed rational number β, where $\alpha^* = d/c$ with $d\equiv a^{-1} \pmod c$ and $A_\pm(\alpha) = A(\pm\alpha)$, $B_\pm(\alpha) = B(\pm\alpha)$, $\ldots$ are algebraic numbers depending only on α modulo 1. (In equations (32) and (33) we have $\alpha=\alpha^*=\beta=0$ and $\alpha=\alpha^*=1/2$, $\beta\equiv N/4 \pmod 1$, respectively, explaining the extra factor $\kappa(N)=\mathbf{J}((N+2)/4)$ in the latter case.)

The factor $\mathbf{J}(\alpha^*\pm\beta)$ in equation (35) looks odd at first sight, but in fact has a simple modular explanation: if we set $\gamma = \left(\begin{smallmatrix} a & b\\ c & d\end{smallmatrix}\right)$, where $b\in\mathbb{Z}$ is chosen so that $\gamma\in\mathrm{SL}(2,\mathbb{Z})$, then we have $\mathbf{J}(\alpha^*\pm\beta) = \mathbf{J}(-\alpha^*\pm 1/c^2\varepsilon) = \mathbf{J}(\gamma^{-1}(\alpha\pm\varepsilon))$, so that (35) can be seen as simply relating the values of $\mathbf{J}(X)$ and $\mathbf{J}(\gamma(X))$ as X $(=-\alpha^*\pm 1/c^2\varepsilon\,)$ tends to infinity through rational numbers with small denominator.

The asymptotic formula (35) is therefore equivalent to the first part of the following conjecture generalizing formulas (32) and (33):

CONJECTURE. *Let* $\alpha \in \mathbb{Q}$ *and choose* $\gamma \in SL(2, \mathbb{Z})$ *with* $\gamma(\infty) = \alpha$. *Then for suitable real numbers* $S_0(\alpha)$, $S_1(\alpha)$, ... *depending only on* α (mod 1) *we have an asymptotic expansion*

$$\frac{\mathbf{J}(\gamma(X))}{\mathbf{J}(X)} \sim \left(\pi/\hbar\right)^{3/2} \exp\left(\sum_{n=0}^{\infty} S_n(\alpha)\, \hbar^{n-1}\right), \qquad \hbar = \frac{\pi/\sqrt{3}}{X - \gamma^{-1}(\infty)} \tag{36}$$

as $X \to \infty$ *through rational numbers with bounded denominators. The value of* $S_0(\alpha)$ *is independent of* α *and is equal to* $\pi C/\sqrt{3}$, *while* $S_1(\alpha) \in \mathbb{Q}\log\bigl(K_a^{\times}\bigr)$ *and* $S_n(\alpha) \in K_\alpha$ *for* $n \geq 2$, *where* K_α *is the maximal real subfield of the cyclotomic field* $\mathbb{Q}\bigl(\sqrt{-3}, e^{2\pi i\alpha}\bigr)$.

We expect that a similar conjecture should hold for any hyperbolic knot complement, with $\hbar$ being defined as $\pi i/(X - \gamma^{-1}(\infty))$ (we divided this by $\sqrt{-3}$ in our special case to make everything real) and K_α being replaced by $K\bigl(e^{2\pi i\alpha}\bigr)$, where K is the trace field of the knot. The case when $\alpha = 0$ and $X \in \mathbb{Z}$ is precisely the arithmeticity conjecture from [**7**] which was cited earlier.

Observe that the correctness of (36) is unchanged by replacing (γ, X) by $(\gamma T, X - 1)$ or $(T\gamma, X)$, where $T = \left(\begin{smallmatrix}1 & 1\\ 0 & 1\end{smallmatrix}\right)$, since these changes do not affect the left-hand side or the quantity $\hbar$, so the quantities $S_n(\alpha)$ really do depend only on α rather than on γ, and are periodic in α.

Here is a table of the numerically obtained values of $S_n(\alpha)$ for some small n and simple α, where in the last line ε_k ($k = 1, 2, 3$) denotes the real cyclotomic unit $\zeta^k + \zeta^{-k}$ with $\zeta = \zeta_3^{-1}\, e^{4\pi i a/5}$ and $\pi_{29} = 2 - \varepsilon_1 + \varepsilon_2 + 2\varepsilon_3$, a prime of $\mathbb{Q}(\zeta)$ of norm -29.

