1974-2

Helvetica Physica Acta
Vol. 47, 1974.

Birkhäuser Verlag Basel

Information Theory and Thermodynamics

by **Olivier Costa de Beauregard**

Laboratoire de Physique Theorique associé an CNRS, Institut Henri Poincaré, 11 rue Pierre et Marie Curie, 75005 Paris, France

and **Myron Tribus**

Xerox Corporation, Webster, New York, USA

(19. XI. 73)

Abstract. In answer to a recent article by Jauch and Baron bearing this same title, the information theory approach to thermodynamics is here upheld. After a brief historical survey and an outline of the derivation as formulated by one of us in previous publications, Jauch and Baron's critique of Szilard's argument pertaining to the 'well-informed heat engine' is discussed.

I. Introduction

In a recent paper Jauch and Baron [1] argue against the identification of the thermodynamic entropy concept as defined by Clausius and the information theory entropy concept as defined by Shannon [2]. In support of their thesis they present a new discussion of Szilard's thought experiment of the 'well-informed heat engine' [4].

We definitely belong to the other school of thought and intend to say briefly why.

While admitting with others [5] that a general mathematical definition of the entropy concept can be produced and used in many fields, Jauch and Baron do not mention the fact that the entropy concept of statistical mechanics has actually been deduced from the information concept. The corresponding papers are not quoted in their article, where Brillouin (1957) [4] is the only author mentioned as asserting the identity of the Clausius and the Shannon entropy concepts.

Jaynes (1957) [6], much inspired [7] by an article of Cox (1946) [8], later expanded in a book [9], is the recognized author of the derivation of the fundamental equations of equilibrium statistical mechanics (or 'thermostatistics' in the terminology of one of us [10]) from the inductive reasoning probability concept as introduced by Bayes [11], Laplace [12] and Hume [13]. In his two pioneering articles [6, 7] Jaynes presents his deduction in terms of, respectively, classical and quantal statistical mechanics. In the latter case he uses of course von Neumann's [14, 15] density matrix.

What is perhaps less known is that the same line of reasoning had been lucidly presented and used, also in von Neumann's quantum statistical formalism, as early as 1937 by Elsasser [16], who quotes Fisher (1929) [17] as one of his inspirers.

One point of interest in the Elsasser–Jaynes quantum information formalism is that the density matrix, and corresponding negentropy or information attached to the system under study, are calculated from the results of a set of 'macroscopic'

measurements simultaneously performed, these being interpreted as the (quantum mechanical) mean values $\langle F_i \rangle$ attached to *not necessarilly commuting* operators F_i.

For the sake of completeness we mention that the history of the inductive reasoning probability concept continues beyond Fisher, Cox and Shannon; Watanabe [18], Carnap [19], Kemeny [20], Jeffreys [21] and others can be added to the list.

In statistical physics Szilard is not the only forerunner of Elsasser, Brillouin and Jaynes. Lewis (1930) [22], in a paper devoted to time symmetry, has the sentence 'Gain in entropy always means loss of information, and nothing more. It is a subjective concept.' Van der Waals (1911) [23] derives the time asymmetry in the H-theorem from the classical time asymmetric use of Bayes' conditional probabilities, an idea also expressed by Gibbs [24] in an often quoted sentence. Jaynes has followed up his pioneering articles by more comprehensive publications [25], as has also one of us [26, 27]. Among other authors using the information theoretical approach in probability theory or in statistical mechanics we quote Kinchin [28], Yaglom and Yaglom [29], Katz [30], Hobson [31], and Baierlein [32].

In Section II we outline the information-theoretical derivation of the laws of equilibrium in statistical mechanics.

As for the more special topic of Maxwell's demon and Szilard's 'well-informed heat engine' we have brief comments in Section III with a reference to Brillouin.

