March 13, 1985

We have been looking at the periodicity map

\[ [0,\pi] \times \mathfrak{I}_k(V) \rightarrow \mathfrak{I}_{k-1}(V) \]

\[(\Theta, J) \mapsto (\cos \Theta) J^k + (\sin \Theta) J^1\]

Here \( V \) is a \( \mathfrak{G}_k \)-module with inner product and \( \mathfrak{I}_k(V) \) is space of p.a. involution \( J \) anti-commuting with \( g_1, \ldots, g_k \). What I want to do now is to find the character forms on \( \mathfrak{I}_k\delta(V) \) and to establish their compatibility under periodicity.

Look at \( k = 1 \), and write \( \epsilon \) for \( \delta^1 \). Then \( V = V^+ \oplus V^- \) where

\[ \epsilon = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad J = \begin{pmatrix} g & g^{-1} \\ 0 & 0 \end{pmatrix}, \quad g \text{ unitary} \]

and

\[ F_\Theta = \cos \Theta \epsilon + \sin \Theta J \]

\[ = \begin{pmatrix} \cos \Theta & \sin \Theta g^{-1} \\ \sin \Theta \cdot g & \cos \Theta \end{pmatrix} \]

\[ c_\Theta = \frac{1 + F_\Theta}{2} = \begin{pmatrix} \frac{1 + \cos \Theta}{2} & \frac{\sin \Theta g^{-1}}{2} \\ \frac{\sin \Theta \cdot g}{2} & \frac{1 + \cos \Theta}{2} \end{pmatrix} \]

is the orthogonal projector onto the subspace

\[ \text{Im} \left( \begin{pmatrix} 1 \\ x \cdot g \end{pmatrix} \right) \]

\[ x = \frac{\sin \Theta}{1 + \cos \Theta} = \frac{2 \sin \frac{\Theta}{2} \cos \frac{\Theta}{2}}{2(\cos \frac{\Theta}{2})^2} \]

\[ = \tan \left( \frac{\Theta}{2} \right) \]

Thus \( \frac{\Theta}{2} \) is the angle in the graph construction.
The character forms on the Grassmannian are

\[ \frac{1}{2^k k!} \text{tr} F (dF)^k \]

because in terms of the projector \( e = \frac{F + 1}{2} \), we know the curvature is \( e (d e)^2 \) and the character form is \( \frac{1}{k!} \text{tr} e (d e)^k \).

Now we want to compute the pull back of the \( k \)th character form under the map \( \Theta \mapsto F_\Theta \) followed by integrating \( \Theta \) from 0 to \( \pi \).

\[ F_\Theta = \cos \Theta e + \sin \Theta J \]
\[ dF_\Theta = d\Theta [\text{sin} \Theta e + \cos \Theta J] + \sin \Theta dJ \]

These two terms in \( dF_\Theta \) commute because \( dJ \) anti-commutes with \( d\Theta, e, J \). Thus

\[ i(\Theta) F_\Theta (dF_\Theta)^k = F_\Theta \ 2k (\sin \Theta e + \cos \Theta J) (\sin \Theta dJ)^{2k-1} \]

\[ = 2k (\cos \Theta e + \sin \Theta J)(-\sin \Theta e + \cos \Theta J)(\sin \Theta)^2(dJ)^{2k-1} \]

\[ = 2k (-\sin^2 \Theta J e + \cos^2 \Theta e J)(dJ)^{2k-1} \]

\[ = 2k (\sin \Theta)^{2k-1} e J (dJ)^{2k-1} \]

Need

\[ 2k \int_0^{\pi} (\sin \Theta)^{2k-1} d\Theta = 2k \int_0^{\pi/2} (\sin \Theta)^{2k-2} (\cos \Theta)^{-1} 2 \sin \Theta \cos \Theta d\Theta \]

\[ = 2k \int_0^1 t^{k-1} (1-t)^{-1/2} dt = 2k \Gamma(k) \Gamma(\frac{1}{2}) \Gamma(k + \frac{1}{2}) \]

\[ = 2k \frac{(k-1)! \sqrt{\pi}}{\Gamma(\frac{1}{2})} = \frac{2^k k!}{1 \cdot 3 \cdot \ldots \cdot (2k-1)} \]
to the $k$th character form pulled back and integrated gives
\[
\frac{1}{2^{2k+1} k!} \cdot \frac{2^{k+1} k!}{1 \cdot 3 \ldots (2k-1)} = \frac{k!}{2^k k! \cdot 1 \cdot 3 \cdot (2k-1)} = \frac{2^k!}{(2k)!}
\]
\[
= \frac{1}{2} \frac{(k-1)!}{(2k-1)!}
\]
times \text{tr} (\varepsilon J (dJ)^{2k-1})

Thus the odd character form of degree $2k-1$ is
\[
\frac{1}{2} \frac{(k-1)!}{(2k-1)!} \text{tr} (\varepsilon J (dJ)^{2k-1})
\]

This checks for if $J = \begin{pmatrix} 0 & g^{-1} \\ g & 0 \end{pmatrix}$, then
\[
\text{tr} (\varepsilon J (dJ)^{2k-1}) = \text{tr} (L \cdot (\begin{pmatrix} 0 & g^{-1} \\ g & 0 \end{pmatrix} (dg^{-1} dg) (dg^{-1} dg)^{2k-1})
\]
\[
= \text{tr} \left\{ g^{-1} dg (dg^{-1} dg)^{2k-1} - g dg^{-1} (dg^{-1} dg)^{2k-1} \right\}
\]
\[
= (-1)^{k-1} 2 \text{tr} (g^{-1} dg)^{2k-1}
\]

Next take $k = 2$. Here the family is
\[
F_\theta = (\cos \theta) \Phi^2 + (\sin \theta) J
\]
and the differential form to be pulled back is
\[
\frac{1}{2} \frac{k!}{(2k+1)!} \text{tr} (\gamma_! F (dF)^{2k+1})
\]
So we calculate as before
\[ dF = d\theta \left[-\sin \theta J^2 + \cos \theta J \right] + \sin \theta \, dJ \]

\[ F \left( dF \right)^{2k+1} = (2k+1) \, d\theta \left( \cos \theta J^2 + \sin \theta J \right) \left(-\sin \theta J^2 + \cos \theta J \right) \]
\[ \times \left( \sin \theta \, dJ \right)^{2k} \]

\[ = (2k+1) \, d\theta \left( \sin \theta \right)^{2k} \, J^2 \, J(dJ)^{2k} \]

so we get

\[ (2k+1) \, \frac{1}{2} \, \frac{k!}{(2k+1)!} \, \frac{\pi}{\Gamma(k+1) \Gamma(\frac{1}{2})} \, \int_0^\pi \left( \sin \theta \right)^{2k} \, d\theta \]
\[ \times \left( \sin \theta \right)^{2k} \, J(dJ)^{2k} \]
\[ = \frac{1}{2^{2k+1} \, k!} \, \pi \, \Gamma(1,3 \ldots ,2k+1) \, \frac{\Gamma(k+1)}{2^k \, k!} \, \text{tr} \left( J(dJ)^{2k} \right) \]

\[ \frac{1}{2} \, \frac{k!}{(2k+1)!} \, \frac{\pi}{\Gamma(k+1) \Gamma(\frac{1}{2})} \, \text{tr} \left( J(dJ)^{2k} \right) \]

But then recall that \( J = \begin{pmatrix} 0 & -i \theta \\ i \theta & 0 \end{pmatrix} \)
\( J^1 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \)
\( J^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \), so

\( J^1 J^2 \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} = iI \)

and so the end answer is

\[ (2i\pi) \, \frac{1}{2^{2k+1} \, k!} \, \text{tr} \left( J(dJ)^{2k} \right) \]
Let me now try to summarize what we have just done. We first described the basic Bott periodicity maps. The model we use for the \( k \)-th representing space for the \( K \)-theory is as follows. We take a large \( \mathcal{Q}_k \)-module with inner product \( V \) and let \( \mathcal{I}_k(V) \) be the space of self-adjoint involutions on \( V \) which anti-commute with \( q_j, \ldots, q_k \). Then one has the periodicity map

\[
[0,\pi] \times \mathcal{I}_k(V) \to \mathcal{I}_{k-1}(V)
\]

\[
(\Theta, J) \mapsto F_\Theta = (\cos \Theta) J^k + (\sin \Theta) J
\]

**Ex.** \( k=1 \):

\[
\gamma^1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad V = V^+ \oplus V^-
\]

\[
J = \begin{pmatrix} 0 & g \\ g^T & 0 \end{pmatrix}
\]

\[
F_\Theta = \begin{pmatrix} \cos \Theta & \sin \Theta g \\ \sin \Theta g & -\cos \Theta \end{pmatrix}, \quad e = \frac{1 + F_\Theta}{2} = \begin{pmatrix} 1 + \cos \Theta \\ 2 & \sin \Theta g \end{pmatrix}
\]

is the projector onto \( \text{Im} \begin{pmatrix} 1 \\ g \end{pmatrix} \), where

\[
x = \frac{\sin \Theta}{1 + \cos \Theta} = \tan(\Theta/2).
\]

So the periodicity map is the graph construction.

**Ex.** \( k=2 \):

\[
\gamma^1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \gamma^2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad V = W \oplus W
\]

and

\[
J = \begin{pmatrix} 0 & -iF \\ iF & 0 \end{pmatrix}
\]

with \( F^2 = 1 \) on \( W \). Then

\[
F_\Theta = \begin{pmatrix} 0 & \cos \Theta \pm i \sin \Theta F \\ \cos \Theta \mp i \sin \Theta F & 0 \end{pmatrix}
\]

corresponds to the path of unitary operators \( \cos \Theta + i \sin \Theta F \) joining
The second thing we did was to describe the Chern character forms on $k^1(V)$, at least for $k=3,4,2$ and check that they are compatible with the periodicity maps.

For $k=0$ the $j$-th character form is

$$\frac{1}{2^{j+1}} \frac{1}{j!} \text{tr}(F(dF)^j).$$

For $k=1$ the character form of degree $2j-1$ is

$$\frac{1}{2} \left( \frac{(j-1)!}{(2j-1)!} \right) \text{tr}(F(dF)^{2j-1}).$$

For $k=2$ I get something like

$$\frac{1}{2^{j+1}} \frac{1}{j!} \text{tr}(g_1 g_2 F(dF)^j).$$

up to a factor of $2i$.

**Summary:**
I have been thinking about Bott periodicity. On the first level it amounts to explicit homotopy equivalences:

$$U \sim \Omega BU \quad BU \sim \Omega U.$$

I construct maps

1) $U \rightarrow \Omega BU \quad BU \rightarrow \Omega U$

using the graph map in the former and the
Cayley transform in the latter.

Then I unify the two constructions using Clifford algebras. Let $V$ be an ungraded $C_k$-module with inner product, and let $\mathcal{D}_k(V)$ be the space of involutions on $V$ which anti-commute with $g^i$, $g^k$. This space is alternately a Grassmannian or unitary group depending on the parity of $k$ because of the algebraic periodicity of the Clifford algebras. The maps $1)$ are part of maps

$$\mathcal{D}_k(V) \longrightarrow \Omega \mathcal{D}_{k-1}(V)$$

so far I have just constructed maps, but I have not explained why they are homotopy equivalences. I have not yet introduced the operators which are essential for the proof of periodicity.
Manifold $E$ real vector bundle over $M$, $X$ section of $E$ transversal to the zero section $Z$, the submanifold of $M$ where $X$ vanishes. Over $M$ we have a complex of vector bundles

$$\Lambda^2 E^* \xrightarrow{X} \Lambda^1 E^* \xrightarrow{X} \Lambda^0 E^*$$

which is acyclic off $Z$. Claim the complex of top. vector spaces (cohomology spaces)

$$\Gamma(M, \Lambda^2 E^*) \xrightarrow{\Gamma(M, X)} \Gamma(M, \Lambda^1 E^*) \xrightarrow{\Gamma(M, X)} \Gamma(M) \xrightarrow{\Gamma(Z)} 0$$

is split exact.

I believe this is true and will try to work out a proof.

The first thing to try to prove is that it is local on $M$. Suppose we have an open covering $\{U_\alpha\}$ of $M$ and contracting homotopies $h_\alpha$ over $U_\alpha$. I can assume that the covering has a subordinate partition of unity $\sum f_\alpha = 1$. Then

$$h = \sum f_\alpha h_\alpha$$

will be a contracting homotopy over $M$.

What is going on here is we have a complex $K(M)$:

$$\Gamma(M, \Lambda^2 E^*) \xrightarrow{\Gamma(M, X)} \Gamma(M, \Lambda^1 E^*) \xrightarrow{\Gamma(M, X)} \Gamma(M) \xrightarrow{\Gamma(Z)} 0$$

and an embedding and projection

$$K(M) \xrightarrow{\cong} \prod_{\alpha} K(U_\alpha).$$
March 15, 1985

Suppose given \( F = \begin{pmatrix} 0 & P \\ 0 & 0 \end{pmatrix} \) in \( \mathbb{H} = \mathbb{H}^+ \oplus \mathbb{H}^- \) where \( \mathbb{H}^+ \) are \( A \)-modules and \( F \) is \( \mathcal{F} \)-summable. Then we have two ways of constructing cyclic cocycles attached to \( F \), \( \varphi_n \), \( n \) even \( n > q \).

1) Grothendieck’s approach: The operator \( F \) determines a map \( K_0 A \to \mathbb{Z} \) as follows. Given an idempotent matrix \( e \) of size \( n \), one forms \( \begin{pmatrix} \beta^e \end{pmatrix} : (\mathbb{H}^+)^n \to (\mathbb{H}^-)^n \) and reduces by \( e \) to obtain \( \begin{pmatrix} \beta^e \end{pmatrix} : e(\mathbb{H}^+)^n \to e(\mathbb{H}^-)^n \). This is Fredholm and it has an index. One has

\[
\text{Ind}(\begin{pmatrix} \beta^e \end{pmatrix}) = (-1)^k \text{tr}(\epsilon [F, e]^{2k}) = \frac{\epsilon 1}{2} \text{tr}(\epsilon F [F, e]^{2k+1})
\]

provided the traces make sense.

This motivates the expression

\[
2\varphi_n(a_0, \ldots, a_n) = \text{tr}(\epsilon F [F, a_0] \cdots [F, a_n])
\]

\[
= \text{tr}(\epsilon (F^2 a_0 F - F a_0 F) [F, a_1] \cdots [F, a_n])
\]

\[
= 2 \text{tr}(\epsilon a_0 [F, a_1] \cdots [F, a_n])
\]

Better:

\[
\varphi_n(a_0, \ldots, a_n) = \text{tr}(\epsilon a_0 [F, a_1] \cdots [F, a_n])
\]

\[
= \frac{1}{2} \text{tr}(\epsilon F [F, a_0] \cdots [F, a_n])
\]

(\text{\( n \) even})

This is defined when \( F \) is at least \( n \)-summable and is a cyclic \( n \)-cocycle.

2) My approach: Let the gauge group \( \mathcal{G} = \text{Maps}(M, \mathcal{U}_n) \) act on \( \mathcal{H} = (\mathbb{H}^+) \oplus (\mathbb{H}^-)^n \) and consider the map

\[
g \mapsto g \begin{pmatrix} a_0 & \cdots & a_n \\ \end{pmatrix} g^{-1} = \begin{pmatrix} 0 & g^{-1} a_0 g \\ g a_n g^{-1} & 0 \end{pmatrix}
\]
Then the natural invariant forms on the space of $\mathbf{F}$'s will pull back to left-invariant forms on $\mathbf{G}$. On the space of invertible operators the basic invariant forms are

$$\text{tr} \left( Q^{-1} dQ \right) \text{odd} \quad (\text{More precisely, } 
\frac{(-1)^{k-1}(k-1)!}{(2k-1)!} \text{tr} \left( Q^{-1} dQ \right)^{2k-1})$$

Here $Q : H^+ \to H^-$ and invariance means with respect to left and right multiplication. We are considering the map $g \mapsto Q_g = gP^{-1}g^{-1}$ (abbreviate $p^\infty$ to $p$); then

$$Q^{-1}dQ = (gPg^{-1})^{-1} (dgPg^{-1} - gPg^{-1}dgg^{-1})$$

$$= g^{-1} \left[ g^{-1}dg, P \right] g^{-1}$$

i.e. at the identity of $G$ we have that

$$\iota_x (Q^{-1}dQ) = p^{-1} [x, p]$$

so we get the left-invariant forms

$$\text{tr} \left( p^{-1} [\theta, p] \right) \text{odd} \quad (\star)$$

I guess I need $F = \begin{pmatrix} 0 & p^{-1} \\ p & 0 \end{pmatrix}$

$$\text{tr} \left( \epsilon F (dF)^{2k-1} \right) = \text{tr} \left( \begin{pmatrix} 0 & p^{-1} \\ p & 0 \end{pmatrix} \left( dp \right)^{2k-1} \right)$$

$$= 2 \text{tr} \left( \begin{pmatrix} 0 & p^{-1} \\ p & 0 \end{pmatrix} \right)^{k-1}$$

$$= (-1)^{k-1} 2 \text{tr} \left( p^{-1} dp \right)^{2k-1}$$

So what happens is that when we convert the left-invariant forms $(\star)$ to cyclic cocycles we end up with the same cyclic cocycles as Connes.
\[ \varphi_n(a_0, \ldots, a_n) = \text{tr} \left( \varepsilon a_0 [F, a_1] \cdots [F, a_n] \right) \quad n=2k \]

\[ = \frac{1}{2} \text{tr} \left( \varepsilon F [F, a_0] \cdots [F, a_{2k}] \right) \]

\[ = (-1)^k \text{tr} \left( p^{-1} [p, a_0] \cdots p^{-1} [p, a_{2k}] \right) \]

**Formulas:**

\[ \varphi(a_0, a_1, \ldots, a_{2k}) = \text{tr} \left( \varepsilon a_0 [F, a_1] \cdots [F, a_{2k}] \right) \]

\[ = \frac{1}{2} \text{tr} \left( \varepsilon F [F, a_0] [F, a_1] \cdots [F, a_{2k}] \right) \]

\[ = (-1)^k \text{tr} \left( p^{-1} [p, a_0] \cdots p^{-1} [p, a_{2k}] \right) \]

\[ \frac{1}{2} \text{tr} \left( \varepsilon F (dF)^{2k+1} \right) = (-1)^k \text{tr} \left( p^{-1} dp \right)^{2k+1} \]

**Ind (ePe) = \**

\[ = (-1)^k \frac{1}{2} \text{tr} \left( \varepsilon F [F, e]^{2k} \right) \]

\[ = (-1)^k \text{tr} \left( \varepsilon e [F, e]^{2k} \right) \]

\[ = \text{tr} \left( p^{-1} [p, e] \right)^{2k+1} \]

The problem to be solved is to relate these two ways of getting the cyclic cocycles. This apparently requires some understanding of the periodicity process.