α	$\exp(S_1(\alpha))$	$S_2(\alpha)$	$S_3(\alpha)$	$S_4(\alpha)$
0	$\frac{1}{3}$	$\frac{11}{2^2 3^2}$	$\frac{2}{3^2}$	$\frac{1081}{2^1 3^5 5}$
$\frac{1}{2}$	$(2^3/3)^{1/2}$	$\frac{41}{2^4 3^2}$	$\frac{19}{2^3 3^2}$	$\frac{71089}{2^7 3^5 5}$
$\frac{1}{3}$	$2 \cdot 3^{2/3}$	$\frac{37}{2^2 3^3}$	$\frac{401}{2^1 3^6}$	$\frac{30767}{2^1 3^8 5}$
$\frac{2}{3}$	$3^{4/3}$	$\frac{25}{2^2 3^3}$	$\frac{182}{3^6}$	$\frac{29027}{2^1 3^8 5}$
$\frac{1}{6}$	$2^{7/2} \cdot 3^{5/6}$	$\frac{193}{2^4 3^3}$	$\frac{24691}{2^7 3^6}$	$\frac{8027957}{2^9 3^8 5}$
$\frac{5}{6}$	$2^{3/2} \cdot 3^{13/6}$	$\frac{67}{2^4 3^3}$	$\frac{1289}{2^3 3^6}$	$\frac{1759883}{2^7 3^8 5}$
$\pm\frac{1}{4}$	$\frac{2^3(2\sqrt{3}\pm 1)}{3(2\pm\sqrt{3})^{1/4}}$	$\frac{1855\pm 360\sqrt{3}}{2^6 3^2 11}$	$\frac{71132\pm 3123\sqrt{3}}{2^8 3^2 11^2}$	$\frac{1499191589\pm 43727850\sqrt{3}}{2^{11} 3^5 5^1 11^3}$
$\frac{a}{5}$, $5 \nmid a$	$\frac{(5^3/3)^{1/2}\lvert\pi_{29}\rvert}{(\varepsilon_1^3\varepsilon_3/\varepsilon_2)^{1/10}}$	$\frac{1}{2^2 3^2 5^3 \pi_{29}}(2678 - 943\varepsilon_1$ $+1831\varepsilon_2 + 2990\varepsilon_3)$	$\cdots$	$\cdots$

Formulas (35) or (36) say that the values of $\mathbf{J}(x)$ as x approaches any given rational number go exponentially rapidly to infinity and lie on certain smooth curves (countably many, all proportional to one another) depending on the rational number in question. This behavior can be seen clearly in the graph of the function $\mathbf{J}$, which looks as follows, where because of the very rapid growth we have plotted $f(x) = \log(\mathbf{J}(x))$ rather than $\mathbf{J}$ itself, so that now the different curves containing the points of the graph with argument near any fixed rational point differ by vertical translations:

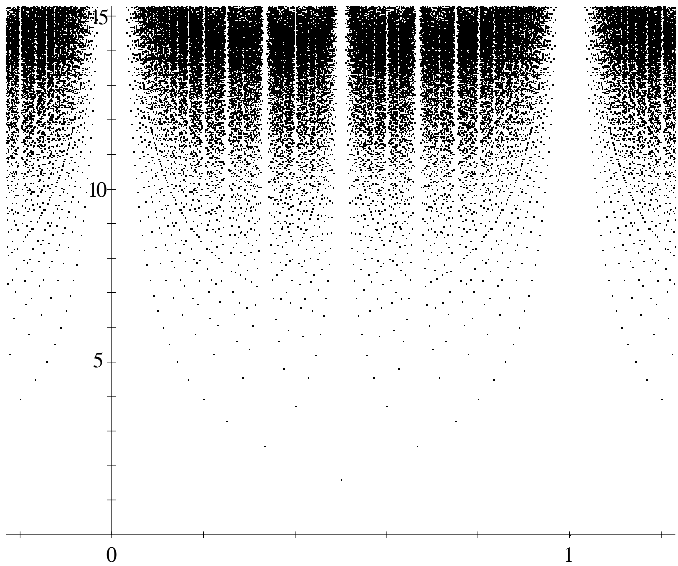


Figure 3. Graph of $f(x) = \log(\mathbf{J}(x))$

To make more sense of this graph, we do as in Examples 1–4 and compare the values of $f(x)$ at x and $1/x$. The graph of the difference indeed looks much better than the graph of f itself:

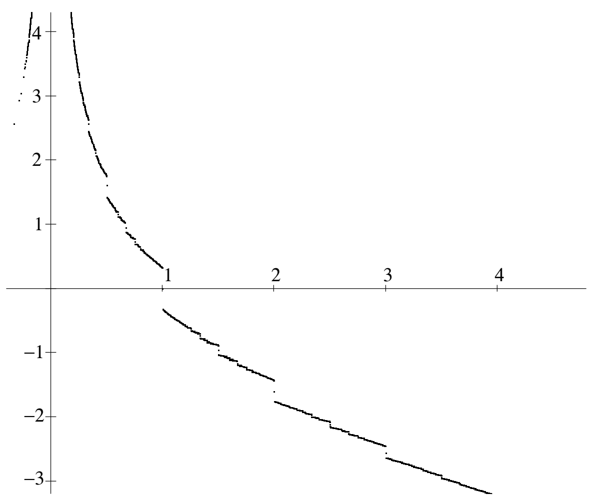


Figure 4. Graph of $h(x) = \log(\mathbf{J}(x)/\mathbf{J}(1/x))$

The behavior that we see here is a consequence of the conjecture above, which can easily be seen to imply that the function $h(x)$ has a jump at every rational point $\alpha = a/c$ but is C^∞ as we approach α from the left or from the right, with limiting values of the form $h^{\pm}(\alpha) = \pm C/ac + \log(\beta_{\pm}(\alpha))$ as x approaches α from the left or from the right, where $\beta_{\pm}(\alpha)$ are real algebraic numbers. This smoothness from the two sides can be seen more clearly by looking more closely at the graph of $h(x)$ in the neighborhood of a rational point α with small denominator, say $\alpha = 3/8$:

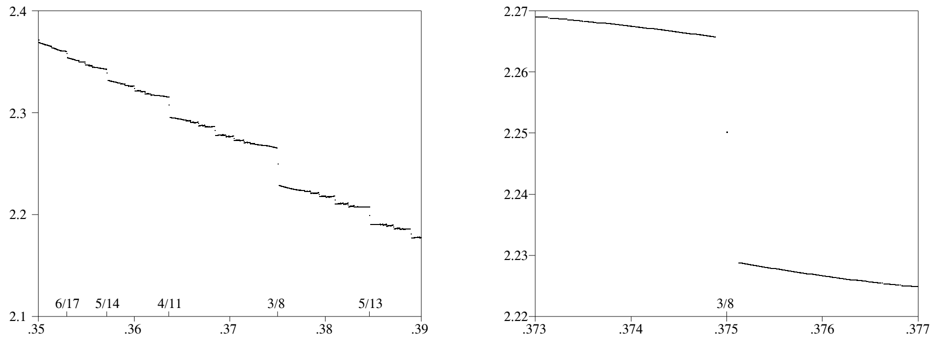


Figure 5. Graphs of $h(x)$ near $x = 3/8$

By contrast, in a small interval around the point $1/\phi$ ($\phi = (1+\sqrt{5})/2$ = golden ratio), where there are no points with particularly small denominator, we get the following picture

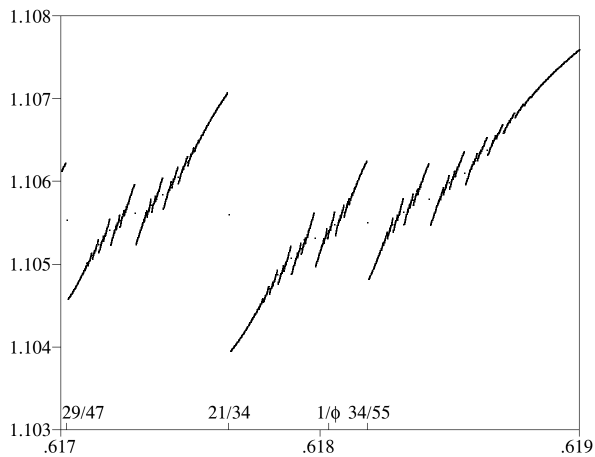


Figure 6. Graph of $h(x)$ near $x = 1/\phi$

showing that, unlike what the picture in Figure 4 might have suggested, $h(x)$ is not monotone decreasing on $\{x > 0\}$ and seeming to indicate that the function $h(x)$ is continuous but in general not differentiable at irrational values of x.

References

[1] J.E. Andersen and S.K. Hansen, *Asymptotics of the quantum invariants for surgeries on the figure* 8 *knot*. J. Knot Theory Ramifications **15** (2006), 479–548.