II. Outline of the Information Theory Basis for Physical Theory

Part of Cox's contribution may be summarized as follows: suppose we wish to inform someone else of our *incomplete* knowledge of a subject. Is there a unique code which enables us to say neither more or less than we really know? Cox saw that instead of trying to *find* such a code, it would be necessary to *design* it. To design implies generation of alternative designs and the selection among them according to criteria. But what criteria should be used for the code? Cox chose criteria equivalent to the following:

1. Consistency
2. Freedom from Ambiguity
3. Universality
4. Honesty

What is surprising is that these criteria are necessary and sufficient to develop unique functional equations (27, Chapter I). For example, one of Cox's functional equations is:

$$\begin{aligned}[AB|E] &= F([A|BE], [B|E]) \\ &= F([B|AE], [A|E])\end{aligned} \tag{1}$$

F is a function to be determined, [ ] is a measure. Cox's solutions are the ordinary equations of mathematical probability theory:

$$p(A|E) + p(\sim A|E) = 1 \quad 0 < p < 1 \tag{2a}$$

$$p(AB|E) = p(A|BE)\, p(B|E) \tag{2b}$$

[A and B are propositions, $\sim A$ is the denial to A, E is the evidence and p is 'a numerical encoding of what the evidence E implies'.] Equations (2a) and (2b) can be used to develop all of probability calculus [27].

Cox's approach is unique in his deliberate attempt to *design a code* for the communication of partial information. The code is constrained to obey certain functional equations, which turn out to yield the equations of the calculus of probabilities. This result sheds new light on an old controversy; namely, the 'meaning' of the concept 'probability'. Because these equations are obtained by design, the interpretation of the function p is clear.

In the notation $p(\bullet|E)$, p is 'an encoding of knowledge about $\bullet$'. E represents the knowledge to be encoded. Only a measure which obeys the above two equations can be used as an encoding to satisfy the desired criteria.

The definition of p, due to Cox, is free of two limitations. On the one hand, it is not defined by reference to physical objects such as balls in urns or frequency of numbers in the toss of dice. On the other hand, it is not developed as an 'element on a measure space', devoid of all reference to anything but mathematical context. In no sense do we wish to minimize the importance of being able to put mathematical probability properly in perspective with respect to the rest of mathematics. But if we are to say what we 'mean' by 'probability', we must go beyond merely stating the mathematical properties of the function $p(\ )$.

The interpretation 'p is a numerical encoding of what the evidence E implies' is critical to all that follows.

If we take the Cox interpretation of p as fundamental and general, the question naturally arises: What are the rules for translating a statement E (normally made in a 'natural' language) into an assignment of a set of numbers represented by p? This question is the central task of statistics.

Jaynes' principle enters as a synthesis of Cox's result, just given, and Shannon's result in communication theory. If the knowledge E has been encoded as a set of p's, the measure S indicates how much is yet left to be learned.

$$S = -k \sum P_i \ln p_i \tag{3}$$

Proofs of the uniqueness and generality of S abound [2, 4, 28].

What interests us here is that Shannon's measure uniquely measures the *incompleteness* of the knowledge represented by E. If E were *complete*, the knowledge would be *deterministic*; i.e., would leave no residue of uncertainty. When E is deterministic the calculus reduces to sets of p's which are either 0 or 1 and the logic becomes purely deductive rather than inductive. For cases in which E is incomplete, Jaynes proposed, therefore, the principle of minimum prejudice as follows [6, 7]:

The minimally prejudiced assignment of probabilities is that which maximizes the entropy

$$S = -k \sum p_i \ln p_i \tag{4}$$

subject to the given information

The particularization of this principle for any field of inquiry depends upon the information to be encoded (i.e., upon E and the set A_i). In common parlance we say the uncertainty measure has to be applied to a well-defined question (which A_i is true?) and well-defined evidence (E). Statisticians say 'define the sample space'; thermodynamicists say 'define the system'. Both are saying the same thing, i.e. define the set A_i and be explicit about the experiment (E).

The entropy of Clausius becomes a special case of Shannon's entropy if we ask the right question. The question is put in the following form: suppose an observer

knows he is dealing with a system in a volume V which may be in a quantum state 'i' characterized by N_{ai} particles of type a, N_{bi} particles of type b, etc., and an energy ϵ_i. He knows his instruments are too crude to say precisely what ϵ_i, N_{ai}, N_{bi}, etc., truly are. All he can usefully observe are the repeatable measurements on ϵ_i, N_{ai}, N_{bi}, etc.

