On the Notions of Entropy and Information

Bo-Sture K. Skagerstam[1]

In this review article we look for some of the historical reasons for the "identification" of the information-theoretical and thermodynamic entropy concepts. We also discuss the Baron–Jauch entropy concept and explicitly show that, for classical systems in thermodynamic equilibrium, there exists a very simple connection between this general definition and the ordinary experimental entropy.

1. INTRODUCTION

R. Clausius introduced the concept of entropy in phenomenological thermo-dynamics when he was studying the so-called heat cycles. This was in 1865. Since then the *word* entropy has been used in many other fields, such as statistical mechanics, communication theory, probability theory, logic linguistics, and such purely mathematical fields as abstract analysis and number theory. In a recent, very interesting review article, Dutta expressed this fact in the following way[1]: "Successful applications of the notion of entropy in different fields have established the usefulness and generality of the concept, which was originally introduced as the basic concept of thermo-dynamics and had its applications mostly in some branches of physics, physical chemistry and mechanical engineering."

---

[1] Institute of Theoretical Physics, Fack, Göteborg, Sweden.

Now we may ask the following question: Are there any "deep" connections among the various applications of the entropy concept? Let us divide this question into two parts of special interest:

1. Can we give a sufficiently general definition of *entropy* that includes many of the interesting applications?
2. What is the connection between the entropy concept used in information theory and that used, e.g., in statistical mechanics?

Very recently Baron and Jauch discussed these two questions and they also constructed an entropy function that gives the standard formulas for entropy as defined by Boltzmann in statistical mechanics or by Shannon in information theory. Baron and Jauch concluded[2]:

> "As has been emphasized, the use of the word entropy should not lead to confusion of the mathematical concept defined here with the physical concept of thermodynamic entropy. The concept introduced here relates two measures, one of which is absolutely continuous with respect to the other and has at this stage of abstraction nothing to do with any particular physical system. The misleading use of the same name for mathematical and for physical entropy is well-entrenched: it is now unavoidable. We use it in the general sense: its special meanings should be clear from the context."

Let us now state their definition. We suppose that (X, Ω, p) and (X, Ω, ν) are two measure spaces, i.e., X is some underlying setstructure, Ω is a σ-algebra generated by subsets from X, and μ and ν are two measures defined on Ω. We also assume that $\mu < \nu$, i.e., μ is absolutely continuous with respect to ν. Then we know[3] that there exists a unique "Radon–Nikodym derivative" f in the function space $L^1(X, \Omega, \nu)$ such that

$$\mu(A) = \int_A f \, d\nu, \qquad \forall A \in \Omega \tag{1}$$

The Baron–Jauch entropy function $H(\mu, \nu)$ is then defined by

$$H(\mu, \nu) = \int_X f \ln(f) \, d\nu \tag{2}$$

It is then possible to show[2] that $H(\cdot, \cdot)$ has the following properties:

$$H(\mu, \mu) = 0 \tag{3}$$

$$H(\mu, \nu) \geqslant 0 \qquad \text{if} \quad \mu < \nu \tag{4}$$

$$H(\mu_1 \otimes \mu_2, \nu_1 \otimes \nu_2) = H(\mu_1, \nu_1) + H(\mu_2, \nu_2) \tag{5}$$

where

$$\mu_1 \otimes \mu_2(A_1 \otimes A_2) \equiv \mu_1(A_1) \cdot \mu_2(A_2), \qquad \forall A_1, A_2 \in \Omega$$

i.e., just the definition of a product measure on $X \otimes X$, the Cartesian product of two identical sets. Moreover, we have that

$$\mu_k(A) = \int_A f_k \, d\nu_k, \qquad \forall A \in \Omega, \quad k \in \{1, 2\}$$

where f_k, $k \in \{1, 2\}$ are unique Radon–Nikodym derivatives corresponding to the relations $\mu_1 < \nu_1$ and $\mu_2 < \nu_2$.

