Vol. 30 (1991) REPORTS ON MATHEMATICAL PHYSICS No. 1

THE MANY FACETS OF ENTROPY

A. WEHRL

Institut für Theoretische Physik, Universität Wien, Baltzmanngane 5, A-1090, Wien, Austria

(*Received December 31, 1990*)

Several notions of entropy are discussed: classical entropies (Boltzmann, Gibbs, Shannon, quantum-mechanical entropy, skew entropy, among other notions as well as classical and quantum-mechanical dynamical entropies.

One should not mean that "entropy" is a uniquely determined quantity. In fact, there exist various concepts that — quite often — refer to different logical levels (and have different merits) called "entropy".

The first notion is due to Clausius (1850) in connection with the Second Law of Thermodynamics (which, at about the same time and apparently independently, had been discovered by Thomson (subsequently Lord Kelvin); both formulations are equivalent). The expression "entropy" itself is an artificial word, it stems from τροπή (changes). (Clausius, 1865).

Yet at that time, Clausius too and a little bit later on, Maxwell, founded the Kinetic Theory (of gases, 1860). A first climax was reached, when Boltzmann (1872) enunciated his celebrated transport equation.

Here we meet another concept of entropy, the *Boltzmann entropy*. It is formulated in terms of the one-particle distribution function (first correlation function) $F(x, v, t)$ $(x, v \in \mathbf{R}^3)$ which is defined on the one-particle space (μ-space, kinetic = mesoscopic level). Roughly speaking, $F(x, v, t) d^3x d^3v$ is the number of particles at time t in the (6-dimensional) volume $d^3x d^3v$, centred around x (position) and v (velocity). The Boltzmann entropy, denoted by the same letter S as the thermodynamic entropy, is

$$S := -k \int F \ln F \, d^3x \, d^3v \, .$$

(This is not quite the original version of Boltzmann. k is Boltzmann's constant, an expression due to Planck, 1902).

Whereas the *thermodynamic entropy* refers to equilibrium states, the Boltzmann entropy is also defined in the non-equilibrium situation. In his 1872 article, Boltzmann was able to give a precise meaning to the phenomenon of irreversibility, too, when he proved the H-theorem: the functional

$$H = \int F \ln F \, d^3x \, d^3v \left(= -\frac{S}{k} \right)$$

is never increasing.

In 1877, Boltzmann discovered the relation between entropy and probability (or the number of states, or complexions, resp.), which is best known in Planck's formulation $S = k \log W$. This is more or less the basis for most future considerations about entropy.

Einstein (1907) coined the word "Boltzmann's Principle" for Boltzmann's basic findings.

J. W. Gibbs, in 1902 (but there exists earlier work) introduced the notion of *Gibbs entropy* (of course not called that by him, but the "average index of the probability-in-phase"). Maybe some explanation is needed at this stage: the Gibbs entropy refers to all phase space (Γ-space, microscopic level) of a system with finitely many degrees of freedom. It is defined — along the lines of Boltzmann — by

$$S(\rho) = -k \int \rho \ln \rho \, d^3x_1 \ldots d^3v_N$$

(N = number of particles — not configurations, ρ = probability on phase space; of course, in contradistinction to the distribution function, which is normalized to N, normalized to 1).

The arguments that led Gibbs to his expression are pretty much as Boltzmann's, nevertheless, there are considerable differences. In a Hamiltonian system (with a compact phase space — which is actually not the most general requirement), the Gibbs entropy always remains constant in time (since, by Liouville's theorem, the Lebesgue measure on the energy shell is invariant under the Hamiltonian flow).

How, after all, can then irreversibility come in? One proposal was made by the Ehrenfests (1911): "coarse-graining". Divide the phase space into "cells" (or "grains") of equal size (after all, later on it turned out that the right size should be $h^{3N} - h$ = Planck's constant). The probability distribution ρ should evolve according to the Poisson equation, but not really, because it is assumed that in every (infinitesimal) step it is averaged over every single cell. This is another version of "molecular chaos", which had already been used by Boltzmann in his derivation of his equation — better known as "Stoßzahlansatz". In any case, for finite systems a perpetual *assumption* on disorder is needed to obtain an irreversible behaviour. (This assumption cannot be proven, it contradicts the fundamental equations of motion. Just in infinite systems only, there are — partly — rigorous results known. Cf., among others, Lanford, 1975, 1976).