[2] G. E. Andrews, F. J. Dyson and D. Hickerson, *Partitions and indefinite quadratic forms.* Invent. Math. **91** (1988), 391–407.

[3] R. Bruggeman, *Quantum Maass forms.* In "The conference on L-functions, Fukuoka, Japan 18–23 February 2006" (eds. L. Weng and M. Kaneko), World Scientific (2007), 1–15.

[4] H. Cohen, *q-identities for Maass waveforms.* Invent. Math. **91** (1988), 409–422.

[5] B. Conrey, *Reciprocity for a cotangent sum, the Riemann hypothesis, and the period function of an Eisenstein series.* Preprint, 2010.

[6] O. Costin and S. Garoufalidis, *Resurgence of 1-dimensional sums of Sum-Product type.* In preparation.

[7] T. Dimofte, S. Gukov, J. Lennells and D. Zagier, *Exact results for perturbative Chern-Simons theory with complex gauge group.* Commun. in Number Theory and Physics **3** (2009), 363–443.

[8] F. J. Dyson, *A walk through Ramanujan's garden.* In "Ramanujan Revisited (Urbana-Champaign, Ill., 1987)," Academic Press, Boston (1988) 7–28. Reprinted in "Selected Papers of Freeman Dyson: With Commentary," Amer. Math. Soc., (1906), 187–208.

[9] S. Garoufalidis and J. Geronimo, *A Riemann-Hilbert approach to the asymptotics of 1-dimensional sums of Sum-Product type.* In preparation.

[10] K. Habiro, *On the quantum sl_2 invariants of knots and integral homology spheres*, Geom. Topol. Monogr. 4 (2002), 5568.

[11] R. Lawrence and D. Zagier, *Modular forms and quantum invariants of 3-manifolds.* Asian J. of Math. **3** (1999) 93–108.

[12] J. Lewis and D. Zagier, *Period functions and the Selberg zeta function for the modular group.* In "The Mathematical Beauty of Physics, A Memorial Volume for Claude Itzykson" (J.M. Drouffe and J.B. Zuber, eds.), Adv. Series in Mathematical Physics **24**, World Scientific, Singapore (1997) 83–97.

[13] J. Lewis and D. Zagier, *Period functions for Maass wave forms. I.* Ann. Math. **153** (2001), 191–258.

[14] D. Zagier, *From quadratic functions to modular functions.* In "Number Theory in Progress, Vol 2" (K. Győry, H. Iwaniec and J. Urbanowicz, eds.), Proceedings of Internat. Conference on Number Theory, Zakopane 1997, de Gruyter, Berlin (1999), 1147–1178.

[15] D. Zagier, *Vassiliev invariants and a strange identity related to the Dedekind eta-function.* Topology **40** (2001), 945–960.

[16] D. Zagier, *Algebraic and asymptotic properties of quantum invariants of knots.* In preparation.

Collège de France, Paris, France
Current address: Max-Planck-Insitut für Mathematik, Bonn, Germany
E-mail address: don.zagier@mpim-bonn.mpg.de

The work of Alain Connes has cut a wide swath across several areas of mathematics and physics. Reflecting its broad spectrum and profound impact on the contemporary mathematical landscape, this collection of articles covers a wealth of topics at the forefront of research in operator algebras, analysis, noncommutative geometry, topology, number theory and physics.

Specific themes covered by the articles are as follows:

- entropy in operator algebras, regular C^*-algebras of integral domains, properly infinite C^*-algebras, representations of free groups and 1-cohomology, Leibniz seminorms and quantum metric spaces;
- von Neumann algebras, fundamental Group of II_1 factors, subfactors and planar algebras;
- Baum-Connes conjecture and property T, equivariant K-homology, Hermitian K-theory;
- cyclic cohomology, local index formula and twisted spectral triples, tangent groupoid and the index theorem;
- noncommutative geometry and space-time, spectral action principle, quantum gravity, noncommutative ADHM and instantons, non-compact spectral triples of finite volume, noncommutative coordinate algebras;
- Hopf algebras, Vinberg algebras, renormalization and combinatorics, motivic renormalization and singularities;
- cyclotomy and analytic geometry over F_1, quantum modular forms;
- differential K-theory, cyclic theory and S-cohomology.

ISBN 978-0-8218-5203-3

9 780821 852033

CMIP/11

www.ams.org
www.claymath.org