To be on safe grounds, it is now very interesting to note that if we define "entropy" by the conditions (3)–(5) then it can be shown,[4] under some minor technical assumptions, that the form (2) is the unique answer (up to some arbitrary positive constant). So from now on we talk about *the* Baron–Jauch (BJ) entropy as defined above. This can now be considered at least as a partial answer to question 1 above. It cannot be complete, however, since we do not know how to treat systems with infinitely many degrees of freedom. Moreover, we will restrict ourselves to classical systems where we know that there exists a joint probability distribution for all observables. From a conceptual point of view this is, however, not an essential loss of generality in the discussion. We should also notice that the problem of giving a precise meaning for entropy so that it can be applied to the description of irreversible processes in physics is perhaps, at present, not a well-defined question. This is so because we do not have any general (and accepted) method for treating such processes. As a matter of fact, Ingarden,[5] among many other authors, has stated that:

"Therefore, the principal question of irreversibility cannot be solved by any approximative method as the perturbation method, for instance. Also finite systems with a finite number of particles), even very large, cannot be approximated (for this purpose) by infinite systems, as is frequently done (cf. Ruelle 1969[6]) simply since the difference between these systems is infinite. Infinite systems are closed and open in the same time, both in the topological and the physical sense of the word (the latter meaning isolating or non-isolation of a system), and therefore they can have a Hamiltonian and being irreversible, which cannot be true for any finite system. Finite systems are irreversible only if they are open (non-isolated), i.e. *non-Hamiltonian*, anyway on the basis of the present quantum (and classical) mechanics: how it will be with future mechanics we cannot judge or even guess now."

It is probable that Kossakowski[7,8] has found a way of giving an axiomatic formulation of the concept of irreversible motions and, in general, a way of classifying the time evolution for general quantum systems. However, we will not do any thorough study of these questions in this paper, but refer to Ref. 8, where an extensive bibliography can be found.

2. ENTROPY AND INFORMATION

What about question 2 above? Baron and Jauch[2] were able to trace the "confusion" among physicists concerning that question back to a

"seemingly paradoxical thought experiment of Szilard ..." and they concluded that "... this experiment cannot be considered as a justification for such identification and there is no paradox." To see the historical steps behind the mentioned identification of Boltzmann's entropy concept in statistical mechanics and Shannon's in information theory, consider Fig. 1.[1,2,9-11] As we mentioned earlier, R. Clausius introduced the concept of

Entropy - Information.

R.CLAUSIUS   1865
Thermodynamics.

J.C.MAXWELL   1871
Violation of the second law.

L.BOTLZMANN & W.GIBBS   1872 & 1902
Statistical mechanics.

J.D.VAN DER WAALS   1911
Entropy,cause and effect.

B.VON SMOLUCHOWSKI   1914
Maxwell's demon.

E.SHANNON   1948
Information theory.

L.SZILARD   1929
Szilard's paradox.

G.N.LEWIS   1930
Entropy and information.

L.BRILLOUIN   1951
Maxwell's demon.

J.VON NEUMANN 1932
The measurement problem in
quantum mechanics.

F.LONDON & E.BAUER   1939
The reduction of the wave
packet.

E.T.JAYNES 1957
Information
theory.