The *Shannon entropy* (first called measure of information — v. Neumann made the suggestion to call it "entropy") — was proposed by Shannon (see Shannon and Weaver, 1949). It refers to sequences of probabilities for the outcome of an experiment, or message, resp. $(p_1, p_2, \ldots)$ $(p_i \geq 0, \Sigma p_i = 1)$. In a message, the average lack of information is (Shannon's formula)

$$-\sum p_i \log_2 p_i \, .$$

(This formula was discovered independently by Wiener. As a matter of fact, it may be traced back as far as to Boltzmann, except for the numerical factor relating logarithms.) We shall postpone the discussion how Shannon was led to his formula because the arguments do not differ too much from v. Neumann's for the quantum-mechanical entropy.

From a *formal* point of view, the Boltzmann and the Gibbs entropy ("classical continuous case", CC) and the Shannon entropy ("classical discrete case", CD) fit in the same frame, the *Baron–Jauch entropy* ("generalized Boltzmann–Gibbs–Shannon entropy"). Its definition is: let (Ω, Σ, μ) be a measure space, ν be a probability measure, being absolutely continuous w.r.t. μ (hence, at least in the separable case, the Radon–Nikodym derivative $d\nu/d\mu$ exists). Then the Baron–Jauch entropy is

$$S := -\int \frac{d\nu}{d\mu} \ln \frac{d\nu}{d\mu}\, d\mu$$

(provided that the integral exists which need not always be the case). Note, however, that in various respects these entropies may have quite different properties, depending on the kind of the measure μ. We will come to this point within short.

Quantum-mechanical entropy was introduced by v. Neumann (1927). It refers to a (generally) mixed state of a quantum-mechanical system, i.e. a density matrix ρ (a positive, a fortiori of course Hermitian operator with trace = 1).

$$S(\rho) := -k \mathrm{Tr}\, \rho \ln \rho$$

(one understands $0 \cdot \ln 0 \equiv 0$). In the following we shall always put Boltzmann's constant equal to 1.

The interpretation or at least one possible of the v. Neumann entropy is as follows: Every density matrix can be diagonalized:

$$\rho = \sum p_i |\varphi_i)(\varphi_i| \,,$$

where φ_i is the normalized eigenvector corresponding to the eigenvalues $p_i (|\varphi_i)(\varphi_i|$ = projection onto φ_i). Note that $p_i \geq 0$, $\sum p_i = 1$ and $S(\rho) = -\sum p_i \ln p_i$ $(0 \cdot \ln 0 \equiv 0)$. p_i is just the probability of finding the system under consideration in the pure state φ_i. If one performs N measurements, one will obtain as a result that (for large N) the system is found $p_i \cdot N$ times in the state $\varphi_1, p_2 \cdot N$ times in the stage φ_2 etc. There are $N!/(p_1 N)!(p_2 N)! \ldots$ possibilities for this. For $N \to \infty$ (by virtue of Stirling's formula), $1/N$ times the logarithm of these possibilities converges to S.

Cum grano salis, this argument applies to the Boltzmann, Gibbs and Shannon entropy (in the latter case up to a numerical factor) as well. However, the Boltzmann and Gibbs entropies depend on the cell size, changing it amounts to changing the value of the entropy by a numerical factor, i.e. in the classical (continuous) case only entropy *differences* can be "measured" — whatever is understood by this. (Instead of $d^3x d^3v$ or $d^3x_1 \ldots d^3v_N$, resp., one should rather integrate over

$$\frac{d^3x d^3v}{A}$$

or

$$\frac{d^3x_1 \ldots d^3v_N}{A^N},$$

resp., where A is on dimension (action)^3; its value cannot be deduced by means of classical arguments.)

The connection between the thermodynamic entropy and the quantum-mechanical (or Boltzmann-, or Gibbs-) entropy is conveyed by the *Maximum Entropy Principle*: given some macroscopic constraints, the equilibrium state is the one compatible with these constraints, that has the biggest (v. Neumann etc.) entropy.

Which properties of the four notions of entropy we have met up to now are different and which ones do they share?

In the quantum-mechanical case as well as in the CD case, entropy is always well-defined, its range is (in the infinite case) the extended interval $[0, \infty]$. (Actually, there are, in a certain sense, "more" density matrices, or discrete probability distributions, resp., with infinite entropy than such with finite entropy — the former ones are a set of second category in $\|\cdot\|_1$-, or l^1-sense, resp., the latter ones are of first category only. But in physics this does hardly matter.) In the CC case, the range of entropy is $[-\infty, \infty]$, but, in contrast to the aforementioned entropies, not only the sets with entropy $= +\infty$ or $-\infty$ are of second category, but even the set of all probability distributions with undefined value of entropy is of second category (so-to-say: what is $(+\infty) - (-\infty)$?).

