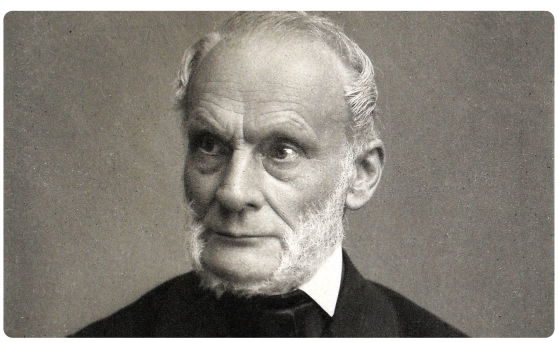
What does entropy measure?

Laure Saint-Raymond

---

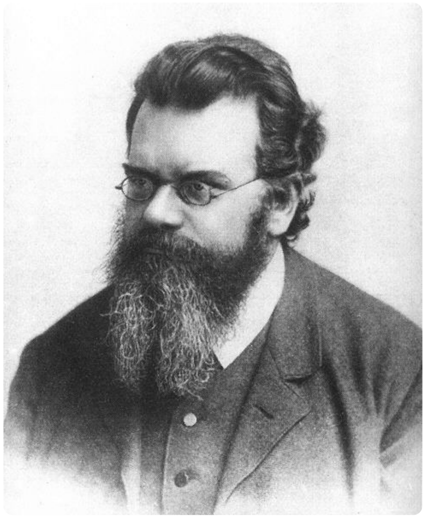
1. From classical thermodynamics to information theory


Rudolf Clausius, 1865

Entropy measures the irreversibility of a transformation

$$\Delta S \geq \frac{\text{heat flux } Q}{\text{temperature } T}$$

I propose, accordingly, to call S the entropy of a body, after the Greek word 'transformation'. I have designedly coined the word entropy to be similar to 'energy', for these two quantities are so analogous in their physical significance, that an analogy of denomination seemed to me helpful. (Clausius)

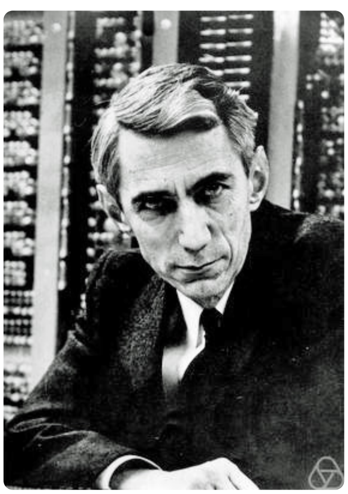


Ludwig Boltzmann, 1872

Entropy is related to the H function measuring the number of microscopic configurations associated to a given state.

$$S = k_B \log W \quad \text{(Planck)}$$

The system will always evolve from the least likely state to the most likely state, i.e. towards thermal equilibrium. By applying this to the second principle, we can identify the quantity generally called entropy with the probability of this state. (Boltzmann)

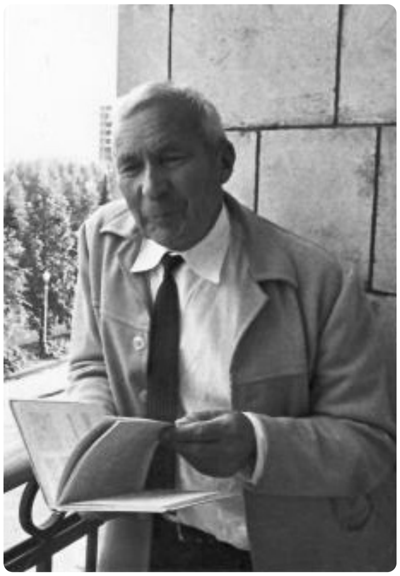


Claude Shannon, 1948

Entropy can be generalized, with a similar definition, to measure the average level of surprise / uncertainty for the possible outcomes of a random variable

$$H(X) = E_p[-\log p(X)]$$

Shannon considered various ways to encode, compress, and transmit messages from a data source, and proved in his famous source coding theorem that the entropy represents an absolute mathematical limit on how well data from the source can be losslessly compressed onto a perfectly noiseless channel.



Andrei Kolmogorov, 1958

Entropy is then extended to dynamical systems to measure their mixing properties

$$H(T) = \sup_{\xi \text{ partition}} H(T,\xi)$$

where $H(T,\xi)$ can be seen as the entropy of the random process $(x, w_1(x), \dots w_k(x) \dots)$ where $T^k x \in \xi_{w_k(x)}$

$$\lim_{n\to\infty} -\frac{1}{n} \sum_{i_1,\dots i_n} \mu(T^{-1}\xi_{i_1} \cap \dots \cap T^{-n}\xi_{i_n}) \log \mu(T^{-1}\xi_{i_1} \cap \dots \cap T^{-n}\xi_{i_n})$$

Kolmogorov proposed the notion of entropy about which it was believed that it will allow to distinguish "probabilistic" dynamical systems and "deterministic" dynamical systems (Sinaï)

Depending on the context, the notion of entropy has different definitions!
These objects share however some features.