I want to describe how to pair an idempotent \( e \) over \( A \) with a left-invariant differential form \( \omega \) on \( \Omega = \text{Map}(M, \Omega) \). This would explain Connes' approach in terms of mine.
The idempotent defines an element of \( K^0(M) \), the form \( \omega \) defines a coh. class on \( \mathcal{A} \). Suppose \( \deg(\omega) = 2k+1 \). Consider the diagram

\[
\begin{array}{ccc}
K^0(M) & \xrightarrow{\sim} & K^0(\mathcal{S}^{2k+2} \mathcal{A} M) = \tilde{\pi}_{2k+2}(BU^M) \\
\downarrow & & \downarrow \\
HC_{2k}(A) & \leftarrow & H_{2k+1}(\mathcal{A}) \leftarrow \tilde{\pi}_{2k+1}(U^M)
\end{array}
\]

where presumably the dotted arrow is Connes' Chern character map. This sends the idempotent \( e \) into a constant multiple of \( (e \ldots e) \in C_{2k}(A) \). The constant is rigged so as to be \( 2^{k+1} \) compatible with the \( S \)-operator. It would seem to be

\[
\frac{1}{2^k 1 \cdot 3 \cdots (2k-1)} \quad e^{\otimes (2k+1)}
\]

\[
= \frac{1}{2} \frac{(k-1)!}{(2k-1)!} \quad e^{\otimes (2k+1)}
\]

based on chasing through the double complex.

So the problem is to show commutativity of the above diagram. But we haven't made precise the map \( H_{2k+1}(\mathcal{A}) \rightarrow HC_{2k}(A) \), which is dual to the process taking a cyclic cocycle to left invariant diff. form.
Let’s review the problem. One start with what Onnes calls a Fredholm module; i.e. a Hilbert space $H$ which is an $A$-module and on which one has an involution $F$ such that $[F,a]$ is compact for all $a \in A$. There are graded and ungraded cases.

To fix the ideas suppose $M$ even, and and $F = (P, P')$ where $P : H^+ \to H^-$ is a $\Psi DO$ of order 0 between two vector bundles.

A Fredholm module which is $g$-summable determines a sequence of cocycles which I understand best using the LQT theorem as follows. We replace $H, F$ by $H \otimes r, F \otimes r$ and let $G = \text{Un}(r, A)$ act in the new $H$. Then $g \mapsto gFg^{-1}$ gives a map from $G$ to the restricted Grassmannian. On this Grassmannian are the standard character forms which pull-back to give left invariant forms on $G$.

This map $G \to \text{Grass}$ represents a $K$-class on $G$. This is the index of the family of operators $g \mapsto gFg^{-1}$ parametrized by $G$ in some tautological sense. So the left-invariant forms on $G$ are supposed to represent the character of the index of the family.

(I don’t know if there is some problem with the character forms on the restricted Grassmannian representing the character. Ultimately we should take a simple attitude and replace $G$ by a map of a finite-dimensional manifold to $G$.)

Now the left-inv. diff. forms on $G$ we have
constructed represent primitive cohomology of $H$, and hence can be detected by homotopy of $H$, i.e., maps $S \to H$ where $S$ is a sphere. By periodicity, such maps are given by idempotent matrices or invertible matrices over $A$.

So the basic problem for me is to carry out the integration of these differential forms.

To be specific, suppose we are in the even case so that $M$ is even, $F$ is graded, and we start with an element $\alpha$ of $K^0(M)$ by

$$K^0(M) = K^0(S^{2k} \wedge M) = [S^{2k} \wedge M, \mathbb{Z} \times BU]$$

$$= [S^{2k-1} \wedge M, U] = [S^{2k-1}, \mathbb{Z}]$$

and my problem is to take the $(2k-1)$-degree form on $H$ and integrate over this $S^{2k-1}$.

Now let's assume this $(2k-1)$-degree form on $H$ represents the character form of the index.

Consider the diagram

$$
\begin{array}{ccc}
K^0(M) & \xrightarrow{c} & K^0(S^{2k} \wedge M) \\
\downarrow & & \downarrow \\
K^0(pt) & \xrightarrow{c} & K^0(S^{2k})
\end{array}
$$

where $c$ represents capping with the $K$-homology class represented by the operator $F$. In this diagram, the image of $\alpha$ under $c$...
should be the index of the family of operators
on the fibres of $S^{2k} \times M$ over $S^k$
which is obtained by mixing $F$ with the
bundle over $S^{2k} \times M$ obtained from $\alpha$.

So it should follow from the commutativity
of this diagram and the standard behavior of the
Chern character with respect to periodicity that
integrating the differential forms over the $S^{2k-1} \rightarrow \Sigma$
belonging to $\alpha$ gives the index of $\alpha$ mixed with
$F$.

Now it should be possible to prove this
analytically.
It is necessary to get a hold on these $F$'s used by Kasparov and Connes. I have treated periodicity using finite-dimensional $F$'s. Recall that given a $C^*$-module $V$ we can consider the space $I_k(V)$ of involutions $F$ on $V$ anti-commuting with $f_1, \ldots, f_k$. This is alternatively a Grassmannian or unitary $G(k,p)$. $I_0(V)$ is the space of subspaces of $V$, $I_1(V)$ is the space of unitary maps $V^+ \sim V^-$. 

I now want to relate the periodicity maps using these $I_k$ spaces to the periodicity maps one gets using Clifford modules.

Consider the Bott class in $K^*_1(C)$. It is represented by the Hopf line bundle on $S^2 = P(C)$ and hence by an idempotent $2 \times 2$ matrix $e_1$ over $C$:

$$e = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \frac{1}{1 + \|z\|^2} \begin{pmatrix} 1 & z^* \\ z & 1 \end{pmatrix}$$

The image of this is the subbundle $S(-1)$.

It is interesting to represent the Bott class as a map to Fredholm operators. Over the disk $|z| < 1$ one has the Fredholm operator $z e_1 e_1^* + e_2 e_2^* + \cdots$ in $l^2$. This has index zero and is invertible for $z \neq 0$. I want to trivialize this map on $|z| = 1$ using Kreps's trick. In the present case this is easy.

Let's recall the basic infinite repetition argument

$$a + (-a) + a + (-a) + a + \cdots = a + 0 + \cdots = a$$

$$a + (-a) + a + (-a) + a + \cdots = 0 + 0 + \cdots = 0.$$

Thus we loop at the map from the unit circle $|z| = 1$ to the unitary group of $l^2$ given by

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \frac{1}{1 + \|z\|^2} \begin{pmatrix} 1 & z^* \\ z & 1 \end{pmatrix}$$
and join the blocks \( \begin{pmatrix} z & z^{-1} \\ \overline{z} & \overline{z}^{-1} \end{pmatrix} \) to the identity to get the required trivialization.

I recall that the proof of Kisyński's theorem proceeds by showing that a map \( X \to U \) can be deformed into the subgroups fixing a subspace of infinite dimension and codimension. Then one shows the contractibility of the map by the above infinite repetition argument.

A third way to represent the Bott class is via Clifford modules. We take the trivial bundle over \( X \) with fibre \( S_2 \) and give it the odd degree self-adjoint endomorphism

\[
L = x z^1 + y \overline{z}^2 = \begin{pmatrix} 0 & z \\ \overline{z} & 0 \end{pmatrix}
\]

Let's go over the problem under consideration. I start with a Fredholm module \((H,F)\) associated to an operator on \( M \). To fix the ideas suppose we are in the even case. Then attached to \((H,F)\) are cyclic cocycles \( \varphi_{2k} \) for \( k \) sufficiently large. I define them as left-invariant forms in \( \mathcal{L} = \text{Map}(M, \mathcal{A}) \) and apply the LQT-theory to write them in terms of cyclic cocycles.
What I learned today:

Let's adopt Connes' viewpoint relative to which a Fredholm module \( F \) is to be regarded as a reduction of an invertible Fredholm module. Then one can define the family of cyclic cocycles attached to a Fredholm module. They are left-invariant differential forms on \( \mathfrak{g} = \text{Maps}(M, U) \).

Let's take then the even case and the cyclic cocycle \( \eta_k \). This corresponds to a form of degree \( 2k+1 \) in \( \mathfrak{g} \).

Next I take \( e \) representing an element of \( K_0(M) \) and use periodicity to associate to \( e \) a map \( S^{2k+1} \times M \rightarrow U \), i.e. a map \( S^{2k+1} \rightarrow \mathfrak{g} \).

Then I can integrate the diff. form over this sphere.

To see what we get we use the fact that we might as well reduce by the idempotent \( e \) first. Thus we can suppose \( e = 1 \) by replacing the Fredholm module by its reduction relative to \( e \). (Define this by taking a direct sum of copies of \( L^2(\mathfrak{g}) \).

But if \( e = 1 \), then the map \( \mathbb{S}^{2k+1} \times M \rightarrow U \) factors \( \mathbb{S}^{2k+1} \times M \rightarrow \mathbb{S}^{2k+1} \rightarrow U \), where the second map is the zero map. \( (\mathbb{S}^{2k+1} \rightarrow U(\mathfrak{g}) \rightarrow \mathfrak{g} \cup \mathfrak{g}) \).

Thus \( \mathbb{S}^{2k+1} \rightarrow \mathfrak{g} \) factors \( \mathbb{S}^{2k+1} \rightarrow U \rightarrow \mathfrak{g} \), where the second is the inclusion of the constant gauge transformations.
So one takes the cyclic cocycle over $A$ and restricts it to $C$ where it becomes a multiple of the unique cyclic cocycle of that degree. It follows that if things are normalized correctly, the integral of the differential form will be just the index.
March 18, 1985

Let's go over the AS proof of periodicity based on Kuijper's theorem. It would be nice to understand why it really works. But I should be able to link it with the Bott periodicity maps constructed above in finite dimensions.

Outline of my previous understanding of this proof:

\[ A = B/\mathbb{K} \] is the Calkin algebra.

\[ F = \text{Fredholm operators} \] is the inverse image of \( A^* \) in \( B \). We know \( F \) deforms to the subgroup of essentially unitary operators, i.e., inverse image of \( U(\mathbb{A}) \).

Similarly \( F_1 = \text{self-adjoint Fredholm operators} \) in \( B \), the self-adjoint operators having essential spectrum on both sides of 0 deforms to \( U(\mathbb{A})_0 \).

This maps onto \( J(\mathbb{A}) = \text{space of non-trivial idempotents} \) in \( A \).

The first step is the exact sequence:

\[ 1 \to U \to \tilde{U} \to U(\mathbb{A})_0 \to 1 \]

\[ 1 \to U(\mathbb{A})_0 \to U(\mathbb{A}) \xrightarrow{\text{index}} \mathbb{Z} \to 1 \]

and because of Kuijper's thm. that \( \tilde{U} \sim pt \), one

\[ BU \sim U(\mathbb{A})_0 \]

\[ \mathbb{Z} \times BU \sim U(\mathbb{A}) \sim F_1 \]

The next step is

\[ J(\mathbb{A}) = U(\mathbb{A})^* \times (\mathbb{A})^* \sim B \otimes U(\mathbb{A}) \]
which should be a version of the standard proof that $BU =$ Grassmannian. I don’t know if Knüpfer’s result is used here.

The key step is to show that the map

$$F_1 \to U, \quad A \to -\exp(iA)$$

is a homotopy equivalence. This uses the natural filtration by number of zero (resp. -1) eigenvalues + Knüpfer’s thm. to identify the homotopy types of the strata. The idea is that the $\pm 1$ spaces for an operator in $F_1$ became the $\pm 1$ eigenspaces of an operator in $U$. Thus we are concerned with the space of splittings of Hilbert space, i.e. with $\tilde{U}/\tilde{U} \times \tilde{U}$

which is contractible by Knüpfer’s thm. *

Let’s now try to fit this with our earlier discussion about periodicity maps using Clifford algebras.

* Idea: We know that by looking at the eigenvalues of a self-adjoint operator in finite dimensions that we get the simplicial structure of the building. There’s a possibility that the simplicial structure is involved in cyclic cohomology.
Recall that if $V$ is an ungraded $C_k$-module with inner product we put

$$I_k(V) = \{ T \in \text{End}(V) \mid T^2 = 1, T = T^*, 1_T + T \leq 0 \}$$

so that

$$\mathbb{I}_0(V) = \text{Grass}(V)$$
$$\mathbb{I}_1(V) = \text{UnitaryIsms}(V^+, V^-)$$

We want to consider the case where $V$ is a Hilbert space and $T$ is Fredholm.

So I am being led to look at the K-theory of $C_k$ in the Kasparov sense:

We know graded $C_k$-modules which are Fredholm do not give $K(R^k)$, but they do in the Kasparov sense. How?

$k = 0$. A graded $C_0$-module is a graded v.s. $V = V^+ \oplus V^-$, and the naive K-theory of these is $\mathbb{Z} + \mathbb{Z}$. In the KK game one looks at $H = H^+ \oplus H^-$ (both infinite dim. in the stable model) equipped with, and $F = (0, \mathcal{B})$ of odd degree such that $F^2 - 1$ is compact. And one works modulo homotopy, getting $\mathbb{Z}$ for the K-theory.

$k = 1$. A graded $C_1$-module is equivalent to an ungraded v.s. The naive K-group is $\mathbb{Z}^\infty$. In the KK game one looks at $F$ on a Hilbert space $H$ such that $F^2 - 1$ is compact. Working stably both the $F = 1$ and $F = -1$ spaces are infinite dim. Working mod homotopy the K-group is zero.
What are the periodicity maps? We want to use the same maps as before. I think we want to exploit the idea that there are three kinds of $F$'s - Fredholm, unitary, unitaries $\equiv I \mod \mathcal{H}$.

Let us now consider the periodicity map

$$I(a) \mapsto \Omega U(a)$$

which recall takes an involution $J$ into the paths from 1 to $-1$ given by

$$(\cos \theta) + i(\sin \theta)J \quad 0 < \theta \leq \pi$$

Where does this come from? We are really dealing with the map from $k=1$ to $\mathcal{K}$ in $k=0$. A graded $C^*$-module with $F$ has the form

$H = H^+ \oplus H^-$ where $H^+ = H^-$ and

$$\varepsilon = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad \chi = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \quad F = \begin{pmatrix} -iJ \\ iJ \end{pmatrix} = \gamma^2 J$$

and where $J^2 \equiv 1 \mod \mathcal{H}$. Then

$$\cos \theta \gamma^2 + i \sin \theta F = \begin{pmatrix} 0 & \cos \theta - i \sin \theta J \\ \cos \theta + i \sin \theta J & 0 \end{pmatrix}$$

so we get the essentially unitary operator

$$\cos \theta + i(\sin \theta)J.$$ 

Note that if $J^2 = 1$ then

$$e^{i\theta J} = \cos(\Theta J) + i \sin(\Theta J)$$

$$= \cos \theta + i \sin(\Theta)J.$$
Thus we have the periodicity map
\[ I(a) \rightarrow \Omega(U(a); 1, -1) \]
which we want to show is a homotopy equivalence. On the other hand the fibration
\[ U \rightarrow \tilde{U} \rightarrow U(a) \]
with the contractibility of \( \tilde{U} \) determines a map
\[ \Omega(U(a); 0, 1) \rightarrow U \]
by lifting paths, which is a bijection.

Now over \( F_1 \) we can do this path lifting explicitly. Given \( A \in F_1 \), we can consider
the path
\[ e^{i\theta A}, \quad 0 \leq \theta \leq \pi \]
in \( \tilde{U} \). This
starts at 1 and covers
\[ e^{i\theta J} = \cos \theta + i \sin \theta J \]
where \( A \mapsto J \). Thus the endpoint
map is
\[ A \mapsto e^{i\theta A} \]
which is \( \tilde{U} \) in \( -U \).

\[ \begin{array}{ccc}
F_1 & \xrightarrow{\exp} & -U \\
\downarrow & & \downarrow \\
I(a) & \rightarrow & \Omega(U(a); 1, -1)
\end{array} \]

commutes up to homotopy. We conclude then
that the Bott periodicity map is an bijection iff \( \exp \) is.

So far we have the part \( F_1 = 2F_0 \) of periodicity. We next look at \( k = 2 \).
March 20, 1985

\[ F_0 = \{ F = (p p^* \mid F^2 - 1 \in K) \} \]

\[ = \{ p \in B(H) \mid pp^* = p^*p = 1 \mod K \} \]

\[ F_1 = \{ p \in B(H) \mid p^2 = 1 \mod K \} \]

are the two spaces taking part in periodicity. They map onto the unitary elements (resp. involutions) in the Calkin algebra A with contractible fibres.