Quite easy to perceive are *invariance properties*: the classical entropies are invariant under measure-preserving transformations, the quantum-mechanical entropy is invariant under unitary transformations.

An important property that is common to all entropies is *concavity*. If a probability distribution, or a density matrix, resp., is a convex combination of two other ones (i.e., speaking in physical terms, an incoherent mixture), mathematically described as

$$\rho = \lambda\rho_1 + (1-\lambda)\rho_2 \quad (0 \leq \lambda \leq 1),$$

then

$$S(\rho) \geq \lambda S(\rho_1) + (1-\lambda)S(\rho_2).$$

Whereas there is not the least problem to show this in the classical case (simply because the function $-x\ln x$ is concave), in quantum mechanics one has to work a bit harder, the reason being that ρ_1 and ρ_2 need not commute (Delbrück and Molière, 1937).

What is common too is *additivity*. By this is meant, in the classical case, that, if the (generalized) "phase space" is a Cartesian product of two other ones, $\Omega = \Omega_1 \times \Omega_2$ (to be more careful, one is concerned with a measure space $(\Omega_1 \times \Omega_2, \Sigma_1 \times \Sigma_2, \mu_1 \times \mu_2)$), and the probability distribution ρ factorizes, $\rho(w) = \rho_1(w_1)\rho_2(w_2)$, $(w \in \Omega,\ w_1 \in \Omega_1,\ w_2 \in \Omega_2)$, then $S(\rho) = S(\rho_1) + S(\rho_2)$. In quantum mechanics one considers a tensor product of two Hilbert spaces, $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2$, and a density matrix $\rho = \rho_1 \otimes \rho_2$ ($\rho_{1,2}$ being density matrices in $\mathcal{H}_1$, or $\mathcal{H}_2$, resp.) to obtain — in a very easy way — the same assertion.

If the probability distribution ρ in $\Omega_1 \times \Omega_2$, or the density matrix (denoted, in a little bit sloppy way, by the same letter) ρ is not a product, then the entropy is *subadditive* only, $S(\rho) \leq S(\rho_1) + S(\rho_2)$. Thereby, in the classical case, it is meant that $\rho_1 := \int \rho(w_1, w_2)\, dw_2$ etc., in the quantum case that ρ_1 is the partial trace of ρ w.r.t. $\mathcal{H}_2$, $\rho_1 := Tr_{\mathcal{H}_2}\rho$ etc.

It is remarkable that (in the CD case) permutation-invariance, or in the quantum case unitary invariance, and additivity and subadditivity characterize the Shannon, or the v. Neumann entropy, almost uniquely: every functional fulfilling these requirements is a linear combination of the Shannon, or the v. Neumann entropy and the (CD), or quantum-mechanical *Hartley entropy* S_0 (Hartley, 1928). It is defined as follows: in the CD case, S_0 is the logarithm of the number of p_i's that are different from zero, in the quantum case, S_0 is the logarithm of the number of eigenvalues $\neq 0$. (The afore-mentioned characterization theorem is — except for the more or less immediate quantum-mechanical generalization — due to Aczel, Forte, and Ng, 1974.)

The Hartley entropy (in the infinite case) is a very sensitive quantity. Although it is true that neither the Shannon nor the v. Neumann entropy are continuous (w.r.t the l^1, or trace norm, resp.) but lower semi-continuous only the continuity properties of the Hartley entropy are by far much worse. However, one could think of some at least formal resemblence to the topological entropy, to be discussed later.

A very interesting problem is that of *monotonicity*. In quantum mechanics, entropy is *not* monotonic, i.e. enlargening of a system can very easily lead to smaller entropies. (One may speculate to which extent this phenomenon has to do with Helmholtz's "Wärmetod", heat death, 1854.) Whereas this is hardly observed in "normal" matter — for reasons that are very difficult to explain — one knows of other situations where this is actually the case.

In this context it is worthwhile to make some remarks on the *relation between quantum-mechanical and CC entropy*. The basic laws of physics are quantum-mechanical; how can one arrive at a classical probability distribution starting from a density matrix? Several proposals have been made (Wigner, Blockhintzev). One possibility is to use coherent states (Wehrl, 1979), i.e. states with minimal uncertainty (thus one is as close as possible to the classical situation). Let, in one dimension for the sake of simplicity, $\varphi(p,q)$ be a coherent state centred at the (classical) values p and q, then one can construct a classical probability distribution $f(p,q)$ out of the density matrix ρ by

$$f(p,q) := (\varphi(p,q), \rho\varphi(p,q)) .$$

This distribution has some nice properties: it is positive, ≤ 1, continuous and, if one fits two spaces together, it *is* monotonic. (Note that of course the set of these distributions is not all of L_1^+.) By the way, the thus obtained classical entropy is $\geq$ than the quantum-mechanical entropy (and ≥ 1), which may be interpreted in the following manner: by passing to the classical limit, information is lost.