* statistical quantities
(make sense only for a complex system)

* additive quantities (for two independent components)

* measure the level of desorganization / uncertainty
(related to mixing properties)

2. $\boxed{\text{Close to equilibrium. the relative entropy method}}$

<u>Example 1</u> : the Boltzmann equation in incompressible inviscid regime

$$\begin{cases} \varepsilon \partial_t f + v \cdot \nabla_x f = \frac{1}{\varepsilon^q} Q(f) \\ f^0 = M(1 + \varepsilon g^0) \end{cases}$$

Knudsen $\varepsilon^q \ll$ Mach ε

- close to local thermodynamic equilibrium $f \sim M_f$
- macroscopic (slow) dynamics governed by the Euler equations
- vanishing dissipation in the absence of singularity

→ Entropy $H(f) = \iint f \log f \, dx\, dv$ is a Lyapunov functional

$$H(f) + \frac{1}{\varepsilon^{q+1}} \int_0^t D(f)\, ds \leq H(f^0)$$

→ Relative entropy $H_\varepsilon(f | M_{1,\varepsilon u,1}) = \frac{1}{\varepsilon^2} \iint \left(f \log \frac{f}{M_{1,\varepsilon u,1}} - f + M_{1,\varepsilon u,1} \right) dx\, dv$

controls the distance $\frac{1}{\varepsilon} \| f - M_{1,\varepsilon u,1} \|_{L^1}$

→ Modulated entropy inequality provides the stability around smooth u

$$H(f | M_{1,\varepsilon u,1}) + \frac{1}{\varepsilon^{q+3}} \int_0^t D(f)\, ds \leq H(f^0 | M_{1,\varepsilon u_0,1}) \exp\left(\int_0^t \|\nabla u\|_{L^\infty} ds \right)$$
$$+ \int (\partial_t u + u \cdot \nabla u + \nabla p) \cdot (u_\varepsilon - u) \exp\left(\int_s^t \|\nabla u\|_{L^\infty} ds' \right)$$

Example 2 : System of interacting particles in mean field regime

$$\begin{cases} \partial_t f_N + \sum_i v_i \cdot \nabla_{x_i} f_N + \frac{1}{N} \sum_{i \neq j} \nabla V(x_i - x_j) \cdot \nabla_{v_i} f_N = 0 \\ f_N^0 = \bigotimes_{i=1}^{N} f^0(x_i, v_i) \end{cases}$$

- close to chaotic distribution $f_N(t) \sim \bigotimes_{i=1}^{N} f(t, x_i, v_i)$
- mean (slow) dynamics governed by the Vlasov equation
$$\partial_t f + v \cdot \nabla_x f + \left(\nabla V * \int f dv\right) \cdot \nabla_v f = 0.$$
- no loss of information in the absence of singularity (in V)

$\to$ Entropy $H_N(f_N) = \frac{1}{N} \int f_N \log f_N \, dX_N \, dV_N$ is a conserved (additive) quantity

$\to$ Relative entropy $H_N(f_N | f^{\otimes N})$ controls the distance of any marginal $f_N^{(k)}$ to the tensor product $f^{\otimes k}$

$\to$ Modulated entropy inequality provides the stability around smooth f

$$H_N(f_N | f^{\otimes N}) \leq \left(H_N(f_N^0 | f_0^{\otimes N}) + \frac{C}{N}\right) \exp\left(C(f) \|V\|_\infty t\right)$$

At equilibrium, the entropy is maximal. Thus, starting close to equilibrium, the evolution is highly constrained.

If the equilibrium is only local, to get some stability, one needs in addition to control the fluxes.

Entropic convergence results from the combination of
- a local equilibrium (maximal entropy)
- a conservative slow dynamics (no dissipation, therefore no possible entropy production)

---

3. $\boxed{\text{Relaxation towards equilibrium - functional inequalities}}$

<u>Example 1</u> : the long time limit of the Boltzmann equation

$$\begin{cases} \partial_t f + v \cdot \nabla_x f = Q(f) \\ H(f^0 | M) < +\infty \end{cases}$$

- collisions induce relaxation towards local equilibrium
- transport homogenizes thermodynamic fields.
- almost exponential relaxation in the absence of pathological velocity profile.