I want now to set up the periodicity maps. One idea is to consider the fibrings

\[ U \longrightarrow U(B) \longrightarrow U(A) \]

\[ \mathbb{Z}_2 \longrightarrow \mathcal{I}(B) \longrightarrow \mathcal{I}(A) \]

defined as follows. In the second case we consider self-adjoint involutions in B and A and the fibre of involutions \( \mathcal{I} \) in B which are congruent to a fixed involution. This is the restricted Grassmannian. In the first case we do the same thing but using odd involution relative to a grading. Thus we take \( F = (p p^* \mid 1 \mod K) \) over \( A \) and try to lift to an \( F \) in \( B \) which is possible over the index zero component. The fibre is then the unitaries congruent to 1 \( \mod K \).
The next step is to consider the Bott maps down in $\mathfrak{g}$. In the second case we have the map

$$U(\mathfrak{g}) \longrightarrow \Omega(\mathfrak{g}(\mathfrak{g}); \varepsilon, -\varepsilon)$$

$$F = (\frac{\partial}{\partial s}) \quad \mapsto \quad F_0 = (\cos \Theta) \varepsilon + (\sin \Theta) F$$

$$= \begin{pmatrix} \cos \Theta & \sin \Theta \varepsilon^* \\ \sin \Theta & -\cos \Theta \end{pmatrix}$$

which we know associates to $g$ the graph of $(\tan \frac{\Theta}{2})^2$ as $\Theta$ runs over $[0, \pi]$. We want to lift a loop in $\mathfrak{g}(\mathfrak{g})$ into $\mathfrak{g}(\mathfrak{b})$ and then take the endpoint getting a map

$$U(\mathfrak{g}) \longrightarrow \mathbb{R}$$

defined up to homotopy. We can do this explicitly by replacing $U(\mathfrak{g})$ by $F_0$ and using

$$F = (\frac{\partial}{\partial s}) \quad \mapsto \quad F_0 = e^{\frac{\partial}{2} F \varepsilon} \varepsilon e^{-\frac{\partial}{2} F \varepsilon} = e^{\Theta F \varepsilon}.$$

What we are doing is to regard $\mathfrak{g}(\mathfrak{b})$ as $U(\mathfrak{b})/\text{centralizer} \varepsilon$ and using the exponential map in $U(\mathfrak{b})$ restricted to the subspace of the Lie algebra complementary to the centralizer of $\varepsilon$.

Let me try to understand things a bit better. The thing about the Atiyah-Singer proof which I missed before is the fact that the restricted Grassmannian has the same kind of
eigenvalue description as the restricted unitary groups. This should have been obvious if we think of the Grassmannian as a symmetric space.

We can look at things as follows. Let $e$ be our given involution on $V$ and let the unitary groups act on these involutions. Thus we have

$$U/U^+ \times U^- \longrightarrow J(V)$$

$$g \longmapsto g \in J_g^{-1}.$$ 

We split the Lie algebra $g$ of $U$ into $k + p$ as usual and then consider the exponential map

$$\exp: p \longrightarrow \text{Grass}.$$ 

This should be onto (at least in ft.dims). Now if $g = e^x$, $x \in p$ then $g \in J = g^{-1}$ so

$$g \exp = g^2.$$ 

Notice that for any $g \in U$

$$(g \in J)^2 = g \in J = 1 \iff g \in J = g^{-1}.$$ 

Thus we can identify the Grassmannian with the subset of $U$ consisting of $g$ with $g \in J = g^{-1}$.

**Proof:** Consider the set of involutions $J$ on $V$.
is a bijection of this space with the subset \( \mathcal{U} \) consisting of \( g \) with \( \varepsilon g \varepsilon = g^{-1} \).

**Proof:** If \( \varepsilon g \varepsilon = g^{-1} \), then \( (\varepsilon g \varepsilon)^2 = 1 \), so \( \varepsilon g \varepsilon \) is an involution and its unitary as \( g \), \( \varepsilon \) are. If \( \mathcal{I} \) is a unitary involution, then \( (\varepsilon \mathcal{I} \varepsilon)^{-1} = \varepsilon \mathcal{I} = \varepsilon (\mathcal{I} \varepsilon) \varepsilon \).

So to describe the Grassmannian one takes unitaries \( g \) such that \( \varepsilon g \varepsilon = g^{-1} \). This implies that the eigenvalues are stable under conjugation.

So you see the spectral structure rather nicely. The eigenvalues not in \( \pm 1 \) will have eigenspaces paired via \( \varepsilon \). The \( +1 \) eigenspace will be stable under \( \varepsilon \); also the \( -1 \) eigenspace.

I now want to work out the geometric meaning of these eigenvalues. Thus I start with a fixed involution \( \varepsilon \) on \( V \). Then by the above, the Grassmannians of \( V \) appear as components of \( \{ g \in \mathcal{U} \mid \varepsilon g \varepsilon = g^{-1} \} \). For example
if \( \varepsilon = 1 \), then we get \( \{ g \in U \mid g^2 = 1 \} \). But this isn't interesting because we really want to consider the case where \( \varepsilon \) belongs to the Grassmannian.

In any case we can look at the spectral structure to get some understanding. \( \varepsilon \) defines a grading \( V = V^+ \oplus V^- \). Clearly given \( W \subset V \) we want to project it onto \( V^+ \) and \( V^- \) to get decompositions of these; this must correspond to the \( \pm 1 \) eigenvalues of \( g \). Leave those for the moment and consider the case where \( W \) sets up an isomorphism between \( V^+ \) and \( V^- \).

Let's first take \( V^+ = V^- = 0 \).

The unitary corresponding to \( \varepsilon \) is

\[
\begin{pmatrix}
\cos \theta + i \sin \theta \\
\sin \theta - i \cos \theta
\end{pmatrix}
\]

so the corresponding unitary is

\[
J_\theta \varepsilon = \begin{pmatrix}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{pmatrix}
\]

This is still not as clear as I would like, because I don't see the unitary transformation very well. The actual eigenspaces are in some complex direction.
Witten's way to attach a subspace to an $F=\begin{pmatrix} 0 & T^* \\ T & 0 \end{pmatrix}$ is to consider the 'massive Dirac of'

$$D = m\mathcal{E} + F = \begin{pmatrix} m & T^* \\ T & -m \end{pmatrix}$$

which satisfies $D^2 = m^2 + T^* T \geq m^2$

and take the decomposition into the positive and negative eigenspaces. To find this we compute the

involution:

$$J = D/|D| = \begin{pmatrix} m & T^* \\ T & -m \end{pmatrix} \begin{pmatrix} m^2 + T^* T & 0 \\ 0 & m^2 + T^* T \end{pmatrix}^{-1/2}$$

$$= \begin{pmatrix} m(m^2 + T^* T)^{-1/2} & T^*(m^2 + T^* T)^{-1/2} \\ T(m^2 + T^* T)^{-1/2} & -m(m^2 + T^* T)^{-1/2} \end{pmatrix}$$

then form the projector

$$e = \frac{J+1}{2} = \begin{pmatrix} m(m^2 + T^* T)^{-1/2} + 1 & T^*(m^2 + T^* T)^{-1/2} \\ T(m^2 + T^* T)^{-1/2} & -m(m^2 + T^* T)^{-1/2} + 1 \end{pmatrix}$$

which should have the positive eigenspace for its image, and then take the image of $H^+$.

We get (at least if $m \neq 0$)

$$\text{Im} \begin{pmatrix} m(m^2 + T^* T)^{-1/2} + 1 \\ T(m^2 + T^* T)^{-1/2} \end{pmatrix} = \text{Im} \begin{pmatrix} m + (m^2 + T^* T)^{1/2} \\ T \end{pmatrix}$$

$$= \text{graph of} T(m + (m^2 + T^* T)^{1/2})^{-1}$$

Suppose $T$ unitary. Then $x = (m + (m^2 + T^* T)^{1/2})^{-1} = (m + \sqrt{m^2 + T})^{-1}$

$= -m + \sqrt{m^2 + T}$ goes from $+\infty$ to 0 as $m$

goes from $-\infty$ to $+\infty$. 
Think in terms of the periodicity map \((\theta = 1)\)
\[
\cos \theta \xi + (\sin \theta) F \quad 0 \leq \theta \leq \pi
\]
which goes from \(\varepsilon\) to \(-\varepsilon\). Now I consider the involution
\[
\frac{D}{|D|}
\]
where \(D = m \varepsilon + F\)
and you see that \(m\) goes from \(+\infty\) to \(-\infty\).
Finally you can consider the graph version
\[
\text{Im} \left( \frac{1 + \cos \theta}{2} \right) = \text{Im} \left( \frac{1}{x^T} \right)
\]
where \(x = \tan \frac{\theta}{2}\) goes from \(0\) to \(\infty\).
In each case I think of a path from \(\varepsilon\) to \(-\varepsilon\) or from \(H^+\) to \(H^-\).
Now we decompose according to the eigenvalues of \(T\) which as we have seen are real nos. \(\lambda \geq 0\). Then there is a difference as to the actual correspondence between the parameters. Say \(T = \lambda > 0\). Then

I have to go over the three possibilities:

1) Graph
\[
\left( \frac{x^2}{\lambda} \right) \quad 0 \leq x < \infty
\]

2) Witten = graph
\[
\left( \frac{\lambda}{m + \sqrt{m^2 + \lambda^2}} \right)
\]

3) \(e^{\theta F \xi} = \begin{pmatrix} \cos(\theta \lambda) + \sin(\theta \lambda) \\ \sin(\theta \lambda) - \cos(\theta \lambda) \end{pmatrix}\) gives graph \(\tan \frac{\theta \lambda}{2}\)
So the speed of the correspondence depends upon \( \lambda \).

Now what I have to do is to understand whether the periodicity map can be done using these variants. The essential object is the null-spaces of \( T \) and \( T^* \).

So let's see what happens to our path when we encounter a zero eigenvalue for \( \Theta (O, T^*) \). Thus we take \( \lambda = 0 \).

1) In the graph construction we get simply the constant path at \( H^+ \).

2) In the Witten case we look at the positive eigenspace of \( (0 \quad 0) \) and this jumps from \( H^+ \) to \( H^- \) as \( m \) passes from \( > 0 \) to \( < 0 \).

3) Constant path at \( H^+ \).

Is it possible that the graph could give a good periodicity map?
Let \( f : \mathbb{N}_{\geq 0} \to \mathbb{R} \) satisfy
\[
f(n+m) \geq f(n) + f(m).
\]
Extend \( f \) to \( \mathbb{N} \) by letting \( f(0) = 0 \); this inequality still holds.
Use the division algorithm:
\[
n = qd + r, \quad 0 \leq r < d, \quad q = \left\lfloor \frac{n}{d} \right\rfloor > 0.
\]
Then by induction
\[
f(n) \geq \left\lfloor \frac{n}{d} \right\rfloor f(d) + f(r)
\]
do
\[
\frac{f(n)}{n} \geq \frac{1}{n} \left\lfloor \frac{n}{d} \right\rfloor f(d) + \frac{f(r)}{n}
\]
and letting \( n \to \infty \) we get
\[
\lim_{n \to \infty} \frac{f(n)}{n} \geq \left( \lim_{n \to \infty} \frac{1}{n} \left\lfloor \frac{n}{d} \right\rfloor \right) f(d) + \lim_{n \to \infty} \frac{f(r)}{n} = \frac{f(d)}{d}
\]
so
\[
\lim_{n \to \infty} \frac{f(n)}{n} \geq \frac{\lim_{d \to \infty} f(d)}{d}
\]
and we see that \( \lim_{n \to \infty} \frac{f(n)}{n} \) exists (possibly it is \( +\infty \)).

**Lemma:** If \( f : \mathbb{N}_{\geq 0} \to \mathbb{R} \) satisfies \( f(m+n) \geq f(m) + f(n) \) then \( \lim_{n \to \infty} \frac{f(n)}{n} \) exists in \( (-\infty, \infty] \).

**Application:** Let \( \mu \) be a prob. measure in a vector space \( V \) and \( \mu^n \) be the image of \( \mu \otimes \cdots \otimes \mu \) under \( (v_1, \ldots, v_n) \mapsto \pm \sum v_i \). Then given a convex set.
\[ A \subset V \] we will show
\[ \lim \frac{1}{n} \log \mu_n(A) \text{ exists in } (-\infty, 0] \]

We have
\[
\mu_n(A) = \int \mu(dx_1) \cdots \mu(dx_n) = \mu^{\otimes n}(S_n)
\]
\[ \frac{1}{n} \sum_{i \in A} \]

where
\[ S_n = \{ (x_i) \in V^n : \frac{1}{n} \sum x_i \in A \} \]

We have
\[ S_m \times S_n \subset S_{m+n} \]

since if
\[ \frac{1}{m} \sum_{i=1}^{m} x_i \in A \]
\[ \frac{1}{n} \sum_{j=1}^{n} x_{m+j} \in A \]

then
\[ \frac{1}{m+n} \sum_{i}^{m+n} x_i = \frac{m}{m+n} \frac{1}{m} \sum_{i}^{m} x_i + \frac{n}{m+n} \frac{1}{n} \sum_{j}^{n} x_{m+j} \]
\[ \in \frac{m}{m+n} A + \frac{n}{m+n} A \subset A \]

as \( A \) is convex. Thus
\[ \mu_{m+n}(A) = \mu^{\otimes (m+n)}(S_{m+n}) \geq \mu^{\otimes (m+n)}(S_m \times S_n) \]
\[ = \mu^{\otimes m}(S_m) \mu^{\otimes n}(S_n) = \mu_m(A) \mu_n(A) \]

and so if
\[ f(n) = \log \mu_n(A) \]

\( f \) satisfies
\[ f(m+n) \geq f(m) + f(n) \]

and we can apply the preceding lemma. Except \( f(n) \) might be \(-\infty\).
Fatou's lemma: If \( f_n > 0 \), then
\[
\int \liminf f_n \leq \liminf \int f_n
\]

**Proof:** Recall first, that if \( f_n > 0 \), \( f_n \leq f_{n+1} \leq \ldots \), then one has
\[
\sup \int f_n = \int \sup f_n.
\]
This is because
\[
\int f_n = \mu \{ (x,y) \in X \times \mathbb{R}_{\geq 0} \mid 0 \leq y < f_n(x) \}
= \mu \text{ (subgraph } f_n) \]
and
\[
\text{subgraph } (\sup f_n) = \bigcup_n \text{ subgraph } f_n.
\]
and
\[
\mu \left( \bigcup_n A_n \right) = \sup_n \mu(A_n) \quad \text{for } A_n \subset A_{n+1}, \ldots
\]
(Notice this is not true for decreasing sequences unless \( \mu(A_n) < \infty \))

Then
\[
\int \inf_{n \geq k} f_n \leq \inf_{n \geq k} \int f_n
\]
so applying monotone convergence
\[
\int \liminf f_n = \int \sup_{n \geq k} \inf_{n \geq k} f_n = \sup_k \int \inf_{n \geq k} f_n
\leq \sup_k \inf_{n \geq k} \int f_n = \liminf \int f_n
\]
Application

\[ Z(J) = \int e^{(J,x)} \mu(dx), \quad x \in (0, \infty] \]

If \( J_n \to J \) then we have

\[ \liminf \int e^{(J_n,x)} \leq \liminf \int e^{(J,x)} \]

or

\[ Z(J) \leq \liminf_n Z(J_n) \]

Since \( \log \) is continuous and monotone it preserves \( \inf \) and \( \sup \), so we have

\[ \log Z(J) \leq \liminf_n \log Z(J_n) \]

Next recall the definition

\[ W(x) = \sup_J (J,x) - \log Z(J) \]

is a \( \sup \) of \( \log \) functions so it has a semi-continuity property. To express this we look at the supergraph \( f = \{(x,y) \mid y \geq f(x)\} \). We ask what it means for this supergraph to be closed. If it is closed, then for \( x_n \to x \) one has that \( (x, \liminf f(x_n)) \) is a limit point of the supergraph, so

\[ \liminf f(x_n) \geq f(x) \quad \text{called lower semi-continuous} \]

Conversely if this holds and \( x_n \to x, \ y_n \to y \) with \( y_n \geq f(x_n) \), then

\[ y = \lim y_n \geq \liminf f(x_n) \geq f(x) \]

so \( (x,y) \) belongs to the supergraph of \( f \).
Summary: if is lower semi-continuous means

\( \inf \{ f(x_n) \} \geq f(x) \quad \text{if} \quad x_n \to x \)

\( b) \quad \{(x,y) \mid y \geq f(x)\} \) is closed.

From Fatou's lemma + Minkowski inequality

\[ |\int fg| \leq \|f\|_p \|g\|_q \quad \frac{1}{p} + \frac{1}{q} = 1 \]

which comes from convexity of the exponential function, we know that

\[ \log Z(J) = \int e^{(x)} \mu(dx) \]

has closed convex supergraph. Hence this fn. is convex and l.s.c.

This implies that \( \{ J \mid Z(J) < \infty \} \) is convex.

In the following we assume it is non-empty (i.e. usually that \( Z(0) = 1 \)). Then when we take the Fenchel transform

\[ W(x) = \sup_J (J,x) - \log Z(J) \]

we have \( W(x) \geq (J_0,x) - \log Z(J_0) \). Otherwise \( W(x) = -\infty \) for all \( x \).

I want now to show that I can recover \( \log Z(J) \) from \( W(x) \):

\[ \log Z(J) = \sup_x \{ (J,x) - W(x) \} \]

Here's the motivation from yesterday.
Let's start with a probability measure $\mu$ on $\mathbb{R}^n$, say, with compact support. We assume it is not supported in a hyperplane. Then we know that the map

$$ J \mapsto x_J = \nabla \log Z(J), \quad V' \mapsto V $$

is a local diffeomorphism. (We know that the Jacobian matrix at $J \in V'$, which is a linear map $V' \to V$, or equivalently a quadratic form on $V'$, is positive definite. It gives the variance of $e^{T_J} \mu / Z(J)$.)

I want to prove that this map is a diffeomorphism of $V'$ with the interior of the convex hull of the support of $\mu$. This result is like the Ahlfors–Millman–Steinberg convexity theorem.