To apply Jaynes' principle requires the definition of the set of all possible answers. For the system postulated the question is therefore:

Q = 'In what quantum state is the system? (V is given)'

The set of possible answers is:

A_i = 'It is in the ith quantum state, for which the energy is ϵ_i, the number of particles of type a is N_{ai}, the number of particles type b is N_{bi}, etc.'

For illustration in this paper we shall confine our attention to systems in which electricity, magnetism and gravity play no part. The generalization to these phenomena has been given [6, 7, 26].

According to the Cox–Shannon–Jaynes development, the observer should 'encode' his knowledge in a probability assignment. We identify the 'repeatable measurements' with a mathematical 'expectation' for within the theory no other quantity can be so identified. The knowledge, E, therefore, is encoded by maximizing

$$S = -k \sum p_i \ln p_i \tag{4}$$

subject to

$$\sum p_i = 1 \tag{5}$$

$$\sum p_i \epsilon_i = \langle \epsilon \rangle \tag{6}$$

$$\sum p_i N_{ci} = \langle N_c \rangle \quad c = a, b, \ldots \tag{7}$$

where p_i is the probability that the system is in state i.

By the usual mathematical methods we find

$$p_i = \exp(-\Omega - \beta\epsilon_i - \alpha_a N_{ai} - \alpha_b N_{bi} - \ldots) \tag{8}$$

which is recognized as Gibbs Grand Canonical Distribution. The application of Jaynes' principle has served to introduce three new constructs, represented by Ω, β and the set $\{\alpha_c\}$.

The system of four equations above may be used to replace the set of probabilities and thereby exhibit a set of necessary and sufficient relations among the four constructs $S, \Omega, \beta, \{\alpha_c\}$.

$$S = \Omega + \beta\langle \epsilon \rangle + \sum_c \alpha_c \langle N_c \rangle \tag{9}$$

$$\Omega = \ln \sum_i \exp\left(-\beta\epsilon_i - \sum_c \alpha_c N_{ci}\right) \tag{10}$$

$$\partial\Omega/\partial\beta = -\langle \epsilon \rangle \tag{11}$$

$$\partial\Omega/\partial\alpha_c = -\langle N_c \rangle \tag{12}$$

Two things should be pointed out here. First, there has been no real use of physics. Thus far $\langle \epsilon \rangle$ and $\langle N_c \rangle$ have been defined only by identifying them with 'repeatable'

measurements of energy and composition without saying what 'energy' and 'composition' are. The mathematical results should be familiar to all who have studied statistical mechanics and their extension to more general cases should be obvious. Since we have not introduced or made use of any of the properties of energy or the particles, the results are all due to the rules of statistical inference inherent in the Cox–Shannon–Jaynes formulation. There is no reference to ensembles, heat baths, 'thermodynamic systems of which ours is an example drawn at random', etc.

This feature of being able to separate clearly which results are due to statistics and which are due to physics is a particular advantage in the information theory approach. The statistical quantities p_i, S, $\langle\epsilon\rangle$ and $\langle N_c\rangle$ generated Lagrange multipliers Ω, β, $\{\alpha_c\}$ *independent of the physical properties* associated with $\langle\epsilon\rangle$ and $\langle N_c\rangle$; the form of the four equations given comes only from the general procedure for inference laid down by Jaynes. That principle comes from logic; not reasoning about physical systems.

To describe the physical behavior of S, Ω, β and $\{\alpha_c\}$ we have to define the rules for changes in $\langle\epsilon\rangle$ and $\langle N_c\rangle$, i.e. put physics into the description. This has been done, for example, in Ref. [26].

The detailed derivation is given in the references. We quote only some general results to illustrate that the maximum entropy encoding of knowledge about the system of volume V to which we attach expectations $\langle\epsilon\rangle$ and $\langle N_c\rangle$ leads to the concepts associated with classical thermodynamics.

In this derivation the zeroth, first, second and third laws become *consequences*, not premises. Such a conclusion is indeed far reaching and it is no wonder the idea has been resisted for the dozen years since it was first put forward [10].