→ Entropy / entropy production inequality

· Cercignani's conjecture $D(f) \geq C\, H(f|M_f)$ is wrong in general, requires technical assumptions on f.

→ Hypocoercive mechanism can be described by a system of differential inequalities on the local and global relative entropies.

$$H(f|M) = H(f|M_f) + H(M_f|M)$$

$$\begin{cases} \frac{d}{dt} H(f|M) \geq D(f) \geq c\, H(f|M_f) \\ \frac{d^2}{dt^2} H(f|M_f) \geq C_+ H(f|M) - C_- H(f|M_f) \end{cases}$$

<u>Example 2</u> : the Boltzmann equation in incompressible viscous regime

$$\begin{cases} \varepsilon \partial_t f + v \cdot \nabla_x f = \frac{1}{\varepsilon} Q(f) \\ f^0 = M(1 + \varepsilon g) \end{cases}$$

Knudsen ε = Mach ε

- close to local thermodynamic equilibrium $f \sim M_f$
- macroscopic (slow) dynamics governed by the Navier-Stokes equations
- dissipation coming from a weak coupling between the relaxation and the transport (reminiscent from hypocoercivity)

→ Hypocoercive mechanism can be described by a multiple scale expansion

$$\begin{cases} g = \left(\rho + u.v + \theta \frac{|v|^2-d}{2}\right) + \varepsilon g_\perp \\ g_\perp = \mathcal{L}^{-1} \Pi^\perp \left(v.\nabla_x \left(\rho + u.v + \theta \frac{|v|^2-d}{2}\right)\right) \end{cases}$$

Second order terms appear in the evolution equations of ρ, u, θ.

→ Modulated entropy dissipation $\frac{1}{\varepsilon^4} D(f \mid g_\perp)$ encodes this process in the stability inequality.

When the relaxation is strong enough (in the sense that it can be quantified by a functional), one can have entropic stability without being really close to equilibrium.

Departures from equilibrium can be tracked by analytical tools (system of ODEs, multiscale analysis)

There is still some rigidity in the system.

4. Out of equilibrium systems. instabilities and mixing

Example 1 : system of hard spheres in low density regime $N\varepsilon^{d-1} = 1$

$$\begin{cases} \partial_t f_N + \sum_i v_i \cdot \nabla_{x_i} f_N = 0 \quad \text{on } |x_i - x_j| > \varepsilon \quad (+\text{specular reflection}) \\ f_N^0 = \bigotimes_{i=1}^{N} f^0(x_i, v_i) \end{cases}$$

- weakly close to weakly chaotic distribution $f_N^{(k)}(t) \sim \bigotimes_{i=1}^{k} f(t, x_i, v_i)$
- mean (slow) dynamics governed by the Boltzmann equation
- loss of information encoded in the correlations at very small scale

---

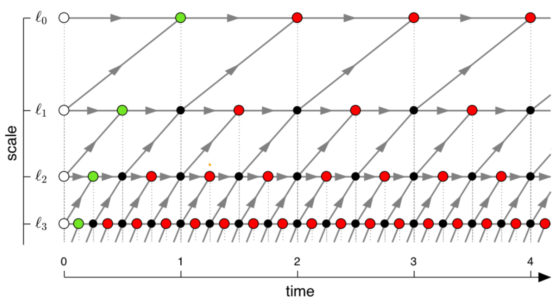
→ Because microscopic trajectories are unstable (in the limit $\varepsilon \to 0$) the system has no rigidity : it is impossible to keep track of all the information.

→ Microscopic uncertainty is transferred in macroscopic randomness. A major issue is to understand whether collisions become all independent.

→ No known suitable functional framework to encode the mixing responsible for the deccorelation.

Example 2 : a toy model based on Arnold's cat map.

$$\begin{cases} u_N(t+2^{-n}, n) = A\, u_N(t,n) + A\, u_N(t, n+1) \pmod 1 \quad n \leq N \\ \quad \text{with } A(x,y) = (2x+y,\ x+y) \pmod 1 \\ u_N(0,n) = u(0,n)\, 1_{n<N} + \xi\, 1_{n=N} \end{cases}$$



- weakly close to a random field
- loss of information encoded in the initial data at very small scale

---

→ The system has no rigidity in the sense that there is no deterministic limit as $N \to \infty$.

→ Microscopic uncertainty is transferred in macroscopic randomness. The limiting law does not depend on the law of ξ. (intrinsic / spontaneous stochasticity)

→ The proof of mixing relies on explicit computations in Fourier, specific to this example.

Spontaneous stochasticity results from the combination of
- instabilities at arbitrary small scales
- stochastic regularization at vanishing scale

Instabilities are expected to create mixing in the phase space which is a weak form of relaxation.

We do not know how to use entropy to encode this weak relaxation.