I can define the Fenchel transform $W(x)$ as above. The first thing is to check that it coincides with the Legendre transform at a point of the form $x_J$. To let $x = x_J$ be the gradient of $\log Z$ at $J_0$. The function

$$ J \mapsto (J, x) - \log Z(J) $$

has negative definite second derivative everywhere and zero first derivative at $J = J_0$. So it follows from calculus that this function has a maximum at $J = J_0$. Thus we see that

$$ W(x) = (J_0, x) - \log Z(J_0) $$

if $x = x_{J_0}$. 
But now recall that where the Legendre transform is defined one has
\[ D W(x) = J \quad \text{when } x = x_f. \]
So this shows that \( J \rightarrow x_f \) is 1-1, and so this map is a diffeomorphism with its image which is an open subset of \( V \). I would like to identify this open subset with the interior of the set where \( W \) is finite.

Actually the fact that \( J \rightarrow x_f \) is an embedding is much simpler to see. We know \( \log Z(J) \) is strictly log-convex so given two points \( J_1, J_2 \) restrict this function to the line containing \( J_1, J_2 \). Then the gradient of \( \log Z(J) \) has to move between \( J_1, J_2 \).

\[ \frac{J_1 - J_2}{\log Z(J_1) - \log Z(J_2)} \]
otherwise over this line \( \log Z(J) \) would be linear.

Fenchel transform stuff. We are interested in subsets of \( V \times \mathbb{R} \) of the form
\[ \hat{Y} = \{ (x,a) | a \geq J \cdot x - c, \forall (J,c) \in Y < V \times \mathbb{R} \} \]
Let's put this into a Galois correspondence setup.

We consider the relation \( Z \subseteq (V \times R) \times (V \times R) \)

\[
Z = \{ (x, a, J, e) \mid a + c \geq J \cdot x \}
\]

Given \( Y \subseteq V \times R \), then \( \hat{Y} \subseteq V \times R \) is the subset of all \( (x, a) \) related by \( Z \) to every elt of \( Y \). We have \( Y \subseteq Y_1 \iff \hat{Y} \supseteq \hat{Y}_1 \).

Also \( \hat{X} \supseteq \hat{Y} \) (means any \( y \in Y \) is \( Z \)-related to any \( x \in X \))

\( \iff \hat{X} \subseteq \hat{Y} \) (means any \( x \in X \) is related to any \( y \in Y \)).

Thus \( \hat{X} \subseteq \hat{X} \implies \hat{X} = \hat{X}_j \) and \( \hat{X} \subseteq \hat{X} \). Thus \( \hat{X} = \hat{X} \) and we have a 1-1 order reversing correspondence between closed subsets of \( V \times R \) and \( V \times R \).

Next we have to identify the closed subsets of \( V \times R \) and \( V \times R \) for this relation.

\[
\hat{Y} = \{ (x, a) \mid for \ all \ (x, c) \in Y \ we \ have \ a + c \geq J \cdot x \\
\implies J \cdot x - a \leq c \}
\]

is a closed convex subset of \( V \times R \) such that \( (x, a) \in \hat{Y} \implies (x, a') \in \hat{Y} \) for \( a' \geq a \).

At this point I need to know that a convex closed set is the intersection of half-spaces. This is proved by some Hahn-Banach type arguments, i.e. by induction on the dimension in the case of finite dimensions.

Let's review this.
Fenchel transform.

Let \( V \) be a real vector space of finite dimension, let \( V' \) be its dual. Let \( X = V \times \mathbb{R} \) and \( Y = V' \times \mathbb{R} \).

Let \( \Gamma = X \times Y \) be
\[
\Gamma = \{ (x, a, y, c) \mid a + c \geq \xi, \eta \}
\]

Given \( A \subseteq X \), define
\[
A' = \{ y \in Y \mid A \times \{ y \} \subset \Gamma \}
\]
which is the largest \( B \subseteq Y \) such that \( A \times B \subset \Gamma \).

Define \( B' \subseteq X \) for \( B \subseteq Y \) similarly. Then clearly
\[
A_1 \subset A_2 \iff A_1' \supset A_2' \quad \text{similarly for } B
\]
\[
A \subset B' \iff A \times B \subset \Gamma \iff B \subset A'.
\]

Hence \( A' \subset A'' \iff A \subseteq A'^{\prime \prime} \), similarly \( B \subset B'' \).

Thus \( A' \subseteq A'' \) and \( A' \supset A'' \), so \( A' = A'' \); simil. for \( B \).

Therefore we have a Galois correspondence between subsets of \( X \) of the form \( B' \) and subsets of \( Y \) of the form \( A' \). Also can say subsets closed for the closure operation \( A \mapsto A' \); simil. for \( B \).

Now our problem is to identify the "closed" sets in the case at hand. First note that if \( B = \emptyset \), then \( B' = X \), whereas if \( B \neq \emptyset \), say \( B \) contains \( (\xi, c) \), then
\[
B' \subset \{ (\xi, c) \}' = \{ (x, a) \mid a + c \geq \xi, \eta \}
\]

supergraph of the linear \( a \mapsto \xi, \eta - c \).
In fact
\[ B' = \{ (v,a) \mid \forall (\xi,c) \in B, a > \xi \cdot v - c \} \]
\[ = \{ (v,a) \mid a > \sup_{(\xi,c) \in B} \xi \cdot v - c \} \]

is the supergraph of
\[ f(\xi) = \sup_{(\xi,c) \in B} \xi \cdot v - c \]

Note that for \( B \neq \emptyset \) this function has values in \( \mathbb{R} \cup \{-\infty\} \) (and for \( B = \emptyset \) it has the constant value \(-\infty\)). Because the supergraph is convex and closed, it follows that the function \( f \) is convex and lower semi-
continuously.

Conversely, given \( f: V \to \mathbb{R} \cup \{-\infty\} \) convex and l.s.c., its supergraph \( S \) is closed and convex. If \( (v_0, a_0) \notin S \), we can separate \( S \) from this point by a hyperplane

\[ S \]

\[ (v_0, a_0) \]

\[ a \]

Case 1. The hyperplane is of the form \( a = \xi \cdot v - c \).

Then \( S \subset \{ (v,a) \mid a > \xi \cdot v - c \} \), \( a_0 > \xi \cdot v_0 - c \)

Case 2. The hyperplane is of the form \( \xi \cdot v = 0 \) with \( \xi \neq 0 \). In this case we have the picture
s somehow we are going to have to juggle this hyperplane.

It seems necessary to understand the separation theorem.
In finite dimension we can use an inner product. Given a closed convex set K, a basic result says K has a unique point k closest to 0.

It follows that K lies outside the tangent hyperplane to the sphere around 0 passing through k. Otherwise you could join a point inside the hyperplane to k and get a contradiction.

Now let's go back to S and \((0,0,0)\) as above.
We draw in the largest open ball around 0 not meeting S and look at the tangent plane at the intersection point of the boundary with S. If this tangent plane is vertical
then we do the same for \( a_0 - t \). As \( t \) increases we watch the radius, i.e. the distance of \((v_0, r_0 - t)\) from \( S \). The distance has to begin to increase at some point; otherwise \( S \) would contain a vertical line. At that point the tangent hyperplane stops being vertical:

\[
\begin{align*}
& (v_0, r_0) \\
& (v_0, r_0 - t)
\end{align*}
\]

and by continuity we will get a hyperplane separating \((v_0, r_0)\) from \( S \) which isn't vertical.

So therefore we see that the supergraph of an \( f: V \to \mathbb{R} \cup \{\infty\} \) which is convex and l.s.c. upward pointing is the intersection of the half spaces containing it. This implies it is closed for the Galois correspondence.

Now we are ready to set up the Feichtel transform. We consider the \( A \subseteq X \) such that \( A = A^\perp \) and such that \( A \neq \emptyset \). Then \( A \) is the supergraph of a unique convex l.s.c. fun. \( f: V \to \mathbb{R} \cup \{\infty\} \) such that \( f(x) < \infty \) for some \( x \). Then

\[
\begin{align*}
A &= \{ (v, a) \mid a \geq f(v) \} \\
A' &= \{ (v, a) \mid \forall (v, a), a \geq f(v) \Rightarrow a \geq f(v) - c \} \\
&= \{ (v, a) \mid \forall (v, a), f(v) \geq f(v) - c \}
\end{align*}
\]
\[ \{ (i,c) \mid V^c, c > \inf_{i \in V} f(v) \} \]
\[ = \{ (i,c) \mid c > \sup_{v \in V} f(v) \} \]

Thus if \( A = \text{supergraph}(f) \), \( A' \) is the supergraph of \( f(i) = \sup_{v \in V} (i,v - f(v)) \).

and we obtain

**Thm:** There is a 1-1 correspondence between \( \text{supersets } f: V \to \mathbb{R}^{\infty} \) such that \( f \neq +\infty \), and similar functions \( g: V \to \mathbb{R}^{\infty} \) given by

\[ g(i) = \sup_{v \in V} (i,v - f(v)) \]
\[ f(v) = \sup_{v \in V} (i,v - g(i)) \]

Now let's return to the original project of finding a measure \( \mu \) on \( \mathbb{R}^n \) and proving that the image of \( \mathbf{f} : \mathbb{R}^n \to \mathbb{R}_+ \) is the interior of the convex hull of \( \text{supp } \mu \). Suppose to simplify that \( \text{supp } \mu \) is compact.

The idea will be to form

\[ W(x) = \sup_{i} i \cdot x - \log Z(i) \]

and to take an interior point of \( \{ x \mid W(x) < \infty \} \).

Call this point \( x_0 \). Let's look at the supergraph of \( W \).

What I really want I think is a point \((x_0, c_0)\) contained in the interior of the supergraph.

In finite dimensions, we can get such a point over \( \mathbb{R}^n \) which is in the interior of a \( \infty \) simplex in \( V \) of the same dimension whose vertices have finite \( W \). Thus
if \( \mathbf{v}_0, \mathbf{v}_1, \ldots, \mathbf{v}_q \in \text{dim } \mathcal{V} \) are affinely independent with \( W(\mathbf{v}_j) < \infty \), then by convexity

\[
W(\mathbf{v}_j) \leq \sum t_j W(\mathbf{v}_j)
\]

and so we get an upper bound on \( W \) near \( \mathbf{x}_0 \).

Next suppose \( W(\mathbf{x}) \leq C \) for \( |\mathbf{x} - \mathbf{x}_0| \leq \varepsilon \). Then

\[
J(\mathbf{x} - \mathbf{x}_0) + J\mathbf{x}_0 - \log Z(J) \leq C
\]

for all \( J \), all \( \mathbf{x} \) with \( |\mathbf{x} - \mathbf{x}_0| \leq \varepsilon \).

Maybe as well suppose \( \mathbf{x}_0 = 0 \), \( \varepsilon = 1 \).

\[
\forall J, |\mathbf{x}| \leq 1 \implies J\mathbf{x} - \log Z(J) \leq C
\]

\[
\implies \forall J, |J| - \log Z(J) \leq C
\]

\[
\implies -\log Z(J) \leq C - |J|
\]

This implies that \( -\log Z(J) \) has a maximum. If \( J_0 \) is the maximum value, then \( \mathbf{x}_{J_0} = 0 \).

In general, given \( \mathbf{x} \) for all \( J \), all \( \mathbf{x} \) with \( |\mathbf{x} - \mathbf{x}_0| \leq \varepsilon \) one gets

\[
|J| + J\mathbf{x}_0 - \log Z(J) \leq C
\]

so \( J\mathbf{x}_0 - \log Z(J) \) has a maximum; if it occurs at \( J_0 \) then \( \mathbf{x}_0 = \mathbf{x}_{J_0} \).
March 26, 1985

Review convex analysis.

$V$ locally convex $\implies$ Hausdorff top. v.s.

$V^*$ its dual.

$f : V \rightarrow \mathbb{R} \cup \{+\infty\}$ convex and l.s.c.

Define $f^*$ on $V^*$ by

$$f^*(\xi) = \sup_{v \in V} (\langle \xi, v \rangle - f(v))$$

Assertion: Assume $f \neq +\infty$. Then $f^*$ maps

$V^*$ to $\mathbb{R} \cup \{+\infty\}$ and it is convex and l.s.c. for the weak * topology on $V^*$. Moreover

$$f(v) = \sup_{\xi} (\langle \xi, v \rangle - f^*(\xi))$$

Proof. Put $\text{epi}(f) = \{ (\xi, a) \in V \times \mathbb{R} \mid a \geq f(\xi) \}$ $\Rightarrow$ graph

$f$ l.s.c. $\Rightarrow$ $\text{epi}(f)$ closed

$f$ convex $\Rightarrow$ $\text{epi}(f)$ convex

Lemma: If $(\xi, a) \notin \text{epi}(f)$, there exists $(\hat{\xi}, c) \in V \times \mathbb{R}$

such that $f(\hat{\xi}) > \langle \hat{\xi}, \xi - c \rangle + a$.

$\alpha < \langle \xi, v \rangle - c$

(This is the non-trivial point: $f^*$ is the intersection of all $\text{epi}(f^* - c)$ containing it.)
Do rest of the proof. As \( \exists v_0 \) with \( f(v_0) < \infty \), clearly \( f^*(\frac{v}{v_0}) > -\infty \). \( f^* \) is the map of linear weak \* cont. functions, it is convex and weak \* l.s.c.

(Criterion: \( f: V \to \mathbb{R} \geq 0 \) convex \( \iff \text{epi}(f) \) convex
\( \iff \text{closed} \))

\( \forall v, f^*(\frac{v}{v_0}) \geq \frac{v}{v_0} - f(v_0) \Rightarrow f(v) \geq \sup_{\frac{v}{v_0}} \left( \frac{v}{v_0} - f^*(\frac{v}{v_0}) \right) \).

To prove equality \( f(x) = f(v_0) \) and let \( a_0 < f(v_0) \). By lemma \( \exists (\frac{v}{v_0}) \Rightarrow \)

\( \forall \frac{v}{v_0}, f(v) \geq \frac{v}{v_0} - c_0 \iff c_0 \geq f^*(\frac{v}{v_0}) \)

and \( a_0 < \frac{v}{v_0} - c_0 \), so

\( a_0 < \frac{v}{v_0} - f^*(\frac{v}{v_0}) \leq \sup_{\frac{v}{v_0}} \left( \frac{v}{v_0} - f^*(\frac{v}{v_0}) \right) \)

for arbitrary \( a_0 < f(v_0) \), we get

\( f(v_0) \leq \sup_{\frac{v}{v_0}} \left( \frac{v}{v_0} - f^*(\frac{v}{v_0}) \right) \)

Proof of Lemma: As \( \text{epi}(f) \subset V \times R \) is closed convex
the Hahn–Banach thm. implies \( \text{epi}(f) \) is the intersection
of the closed affine half-spaces containing it. A
closed affine half-space is described of the form

\[ H = \{(v, a) \mid \frac{v}{v_0} - ba \leq c \} \]

with \( (\frac{v}{v_0}, b, c) \in V \times R \times R \). One can suppose \( b = -1, 0, 1 \).

Using the assumption \( \exists v_0 \) with \( f(v_0) < +\infty \) we
see that \( H \cap \text{epi}(f) \) forces \( b = 0 \) or \( 1 \). If \( b = -1 \)
then \( H \) contains \( (v_0, a) \) for \( a \geq f(v_0) \), so

\[ \frac{v_0}{v_0} + ba \leq c \]
for arbitrarily large $a$ which is impossible.

Divide the $H$ containing $\text{epi}(f)$ into $H^+, H^-$

according as $b = 1$ or $b = 0$. The lemma says

\[ \text{epi}(f) = \bigcap_{H \in H^+} H \quad \text{and we know} \quad \text{epi}(f) = \bigcap_{H \in H^-} H. \]

So it is enough to take $(v_0, a_0) \notin H_0$ with $H_0 \in H^-$

and show $\exists H \in H^+$ with $(v_0, a_0) \notin H$. Picture:

\[ H_0 \]
\[ \text{epi}(f) \]
\[ (v_0, a_0) \]

Claim $H^+ \neq \emptyset$. If with

\[ f(v) < +\infty, \quad \exists \quad a \in (v_0)^{*} \quad \text{epi}(f), \quad \text{so there is an} \quad H_1 \quad \text{containing} \quad \text{epi}(f) \quad \text{but not} \quad (v_0, a), \quad \text{and clearly} \]

\[ H_1 \notin H_0, \quad \text{so} \quad H_1 \in H^+. \]

Let $H_1 = \text{epi} \left( \frac{\lambda}{\beta} \cdot \nu - c \right)$. Now consider

\[ H_2 = \text{epi} \left( (\lambda v_0 + \frac{\lambda}{\beta}) \cdot \nu - c_1 \right) \]

where

\[ H_0 = \{ (v, a) \mid \beta \cdot v \leq 0 \} \quad \text{and} \quad \lambda > 0. \]

Then $H_2 \supset \text{epi}(f)$.

\[ (v, a) \in \text{epi}(f) \Rightarrow \beta \cdot v \leq 0 \]
\[ \Rightarrow \beta \cdot v - c \leq 0 \]
\[ \Rightarrow (v_0, a) \in H_1 \]

On the other hand as $\lambda \to +\infty$, $H_2$ approaches $H_0$ which doesn't contain $(v_0, a_0)$. So for large $\lambda$, $H_2$ doesn't contain $(v_0, a_0)$. QED.
Remarks: 1) This Hahn-Banach theorem for closed convex sets implies that a closed convex set is weakly closed. So a l.s.c. convex $f$ is l.s.c. for the weak topology.

2) To finish the above, one would like to start with a $g: V^* \to \mathbb{R} \cup \{+\infty\}$ and define

$$g^*(v) = \sup_{v^* \in V^*} \langle v^*, v \rangle - g(v^*)$$

and then recover $g$ as the conjugate of $g^*$. Thus we need to identify $V$ with the dual of $V^*$.

Question: Is it true that $V$ is the dual of $V^*$ for the $\sigma(V, V^*)$ (or weak star) topology?