Both the v. Neumann (or the Shannon) entropy and the Hartley entropy belong to a family of entropies, which we shall describe for the quantum case only, since the CD case is entirely obvious — the (quantum-mechanical) *Renyi-entropies*, also often called *α-entropies*. Their definition is

$$S^\alpha(\rho) = \frac{1}{1-\alpha} \ln \operatorname{Tr} \rho^\alpha$$

for $0 < \alpha < \infty$, $\alpha \neq 1$; the limiting cases $S_0, S_1 \equiv S$ or S_∞, resp., ($\alpha \to 0$, $\alpha \searrow 1$, $\alpha \to \infty$) are the Hartley, or the v. Neumann entropy, or — less $\ln \|\rho\|$, resp. (Note that $S_\alpha(\rho)$ may be $= +\infty$ for $0 \leq \alpha \leq 1$, otherwise it is finite). Simple properties: $S_\alpha(\rho)$ is — for fixed ρ — decreasing and convex in α. (It can happen that $S(\rho)$ is finite, but $S_\alpha(\rho)$ "jumps" suddenly, i.e. $S_\alpha = \infty$ for all $\alpha < 1$). Just a remark: the case $\alpha = 2$ was favoured by Fano (1957) and Prigogine, among others, apparently because S_2 can — comparatively —

be computed in a simpler way, since there is no need to diagonalize the density matrix (or doing something equally difficult).

α-entropies are of course unitarily invariant, additive (as it is easy to see), but except for $\alpha = 0$ and $\alpha = 1$ not subadditive.

There are some other concepts that are somehow related to the α-entropies, for instance — $\ln f^{-1}(\mathrm{Tr}\,\rho f(g))$, f being an increasing (convex or concave) function,

Aczel-Daroczy entropies (1963) :

$$f^{-1}(\mathrm{Tr}\,\rho f(-\ln\rho))$$

Daroczy entropies:

$$\frac{1}{1-\alpha}(\mathrm{Tr}\,\rho^{\alpha} - 1)\,.$$

Only few practical applications are known.

The *quasi-entropies* $S_f(\rho) := \mathrm{Tr}\, f(\rho)$ (f being concave, $f'(0) = +\infty$) share with the v. Neumann entropy the property of being concave in ρ as well as continuity (plus metric, i.e. category) and algebraic properties. One can even show examples of non-linear evolution equations, where all S_f are increasing in time.)

Up to now all kinds of (quantum-mechanical versions of) entropies were unitarily invariant, the next quantity in *not*. It may be traced back as far as to Pauli (1928) in his work on the ergodic hypothesis, later on, Ingarden (beginning from 1962) was studying it (cf. also Wehrl, 1977): *A-entropy* (this expression is not universally accepted, sometimes it is also called the Ingarden-Urbanik entropy). It reminds a little bit of coarse-graining. Let $\mathcal{P}$ be a partition of unity, i.e. a family of (finite-dimensional) pair wise orthogonal projections (P_i) with $\sum P_i = 1$. The A-entropy of a density matrix ρ (w.r.t. $\mathcal{P}$) is the v. Neumann entropy of $\rho' := \sum \lambda_i P_i$, where $\lambda_i := \mathrm{Tr}\rho P_i/\mathrm{Tr}\, P_i$. It is plain that $S(\rho') \geq S(\rho)$. In considerations about irreversibility this concept is — at least partly — useful.

The more "skew" ρ and $\mathcal{P}$ lie, the bigger will be the difference between $S(\rho)$ and $S(\rho')$.