?
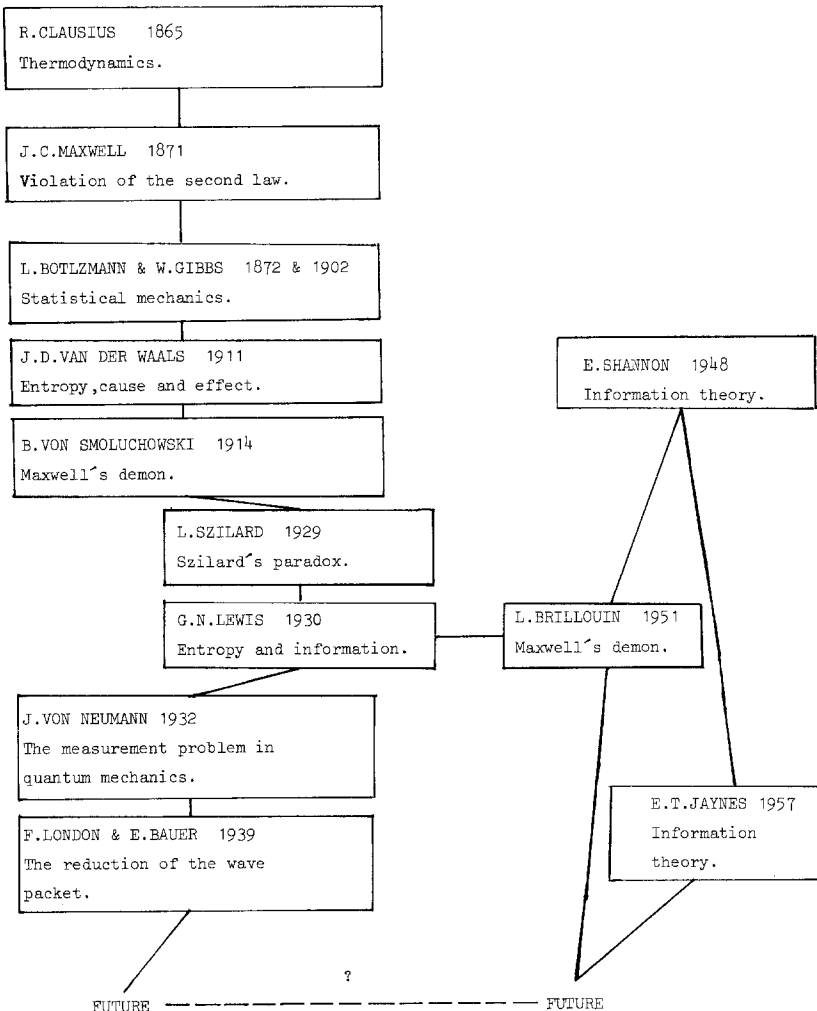FUTURE — — — — — — — — — — — FUTURE

Fig. 1

entropy in his study on heat cycles. In thermodynamics we know this today in the following form: The entropy change S_C from a thermodynamic state A to a state B is given by

$$\Delta S_C \stackrel{\text{def}}{=} S_B - S_A = \int_A^B \delta Q / T \tag{6}$$

where S_A and S_B are the entropies for the states A and B and δQ is the heat absorbed by the system under consideration at temperature T. The path of integration is such that each infinitesimal process (along the path) is reversible. The temperature is then essentially an integrating factor making the differential form δQ exact.[12] We can then express the *entropy principle* or what is more generally known as *the second law of thermodynamics* in the following well-known way: The entropy for an isolated system can never decrease, i.e.,

$$\Delta S_C \geqslant 0 \tag{7}$$

In 1871 J. C. Maxwell introduced the well-known "Maxwell's demon," a hypothetical "being" acting as a violator of the second law.[11] In trying to find a microscopic interpretation of the entropy concept L. Boltzmann introduced the following formula for the entropy[2]:

$$S_B \stackrel{\text{def}}{=} k \ln(W) \tag{8}$$

where k is Boltzmann's constant and W represents some probability distribution. This expression is perhaps one of the most common formulas for entropy used in modern textbooks on classical and/or quantum statistical mechanics. For practical use of the expression (8) one has the following scheme[2]:

1. For the probability distribution W one assumes an a priori given expression. For classical systems this is, e.g., given by some convenient choice of measure on the phase space.
2. One then imposes constraints on W that are in agreement with a fixed number of external parameters defining the thermodynamic state under consideration.
3. On the basis of the assumed a priori probability distribution one finds an explicit expression for W.
4. Assuming that the equilibrium states are characterized by a maximal entropy, one finds the corresponding probability distribution by calculating the extreme values for W under the assumed constraints on the system.

From standard textbook arguments one then obtains, for finite quantum systems, the Bose–Einstein and the Fermi–Dirac distribution formulas and

from these the classical distribution of Boltzmann.[13] In this context it is now interesting to note that the expression (8) can be deduced from (6) using some reasonable interpretation, in a quantum mechanical sense, of work and heat.[14] The reverse process could also, of course, be done by some proper identification of parameters.