Yes. Suppose $\lambda: V^* \to \mathbb{R}$ is continuous w.r.t. the $\sigma(V^*, V)$-topology. Then $\exists u_1, \ldots, u_k \in V$ such that

$$|v_i \cdot \frac{\lambda}{\lambda(v)}| < 1, \quad i = 1, \ldots, k \Rightarrow |\lambda(v)| < 1.$$ 

It follows that $\lambda$ kills $\text{Ker} \{V^*, u_i \to \mathbb{R}^k\}$, so $\lambda$ is a linear combination of the $u_i$.

So we can identify the transforms $f^*$ on $V^*$ for $f: V \to \mathbb{R} \cup \{+\infty\}$ l.s.c. convex, $f \neq +\infty$ with the set of similar $g: V^* \to \mathbb{R} \cup \{+\infty\}$, but where l.s.c. refers to the weak * topology on $V^*$.
The problem is the weak law of large numbers in infinite dimensions, in particular, to show that if \( G \) is an open set containing \( \bar{x} \), then \( \mu_n(G) \to 1 \). Already this is interesting in the Gaussian case in which \( \mu_n \) is simply a rescaling of \( \mu \); one is interested in showing \( \mu(V_n G) \to 1 \), and more specifically the rate of convergence.

Let's first check the scaling. Suppose we have a Gaussian measure on \( V \); say we write formally

\[
\mu(dx) = e^{-\frac{1}{2} x \cdot A x} \, dx / \eta
\]

and then rigorously we have

\[
Z(T) = \int e^{T x} \mu(dx) = e^{-\frac{1}{2} T \cdot A^{-1} T}
\]

with \( A \) positive definite (formally \( Q(T) = T \cdot A^{-1} \cdot T \)). Then \( \mu_n \) has the partition function

\[
Z(T_n) = e^{-\frac{1}{2} n Q(T/n)} = e^{-\frac{1}{2} n Q(T)}
\]

which formally means

\[
\mu_n(dx) = e^{-\frac{1}{2} n x \cdot A x} \, dx / \eta
\]

and so \( \mu_n \) is \( \mu \) rescaled under \( x \to \sqrt{n} x \).

Thus

\[
\mu_n(G) = \mu(V_n G)
\]
Next I want to take $G$ to be a ball $\{ \|x\| \leq 1 \}$ where $\|x\|$ is a norm on $V$. Actually $\|\|$ can be just densely defined; if one shows

$\mu(\cap_n G) \to 1$ as $n$ goes to infinity, then the measure is supported on the Banach space $U \cap G$ defined by this norm.

Let's get to the essential example where everything is Gaussian. Suppose $V = \mathbb{R}^{n \geq 0}$ with

$$\mu(dx) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2} a_n x_n^2} \frac{dx_n}{\sqrt{2\pi}}$$

and

$$\|x\|^2 = \frac{1}{2} \sum_n x_n^2.$$ Thus I want to compute

$$\nu(r) = \int_{\frac{1}{2} \|x\|^2 \leq r} \mu(dx)$$

It's simpler to compute the Laplace transform

$$\int e^{-sr} \nu(r) = \int e^{-s\frac{1}{2} x_B} \mu(dx)$$

$$= \det \left( \frac{A + sB}{A} \right)^{-1/2} = \frac{1}{\sqrt{\det (A + s\frac{b_n}{a_n})}}$$

(Incidently this same calculation says something when $B$ isn't positive definite, i.e. we are computing the measure of a hyperbolic ball.)
Now let's use the estimates involved in the 1-dimensional Cramer form.

\[
\int_0^\infty e^{sr} \, d\nu(r) = \sqrt{\pi} \left( 1 - s \frac{b_n}{a_n} \right)^{-1/2}
\]

\[
\Rightarrow \quad \int_0^{s \geq 0} e^{sr} \, d\nu(r) \geq \int_{\mathbb{R}, \mathbb{R}} e^{sr} \, d\nu(r) = e^{sR} (1 - \nu(R))
\]

Thus

\[
1 - \nu(R) \leq e^{-sR} \left[ \sqrt{\pi} \left( 1 - s \frac{b_n}{a_n} \right)^{-1/2} \right]
\]

assuming \( s \geq 0 \).

The first condition is

1) \[ \sum \frac{b_n}{a_n} < \infty \]

This implies that the infinite product converges for small \( s \). It is necessary for the ball \( \frac{1}{2} \times \mathbb{B} x \leq r \) to have positive measure. Assume 1) holds.

Then the infinite product in (x) converges for all \( s \), so we can use (x) for all \( s \geq 0 \) such that

\[ 1 - s \frac{b_n}{a_n} > 0 \quad \text{or} \quad s < \inf \left( \frac{a_n}{b_n} \right) \]

Thus we have exponential convergence

\[ 1 - \nu(R) \leq e^{-sR} \text{ const} \]

for any \( 0 \leq s < \inf \left( \frac{a_n}{b_n} \right) \).
\[ \lim_{R \to \infty} \frac{1}{R} \log (1 - v(R)) \leq -s \quad \forall s > 0 \iff s < \inf \left( \frac{a_k}{b_k} \right) \]

or

\[ \lim_{R \to \infty} \frac{1}{R} \log (1 - v(R)) \leq -\inf \left( \frac{a_k}{b_k} \right) \]

Note this is exactly the Cramér estimate. If \( R = n \)

\[ v(n) = \mu( \{ x | \frac{1}{2} \times B x \leq n \}) = \mu( \sqrt{n} \cdot \{ x | \frac{1}{2} \times B x \leq 1 \} ) \]

\[ = \mu_n( \{ x | \frac{1}{2} \times B x \leq 1 \} ) \]

\[ 1 - v(n) = \mu_n( \{ x | \frac{1}{2} \times B x > 1 \} ) \]

\[ W(\gamma) = \frac{1}{2 \pi b_k^2} \sum_{k} a_k x_k^2 \; ; \; \text{we want the inf of this over} \; \{ x | \sum_{k} a_k x_k^2 > 1 \} \]

Rescale \( x \rightarrow \sqrt{2 / b_k} \; ; \; \text{we want} \)

\[ \inf \left\{ \sum_{k} \frac{a_k x_k^2}{b_k} \left| \sum_{k} x_k^2 \geq 1 \right\} \right. \]

and this is clearly \( \inf \left( \frac{a_k}{b_k} \right) \).

---

Let's consider a Gaussian density \( \mu \) on a top vector space \( V \) (subject to some later hypothesis). Gaussian means the image under any linear map \( V \to R \) is Gaussian and this should be the same as

\[ Z(J) = \int e^{J \cdot x} \mu(dx) = e^{Q(J)} \]

where \( Q \) is a non-negative quadratic form on \( V^* \)

Let's consider the standard form for such a measure. The form \( Q \) is equivalent to a map

\[ G: (V) \to (V)^* \]

which is symmetric: \( G^T = G \). We suppose \( G \)
Continuous and $V$ reflexive.

Better, we suppose $Q$ is given by a continuous linear map $G: V^* \to V$ which is symmetric $G^t = G$ and injective with dense image. (One knows $Q(J)$ is convex, quadratic, and l.s.c. Perhaps one could show $Q$ continuous.) Formula

$$Q(J) = \frac{1}{2} J \cdot G J.$$  

Continue with the standard picture: We use $Q$ to define an inner product on $V^*$ and form the completion $H$ of $V^*$ with respect to this inner product. We let $W \subset V$ be the subspace of $x \in V$ such that the associated linear functional $V^*$ is continuous for the inner product, $W$ can be identified with $H$.

The inner product is

$$\langle J, J' \rangle = J \cdot G J'.$$

So let us try to describe the picture. One has maps $V^* \hookrightarrow H \hookrightarrow V$ which are embeddings with dense range such that the first and second are transposes of each other (assuming $V$ reflexive). The composition is the map $G$.

Next try to compute the transform of $Q(J) = \frac{1}{2} J G J$.

$$W(x) = \sup_{J} \left( J \cdot x - \frac{1}{2} J \cdot G J \right)$$

If $x = G J_0$, then $J = J_0$ should be the critical point and we would get

$$W(x) = \frac{1}{2} J_0 \cdot G J_0 = \frac{1}{2} x \cdot G^{-1} x$$
But actually $w(y)$ should be finite for $x \in H$. and then $w(x)$ should be a natural extension.
Let consider a Gaussian discrete random process and analyze what it means for it to be a martingale.

Let $(\Omega, \mu)$ be a probability space, let $f_1, f_2, \ldots$ be a sequence of random variables in $\Omega$ (real valued). One says $\{f_n\}$ is a Gaussian process if the map

$$(f_1, \ldots, f_n) : \Omega \rightarrow \mathbb{R}^n,$$

pushes $\mu$ forward to a Gaussian measure $\mu_n$ in $\mathbb{R}^n$, for each $n$. (Assume centered at 0)

One might as well suppose $\Omega = \mathbb{R}^\infty$ and the $f_n$ are the coordinate functions $x_n$.

Consider $H = L^2(\Omega, \mu)$, and let $V \subset H$ be the closed subspace spanned by the $\{f_n\}$. I think we know that if $\mu_n = (f_1, \ldots, f_n)_* \mu$, then

$$L^2(\Omega, \mu) = \lim_{n \to \infty} L^2(\mathbb{R}^n, \mu_n)$$

and the polynomial functions $\text{S}(V)$ is dense.

Assertion: A discrete Gaussian random process $\{f_n\}$ is equivalent to a Hilbert space $V$ together with a sequence $f_1, \ldots, f_n, \ldots$ in $V$ spanning $V$. Equivalently, a non-negative symmetric matrix $\langle f_i, f_j \rangle$.

Next recall that for Gaussian r.v.'s independence is equivalent to orthogonality. $\{f_n\}$ will be an $L^2$ martingale when $f_n - f_{n-1}$ is orthogonal to $L^2(\mathbb{R}^n, \mu_n)$. 

April 5, 1985
It should be enough that \( f_n - f_{n-1} + V_{n-1} = \langle f_n; f_{n-1} \rangle \) (think of the Wick process)

Concluded: A Gaussian martingale is of the form

\[
f_n = g_1 + g_2 + \cdots + g_n
\]

where the \( g_n \) are a sequence of independent Gaussian r.v.'s.

The martingale convergence theorem says that

such a sequence \( f_n = g_1 + \cdots + g_n \) converges almost everywhere.

Let's now consider the Hilbert space \( V = l^2 \)

which we identify with a space of functions on \( \mathbb{R}^\infty \).

Better: I want to take the standard Gaussian measure \( e^{-x^2/2} \, dx / \sqrt{2\pi} \) on \( \mathbb{R} \) which describes a Gaussian r.v. with standard deviation 1.

Then take the product measure

\[
\mu = \prod_{n=1}^{\infty} e^{-\frac{1}{2}x_n^2} \frac{dx_n}{\sqrt{2\pi}} \quad \text{on} \quad \mathbb{R}^\infty.
\]

We want \( \Omega \subset \mathbb{R}^\infty \) to be subspace carrying \( \mu \).

The idea is that \( \Omega \) should be a Banach space on top r.v. such that

there is an embedding \( l^2 \subset \Omega \) carrying Gaussian cylinder measure on \( l^2 \) to a measure on \( \Omega \).

Now take \( g_n = b_n x_n \) where \( x_n \) is the \( n \)th coordinate function and \( f_n = \sum_{k \leq n} b_k x_k \). This is a Gaussian martingale, so we see that provided
the series $\sum b_k x_k$ converges almost everywhere on $\Omega$.

April 6, 1985

I want to now try to prove that if I have a sequence of random variables $g_1, g_2, \ldots$ which are independent and of mean zero and such that $\Sigma \langle g_i^2 \rangle < \infty$, then the series $\Sigma g_k$ converges almost everywhere.

Put $f_n = g_1 + \ldots + g_n$. We want to prove this sequence converges a.e. so we look at those $\omega$ such that $f_n(\omega)$ fails to converge. If the sequence $f_n(\omega)$ fails to converge, we know it is unbounded or has at least 2 limit points. If it has at least two limit points, then there is an interval $[a, b)$ with rational endpoints such that $f_n(\omega)$ is infinitely often below $a$ and above $b$. Let

$$S_{(a,b)} = \{ \omega \mid f_n(\omega) > b \text{ i.o.} \}$$

Then the subset of $\Omega$ for which $f_n(\omega)$ doesn't converge is

$$\bigcup_{a < b} S_{(a,b)} \cup \{ \omega \mid f_n(\omega) \to +\infty \} \cup \{ \omega \mid f_n(\omega) \to -\infty \} = S_{+\infty} \cup S_{-\infty}$$

It suffices to show $S_{(a,b)}$ and $S_{+\infty}, S_{-\infty}$ are null sets.

The key point in the proof will be to estimate the probability that with $k$ fixed one has

$$f_n - f_k = g_{k+1} + \ldots + g_n \geq \varepsilon$$

for some $n$. Better still...
Better: Given \( k \), find

\[
P \{ x \mid f_n \text{ with } (g_{k+1} + \ldots + g_n)(x) \geq \varepsilon \}
\]

Let's analyze this for \( k = 0 \). We are interested in the set where \( f_n \geq \varepsilon \) for some \( n \). We decompose this into

\[
B_k : f_1, \ldots, f_{k-1} < \varepsilon, f_k \geq \varepsilon
\]

Then for \( n \geq k \)

\[
\int_{B_k} f_n = \int_{B_k} f_k \geq \varepsilon P(B_k)
\]

Cantelli-Schroder + \( \int_{B_k} f = \int_{B_k} f \).

So

\[
\varepsilon P(U_{B_k}) \leq \int_{U_{B_k}} f_n \leq P(U_{B_k})^{1/2} \cdot \| f_n \|
\]

\[
\varepsilon P(U_{B_k})^{1/2} \leq \| f_n \|
\]

\[
P(U_{B_k}) \leq \frac{1}{\varepsilon^2} \| f_n \|^2
\]

\[
P(U_{B_k}) \leq \frac{1}{\varepsilon^2} \| f \|^2
\]

(\( \varepsilon > 0 \) needed)

\[
\| f \|^2 \geq \int f^2 \geq \varepsilon^2 P \{ |f| \geq \varepsilon \} \quad \Rightarrow \quad P \{ |f| \geq \varepsilon \} \leq \frac{1}{\varepsilon^2} \| f \|^2
\]

so this is the same kind of inequality.)

Thus we have the estimate
\[ \forall \varepsilon > 0 \quad P \left\{ \exists n \text{ with } (g_{k+1} + \ldots + g_n) \geq \varepsilon \right\} \leq \frac{1}{\varepsilon^2} \sum_{j=k+1}^{\infty} \| g_j \|^2 \]

Now we should be able to show that the set where \( f_n \to +\infty \) is a null set. But if \( f_n(x) \to +\infty \), then for any \( \varepsilon \) one has \( f_n(x) \geq \varepsilon \) for \( n \) large enough. Thus \( \forall \varepsilon \)

\[ S_{\infty} = \{ x \mid f_n(x) \to +\infty \} \subset \{ x \mid \exists n \quad f_n(x) \geq \varepsilon \} \]

\[ P(S_{\infty}) \leq \frac{1}{\varepsilon^2} \sum_{j=1}^{\infty} \| g_j \|^2 \quad \text{for all } \varepsilon, \text{ etc.} \]

Next let us take a further bounding \( P(S_{a,b}) \). Here the idea is to note

\[ S_{a,b} = \bigcap_{n} S_{a,b}^n \]

where \( S_{a,b}^n \) is the set of \( x \) such that we have \( \geq n \) crossings. Thus \( S_1 \) means \( f_k(x) \leq a \) for some \( k \), and \( S^2 \) means \( f_k(x) \leq a \) and \( f_k(x) \geq b \), etc.

Introduce stopping times

\[ N_1(x) = \inf \left\{ k \mid f_k(x) \leq a \right\} \]

\[ N_2(x) = \inf \left\{ l \mid l > N_1(x) \text{ and } f_l(x) \geq b \right\} \]

Then \( S_{a,b}^n = \{ x \mid N_n(x) < \infty \} \). Now what we have to do is estimate \( P(S_{a,b}^{n+1}) \) in terms of \( P(S_{a,b}^n) \).

First observe that \( S_{a,b}^{n+1} \) decomposes according to the values of \( N_1, \ldots, N_{n+1} \). Similarly for \( S^2 \). Let's fix
the values of \( N_1, \ldots, N_k \), i.e., we take one of the components of \( S^n \), denote it \( S^n(k_1, \ldots, k_r) \).

\[
S^n(k_1, \ldots, k_r) : \quad f_1 \cdots f_{k-1} > a, \quad f_k < a
\]

\[
f_{k+1} \cdots f_{k+1} < b, \quad f_{k+2} > b
\]

Next we want the part of \( S^{n+1} \) which is contained in \( S^n(k_1, \ldots, k_r) \). This is

\[
S^{n+1} \cap S^n(k_1, \ldots, k_r) = \bigcup_{l > k_r} S^{n+1}(k_1, \ldots, k_r, l).
\]

We look at a point of this intersection, and note that the sequence

\[
f_{k+1} < f_{k+2}, \quad f_{k+2} > f_{k+3}, \quad \ldots
\]

has a jump of at least \( b - a \).