Concerning "skew": In 1963 Wigner and Yanase introduced a notion of entropy, called *skew entropy,* measuring the amount of non-commutativity of a density matrix ρ w.r.t. a fixed (bounded) observable K:

$$S(\rho, K) := \tfrac{1}{2}\mathrm{Tr}[\rho^{\frac{1}{2}}, K]^2\,.$$

They were able to show that this quantity — though obviously not unitarily invariant — is concave in ρ, i.e. that one of the most important properties of entropy holds. Dyson, later on (1067) proposed the generalization

$$S_p(\rho, K) := \tfrac{1}{2}\mathrm{Tr}([\rho^p, K][\rho^{1-p}, K])$$

($0 < p < 1$; the degenerate case $p \to 0$ reads as $\frac{1}{2}\mathrm{Tr}([\rho, K][\ln\rho, K])$). It is true that all expressions are concave in ρ, too, however, the proof was given later on only (Lieb, 1973). We will discuss within short why this fact became so important in the further deliberations on quantum-mechanical entropy. Before doing this, let us introduce another notion:

Relative entropy. It "compares" two density matrices, σ and ρ : $S(\sigma \mid \rho) := \operatorname{Tr} \rho(\ln \rho - - \ln \sigma)$ (Umegaki, 1062; Lindblad, 1974). The classical predecessors are (in the CC case) the "Kullback entropy" and (in the CD case) Renyi's "information gain".

Main properties: $S(\sigma \mid \rho) \geq 0$ ($= 0$ iff $\sigma = \rho$) and, what is most important, it is *jointly convex* in σ and ρ. This is by far not easy to prove, first Lieb (1973) succeeded. (At present, some seven different proofs are already known.)

The main ingredient in this proof, called *Lieb concavity*, can be formulated in various (equivalent) ways, two of them shall be mentioned:

(i) The mapping $(A, B) \to \operatorname{Tr} K A^p K^* B^q$ (K Hilbert–Schmidt, A and $B \geq 0$, $0 \leq p$, $q \leq 1$, $p + q \leq 1$) is jointly concave in A and B,

(ii) (Ando's formulation): The mapping $A \to A^p \otimes A^q$ ($A \geq 0$, p, q as before) is concave.

This yields the concavity of

$$\operatorname{Tr} B(\ln A - \ln B) = \frac{d}{dt} \operatorname{Tr} A^t B^{1-t} \,|_{t=0}$$

and thus the convexity of the relative entropy (at first in finite dimensions, but the transition to infinite dimensions merely demands standard techniques). The concavity of $S_p(\rho, K)$ in ρ is an immediate consequence, too.

This is also the key to the proof of *strong subadditivity* (Lieb and Ruskai, 1973). In the classical case this is not at all a stronger property than subadditivity (it is just a simple consequence only), in quantum mechanics, however, this was an open problem for a very long time. It refers to three Hilbert spaces. Let $\mathcal{H} = \mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \mathcal{H}_3$, ρ be a density matrix in $\mathcal{H}$ and $\rho_{12} := \operatorname{Tr}_{\mathcal{H}_3} \rho$, $\rho_2 := \operatorname{Tr}_{\mathcal{H}_1} \otimes \mathcal{H}_2 \rho$ etc. be the partial traces, then

$$S(\rho) + S(\rho_2) \leq S(\rho_{12}) + S(\rho_{23}) .$$

What is this good for? For instance, one can prove the existence of a mean entropy and the monotonicity of the entropy for translationally invariant quantum systems.

(*Remark*: One cannot proceed further to more than three Hilbert spaces, at least not in an obvious, "canonical" way.)

Quite different from all the previous notions of entropy are the *dynamical entropies.* These are functions of the *dynamics* rather than of a state. The best known, in the classical case, is the *Kolmogorov–Sinai* entropy (Kolmogorov, 1953, 1959; Sinai, 1961, 1965). Consider a partition $\mathcal{P} = (\Omega_i)$ of the "phase space" Ω (i.e. $\Omega_i \cap \Omega_j = \oslash$ for $i \neq j$, $\cup \Omega_i = \Omega$). The entropy of this partition is $S(\mathcal{P}) := -\Sigma \mu(\Omega_i) \ln \mu(\Omega_i)$. Given two partitions $\mathcal{P}_l$ and $\mathcal{P}_2$, ($\mathcal{P}_1 = (\Omega_i^{(1)})$, $\mathcal{P}_2 = (\Omega_2^{(2)})$), for the refinement $\mathcal{P}_1 \vee \mathcal{P}_2$, i.e. the family $(\Omega_i^{(1)} \cap \Omega_j^{(2)})$, $S(\mathcal{P}_1, \mathcal{P}_2) := -\Sigma_{i,j} \mu(\Omega_i^{(1)} \cap \Omega_j^{(k)}) \cdot \ln \mu(\Omega_i^{(1)} \cap \Omega_j^{(2)})$. Given a measure-preserving transformation (automorphism) $T : \Omega \to \Omega$, one can — in an obvious manner — define the partitions $T\mathcal{P}$, $T^2\mathcal{P}$, etc. and the entropy $S(\mathcal{P}, T\mathcal{P}, \ldots, T^{n-1}\mathcal{P})$. Due to subadditivity,