In 1872 Boltzmann introduced his well-known H_B-function defined by[15]

$$H_B \overset{\text{def}}{=} \int_\Gamma \rho \ln(\rho) \, d\Omega \tag{9}$$

where ρ is a probability measure on the classical phase space Γ and $d\Omega$ is the ordinary Lebesgue measure on Γ, i.e., we have that

$$\rho(A) \geqslant 0 \quad \forall A \in \Gamma \qquad \text{and} \qquad \int_\Gamma \rho(\Omega) \, d\Omega = 1 \tag{10}$$

Using the assumption of "molecular chaos" or what is also called the "Stosszahlansatz" (Ref. 13, §1.4), Boltzmann was able to show that H_B, as a function of time, never increases, i.e.,

$$dH_B(t)/dt \leq 0 \tag{11}$$

which is the celebrated H-theorem of Boltzmann. This means that (9) is a good candidate for the entropy, because it really expresses the fact stated in the second law of thermodynamics. We also note that if we have systems with many particles [so that $n! \cong n^n \exp(-n)$ is valid], it is possible to show that H_B is related to the expressions (6) and (8) in the following way:

$$-kH_B = S_C = k \ln(W) \tag{12}$$

at least in the case of gaseous systems with noninteracting particles.[13] Under the constraint of constant energy E, i.e.,

$$E = \text{const} = \int_\Gamma \epsilon(\Omega) \rho \, d\Omega \tag{13}$$

it is straightforward to prove that the maximum value for the entropy S (where $S = -kH_B$) is attained when ρ is given by the canonical distribution of Gibbs, namely

$$\rho = \exp[(F - \epsilon/kT] \tag{14}$$

where F, k, and T are constants with an immediate physical interpretation.

Let us now investigate the generality of expression (11). We assume that we have some "adiabatic" physical system, which is not in an equilibrium state and which is such that at time $t = 0$ the thermodynamic state is described by an ensemble for which the distribution $\rho(\Gamma)$ has support only for

$$E < E_{\text{tot}} < E + \delta E \tag{15}$$

In (15) E_{tot} is the total energy for the system under consideration and δE some positive number. We now make a Riemann approximation of the entropy function $S(t)$ in the sense that

$$S(t) \stackrel{\text{def}}{=} -kH_{\text{B}} = -k \lim_{\Delta\Omega_i \to 0} \sum_i \rho(\Omega_i; t) \ln[\rho(\Omega_i; t)] \, \Delta\Omega_i \qquad (16)$$

But, by the Liouville theorem, we have that

$$d\rho(\Omega; t)/dt = 0 \qquad (17)$$

from which we can conclude that the expressions $\{\rho(\Delta\Omega_i)\}$ are constant in time if the subsets $\{\Delta\Omega_i\}$ are properly chosen. We conclude that

$$S(t) = S(0) \qquad (18)$$

Hence we can say that the Boltzmann H_{B}-function cannot be the proper candidate for describing what we call entropy, since experience tells us, as expressed by the second law of thermodynamics, that

$$S(t) \geqslant S(0) \qquad \text{and} \qquad \lim_{t \to \infty} S(t) \equiv \text{some equilibrium value} \qquad (19)$$

According to Gibbs, however, the Riemann sum in (16) has no meaning because the ergodic properties of the system have the effect of distributing the subsets $\{\Delta\Omega_i\}$ over the whole phase space Γ. Instead, the observer of the system is constrained to so-called "coarse-grained" experiments. The physical meaning of this is that it is impossible to "resolve" the phase-space volume elements to an arbitrary degree. Hence we cannot give a physical meaning to the limit procedure in (16). Gibbs suggested that instead we should "coarse-grain" the phase space, i.e., divide it into a number of finite (i.e., the measure of the coarse-grained parts is greater than zero but finite) subsets $\{\Omega_k\}$ such that (Ref. 13, §5.14)

$$\bigcup_k \Omega_k = \Gamma \qquad (20)$$

We can then define the probabilities $\{p_k\}$ in the following way:

$$p_k = (1/\Omega_k) \int_{\Omega_k} \rho \, d\Omega \qquad (21)$$

where we use the same notation for the set Ω_k and the measure of Ω_k. From this definition we immediately conclude that

$$\sum_k p_k \cdot \Omega_k = 1 \qquad (22)$$

We are then able to define the following coarse-grained probability density:

$$\rho_{\text{G}}(x) \stackrel{\text{def}}{=} p_k \qquad \text{if} \quad x \in \Omega_k, \quad \text{etc.} \qquad (23)$$

and the corresponding entropy S_G by

$$S_G \overset{\text{def}}{=} -k \int_\Gamma \rho_G \ln(\rho_G)\, d\Omega = -k \sum_k p_k \ln(p_k) \tag{24}$$