The properties of the Grand Canonical Distribution were first given by Gibbs who referred to his distributions as 'analogues' [24]. Denbigh [33], in his famous textbook on thermodynamics, also makes it quite clear that his statistical descriptions are *analogies* to classical thermodynamics. All workers who have dealt with statistical mechanics without basing their work squarely upon Shannon's information theory, as used by Jaynes and Cox, have either been silent on the connection between Clausius' entropy and the statistically defined entropy or been careful to disclaim any *necessary* connection between the two.

The important clue to understanding why the results given are more than a mere analogy is the recognition that we are actually defining that elusive state called 'equilibrium'. In this treatment, the encoding and all deductions from it, are valid only for those systems for which 'repeatable' measurements on ϵ_i and N_{ci} are possible and for which knowledge of $\langle\epsilon\rangle$ and $\langle N_c\rangle$ *are sufficient to define the macrostate of the system in volume* V. We can use the derivation to describe the mathematical properties of this state, which we call 'equilibrium', i.e. how changes occur on passage from one equilibrium state to another, how two or more systems interact, etc. These results are valid for systems which satisfy the premises of the theory, i.e. for which 'equilibrium' exists. It remains for experiment to decide if there exist states of physical systems for which the premises can be met. Such a situation is quite common in physics. For example, in mechanics it is postulated that in a 'Newtonian frame': force = mass × acceleration. The question of whether Newtonian frames exist or if a particular frame of reference is Newtonian is settled by seeing if all the deductions from 'force = mass × acceleration' are satisfied.

This circularity in all physical theories is usually glossed over in the education of physicists. It is the subject of very careful scrutiny in Norwood Hanson's illuminating

inquiry into theory building [34]. It is in a similar way that the general rules of inference (the maximum entropy encoding) are used to develop a description of 'equilibrium'. Mathematical relations are shown to be a consequence of the encoding process; behavior consistent with these equations proves that 'equilibrium' exists.

It is not generally understood that the concept of equilibrium is ill-defined in the literature of classical thermodynamics just as 'Newtonian frame' was glossed over in pre-relativity days. Attempts to define equilibrium, when they are made at all, are usually based on the idea of waiting for a long time. This idea of waiting a long time is not useful. Geological samples from the earth's interior have been found to be in thermodynamic disequilibrium. If substances that have been around for times comparable to the life of the earth are not in equilibrium, surely 'waiting' doesn't guarantee equilibrium. The proof of disequilibrium is the failure of the material to satisfy the phase rule. And the phase rule is, of course, a consequence of thermodynamics.

There is no way out of the dilemma that equilibrium is defined via thermodynamic constructs which constructs were in turn defined for the equilibrium state. We have no way of telling if a system is 'at equilibrium' except by making experiments which rely on the constructs which are defined by the theory of equilibrium. It is this dilemma which has inspired the comment 'there is no such thing as an immaculate perception'. What we see does depend on what we think.

This is not a special weakness of the information theory approach. It is inherent in the work of Gibbs, who defined equilibrium as the condition of maximum entropy.

The definition of 'equilibrium' thus given does not depend on the physics associated with ϵ_i and N_{ci}. Indeed, these symbols could stand for *anything* and the results would satisfy the desiderata. In references [10] and [27] it is demonstrated that the needed physics is obtained by introducing the following ideas:

1. The ϵ_i are additive, conserved.
2. The ϵ_i depend on i and the dimensions which determine V.

From these ideas and the previous equations we may derive all the known relations of classical thermodynamics.

Because the information theory's simplicity and freedom from such artificial constructs as ensembles, heat baths, etc., enables us to keep separate which results come from physics and which from statistics; important differences, often lost in non-information theory treatments are kept in the foreground. For example, irreversibility is traced to the difference between 'force' and 'expected (or equilibrium) force'. There is maintained a distinction between the principle of conservation of energy (a deterministic addition of energies of Newtonian systems) and the First Law of Thermodynamics (a statistical treatment of energy). The roles of $\beta(=1/RT)$ in diffusion of energy and of $\alpha(=-\mu/RT)$ in diffusion of particles are seen to be identical.