$$S(\mathcal{P}, T) := \lim_{n \to \infty} \frac{1}{n} S(\mathcal{P}, T\mathcal{P}, \ldots, T^{n-1}\mathcal{P})$$

exists (and in fact, because of strong subadditivity, equals $\lim[S(\mathcal{P}, T\mathcal{P}, \ldots, T^n\mathcal{P}) - S(\mathcal{P}, T\mathcal{P} \ldots, T^{n-1}\mathcal{P})]$).

The Kolmogorov–Sinai entropy $s(T)$ is $:= sup_{\mathcal{P}} s(\mathcal{P}, T)$. It is (even a conjugacy-) invariant. (Two transformations T_1 and T_2 are called "conjugate", if there exists a measure-algebra isomorphism, say, φ, such that $\varphi \circ T_2^{-1} = T_1^{-1} \circ \varphi$.)

Among other possible interpretations, one can think of the Kolmogorov–Sinai entropy as the measure of the speed of mixing of a dynamical system. (It would take too much time to discuss all the remarkable properties of this entropy at this place.)

Similar (and conjugacy invariants too) are the *sequence entropies* (also called A-entropies, Kouchnirenko, 1967): Let $A = (a_1, a_2, \ldots)$ be an increasing sequence of integers.

$$s_A(\mathcal{P}, T) := \limsup_n \frac{1}{n} S(T^{a_1}\mathcal{P}, \ldots, T^{a_n}\mathcal{P}), \quad s_A(T) := \sup_{\mathcal{P}} s_A(\mathcal{P}, T).$$

Coming back to the Kolmogorov–Sinai entropy, it is, "in general, far from being complete" (Walters, 1975). For instance, examples are known where two transformations are *not* conjugate but have the same entropy.

In infinite systems, one is almost everytime faced with the situation that the Kolmogorov–Sinai entropy is infinite (we are here thinking of the automorphism group qua time-translations). This can — at least to a certain extent — be remedied by introducing the *space-time-entropy* (Hudetz, 1988), which, vaguely speaking, is obtained by "adding" the spatial translation group to the time evolution, thus yielding a finite value for the entropy.

A different notion, not based on measure theory, is the *topological entropy*, which refers to compact topological (or metric) spaces. It is a topological conjugacy invariant. By this is meant: let X and Y be compact spaces and $T: X \to X$, or $S: Y \to Y$ be homeomorphisms. T is topologically equivalent to S if there exists a homeomorphism $\varphi: X \to Y$, such that $\varphi T = S\varphi$. (Remember the measure-theoretical definition of conjugacy).

Now the construction goes — described very cursory — as follows: Let α be an open cover of X. The join, $\alpha \vee \beta$, of two open covers consists of all intersections $A \cap B$ ($A \in \alpha$, $B \in \beta$). A refinement β of an open cover, denoted by $\alpha < \beta$, means that any set of β is a subset of a set in α.

The entropy of α, $H(\alpha)$, is defined as $\log N(\alpha)$ (conventionally the logarithm is to base 2), where $N(\alpha)$ is the number of sets in a finite subcover of α with smallest cardinality. Two properties:

(i) $\alpha < \beta \to H(\alpha) \leq H(\beta)$,
(ii) ("subadditivity") $H(\alpha \vee \beta) \leq H(\alpha) + H(\beta)$.

Let $T: X \to X$ be continuous (not necessarily an automorphism). $H(T^{-1}\alpha) \leq H(\alpha)$. One can proceed in more or less the same manner as we have already used to establish the existence of

$$\lim_{n\to\infty} \frac{1}{n} H(\alpha \vee T^{-1}\alpha \vee \ldots \vee T^{-(n-1)}\alpha) := h(T, \alpha),$$

the entropy of T relative to α. (Just use subadditivity; that one has to use the negative powers of T is only for technical reasons.)

Finally, the *topological entropy* of T is $h(T) := \sup_\alpha h(T, \alpha)$.

A very challenging problem was, for many years, to find the *quantum analogue of the Kolmogorov–Sinai entropy*. Several attempts failed, a few partial results had been obtained only, none of them was completely satisfactory.