It is then a straightforward exercise to show that $S_G(t)$ is an increasing function in time and that its maximal value is given when the probabilities $\{p_k\}$ define a canonical distribution (when the mean energy is kept constant).

This was then a success for Gibbs, but of course he and Boltzmann were aware of the ambiguities in these definitions. Indeed Gibbs, for example, stated that[2]; "It is evident that there may be more than one quantity defined for finite values of degrees of freedom, which approach the same limiting form for infinitely many degrees of freedom. There may be, therefore, and there are, other quantities which may be thought to have some claim to be regarded as temperature and entropy with respect to systems of a finite number of freedoms."

We can now ask what sort of philosophy underlies the Gibbs, or, in general, any "coarse-grained" approach for describing essentially irreversible processes, which we indeed know exist. To let the actual physical processes be "dependent on the accidental shortcomings of the observer who makes the measurements on the system"[16] can be somewhat disconcerting. By referring to a spin-echo experiment of Hahn,[17] it is indeed claimed in Ref. 16 that the observers are not constrained to coarse-grained experiments and, "*It is therefore not permissible to base fundamental arguments in statistical mechanics on coarse-graining.*" This is essentially in the same spirit as expressed by the Toru'n school, as mentioned above.

But what about the origin of the irreversibility in physical processes? Some authors derive this property in statistical mechanics by letting the number of degrees of freedom tend to infinity or by some coarse-graining procedure as stated above. Others claim that it is necessary to use some stochastic or random concepts, as in Ref. 16. It has also been suggested that the origin of the irreversibility can be traced back to some fundamental processes in elementary particle physics (T-noninvariance) or to the structure of the universe on a global scale.[18] These questions will be the subject for another study and we only remark that there exist simple, isolated, finite, and reversible models where we can give a meaning to "irreversibility" by using ideas from information theory.[19,20]

Thus we see that perhaps there exists some deep connection between *observation* and *information* of a physical system and the standard concept of *entropy*. Indeed since the days of Maxwell there has been a continuous study of these questions. For example, in 1911 J. D. van der Waals speculated on the relation between entropy change and cause and effect.[10] In 1914, B. von Smoluchowski discussed the Maxwell's demon and concluded[2]:

"As far as our present knowledge is concerned there does not exist a permanently working automatic perpetuum mobile in spite of molecular fluctuations, but such a contraption could function if it were operated by intelligent beings in a convenient manner...."

L. Szilard then constructed a similar *Gedanken* experiment as that of Maxwell's demon, based essentially on the remark by Smoluchowski. From a methodological point of view something very interesting now happened, which we briefly discuss. G. N. Lewis concluded in 1930 that "... Gain in entropy means loss of information, nothing more...."[10] At that time, however, the word "information" had no quantitative meaning, so Lewis' statement was just an idea of a possible interpretation of entropy. In his famous book on quantum mechanics von Neumann[21] transferred the *Gedanken* experiment of Szilard's to the measurement problem in quantum mechanics. F. London and E. Bauer thus claimed in 1939 that the measurement process in quantum mechanics "is connected in an essential manner with the presence of a conscious observer who registers in his mind an effect, and that this conscious awareness is responsible for the oft-discussed, paradoxical 'reduction of the wave packet.'"[2] We mention this because it is very interesting to see that "old problems" in the theory of the measurement process are connected, at least historically, to the problem of giving entropy some "deeper" meaning on a microscopical level.