Heat is usually treated as a pre-existing idea to be 'explained' by physics. In the information theory treatment it is seen that someone who believes in the principle of conservation of energy (and who wishes to retain consistency in his ideas) *must invent* the concept of heat if he tries to make statistical descriptions. In this treatment entropy is taken as *primitive* and heat *derived* from the resulting statistical descriptions. Also, heat is introduced without the need to define it in terms of temperature or adiabatic walls (and not developed as if 'adiabatic' were understood before heat)!

The information theory treatment thus inverts the usual procedure, in which heat, temperature, and work are taken as primitive, and in which energy and entropy are derived. This break in tradition is hard to accept for those steeped in a tradition that

treats the laws of thermodynamics as though they were 'discovered' by experimentalists who knew instinctively about heat, temperature, work, equilibrium and equations of state. But the laws of physics are not 'discovered', they are 'invented' or – better yet – 'designed' according to criteria such as:

1. Consistency
2. Freedom from Ambiguity
3. Universality
4. Honesty

and are, therefore, constrained to the same results as produced simply, elegantly and directly by the information theory approach.

The information theory approach tells us *why* our ancestors *had* to invent temperature, heat, and reversible processes. It elucidates the 'paradoxes' of Szilard, Gibbs and many others. But most of all, it unifies our understanding of many phenomena. And it does so by showing that the entropy of thermodynamics is but a special case of the entropy of information theory.

III. Critique of Jauch and Baron's Discussion of Szilard's Thought Experiment

We definitely do not follow Jauch and Baron in their rebuttal of Szilard's argument. We understand the question as follows:

1. *A matter of semantics.* Jauch and Baron borrow from von Neumann the statement that 'in phenomenological thermodynamics each conceivable process constitutes valid evidence, provided that it does not conflict with the two fundamental laws of thermodynamics', and then add from their own that 'If an idealization is in conflict with a law that is basic to the second law, then it cannot be used as evidence for violation of the second law'.

It is certainly obvious that a self-consistent theory (and macroscopic, phenomenological thermodynamics is a consistent theory) cannot be criticized from within. But this does not forbid the production of thought experiments based on knowledge from *outside* the domain of the theory in order to criticize it. In his criticism of Aristotelian mechanical conceptions Galileo largely used thought experiments based on information that was at hand to everybody, *but* was outside the realm of Aristotelian physics.

The very concept of bouncing point molecules, and of pressure as integrated momentum exchange per sec cm^2, certainly is *outside* the domain of phenomenological, macroscopic, thermodynamics – not to speak of the consideration of one single molecule, and of learning in what half of the cylinder it is found.

Therefore, let us imitate Socrates in his criticism of Zeno, and proceed.

2. With Jauch and Baron, let us not get involved in quantum subtleties, but instead use the point particle concept of classical mechanics and of the classical kinetic theory of gases.

Let us first discard Jauch and Baron's large heat reservoir; we will bring it in later on.

There is no question that if Szilard's single molecule is (nineteenth-century language) or is known to be (post Szilard and Lewis language) in one definite half of the cylinder, the entropy of the single molecule gas is smaller than when the molecule is allowed to move all through the whole cylinder. If the transition from one state to the

other is secured by opening the door in the piston, then the entropy of the gas inside the cylinder increases by $k\mathrm{Ln}2$.

Does this statement contradict the classical thermodynamical statement that (in Jauch and Baron's terms) 'The entropy of the...system piston, gas...remains constant because it is closed and the changes are reversible'? In fact it does not, because the mere (frictionless) opening of the door entails an 'irreversible' change. Though energetically isolated, our (limited) system is not informationally isolated. However, with this distinction, we are certainly stepping outside the domain of classical thermodynamics.

As a result of opening the door, the pressure inside the cylinder falls down to one-half of what it was; this is because the molecule, while retaining by hypothesis its kinetic energy, now spends only half of its time in each half of the cylinder. Whence the classical expression for the change in entropy. (Note that while the molecule is not 'aware' that the door has been opened, the observer is, hence the entropy increase is computed as of the opening of the door.)

3. Jauch and Baron write: "However, at the exact moment when the piston is in the middle of the cylinder and the opening is closed, the gas violates the law of Gay Lussac because it is compressed to half its volume without expenditure of energy. We therefore conclude that the idealizations in Szilard's experiment are inadmissible in their actual context."