The problem is that one cannot mimick the classical construction; this procedure fails at the very first step. If one starts with two (quantum-mechanical) partitions of unity, $\mathcal{P}_1$ and $\mathcal{P}_2$, then what should $\mathcal{P}_1 \vee \mathcal{P}_2$ mean? In general (because of non-commutativity), this will be tremendously large (even in very simple, seemingly trivial cases); thus the construction cannot be carried through.

The first major progress was achieved by Connes and Størmer (1975), when they studied the conjugacy problem for the (unique) hyperfinite II_1-factor. (Two automorphisms T_1 and T_2 of a v. Neumann algebra are called conjugate, if there exists a^*-automorphism σ, such that $T_2 = \sigma^{-1} \circ T_1 \circ \sigma$.) The final solution is due to Connes, Narnhofer, and Thirring (1987).

Their method is quite intricate and can, on this occasion, be sketched only. Since the "obvious way" merely leads to trivialities, one has to pursue another route. There are — so-to-say — two main ingredients entering in the proof. First, one considers the relative entropy, or the conditional entropy, resp., rather than the "normal" one. The advantage of this detour is that — in contrast to the v. Neumann entropy — the relative entropy has some monotonicity properties that can be exploited (though in a very sophisticated way). Secondly, one has to arrive — all in all — at a commutative situation ("Abelian model"). Just a few hints:

To begin with, let us consider the *Connes–Størmer entropy* for the hyperfinite II_1-factor. It is a functional $H(\mathcal{A}_1, \dots, \mathcal{A}_n)$ of *several* finite-dimensional v. Neumann algebras (this set is denoted by $\mathcal{F}$), symmetric in its arguments, which has nice monotonicity and subadditivity properties. These allow, once more, to perform an analogue of the Kolmogorov–Sinai construction: for an automorphism T, the limit

$$\lim_{n \to \infty} \frac{1}{n} H(\mathcal{A}, T\mathcal{A}, \dots, T^{n-1}\mathcal{A}) =: h(\mathcal{A}, T)$$

exists; the Connes–Størmer entropy is

$$h(T) =: \sup_{\mathcal{A} \in \mathcal{F}} h(\mathcal{A}, T)\,.$$

Unfortunately, the explicit expression for H is rather complicated and by no means (at least not at a first glance) suggestive. (Just a cryptic remark: the "conditional expectation" plays a prominent role.)

The *Connes–Narnhofer–Thirring* entropy refers to *C^*-algebras*. Let $\mathcal{A}$ be a C^*-algebra, ω be a state and γ_k be completely positive maps $\mathcal{A}_k \to \mathcal{A}$, where the $\mathcal{A}_k$ are finite-dimensional algebras. In order to define the entropy of $\gamma_1, \dots, \gamma_n$ (w.r.t. ω), $H_\omega(\gamma_1, \dots, \gamma_n)$ (in the Connes–Størmer-prescription there is of course no subscript ω), one needs the concept of an Abelian model. It consists of a mapping $P:\ \mathcal{A} \to \mathcal{B}$, $\mathcal{B}$ Abelian, and a state μ such that $\mu \circ P = \omega$. Due to this concept, it is possible (in a surely complicated manner, more complicated than in the Connes–Størmer case) to define this entropy. Then subadditivity guarantees, for an automorphism T and an invariant state ω (i.e. $\omega = \omega \circ T$) the existence of

$$h_\omega(\gamma, T) := \lim_{n\to\infty} H_\omega(\gamma, T \circ \gamma, \ldots, T^{n-1} \circ \gamma).$$

Let $\mathcal{P}$ be the family of all completely positive maps of finite-dimensional C^*-algebras into $\mathcal{A}$. Then the Connes–Narnhofer–Thirring entropy is $h_\omega(T) := \sup_{\gamma\in\mathcal{P}} h_\omega(\gamma, T)$.

Final remark: although the derivation of the quantum-mechanical Kolmogorov–Sinai entropy is difficult enough, one could even be still more ambitious, for instance to try to find an analogue of the Kolmogorov–Sinai entropy for automorphisms of lattices. Up to now (at least to my knowledge) most attempts failed, yielding trivial results only. (One should bear in mind that the structure of a lattice can be very complicated, by far more intricate than as it is the case in the quantum-mechanical situation.)

Acknowledgments

Best thanks to Thomas Hudetz for his advice.