Lewis' "guess" about the connection between "information" and entropy did not get a more precise meaning until C. L. Shannon published his paper in 1948.[22] In this famous paper he introduced the concepts of "information-theoretical" *entropy* and *information* in essentially the following way. One introduces the concept of "discrete information source" as a discrete random process. Now suppose that p_k is the probability of the kth outcome of the random event. The entropy-information measure S_I is then defined in the following way:

$$S_I \overset{\text{def}}{=} -k' \sum_k p_k \ln(p_k) \tag{25}$$

where k' is some arbitrary, positive constant. The *information I* is then defined to be the difference of the information-theoretical entropy (25) for two given probability distributions $\{p_k\}$ and $\{p_k'\}$ for the random event, i.e.,

$$I \overset{\text{def}}{=} S_I(p_1,...,p_n) - S_I(p_1',...,p_n') \tag{26}$$

We see that (26) is similar to the definition of the "entropy for a finite scheme" in probability theory.[23] In this context it is now of some interest to ask why Shannon used the name "entropy" for the expression (25). Shannon himself answered this question: It was a suggestion from von Neumann, who in fact said,[10] "You should call it [(25)] entropy for two reasons. In the first place your uncertainty function has been used in

statistical mechanics under that name, so it already has a name. In the second place, and more important, no one knows what entropy really is, so in a debate you will always have the advantage." According to some authors one should take the "unification" of these two entirely different concepts (thermodynamic and information-theoretical entropy) seriously and, for example, L. Brillouin "resolved" the problem of Maxwell's demon in 1951 by using this identification.[10] In this same reference it is also stated that, "The proof that they are indeed the same (and not mere analogues) has been dealt with extensively elsewhere and will not be treated here" (Ref. 10, p. 181).

Baron and Jauch traced this identification back to the Szilard paradox and they concluded that there indeed exists no paradox. At least there is no clear reason from a purely historical point of view for such an identification. One should really observe that this way of connecting the thermodynamic entropy with Shannon's expression (25) in no way is generally accepted by physicists. Jaynes, who has tried to find another basis for such a discussion, states, for example[24]:

"... It is now amply demonstrated by many workers that the 'information measure' introduced by Shannon has special properties of consistency and uniqueness which makes it *the* correct measure of 'amount of uncertainty' in a probability distribution. This is, of course, the expression (25) which, for *some distributions* and *in some physical situations*, has been recognized as representing entropy. However, we have to emphasize that 'information-theory entropy' S_I and the experimental thermodynamic entropy S_E are entirely different concepts. Our job cannot be to *postulate* any relation between them, it is rather to *deduce* whatever relations we can from known mathematical and physical facts. Confusion about the relation between entropy and probability has been one of the main stumbling blocks in developing a general theory of irreversibility."

3. THE CANONICAL DISTRIBUTION

Finally let us now try to investigate the connection between information-theoretical entropy and the thermodynamic concept for some simple physical system in thermodynamic equilibrium. In doing this we will use some elements from Jaynes' work on information theory and statistical mechanics[24] and also the Baron–Jauch definition of entropy (2). First we make the following very reasonable assumption (Ref. 24, p. 196): The equilibrium thermodynamic properties of a system, as measured experimentally, agree with the results calculated by the usual methods of statistical mechanics, i.e., from the canonical or grand canonical ensemble appropriate to the problem.

We then coarse-grain the phase space for the system into some number of cells $\{\Omega_k\}$ such that (20) is fulfilled. For the a priori measure ν in the BJ definition of entropy, we put

$$\nu(\Omega_k) \stackrel{\text{def}}{=} g_k \qquad (27)$$

in such a way that

$$\sum_k g_k = \int_\Gamma d\nu = 1 \tag{28}$$

To calculate the probabilities $\{g_k\}$, we can use the "first principle" of statistical mechanics, which states that equal a priori probability is assigned to equal volumes of phase space. Now suppose that $\{p_k \equiv \text{prob}(E_k)\}$, where E_k is some characteristic energy for the phase space volume Ω_k, is some probability assignment such that it conforms to the given data, which in our case is taken to be the mean energy $\langle E \rangle$, i.e.,