Socrates' walking and walking certainly was inadmissible in the context of Zeno's arguments. Nevertheless he *could* walk. If, by definition, our single molecule is a Newtonian point particle, nothing on earth can prevent us from suddenly closing our frictionless, massless, impenetrable door, and learning afterwards in what half of the cylinder we have trapped our molecule. There is *absolutely nothing* self-contradictory in this, not even the idealized concept of the door, which can be approached at will without even contradicting *classical* thermodynamics.

Moreover, no *classical* thermodynamicist would object to the reversed procedure: opening the door. This is adiabatic expansion, the subject of an experiment by Joule.

Why then had we termed adiabatic expansion an 'irreversible' process, and are we now speaking of it as time symmetric to the process of trapping the single molecule by closing the door? Because macroscopic irreversibility is one thing, and microscopic reversibility another thing, which have to be reconciled. It would lead us too far astray to delve here in this problem; we simply refer the reader to a recent discussion of it in terms that are consonant to those we are using here [35].

4. Finally we bring in Jauch and Baron's pulleys, strings and scales, which will restore the macroscopic irreversibility of the whole process.

From now on, as soon as we suddenly close the door, we do *not* allow the (single molecule) gas to expand adiabatically: we harness it, and have it lift (reversibly) a weight. This will cause its temperature to drop down, so that the heat reservoir becomes extremely useful if we want to lift many weights. So we bring it in also.

And *now* the *whole* system, supposed to be isolated, undergoes a (macroscopically) *reversible* change, so that its entropy remains constant.

However, *potential energy has been gained at the expenditure of heat* so that, at first sight, there is something wrong in the entropy balance. What are we overlooking?

We are overlooking that the observer has to *know*, that is, to *learn* which way the piston begins moving, and *then* push a weight on the right scale. This is his *free decision*, in accord with the *general decision* he has made, to use the heat in the reservoir for lifting weights. However, we are not delving here in a discussion of the twin Aristotelian aspects of information: cognizance and will [36].

Suffice it to say that, in order to have the entropy balance right, we *must* include in it the *information* gained by the observer. But this was Szilard's statement.

5. Finally, together with Brillouin [4] and others, Jauch and Baron point out rightly that the preceding 'system' can be transformed into a 'robot', the functioning of which, however, will require drawing negentropy from an existing source, by an amount at least as large as the negentropy created by lifting the weights.

The point is that the robot does not just step out of nowhere – no more than does a refrigerator or a heat engine. This brings into the problem Brillouin's 'structural negentropy', also Brillouin's 'information contained in the expression of the physical laws' that the engineer has used, and, finally, the problem of all the thinking and decision-making on the engineer's part.

In other words, if you push information out of the door by appealing to the robot, then it will come back right through the adiabatic wall.

6. *Concluding* this section, we believe that, while it is possible, and very informative, to deduce Thermodynamics from a general theory of Information, the converse is not possible. Thermodynamics is too rooted in specifics of physics to produce a general theory of information.

But this is a mere instance of the way scientific progress goes on...