Now I want to use the independence. \( S^n(k_1, \ldots, k_r) \) is a subset coming from \( \Omega_k \), which means that it is equivalent to the product of a subset of \( \Omega_k \) and the space of sequences \( (g_j) \), \( j > k_r \). It seems that \( S^{n+1} \cap S^n(k_1, \ldots, k_r) \) is the product of this subset of \( \Omega_k \) and the space of sequences such that \( g_{k+1} + g_{k+2} \) has a jump of \( \varepsilon = 16 - 2/\varepsilon \) of the right sign. Thus we should get the estimate

\[
P(S^{n+1} \cap S^n(k_1, \ldots, k_r)) \leq P(S^n(k_1, \ldots, k_r)) P(\varepsilon \leq \sum_{j > k_r}^2 g_j l_j)
\]
and finally because $k_2 > 2$, we get
\[ P(S^m \cap S^m(k_1, k_2)) \leq P(S^m(k_1, k_2)) \cdot \frac{1}{\varepsilon_2} \sum_{j \neq 2} ||g_j||_2^2 \]
which we can add up to get
\[ P(S^{m+1}) \leq P(S^m) \cdot \frac{1}{\varepsilon_2} \sum_{j \neq 2} ||g_j||_2^2 \]
Then for large $n$, the factor $\frac{1}{\varepsilon_2} \sum_{j \neq 2} ||g_j||_2^2$ is $< 1$, so one sees that $P(S^n) \downarrow 0$. 
April 8, 1985

Goal: Proof of the martingale convergence thm.

Recall the definition: The probability space \( \Omega \) is assumed to be the limit of an inverse system

\[ \Omega_n \xrightarrow{\pi_{n-1}} \Omega_{n-1} \xrightarrow{\pi_{n-2}} \cdots \xrightarrow{\pi_1} \Omega_1 \]

and the measure \( \mu \) on \( \Omega \) induces \( \mu_n \) on \( \Omega_n \). Then \( \{ f_n \} \) is a sequence of random variables on \( \Omega \) such that \( f_n \) comes from \( f_n \) on \( \Omega_n \). This sequence is a martingale when

\[ (\pi_k)_* f_n = f_k \quad n \geq k \]

Here \( (\pi_\ast)_* \) denotes conditional expectation:

\[ \int_A \pi_\ast (f) \, d\pi = \int_A f \, d\pi^{-1}(A) \]

It is defined by virtue of Radon-Nikodym.

Consequence:

\[ \left| \int_A \pi_\ast (f) \, d\pi \right| \leq \int_A |f| \, d\pi \Rightarrow \left| \pi_\ast (f) \right| \leq \pi_\ast |f| \]

holds for all \( A \Rightarrow \left| \pi_\ast (f) \right| \leq \pi_\ast |f| \).

So for a martingale

\[ |f_k| = |(\pi_k)_* (f_n)| \leq |(\pi_k)_* f_n| \Rightarrow |f_{k+1}| \leq |f_n| \quad \text{for } k \leq n \]

The martingale convergence thm. says that if \( |f_n| \) is bounded, then \( f_n \) converges almost everywhere.
Recall the concept of stopping times. This is a measure \( f_n : \sigma \rightarrow \mathbb{N}_0 \) such that for each \( n \in \mathbb{N} \), the set \( \{ x \mid N(x) \leq n \} \) comes from \( \sigma_n \). What this gives us is a decomposition of \( \sigma \)

\[
\sigma = \bigcup_{n=1}^{\infty} \sigma_n \quad \text{where} \quad \sigma_n : N = n
\]

where \( B_n \) comes from \( \sigma_n \).

A stopping time \( N \) allows us to construct two martingales associated to a given martingale. The first represents running the process to the stopping time and then fixing the value at this point. This is

\[
n \mapsto f_{n \wedge N} \quad \text{i.e.} \quad f_{n \wedge N}(x) = \begin{cases} f_n(x) & n \leq N(x) \\ f_{N(x)}(x) & n > N(x) \end{cases}
\]

The other martingale is defined on \( \{ x \mid N(x) < \infty \} \) assuming this has positive probability and is

\[
n \mapsto f_{n+N} \quad \text{i.e.} \quad f_{n+N}(x) = f_{n+N\omega}(x).
\]

We can describe this process as being \( f_{n\wedge N} \) over \( B_n \).

(One probably has to write out some details about how the \( \sigma_n \)'s are defined for this process. However it is clear that the basic averaging property holds.)

\[\text{Remarks: From the Hilbert space viewpoint one has an increasing sequence of von Neumann algebras} \]

\[L^\infty(\mathcal{H}_1) \subset L^\infty(\mathcal{H}_2) \subset \cdots \]

and the measure gives compatible traces. Continued
Let go back to the martingale \( \{ s_n \} \) and the stopping time \( N \). I assume we have an \( L^2 \)-martingale i.e. each \( s_n \) is in \( L^2 \) and \( \| s_n \|_2 \) is bounded. In this case we know \( f_n \) converges to \( f \) in \( L^2(\Omega) \), because the differences \( f_n - f_{n-1} \) are mutually orthogonal.

From now on I want to think of this setup as being determined by the limit function \( f \).

Now consider the stopping time \( N \) and the family \( \{ f_{n\wedge N} \} \). I want to see clearly that this is a martingale. To do this we decompose everything according to the values of \( N \). Look at \( B_k \); \( N = k \).

At a point \( x \) of \( B_k \) (i.e. \( N(x) = k \)) we have:

\[
    f_{n\wedge N}(x) = \begin{cases} f_n(x) & n \leq k \\ f_k(x) & n > k \end{cases}
\]

Suppose I introduce the orthogonal sequence \( g_j = f_j - f_{j-1} \). Then \( f_{n\wedge N} \) is over \( B_k \) the partial sums of the series \( \sum_{j=0}^{k} g_j, j \leq k \) and \( \sum_{j=k+1}^{\infty} g_j, j > k \)

Let's carefully discuss the decomposition of \( L^2 \) that a stopping time gives.

\( \Omega_1 \) decomposes into \( \Pi_1 B_1 \) and its complement \( \Omega_2 \) into \( \Pi_2 B_1, \Pi_2 B_2 \) and the rest.
Picture: We have this sequence of von Neumann algebras
\[ L^0(\Omega) \leq L^0(\Omega') \leq \cdots \]
with limit \( L^0(\Omega) \) and we have the measure on \( \Omega \) giving rise to traces on the \( L^0(\Omega_n) \) and the conditional expectation maps
\[ L^1(\Omega) \leftarrow L^1(\Omega') \leftarrow \cdots \]

Now any \( L^2 \)-martingale associated to the filtration \( \mathcal{F} \) is equivalent to an \( f \in L^2(\Omega) \); then \( f_n \) is the orthogonal projection of \( f \) onto \( L^2(\Omega_n) \).

The point is that any compatible family of \( f_n \in L^1(\Omega_n) \) converges almost everywhere provided \( \|f_n\|_1 \) are bounded.

I would like to write out the proof at least in the case where \( f_n \in L^2(\Omega_n) \). We fix an interval \((a, b)\) and let \( S \) be the subset of \( \Omega \) where \( f_n \leq a \) for m.f. many \( n \) and \( f_n \geq b \) for infinitely many \( n \). We must show \( S \) has measure zero.

We define stopping times
\[ N_1 = \text{first } n \text{ such that } f_n \leq a \quad \text{ or } \quad \infty \]
\[ N_2 = \text{first } n > N_1 \quad \text{ such that } f_n \geq b \quad \text{ or } \quad \infty \]

Better
\[ N_1(x) = \inf \{ n \mid f_n(x) \geq a \} \]
\[ N_2(x) = \inf \{ n \mid n > N_1(x) \text{ and } f_n(x) \leq b \} \]
\[ N_3(x) = \inf \{ n \mid n > N_2(x) \text{ and } f_n(x) \leq a \} \]
e.tc.

Then if \( S_2 = \{ x \mid N_2(x) < \infty \} \) we have \( \cap S_n = S \), so we want to show \( P(S_2) \rightarrow 0 \).
Actually I want to first estimate $P\{N_2 < \infty \}$. On $N_2 < \infty$ we have $f_{N_2} \geq b$, $f_{N_1} \leq a$ so
\[ \int f_{N_2} - f_{N_1} \geq (b-a) P\{N_2 < \infty \}. \]

On the other hand
\[ \int f_{N_2} - f_{N_1} = \int f_{N_2} - f_{N_1} \leq P\{N_2 < \infty \} \frac{1}{2} \| f - f_{N_1} \|_2^2. \]

So putting these together yields
\[ P\{N_2 < \infty \} \leq \frac{1}{(b-a)^2} \| f - f_{N_1} \|_2^2. \]

It seems the same argument shows more generally
\[ P\{N_k < \infty \} \leq \frac{1}{(b-a)^2} \| f_{N_k} - f_{N_{k-1}} \|_2^2. \]

A similar argument should yield
\[ P\{N_k < \infty \} \leq \frac{1}{(b-a)^2} \| f_{N_k} - f_{N_{k-1}} \|_2^2. \]

And on the other hand
\[ \| f_{N_k} - f_{N_{k-1}} \|_2^2 \leq \| f - f_{N_{k-1}} \|_2^2. \]

Which goes to zero as $k \to \infty$.

Now I want to check that the above is correct. I am using that for an $L^2$ martingale and stopping time $N_2 > N_1$ that
\[ f = (f - f_{N_2}) + (f_{N_2} - f_{N_1}) + f_{N_1}, \]
is an orthogonal direct sum. We want to check
We begin with the increasing sequence
\[ L^2(\Omega_1) \subset L^2(\Omega_2) \subset \cdots \]
and the orthogonal projections backward. Recall that if \( N \) is a stopping time, then \( N: \Omega \to \mathbb{N} \cup \{+\infty\} \) and \( \{ \omega \mid N(\omega) = n \} \) comes from \( \Omega_n \). We define for \( f \in L^2(\Omega) \)
\[
(f_N)(\omega) = \begin{cases} 
  f_{N(\omega)}(\omega) & \text{if } N(\omega) < \infty \\
  f(\omega) & \text{if } N(\omega) = \infty
\end{cases}
\]
Put \( B_n = \{ \omega \mid N(\omega) = n \} \). Now I want to describe \( f_N \) more precisely with the goal of showing \( f - f_N \) and \( f_N \) are orthogonal. Over \( B_\infty \) this is clear. First we should point out that we propose to compute as follows:
\[
\int (f - f_N)f_N = \sum \int (f - f_N)f_N = \sum \int (f - f_n)f_n
\]
But by assumption \( B_n \) comes from \( \Omega_n \) and
\[
\int f_n \chi_{B_n} = \int (f_n)_\Omega f_n \chi_{B_n} = \int f_n \chi_{B_n}
\]
so it's clear that \( \int (f - f_N)f_N = 0 \).

Similarly if \( N_1 \leq N_2 \) are stopping times, then \( f_{N_2} - f_{N_1} \perp \) anything occurring before \( N_1 \).

Proof. Decompose according to the values of \( N \); it's enough to assume \( N \) is constant say \( N = k \) and then we want to prove that projecting onto \( L^2(\Omega_k) \) sends \( f_k - f_k \)
into zero. But this should be the same as projecting \( f - f_k \).
which we know goes to zero.

To really be convincing I ought to assign to \( N \) a quotient \( \Omega_N \) of \( \Omega \) such that

i) \( f_N \) is the orthogonal projection of \( f \) in \( L^2(\Omega_N) \)

ii) if \( N > N' \) then \( L^2(\Omega_{N'}) \subset L^2(\Omega_N) \).

Let \( \Omega = \Omega_1 \cup \cdots \cup \Omega_\infty \) be the decomposition by the values of \( N \). We want to define a quotient tower

\[
\begin{align*}
\Omega_1 &= \Omega_1 - B_1 \\
\Omega_2 &= \Omega_2 - (B_1) \\
\Omega_3 &= \Omega_3 - (B_1, B_2) \\
& \vdots
\end{align*}
\]

Let

\[
\begin{align*}
\Omega_1^* &= B_1^* \cup \Omega_1 - B_1 \\
\Omega_2^* &= B_1^* \cup B_2^* \cup \Omega_2 - (B_1) \\
& \vdots
\end{align*}
\]

etc.

Thus

\[
\Omega_N = \lim_{N \to \infty} \Omega_N^* = B_1 \cup B_2 \cup \cdots
\]

Now corresponding to the splitting of \( \Omega \) into \( B_1 \cup \cdots \cup B_\infty \) one has a decomposition of \( f \) into \( f_1 + \cdots + f_\infty \). The projection of \( f \) onto \( L^2(\Omega_N) \) should be \( f_1 + \cdots + f_\infty \) where \( f_i \) is the projection of \( L^2(B_i) \) onto \( L^2(B_i) \). To compare this with \( f_N \). Over \( B_i \), \( f_N \) is \( f_i \) and it should be OK.

still not clear,
Let us begin with our tower
\[ \Omega_1 \subset \Omega_2 \subset \Omega_3 \subset \cdots \subset \Omega_\infty = \Omega \]
and let the σ-fields of measurable subsets be
\[ \mathcal{F}_1 \subset \mathcal{F}_2 \subset \mathcal{F}_3 \subset \cdots \subset \mathcal{F}_\infty = \mathcal{F} \]
(\(\mathcal{F}_n = \text{Boolean algebra of projectors in } L^n(\Omega_n) \) up to null set equivalence)

A stopping time \(N\) is a map \(N: \Omega \rightarrow \{1 \leq n \leq \infty\}\) such that \(\{N \leq n\} \in \mathcal{F}_n\) for each \(n\). Such an \(N\) is the same thing as a decomposition
\[ \Omega = B_1 \cup \cdots \cup B_\infty \quad \text{with } B_n \in \mathcal{F}_n. \]

Let's define \(\mathcal{F}_N\) to be the set of all \(A \in \mathcal{F}\) such that \(A \cap B_n \in \mathcal{F}_n\) for all \(n\). \(\mathcal{F}_N\) is a σ-algebra.

(What means that \(\mathcal{F}_N\) is a Boolean algebra - closed under \(\cap, \cup, \) - in \(\mathcal{F}\) and also countable unions.)

Next note that
\[ A \cap \{N \leq n\} = \bigcap_{k \leq n} A \cap B_k \]
so \(\forall n (A \cap B_n \in \mathcal{F}) \implies \forall n (A \cap \{N \leq n\} \in \mathcal{F}_n). \) And as
\[ A \cap B_n = A \cap \{N \leq n\} \cap \{N \leq n-1\} \]
the converse is true.

Thus
\[ \mathcal{F}_N = \{ A \in \mathcal{F} \mid \forall n (A \cap \{N \leq n\} \in \mathcal{F}_n) \} \]

If \(N' \leq N\), then \(\forall n \ A \cap \{N' \leq n\} = A \cap \{N \leq n\} \cap \{N \leq n\} \cap \{N' \leq n\} \in \mathcal{F}_n\) implies \(A \cap \{N' \leq n\} \in \mathcal{F}_n\), and as \(\{N_n\} \in \mathcal{F}_n\), one has \(A \cap \{N_n\} \in \mathcal{F}_n\) for all \(n\), so \(A \in \mathcal{F}_N\).
Finally, what stopping time gives $\mathcal{F}_n$? It is $N \equiv n$, so that $B_n = \Omega$ and $A \in \mathcal{F}_n \iff A \in \mathcal{F}_n$. 
Relation with Markov chains. State space $E$, the probability space is $\Omega = E^\mathbb{N}$, the $\sigma$-algebra $\mathcal{B}$ is generated by $X_n : \Omega \to E$, the $\sigma$-algebra $\mathcal{B}_p$ is generated by $X_0, \ldots, X_p$. For each $x \in E$ one has $P_x$ on $(\Omega, \mathcal{B})$ given by

$$P_x (X_0 = x_0, \ldots, X_p = x_p) = \delta_x (x_0) P(x_0, x_1) \cdots P(x_{p-1}, x_p)$$

where $P(x, y)$ is a Markov matrix ($E$ supposed discrete, i.e. a countable set; $P(x, \cdot)$ is a prob. measure in $E$ for each $x$). We have a stationary Markov chain.

Given $f : E \to \mathbb{R}$, consider $f_n$ on $\Omega$ given by $f_n = f(X_n)$. The conditional expectation of $f_n$ rel. to $\mathcal{B}_n$, is

$$\mathbb{E}_{\mathcal{B}_n} [f_n] (x_0, \ldots, x_{n-1}) = \sum_{x_n} p(x_0, \ldots, x_n) f(x_n)$$

$$= \sum_{x_n} p(x_{n-1}, x_n) f(x_n) = (Pf)_{n-1} (x_0, \ldots, x_{n-1})$$

In general use $P(x, y)$ to define an operator on func. $u : E \to \mathbb{R}$ by

$$(Pf)(x) = \sum_y P(x, y) f(y).$$

Thus

$$f_n = f(X_n) \text{ is a martingale} \iff Pf = f \quad (\text{f harmonic})$$

$$\text{submartingales} \iff Pf \geq f \quad (\text{f subharmon})$$
General definitions. Given \((A, B, P)\) and \(C_B < \cdots < C_B\).
A process \(\{f_n\}\), i.e., sequence of r.v. (real random variables), is called adapted if \(f_n\) is measurable rel. \(B_n\), and predictable if \(f_n\) is measurable rel. \(B_{n-1}\). (2\(n\)th term \(f_0 = 0\).

Proof: Decompose any adapted process \(\{f_n\}\) as a uniquely a sum \(m_n + a_n\), where \(m_n\) is a martingale and \(a_n\) is predictable.

Why:
\[
\begin{align*}
\Pi_{n-1} f_n &= m_{n-1} + a_n \\
f_{n-1} &= m_{n-1} + a_{n-1}
\end{align*}
\]
\[
\Rightarrow a_n - a_{n-1} = \Pi_{n-1} f_n - f_{n-1}
\]

This tells you how to define the \(a_n\) inductively, starting with \(a_0 = 0\).

Note: \(\{f_n\}\) submartingale \(\Rightarrow a_n \geq a_{n-1}\), so the predictable process is increasing.

Integration with a martingale (also called martingale transform).

Let \(\{f_n\}\) be a martingale, let \(\{g_n\}\) be predictable. Then
\[
h_n = \sum_{k=1}^{n} g_k (f_k - f_{k-1})
\]
is a martingale:
\[
\Pi_{n-1} (h_n) = h_{n-1} + \Pi_{n-1} \left( g_n (f_n - f_{n-1}) \right)
\]
\[
g_n \Pi_{n-1} (f_n - f_{n-1}) = 0
\]
(hypothesis: \(g_k \in L^\infty\), \(f_k \in L^1\) or maybe \(L^p\) and \(L^q\)?)