$$\langle E \rangle \overset{\text{def}}{=} \sum_k p_k E_k \tag{29}$$

The probability distribution $\{p_k\}$ is associated with the "observation" and hence is an immediate candidate for the measure μ in the definition (2). We then have that

$$\mu(\Omega_k) \overset{\text{def}}{=} p_k$$

with the normalization condition

$$\sum_k p_k = \int_\Gamma d\mu = 1 \tag{30}$$

Moreover, it is very natural to assume that if $g_k = 0$, then we also have that $p_k = 0$. From this assumption we conclude that $\nu(\Omega) = 0 \Rightarrow \mu(\Omega) = 0$, i.e., $\mu < \nu$. It is then a straightforward exercise to calculate the Radon–Nikodym derivatives. One finds the result

$$(d\mu/d\nu)(\Omega_k) = p_k/g_k, \quad \forall g_k \neq 0, \quad \text{and zero otherwise} \tag{31}$$

The expression (2) for the BJ entropy then gives us that

$$H(\mu, \nu) = \sum_k p_k \ln(p_k) - \sum_k p_k \ln(g_k) \tag{32}$$

where we exclude terms where $g_k = 0$ from the last summation. We now put $k' = 1$ in Shannon's definition of entropy (25), which implies that

$$H(\mu, \nu) = -\sum_k p_k \ln(g_k) - S_I \tag{32a}$$

To connect this with the usual experimental entropy S_E we use the following theorem of Jaynes' (Ref. 24, p. 197):

Theorem. Let $\{p_k \equiv \text{prob}(E_k)\}$ be any probability assignment that conforms to the data in the sense that

$$\langle E \rangle \overset{\text{def}}{=} \sum_k p_k E_k \tag{33}$$

is the measured energy. Let

$$S_I \overset{\text{def}}{=} - \sum_k p_k \ln(p_k) \tag{34}$$

be the corresponding information entropy, and S_E be the experimentally measured entropy for the system. The arbitrary additive constant in the definition of S_E is chosen so that at zero temperature $S_E = \ln(n)$, where n is the degeneracy of the ground state. Let S_E be expressed in units such that Boltzmann's constant $k \equiv 1$. Then

$$S_I \leqslant S_E \tag{35}$$

with equality if and only if $\{p_k\}$ is chosen as the canonical distribution, i.e.,

$$p_k = (1/Z) \exp[-\beta E_k] \tag{36}$$

where β is a Lagrangian multiple and Z is a normalization factor.

Hence we conclude that in the case of a canonical distribution we have that

$$S_E = -\langle \ln(g_k) \rangle - H(\mu, \nu) \tag{37}$$

From the theorem above we also have that

$$H(\mu, \nu) \geqslant -\langle \ln(g_k) \rangle - S_E \tag{38}$$

Let us now investigate the extreme values for the expression (32) under the condition (33). Introducing the Lagrangian multiplier β', one finds that

$$\sum_k [\ln(p_k) - \ln(g_k) + \beta' E_k + \ln(Z)] \, \delta p_k = 0$$

i.e.,

$$p_k = (1/Z) g_k \exp(-\beta' E_k) \tag{39}$$

where Z is a normalization factor defined by

$$Z \overset{\text{def}}{=} \int_\Gamma d\mu = \sum_k g_k \exp(-\beta' E_k) \tag{40}$$

This is essentially of the same form as (39) but is somewhat more general since this canonical distribution includes the probabilities $\{g_k\}$. Using the expressions (39), we find, using (32), the following expression for the extreme value of the BJ entropy:

$$H_E(\mu, \nu) = - \sum_k \beta' p_k E_k - \ln(Z) \tag{41}$$

which up to some constant is equal to the experimental entropy S_E. This

gives a connection between the entropy concept introduced by Baron and Jauch and that used in thermodynamics.

4. CONCLUSIONS

In this review article we have looked for some of the historical reasons for the "identification" of the information-theoretical and thermodynamic entropy concepts. We have seen that in general they are different both as fundamental concepts and quantitative instruments for analyzing the properties of physical systems.