REFERENCES

[1] Aczel, J., and Z. Daroczy: *C.R. Acad. Sci. Paris* **257** (1963), 1581.
[2] Aczel, J., B. Forte, and C.T. Ng: *Adv. Appl. Prob.* **6** (1974), 131.
[3] Ando, T.: *Topics on Operator Inequalities*, Sapporo University, Hokkaido, 1979.
[4] Baron, J.G., and J.J. Jauch: *Helv. Phys. Acta* **45** (1972), 20.
[5] Blokhintzev, D.I.: *J. Phys.* **2**, (1940), 71.
[6] Boltzmann, L.: *Wiener Ber.* **66** (1872), 275.
[7] Boltzmann, L.: *Wiener Ber.* **75** (1877), 67; *ibid.* **76** (1877), 373.
[8] Clausius, R.: *Pogg. Ann.* **368** (1850), 500.
[9] Clausius, R.: *Pogg. Ann.* **125** (1865), 390.
[10] Connes, A., and E. Størmer: *Acta Math.* **134** (1975), 289.
[11] Connes, A., H. Narnhofer, and W. Thirring: *Commun. Math. Phys.* **112** (1987), 691.
[12] Daroczy, Z.: *Inf. Control* **16** (1970), 36.
[13] Delbrück, M., and G. Molière: *Abhandl. Preuss. Akad.* **P** (1937), 1.
[14] Dyson, F. J.: *J. Math. Phys.* **8** (1967), 1538.
[15] Ehrenfest, P. and T.: *Encyklopädie der Mathematischen Wissenschaften* **IV** (6)(1911), Art. 32.
[16] Fano, U.: *Rev. Mod. Phys.* **29** (1957), 74.
[17] Gibbs, J.W.: *Elementary Principles in Statistical Mechanics*, Yale University, New Haven, Conn., (1902).
[18] Hartley, R.V.: *Bell Syst. Techn. J.* **7** (1928), 535.
[19] Hudetz, T.: *Lett. Math. Phys.* **16** (1988), 151.
[20] Ingarden, R.S.: *Fortschr. Phys.* **13** (1965), 755.
[21] Kolmogorov, A.N.: *Dokl. Akad. Nauk SSSR* **119** (1953), 861; *ibid.* **124** (1953), 754.
[22] Kouchnirenko, A.G.: *Funkz. Analys i ego Prilojenija Moscow* **1** (1967), 103.
[23] Lanford, O.E. in: *Dynamical Systems, Theory and Applications* (J. Moser, ed.), Springer, Berlin, 1975.
[24] Lieb, E.H.: *Adv. Math.* **11** (1973), 267.
[25] Lieb, E.H., and M.B. Ruskai: *Phys. Rev. Lett.* **30** (1973), 434; *J. Math. Phys.* **14** (1973), 1938.
[26] Lieb, E.H.: *Bull. Am. Math. Soc.* **81** (1975), 1.
[27] Lindblad, G.: *Commun. Math. Phys.* **39** (1974), 111.
[28] Maxwell, J.C.: *Philos. Mag.* **19** (1860), 19; *ibid.* **20** (1860), 21.
[29] v. Neumann, J.: *Gött. Nachr.* **1927**, 273.
[30] Pauli, W, in: *Festschrift zum 60. Geburtstage A. Sommerfelds* (P. Debye, ed.), Hirzel, Leipzig, 1928.

[31] Planck, L.: *Vorlesungen über die Theorie der Wärmestrahlung*, Barth, Leipzig, 1906.
[32] Renyi, A.: *Wahrscheinlichkeitsrechnung*, Deutscher Verlag der Wissenschaften, Berlin, 1966.
[33] Shannon, C., and W. Weaver: *The Mathematical Theory of Communication*, University of Illinois, Urbana.
[34] Sinai, Ya.G.: *Dokl. Akad. Nauk SSSR* **25** (1961), 899; *Usp. Math. Nauk.* **20** (1965), 232.
[35] Thomson, W. (Lord Kelvin): *Philos. Mag.* **4** (1852), 304.
[36] Thomson, W. (Lord Kelvin): *Proc. R. Soc. Edinb.* **3** (1857), 139.
[37] Umegaki, H.: *Kodai Math. Sem. Rep.* **14** (1962), 59.
[38] Walters, P.: *Ergodic Theory — Introductory Lectures*, Springer, Berlin–Heidelberg–New York, 1975.
[39] Wehrl, A.: *Rep. Math. Phys.* (1977).
[40] Wehrl, A.: *Rev. Mod. Phys.* **50** (1978), 221.
[41] Wehrl, A.: *Rep. Math. Phys.* **16** (1979), 353.
[42] Wigner, E.P.: *Phys. Rev.* **40** (1932), 749.
[43] Wigner, E.P., and M.M. Yanase: *Proc. Natl. Acad. Sci. USA* **49** (1963), 910.