Example: This is a discrete version of
\[
h(t) = \int_t^\infty \delta(\omega, s) \, d\beta(s)
\]
Quadratic variation of a sequence $f_n$, now is

$$V = f_0^2 + \sum_{k=1}^{\infty} |f_k - f_{k-1}|^2$$

More generally we are interested in the sequence

$$V_n = f_0^2 + \sum_{1 \leq k \leq n} |f_k - f_{k-1}|^2$$

If $f_n$ is a $L^2$ martingale, then $f_n^2$ is a submartingale (Cauchy-Schwartz $\Rightarrow |f|^2 \leq |f|^2$ for a prob. measure - use this for conditional exps which is roughly a family of prob. measures on the filters)

$f_n^2$ has a Doob decomposition $f_n^2 = m_n + a_n$, where $a_n$ is increasing predictable.

$$a_n - a_{n-1} = \pi_{n-1}(f_n^2) - f_{n-1}^2$$

Also

$$\pi_{n-1}(f_n - f_{n-1})^2 = \pi_{n-1}(f_n^2) - f_{n-1}^2$$

so

$$a_n = \sum_{k=1}^{n} \pi_{k-1}(f_k - f_{k-1})^2$$

If $f_k - f_{k-1}$ are independent r.v. then $\pi_{k-1}(f_k - f_{k-1})^2$ is a constant, namely, the variance of this variable.

In this case the predictable process is just an increasing sequence of constants $a_n = \|f_n\|^2$.

We need the quadratic variation sequence to control the integral.

$$h_n = \sum_{1 \leq k \leq n} g_k (f_k - f_{k-1})$$

if mart $g$ pred.

then $h_n - h_{n-1} = h_{n-1}$ so

$$\|h_n\| - \|h_{n-1}\|^2 = \int g_k^2 (f_k - f_{k-1})^2$$
\[ \pi_{n-1} \frac{h^2}{h_{n-1}^2} = \pi_{n-1} \left( \frac{h_{n-1}}{h_n} \right)^2 \]
\[ = \pi_{n-1} \frac{g_n^2}{g_{n-1}^2} \left( f_n - f_{n-1} \right)^2 \]
\[ = g_n^2 \pi_{n-1} \left( f_n - f_{n-1} \right)^2 = g_n^2 \left( a_n - a_{n-1} \right) \]

Thus
\[ \| h \|^2 = \sum \int g_n^2 \left( a_n - a_{n-1} \right). \]

**Bell's problem:** Let \( Z_n \) \( n \in \mathbb{Z} \) be a sequence of i.i.d. random variables representing winnings of a gambler. For any stopping time \( \nu \) we get an expectation \( E(Z_\nu) \). The problem is to find the optimal stopping time, i.e. which realizes \( \sup E(Z_\nu) \).

Analogy: suppose \( z_n \) is a sequence in \( \mathbb{R} \) and we want to find \( p \) with \( z_p = \sup z_n \). Then we introduce \( x_p = \sup z_{n \geq p} \), which is a decreasing sequence and look for the first \( p \) with \( x_p = z_p \).

If this \( p \) is \( < \infty \), one wins.

In general one sets \( X_n = \sup \pi_{n-1} (Z_\nu) \).

\[ X_n = \sup \pi_{n-1} (Z_\nu) \]

\[ \nu \in \Lambda_n \text{ means } h \leq \nu \text{ a.e. and } E(Z_\nu) < \infty. \]

**Bell's theorem:** says that if \( \nu_\circ = \inf \{ n \mid X_n = Z_n \} \)

then there is an optimal \( \nu \equiv \nu_\circ < \infty \text{ a.e. Also } \nu_\circ \in \{ n \mid X_n < Z_n + \varepsilon \} \) is finite a.e. and satisfies \( E(Z_{\nu_\circ}) + \varepsilon \geq \sup E(Z_\nu) \).
April 19, 1985

**Positive supermartingales and potentials**

\[ X_n \geq 0 \text{ for } n \geq 0; \ X_n \text{ supermart: } \mathbb{P}_n X_n \leq X_{n-1}. \]  
This is an analogue of a positive decreasing sequence. The basic convergence then implies \( X_n \) converges a.e. to \( X_\infty \) and \( \mathbb{P}_n X_\infty \leq X_n \) for all \( n \).

Next one has the Doob decomposition.

\[ X_n = M_n - A_n \quad A_0 = 0 \]

\[ \mathbb{P}_n \mathbb{P}_n X_n = M_{n-1} - A_{n-1} \quad A_n - A_{n-1} = X_{n-1} - \mathbb{P}_n X_n \geq 0 \]

so \( A_n \) is an increasing predictable process. Put \( A_\infty = \sup_n A_n \).

Then \( \mathbb{P}_0 A_n \leq \mathbb{P}_0 A_n + \mathbb{P}_0 X_n = \mathbb{P}_0 M_n = X_0 \)

\[ \Rightarrow \ \mathbb{P}_0 A_\infty \leq X_0 \quad \Rightarrow \ A_\infty < \infty \text{ a.e.} \]

More generally \( \mathbb{P}_n A_\infty \leq M_n \) for all \( n \), so we get the decomposition

\[ X_n = (M_n - \mathbb{P}_n A_\infty) + (\mathbb{P}_n A_\infty - A_n) \]

pos. mart. pos. supermart

The **positive supermart** \( \bar{X}_n = \mathbb{P}_n A_\infty - A_n \) is called the potential of the increasing predictable process \( A_n \).

Such potentials are characterized by \( \mathbb{P}_0 \bar{X}_n \downarrow 0 \text{ a.e. as } n \to \infty \).

The above is called the **Riesz decomposition**.

Ex: Given \( B_n \) an increasing sequence of positive r.v. (not nec. adapted) such that \( \mathbb{P}_0 B_\infty < \infty \) a.e. put

\[ \bar{X}_n = \mathbb{P}_n (B_\infty - B_n) \]

This is a potential.
Ex: The Markov chain $\Omega = E^n$, $E$ = state space, defined by transition probability $p(x, \cdot)$ for each $x \in E$. I am thinking of $E$ as discrete. On $\Omega$ goes the probability $P_x$ such that on $\Omega$ it is 

$$p(x_0, \ldots, x_n) = \delta(x_0) p(x_0, x_1) \cdots p(x_{n-1}, x_n).$$

Given $f : E \to \mathbb{R}$ we get a process

$$f_n = f(X_n), \quad X_n : \Omega \to E \quad \text{with proj}.$$

and one has

$$\pi_n f_n = (Pf)_{n-1}, \quad (Pf)(x) = \int p(x, y) f(y) dy.$$

If $f$ is superharmonic: $Pf \leq f$, then $f_n$ is a supermart. Assume $t > 0$ and denote $P_t$ compute the Doob decompo.

$$f_n = M_n - A_n$$

$$A_n - A_{n-1} = f_{n-1} - \pi_{n-1} f_n = (f - Pf)_{n-1}$$

$$A_\infty - A_n = (f - Pf)_n + (f - Pf)_{n+1} + \cdots$$

$$\bar{X}_n = \pi_n (A_\infty - A_n) = (f - Pf)_n + (P(f - Pf))_n + (P^2(f - Pf)) + \cdots$$

$$= \left( \sum_{n=0}^{\infty} P^n \right) (1 - P) f$$

We have to be a bit careful with the limits.

$f \geq Pf \quad \text{positive} \quad \Rightarrow 

f \geq Pf \geq P^2 f \geq \cdots \quad \Rightarrow \lim_{n \to \infty} P^n f$ exists

and it is harmonic. Then

$$\sum_{k=0}^{\infty} P^k (f - Pf) = f - P^\infty f \quad \text{pos harm.}$$

Note that the Riesz decomposition amounts to

$$f = (P^\infty f) + (f - P^\infty f)$$

\text{pos. superharm. \quad \& killed by} \lim_{n \to \infty} P^n f$
Dubin's inequality. 1) in positive supermartingale, a > 0

$$\Rightarrow \quad \pi_0 \mathbb{P}\left(\cap_{n \geq 1} f_n \geq a\right) \leq \min \left(\frac{f_0}{a}, 1\right)$$

2) f_n is a supermartingale, 0 < a < b < \infty

$$\pi_0 \mathbb{P}\left(\exists \kappa \text{ uppercrossing of } (a, b)\right) \leq \left(\frac{a}{b}\right)^k \min \left(\frac{f_0}{a}, 1\right)$$

Why this is true: Think in terms of discrete prob. spaces, and suppose we have a tower

$$\Omega_0 \leftarrow \Omega_1 \leftarrow \Omega_2 \leftarrow \cdots \leftarrow \Omega$$

with f_n defined on \( \Omega_n \). When we compute \( \pi_0 \) we get a f_n on \( \Omega_0 \) which gives the relative or conditional expectation. The first inequality says that given \( x_0 \in \Omega_0 \), the probability of \( \cup_{n \geq 1} f_n \geq a \) over events \( x \) beginning with \( x_0 \) is \( \leq \frac{1}{a} f_0(x_0) \) (and \( \leq 1 \) of course).

Now suppose we want the probability of \( \geq 1 \) uppercrossings again over events x starting with \( x_0 \). To introduce stopping times

$$\nu_1(x) = \inf \left\{ n \mid f_n(x) \leq a \right\}$$

$$\nu_2(x) = \sup \left\{ n \mid n > \nu_1(x), f_n(x) \geq b \right\}$$

I have to think in terms of \( \Omega \) consisting of \( (x_n) \), \( x_n \in \Omega_n \), i.e. \( \Omega = \varinjlim \Omega_n \). I want to calculate the probability that \( \nu_2 < \infty \). Actually it seems to be possible to define \( \Omega_{\nu_2=\infty} = \bigcup_{n=1}^{\infty} \text{ Image of } \nu_2=n \text{ in } \Omega_n \) and similarly \( \Omega_{\nu_1} \). Then one has a map \( \Omega_{\nu_2=\infty} \rightarrow \Omega_{\nu_1} \rightarrow \Omega \) and we want to compute the probability of \( \nu_2 < \infty \) by bounding
the relative probability of the fibre, Thus given \( x_0, \ldots, x_k \) with \( f_k(x_k) \leq a \), the probability that the sequence gets above \( b \) later on is bounded by (from 1)

\[
\min \left( \frac{1}{b} f_k(x_k), 1 \right) \leq \frac{a}{b}
\]

This certainly will handle all the future upcrossings however at the first upcrossing we can argue as follows. On the set where \( f_0(x_0) \leq a \) we use the bound \( \min \left( \frac{1}{b} f_0(x_0), 1 \right) \) and on the set where \( f_0(x_0) > a \) we just use \( \frac{a}{b} \). Thus the relative prob. is

\[
\begin{cases}
\frac{f_0(x_0)}{b} & f_0(x_0) \leq a \\
\frac{a}{b} & f_0(x_0) > a
\end{cases}
= \frac{a}{b} \min \left( \frac{1}{b} f_0(x_0), 1 \right)
\]
Let's review the basics concerning upcrossing inequalities for martingales.

One starts with an increasing family $B_0 \subset B_1 \subset \ldots$ of $\sigma$-fields and an adapted process $\{X_n, n \geq 0\}$ which is a sequence of r.v. with $X_n$ measurable to $B_n$.

A basic idea is that given a stopping time $\tau$ one can define a $\sigma$-field $B_\tau$, and a r.v. $X_\tau$.

Then one extends notions from integers to stopping times.

(To be more precise about the value $\infty$, suppose $B_\infty$ is given containing all $B_n$. Then

$$B_\tau = \{ A \mid A \cap \{ \tau \leq n \} \in B_n \text{ for all } n \}$$

is defined without problem. However $X_\nu(x) = X_{\nu\omega}(x)$ is defined only if $\nu(x) < \infty$ a.e. or if and $X_\infty$ is given, or if $X_\nu$ has a limit $X_\infty$.)

Then the definition of martingale:

$$\Pi_n X_m = X_n$$  \quad \text{for } n < m$$

can be extended to stopping times

$$\Pi_S(X_T) = X_S \quad \text{if } S \leq T.$$
\[
\int U^n_a \leq \frac{1}{b-a} \lim_{n \to \infty} (X^n_n - a)^+
\]

Proof: One can suppose \( a = 0 \), and one can replace \( X^n_n \) by \( X^n_n^+ \) (where \( x^+ \) is convex and a convex fn.

preserves submartingales (modulo int. hypothesis) by Jensen).

so one can suppose \( X^n_n \) is a pos. submart.

Think of a submartingale as a game. Consider the following betting strategy.

Better think of \( X^n_n \) as the value of a stock at time \( n \). Here's the strategy. If \( X^n_n \) falls below \( a \) we place an order to buy 1 share which we then hold on \( b \) until the value rises above \( b \). This isn't quite right.

Maybe I should think of \( X^n_n - X^n_{n-1} \) as the result of the \( n \)-th play. Each time we have to decide how much to bet. If I want to think in terms of stocks, I buy at the beginning and sell at the end of each day. I then have to decide whether to do this process for the coming day.

So the strategy is to begin buying when \( X^n_n \) first drops below \( a \), and to stop when it first rises above \( b \). This gives me a total amount \( Z^n_n \) of the number of upcrossings. Here use \( X^n_n \geq 0 \), \( a = 0 \).

\[
\therefore b \int (U^n_a) = \int Z^n
\]

But now because the game is favorable, the more often one plays the better the expectation. This means that

\[
\int Z^n \leq \int X^n
\]

Q.E.D. by letting \( n \to \infty \).
Doob's upcrossing inequality for submartingales (assumed integrable, i.e. $X_n \in L^1$ for each $n$, so that the conditional expectation $\mathbb{E}_n$ has a meaning.)

Given an open interval $(a, b)$ define functions $\nu_1, \nu_2, \ldots$ from $\Omega$ to $\mathbb{N} \cup \{\infty\}$ by

\[\nu_1 = \inf \{ n \geq 0 \mid X_n \leq a \}\]
\[\nu_2 = \inf \{ n \geq \nu_1 \mid X_n \geq b \}\]
\[\nu_3 = \inf \{ n \geq \nu_2 \mid X_n \leq a \}\] etc.

Then the $\nu_k$ are stopping times, i.e., $\{\nu_k \leq n\} \in \mathcal{F}_n$ (in words: to determine whether $\nu_k \leq n$ one needs only look at the values of $X_0, \ldots, X_n$)

Now put

\[\beta_n = \text{largest } k \text{ such that } \nu_{2k} \leq n\]

Then $\beta_n$ is the number of upcrossings of $(a, b)$ in time $\leq n$.

The result is then
Doob's preserving inequality for submartingales:
\[ \int \beta_n \leq \frac{1}{b-a} \int (X_n-a)^+ - (X_0-a)^+ \]

I want next to describe Doob's proof which is based on the idea that a submartingale is a favorable game and betting. First note that by replacing \( X_n \) by \( X_n - a \) we can suppose \( a = 0 \).

Also because \( x \mapsto x^+ \) is convex and Jensen's inequality \( X_n^+ \) is also a submartingale, (Or more elementary is:
\[
\frac{\tau_{n+1}(X_n^+)}{\tau_{n+1}(X_n^-)} = \frac{\tau_{n+1}(X_n)}{\tau_{n+1}(X_n^+)} > X_{n+1} \Rightarrow X_n^+ \leq \tau_{n+1}(X_n^+) \).
\]

Replacing \( X_n \) by \( X_n^+ \) doesn't change the functions \( \nu_i^* \), so we can suppose \( X_n \geq 0 \).

Now consider the following betting process:
We think of the process \( X_n \) mainly in terms of its increments \( X_n - X_{n-1} \); thus \( X_n - X_{n-1} \) represents the change in going from time \( n-1 \) to time \( n \). A betting scheme consists of giving \( n \) functions \( V_n, \ldots, V_2, V_1 \). \( V_n \) is the amount one bets on the increment \( X_n - X_{n-1} \), and one assumes \( V_n \) is measurable relative to \( F_{n-1} \), i.e. the amount bet on \( X_n - X_{n-1} \) depends only upon one's knowledge of the process up through time \( n-1 \).

(by convention it is useful to allow \( V_0 \) to be a constant).

The result of the betting is the process:
\[ (V \cdot X)_n = V_0 X_0 + V_1 (X_1 - X_0) + \cdots + V_n (X_n - X_{n-1}) \]
This gives the net gain or loss at time \( n \).
Example: Let \( \mathcal{N} : \Omega \to [0, \infty] \) be a stopping time. 

and take 

\[
V_n(\omega) = \begin{cases} 
1 & \text{if } n \leq \mathcal{N}(\omega) \\
0 & \text{if } n > \mathcal{N}(\omega)
\end{cases}
\]

or 

\[
V_n = 1 - 1_{\mathcal{N} < n} = 1 - 1_{\mathcal{N} \leq n-1}
\]

is measurable relative to \( \mathcal{B}_{n-1} \). Thus \( \{V_n\} \) is predictable. Clearly 

\[
(V \cdot X)_n = \sum X_0 + (X_1 - X_0) + \cdots + (X_{\mathcal{N} n} - X_{\mathcal{N} n-1}) = X_{\mathcal{N} n}.
\]

This process represents betting \( \mathbb{I} \) until the stopping time is reached and then stopping.