We have also discussed the Baron–Jauch definition of a general entropy concept and explicitly shown that for classical systems in thermodynamic equilibrium there is a very simple connection with the thermodynamic and experimental entropy. Moreover, we see from the arguments above that it is a trivial exercise to extend the formalism to quantum systems with a finite number of particles. One then obtains the standard expressions for the Einstein–Bose and the Fermi–Dirac distributions.

For infinite classical systems the use of the BJ definition of entropy turns out to be more complicated. In a recent study the present author has tried to find a general scheme for a discussion of "irreversible" processes using contractive semigroups to represent the time evolution. We have found that the HJ function can describe the characteristic behavior for "entropy," as expressed by the second law of thermodynamics, if one includes contractive parts in the generator for the time evolution.[25] Whether this has an interesting extension to the quantum dynamics for infinite systems is a question which we will discuss elsewhere.

REFERENCES

1. M. Dutta, A hundred years of entropy, *Physics Today* **1968** (January).
2. J. G. Baron and J. M. Jauch, Entropy, information and Szilard's paradox, *Helv. Phys. Acta* **45**:220 (1972).
3. W. Rudin, *Real and Complex Analysis*, McGraw-Hill, New York (1970).
4. B. Skagerstam, On the mathematical definition of entropy, *Z. Naturforsch.* **29a**(9) (1974).
5. R. S. Ingarden, Generalized irreversible thermodynamics and its applications to lasers, *Acta Phys. Polon.* **43**:3 (1973).
6. D. Ruelle, *Statistical Mechanics*, W. A. Benjamin, New York (1969).
7. A. Kossakowski, On quantum statistical mechanics of non-Hamiltonian systems, *Rep. Math. Phys.* **3**(4):247 (1972).
8. R. S. Ingarden and A. Kossakowski, On the connection of non-equilibrium information thermodynamics with non-Hamiltonian quantum mechanics of open systems, *Ann. Phys. (N.Y.)* **89**:451 (1975).
9. J. L. Lebowitz and O. Penrose, Modern ergodic theory, *Physics Today* **1973** (February).

10. M. Tribus and E. C. McIrvine, Energy and information, *Sci. Am.* **1969** (September).
11. W. Ehrenberg, Maxwell's demon, *Sci. Am.* **1967** (November).
12. J. M. Jauch, On a new foundation of equilibrium thermodynamics, *Found Phys.* 2(4):271 (1972); see also R. Hermann, *Lectures in Mathematical Physics*, Vol. II. Chapter VI, W. A. Benjamin, New York (1972).
13. D. ter Haar, *Elements of Thermostatistics*, Holt, Rinehard and Winston, New York (1966).
14. D. H. Frisch, The microscopic interpretation of entropy, *Am. J. Phys.* **34**:1171 (1966).
15. L. Boltzmann, *Wien Ber.* **66**:275 (1872).
16. J. M. Blatt, An alternative approach to the ergodic problem, *Prog. Theor. Phys.* **22**:745 (1959).
17. E. L. Hahn, *Phys. Rev.* **80**:580 (1950).
18. B. Gal-Or, *Modern Developments in Thermodynamics*, Israel University Press (1974).
19. A. Hobson, Irreversibility in simple systems, *Am. J. Phys.* **34**:411 (1966).
20. H. Grad, The many faces of entropy, *Commun. Pure Appl. Math.* **14**:323 (1961).
21. J. von Neumann, *Die Matematischen Grundlagen der Quantenmechanik*, Springer-Verlag, Berlin (1932).
22. C. L. Shannon, The mathematical theory of communication, reprinted in *The Mathematical Theory of Communication*, E. Shannon and W. Weaver (eds.), Urbana, Illinois (1949).
23. A. I. Khinchin, *Mathematical Foundations of Information Theory*, Dover, New York (1957).
24. E. T. Jaynes, Information theory and statistical mechanics, in *Statistical Physics*, 1962 Brandeis Lectures, Vol. 3, K. W. Ford (ed.), W. A. Benjamin, New York (1963).
25. B. Skagerstam, Time-evolution induced by contractive semi-groups, Institute Report No. 74–20, Gothenburg (1974).