REFERENCES

[1] J. M. Jauch and J. G. Baron, Helv. Phys. Acta *45*, 220 (1972).
[2] C. E. Shannon, Bell Syst. Techn. Journ. *27*, 379, 623 (1948); C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication* (Univ. of Illinois Press, Urbana 1949).
[3] L. Szilard, Zeits. für Phys. *53*, 840 (1929).
[4] L. Brillouin, *Science and Information Theory* (Academic Press, New York 1957).
[5] H. Grad, Comm. Pure Appl. Math. *14*, 323 (1971).
[6] E. T. Jaynes, Phys. Rev. *106*, 620 (1957); *108*, 171 (1957).
[7] E. T. Jaynes, Amer. Journ. Phys. *31*, 66 (1963).
[8] R. T. Cox, Amer. Journ. Phys. *14*, 1 (1946).
[9] R. T. Cox, *The Algebra of Probable Inference* (John Hopkins Press, Baltimore 1961).
[10] M. Tribus, J. Appl. Mech. March 1961, pp. 1–8; M. Tribus and R. B. Evans, Appl. Mech. Rev. *16* (10), 765–769 (1963); M. Tribus, P. T. Shannon and R. B. Evans, AEChI Jour. March 1966, pp. 244–248; M. Tribus, Amer. Sci. *54* (2) (1966).
[11] R. T. Bayes, Philos. Trans. *53*, 370 (1763).
[12] S. P. de Laplace, *Essai Philosophique sur les Probabiliités*, 1774. See also *A Philosophical Essay on Probabilities* (Dover, New York 1951).
[13] D. Hume, *An Enquiry Concerning Human Destiny*, 1758.
[14] J. von Neumann, Nachr. Akad. Wiss. Goettingen, Math.-Phys. Kl., 245 and 273 (1927).
[15] J. von Neumann, *Mathematische Grundlagen der Quantenmachanik* (Springer, Berlin 1972). See also *Mathematical Foundations of Quantum Mechanics* (Princeton Univ. Press, Princeton, N.J. 1955).
[16] W. M. Elsasser, Phys. Rev. *52*, 987 (1937).
[17] R. A. Fisher, Proc. Camb. Phil. Soc. *26*, 528 (1929); *28*, 257 (1932).
[18] S. Watanabe, Zeits. für Phys. *113*, 482 (1939).
[19] R. Carnap, *The Continuum of Inductive Methods* (Univ. of Chicago Press 1952).
[20] J. Kemeny, Journ. Symbolic Logic *20*, 263 (1955).
[21] H. Jeffreys, *Theory of Probability* (Oxford, Clarendon Press 1939).
[22] G. N. Lewis, Science *71*, 569 (1939).
[23] J. D. Van der Waals, Phys. Zeits. *12*, 547 (1911).
[24] J. W. Gibbs, *Elementary Principles in Statistical Mechanics* (Yale Univ. Press, New Haven, Conn. 1914), p. 150.
[25] E. T. Jaynes, *Foundations of probability theory and statistical mechanics*, in *Studies in the Foundations of Methodology and Philosophy of Science*, Vol. 1, Delaware Seminar in the Foundations of Physics (Springer, New York 1967); *Information theory and statistical mechanics*, in *Brandeis Lectures* (Benjamin, Inc., New York 1963), p. 181.

[26] M. TRIBUS, *Thermostatics and Thermodynamics* (D. van Nostrand Co., Princeton, N.J. 1961). See also N. TEPMOCTATNKA, *Tepmoznhamnka* (Moscow 1970).
[27] M. TRIBUS, *Rational Descriptions, Decisions and Designs* (Pergamon, Oxford and New York 1970). See also *Decisions Rationnelles dans l'Incertain* (Masson, Paris 1972).
[28] A. I. KHINCHIN, *Mathematical Foundations of Information Theory* (Dover, New York 1957).
[29] A. M. YAGLOM and I. M. YAGLOM, *Probabilité et Information* (Dunod, Paris 1959).
[30] A. KATZ, *Principles of Statistical Mechanics* (W. H. Freeman and Co., San Francisco and London 1967).
[31] A. HOBSON, *Concepts in Statistical Mechanics* (Gordon and Breach, New York 1971).
[32] R. BAIERLEIN, *Atoms and Information Theory* (W. H. Freeman and Co., San Francisco and London 1971).
[33] K. G. DENBIGH, *The Principles of Chemical Equilbrium* (Cambridge University Press 1955).
[34] N. R. HANSON, *Patterns of Discovery* (Cambridge University Press 1958).
[35] O. COSTA DE BEAUREGARD (reporter): *Discussion on temporal asymmetry in thermodynamics and cosmology*, in *Proceedings of the International Conference on Thermodynamics*, held in Cardiff, edited by P. T. LANDSBERG (Butterworths, London 1970).
[36] O. COSTA DE BEAUREGARD, *Is there a paradox in the theory of time anisotropy*, in *A Critical Review of Thermodynamics*, edited by E. B. STUART, B. GAL-OR and A. J. BRAINARD (Mono Book Corps., Baltimore 1970), pp. 463–472. See also Studium Generale *54*, 10 (1971).