Next we want to derive the fact if we are going to bet \( \mathbb{I} \) amounts \( V_n \) with \( 0 \leq V_n \leq 1 \) on a submartingale (which is supposed to be a favorable game), then our expectation is greatest when we bet \( \mathbb{I} \) all the time. Put \( Y_n = (V \cdot X)_n \) whence 

\[
Y_n - Y_{n-1} = V_n(X_n - X_{n-1}).
\]

Then 

\[
\Pi_{n-1} Y_n = Y_{n-1} \quad = \quad \Pi_{n-1} \{V_n(X_n - X_{n-1})\}
\]

\[
= V_n \cdot (\Pi_{n-1}(X_n) - X_{n-1}) \leq \Pi_{n-1} \frac{V_n}{0 \leq V_n \leq 1} \geq 0
\]

so 

\[
\int Y_n - Y_{n-1} \leq \int X_n - X_{n-1}
\]

or 

\[
\int Y_n - \mathbb{I} Y_0 \leq \int X_n - X_0
\]
Now apply this to the following betting process in the case of a positive submartingale. We bet nothing until the time 
\[ \nu_1 = \inf \{ n \mid X_n = 0 \} \]
then we bet 1 each time until the time 
\[ \nu_2 = \inf \{ n > \nu_1 \mid X_n = b \} \]
then zero until the time \( \nu_3 \), etc. Thus 
\[ V_n = \begin{cases} 1 & \text{if } \nu_{2k-1} < n \leq \nu_{2k} \text{ for some } k \\ 0 & \text{otherwise} \end{cases} \]

\[ V_n = \left( \mathbb{1}_{\nu_2 \leq n} - \mathbb{1}_{\nu_1 \leq n} \right) + \left( \mathbb{1}_{\nu_3 \leq n} - \mathbb{1}_{\nu_2 \leq n} \right) + \cdots \]

\[ Y_n = \left( X_{\nu_2 \wedge n} - X_{\nu_1 \wedge n} \right) + \left( X_{\nu_3 \wedge n} - X_{\nu_2 \wedge n} \right) + \cdots \]

\[ = \begin{cases} 0 & n \leq \nu_1 \\ X_n - X_{\nu_1} & \nu_1 < n \leq \nu_2 \\ X_{\nu_2} - X_{\nu_1} & \nu_2 < n \leq \nu_3 \\ X_n - X_{\nu_3} + (X_{\nu_2} - X_{\nu_1}) & \nu_3 < n \leq \nu_4 \\ & \cdots \end{cases} \]

Note \( X_{\nu_1} = 0 \), \( X_{\nu_2} \geq b \) (when these \( \nu \)'s are finite).

Thus \( Y_n \geq b \beta_n \) for all \( n \). (In words, we wait to the process drops to 0 then we bet 1 until it rises \( \geq b \) and stop, then start when it next falls to zero, etc. Our gain is \( \geq b \cdot \text{no. of up-crossings} \)).

Putting things together we get
\[ b \int \beta_n \leq \int Y_n \leq \int X_n - X_0 \]
which is exactly Doob's upcrossing inequality.

Next I want to describe Neveu's proof which is based on the following switching lemma.

**Lemma:** Let $X_n$ and $Y_n$ be submartingales and let $\nu$ be a stopping time. Assume $X_\nu \leq Y_\nu$ on $\nu < \infty$. Then

$$Z_n = X_n 1_{n < \nu} + Y_n 1_{n \geq \nu}$$

is a submartingale.

**Proof.** $Z_n = X_n 1_{n < \nu} + (Y_n - X_n) 1_{n = \nu} + Y_n 1_{n > \nu}$

$$\in B_{n-1} \Rightarrow 0$$

So

$$\Pi_{n-1} Z_n = (\Pi_{n-1} X_n) 1_{n < \nu} + \Pi_{n-1} [(Y_n - X_n) 1_{n = \nu}] + \Pi_{n-1} (Y_n) 1_{n > \nu} \geq X_{n-1}$$

$$\geq X_{n-1} 1_{n < \nu} + Y_{n-1} 1_{n > \nu}$$

$$= X_{n-1} 1_{n-1 < \nu} + Y_{n-1} 1_{n-1 \geq \nu} = Z_{n-1}$$

Now define the stopping times as before on p.448:

$$\nu_1 = \inf \{ n \geq 1 : X_n \leq a \}$$

$$\nu_2 = \inf \{ n \geq 1 : X_n \geq b \}$$

and define $Z_n$ as follows:

$$Z_n = X_n - a \quad 0 \leq n < \nu_1$$

$$= 0 \quad \nu_1 \leq n < \nu_2$$

$$(X_n - a) - (b-a) \quad \nu_2 \leq n < \nu_3$$
in general
\[ Z_n = \begin{cases} (x_n - a) - (b-a)\beta_n & \text{in } U \{ \nu_{2k} \leq n < \nu_{2k+1} \} \\ -(b-a)\beta_n & \text{in } U \{ \nu_{2k+1} \leq n < \nu_{2k+2} \} \end{cases} \]

The switching lemma implies that \( Z_n \) is a submartingale. Note
\[ Z_0 = (x_0 - a)^+ \]
\[ Z_n \leq (x_n - a)^+ - (b-a)\beta_n \]
and so we have
\[ \int Z_n \geq \int Z_0 \]
or
\[ (b-a)\beta_n \leq \int (x_n - a)^+ - (x_0 - a)^+ \]

Dubins' inequality for positive supermartingales. Let
\[ 0 < a < b < \infty. \]
Let \( \beta \) be the number of upcrossings of \((a, b)\). Then
\[ \Pi_0(\{ \beta \geq k \}) \leq \left( \frac{a}{b} \right)^k \min \left( \frac{x_0}{a}, 1 \right) \]

Proof after Neveu: Introduce stopping times
\[ \nu_1 = \inf \{ n \geq 0 \mid x_n \leq a \} \]
\[ \nu_2 = \{ n \geq \nu_1 \mid x_n \geq b \} \]

e tc. down thru \( \nu_{2k} \), and then using switching to see...
that
\[ Z_n = \begin{cases} 
1 & 0 \leq n < \nu_1 \\
\frac{X_n}{a} & \nu_1 \leq n < \nu_2 \\
\frac{b}{a} & \nu_2 \leq n < \nu_3 \\
\frac{b \cdot X_n}{a} & \nu_3 \leq n < \nu_4 \\
\ldots & \nu_{2k} \leq n \\
\left(\frac{b}{a}\right)^k & \nu_{2k} \leq n
\end{cases} \]

is a positive supermartingale. So
\[ \min\left(\frac{X_0}{a}, 1\right) \leq Z_0 \geq \prod_0 Z_n \geq \prod_0 (Z_n : \beta \geq k) = \left(\frac{b}{a}\right)^k \prod_0 (1 : \beta \geq k) \]

Next consider the maximal inequality for positive supermartingales
\[ \prod_0 1 \left\{ \sup_{k \leq n} X_k \geq a \right\} \leq \min\left(\frac{X_0}{a}, 1\right) \]

By a limiting argument it is enough to show
\[ \prod_0 1 \left\{ \sup_{k \leq n} X_k > a \right\} \leq \frac{X_0}{a} \]

I want to do this the way I understood earlier and then by Neveu's switching method. Let
\[ \nu = \inf \{ n > 0 \mid X_n > a \} \]
so that \[ \{ \sup_{k \leq n} X_k > a \} = \{ \nu \leq n \} \]. Then
we argue as follows

$$X_0 \geq \pi_0(X_{\nu \wedge n}) \geq \pi_0(X_{\nu \wedge n}, \nu \leq n) \geq \pi_0(1_{\nu \leq n}).$$

The middle \(\geq\) results from the positivity of \(X_{\nu \wedge n}\), the last inequality from the fact that \(X_{\nu} \geq a\) on \(\nu \leq n\).

The key step then is the first inequality where the fact that \(X_n\) is a supermartingale is used. It suffices to show that

$$\pi_{n-1}(X_{\nu \wedge n}) \leq X_{\nu \wedge n(n-1)}$$

for any stopping time \(\nu\), i.e. that \(X_{\nu \wedge n}\) is a supermartingale. But this follows from the fact that \(X_{\nu \wedge n} = (V_n X)\), where \(V_n = 1_{n \leq \nu}\) and that such a transform for \(V \geq 0\) is always a supermartingale.

Specifically, in the case of interest

$$X_{\nu \wedge n} - X_{\nu \wedge n(n-1)} = 1_{n \leq \nu} (X_n - X_{n-1})$$

so

$$\pi_{n-1} X_{\nu \wedge n} - X_{\nu \wedge n(n-1)} = 1_{n \leq \nu} (\pi_{n-1} X_n - X_{n-1}) \leq 0.$$  

Next consider Neveu's proof. With \(\nu\) as before set

$$Z_n = \begin{cases} X_n & \text{if } n < \nu \\ a & \text{if } n \geq \nu. \end{cases}$$

By switching this is a supermartingale so

$$Z_0 \geq \pi_0(Z_n) \geq \pi_0(Z_n, \nu \leq n) = a \pi_0(1_{\nu \leq n}) \geq \min(X_0, a) \quad \text{so} \quad \pi_0(1_{\nu \leq n}) \leq \min(\frac{X_0}{a}, 1).$$
The stopping inequalities for positive supermartingales leads easily to the fact they converge a.e. set $X_\infty = \lim X_n$; Fatou's lemma says

$$T_n(x_\infty) = \lim_{p \to \infty} T_n(x_p) \leq \lim_{p \to \infty} T_n(x_p) \leq X_n$$

and so $X_\infty < \infty$ a.e. outside the set where all $X_n$ are $\infty$, to some $X_n \in L^1 \Rightarrow X_\infty \in L^1$.

Now let us turn to the supermartingale convergence thm:

Thm: Assume $X_n$ is an integrable supermartingale such that $\sup_0 \int x_n^+ < \infty$. Then $X_n$ converges a.e. to an integrable $X_\infty$.

Proof. $x_n^+$ is a positive submartingale, so $T_n(x_n^+)$ is increasing for $p \geq n$ and we can define

$$M_n = \limsup_{p \to \infty} T_n(x_p^+) = \lim_{p \to \infty} T_n(x_p^+) = \lim_{p \to \infty} T_n(x_p^+) = M_{n-1}$$

By monotone convergence

$$\int M_n = \lim \int x_n^+ < \infty$$

so each $M_n \in L^1$. Also

$$\pi_n(M_n) = \lim_{p \to \infty} \pi_n(x_n^+) = \lim_{p \to \infty} \pi_n(x_n^+) = M_{n-1}$$

so $M_n$ is a positive martingale.

By the convergence theorem for pos. supermartingales $M_n \to M_\infty$ a.e. where $\int M_\infty \leq \int M_n = \lim \int x_n^+ \Rightarrow M_\infty \in L^1$.
Next we have
\[ M_n = \lim_{P_n} (X^+_n) \geq X^+_n \geq X_n \]
so if we put \( Y_n = M_n - X_n \), then \( Y_n \in L^1 \) and \( Y_n \geq 0 \).
Clearly because \( M_n \) is a mart., \( + X_n \) is a submart., \( Y_n \) is a supermart. Thus by the supermart. conv. thm. \( Y_n \to Y_\infty \)
a.e. where \( Y_\infty \in L^1 \). Thus \( X_n = M_n - Y_n \to M_\infty - Y_\infty \) a.e.
and this belongs to \( L^1 \). \( \text{QED.} \)

In general
\[
\int X = \int x^+ - \int x^-
\]
\[
\int |x| = \int x^+ + \int x^-
\]
\[
2 \int x^+ = \int |x| + \int x
\]
For a martingale \( \int X_n \) is constant + finite so
\[
\sup \int x^+_n < \infty \iff \sup \int |x_n| < \infty.
\]
Thus a martingale bounded in \( L^1 \) converges a.e. to an element of \( L^1 \).

Doob proves: If \( X_n \) is a martingale bold in \( L^p \), then \( X_n \) converges in \( L^p \) for \( p > 1 \). For \( p = 1 \) false, but true if \( X_n \) is a regular integrable martingale:
\[
\sup \int_{|X_n| > a} |X_n| \to 0 \quad \text{as} \quad a \to \infty
\]
(satisfied if \( \sup |X_n| \in L^1 \) and this follows from \( \sup \int_{|X_n| > a} |X_n| < \infty \))
What is a martingale? To simplify we consider the discrete case where the set of times is \( N \). A martingale is a stochastic process \( \{X_n, n \in N\} \). This means we are given a probability space \((\Omega, \mathcal{B}, P)\) on which the \( X_n \) are random variables. Further we are given a filtration \( B_0 \subset B_1 \subset \cdots \) of the \( \sigma \)-algebra \( \mathcal{B}_n \) where \( \mathcal{B}_n \) is the algebra of events up to time \( n \). I like to think of \( \Omega \) as the inverse limit of a tower

\[ \Omega_0 \subset \Omega_1 \subset \Omega_2 \subset \cdots \]

of probability spaces and of \( \mathcal{B}_n \) as the measurable subsets in \( \Omega_n \). A point of \( \Omega_n \) is a possible history up to time \( n \).

With this notation a martingale is sequence of real r.v.'s \( X_n \) which is adapted to the filtration \( \{B_n\} \) i.e. \( X_n \) is measurable relative to \( B_n \), such that the conditional expectation of \( X_n \) relative to \( B_m \) when \( m < n \) is \( X_m \).

Equivalently, the conditional expectation relative to \( B_{n-1} \) of the \( n \)-th increment \( X_n - X_{n-1} \) is zero. Thus if we know the history up to time \( n-1 \), the average value of \( X_n - X_{n-1} \) is zero.

An example of an integrable martingale is

\[ X_n = Y_0 + \cdots + Y_n \]

where the \( Y_j \) are independent integrable r.v.'s of mean zero.

Here are the martingale convergence theorems.

1) A bounded \( L^1 \)-martingale \( X_n \) (mean sup \( \mid X_n \mid \) < \( \infty \)) converges a.e. The limit \( X_\infty \) is in \( L^1 \).
Recall $L^p \subseteq L^1$ for $1 \leq p$, so that a bounded $L^p$ martingale $X_n$ has an a.e. limit $X_\infty$ by 1).

2) If $1 < p < \infty$, and $X_n$ is a bounded $L^p$ martingale, then $X_n \rightarrow X_\infty$ in $L^p$.

Remarks:

a) For $p = 2$ it is very easy to see that an $L^2$ bounded martingale converges in $L^2$, because the conditional expectation operator coincides with the orthogonal projection: $L^2(\Omega) \rightarrow L^2(\Omega^n)$. Thus for an $L^2$ martingale the increments $X_n - X_{n-1}$ are mutually orthogonal and

$$\|X_n\|_2^2 = \|X_0\|_2^2 + \sum_{k=1}^{n} \|X_k - X_{k-1}\|_2^2$$

and so if this is bounded the series

$$X_\infty = X_0 + \sum_{k=1}^{n}(X_k - X_{k-1})$$

converges in $L^2$.

b) Since $x \mapsto |x|^p$ is convex for $p > 1$, it follows from Jensen that $|X_n|^p$ is a submartingale, hence $\{ |X_n|^p \}$ also $\|X_n\|_p$ is an increasing sequence.

The proof of 2) is based on Doob's "maximal" inequality: if $\{X_n\}$ is a positive (integrable) submartingale, and $X^*_n = \sup_{k \leq n} X_k$, then

$$P\{X^*_N \geq a\} \leq \frac{1}{a} \int_{X_N \geq a} dP$$

and hence

$$P\{X^*_\infty \geq a\} \leq \frac{1}{a} \sup_{N} \int_{X_N \geq a} dP$$

and

$$X^*_\infty = \sup_{n} X_n$$

The reason I record this is that a key idea is whether $X^*_\infty$ is integrable.
Let $X_n$ be an $L^p$ bounded martingale, $1 \leq p < \infty$.

From the maximal inequality

$$
P \left\{ \sup_{k \in \mathbb{N}} |X_k| > a \right\} \leq \frac{1}{a} \int \mathbb{1}_{\left\{ \sup_{k \in \mathbb{N}} |X_k| > a \right\}} |X_n| \, \text{d}n
$$

one can deduce that $\sup |X_n| \in L^p$ for $p > 1$. (In fact $\| \sup |X_n| \|_p \leq \frac{p}{p-1} \| X_n \|_p$)

By the martingale convergence theorem, one has $X_n \to X_\infty$ in $L^p$, i.e. By dominated convergence $|X_n| \leq \sup |X_n| \in L^p$, we see that $X_n \to X_\infty$ in $L^p$.

Thus even when $p = 1$ we have

$$\sup |X_n| \in L^1 \Rightarrow X_n \to X_\infty \text{ in } L^1$$

The actual behavior at $p = 1$ is very delicate.

Variant of $\odot$

$$\int \sup_{k \in \mathbb{N}} |X_k| \leq \frac{e}{e-1} \left( 1 + \int |X_n| \log^+ |X_n| \right)$$

which shows that $\int |X_n| \log^+ |X_n|$ bounded $\Rightarrow X_n \to X_\infty$ in $L^1$.

The actual behavior at $p = 1$ is very delicate and there is an extensive development analogous to Fefferman's theorem that $H^1$ and $BMO$ are dual. Also Littlewood–Paley theory is involved.
I make an attempt to summarize the theory.

First we must describe the space $M_p$ (or $H^p$).

For $p > 1$ it consists of bounded $L^p$ martingales with norms

$$\sup \|X_n\|_p \leq \sup \|X_n\|_p$$

Thus $M_p$ should be the same as $L^p(\Omega)$ (assuming $L^p(\Omega_n)$ is dense in $L^p(\Omega)$).

Now we take up $H^1$. This consists of $L^1$-bounded martingales $X$ (with $X_0 = 0$ to simplify) with either of the equivalent norms. (The equivalence is non-trivial)

The first is the "maximal" norm:

$$\int \sup_N |X_n|$$

and the second is the "quadratic" norm

$$\int \sqrt{\mathbb{V}} = \int \left( \sum_N \tau_n (X_{n+1} - X_n)^2 \right)^{1/2}$$

(Remark: Littlewood - Paley theory concerns Fourier series being summed in a "dyadic" way. There is an analogue for martingales in which the interesting $L^1$-gadget is

$$\left( \sum_n \tau_n (X_n - X_{n-1})^2 \right)^{1/2} = \sqrt{\mathbb{V}}$$

In any case there is an equivalence of norms

$$\|\sqrt{\mathbb{V}}\|_p \quad \text{and} \quad \sup \|X_n\|_p$$

for all $p > 1$. )