1971-8

INFORMATION AND IRREVERSIBILITY PROBLEMS

O. Costa de Beauregard
INSTITUT HENRI POINCARÉ, PARIS

Reprinted from
Time in Science and Philosophy
An International Study of Some Current Problems
Editor: Jiří Zeman

Published in co-edition with
Elsevier, Publishing Company Amsterdam London New York
Academia, Publishing House
of the Czechoslovak Academy of Sciences
Prague 1971

quantity in systems. This question, which is essentially the same one as explaining how the subjective probability comes out as objective frequency, is by no means a trivial one, as unsophisticated commonsense would have it; this point has been strongly felt by Popper[9] and by Landé.[10]

The transition from subjective to objective probabilities occurs, loosely speaking, when one goes from small to large numbers; in physics, the entropy concept as used by Smoluchowski,[11] Szilard,[12] Lewis,[13] in problems involving very few molecules at a time is clearly "missing information"; measurable, objective entropy emerges only on the macroscopic level. Similarly, in quantum mechanics, where the wave function is a probability computing device, this ψ has a pronounced subjectivistic flavour in problems involving few quantum transitions; the famous "wave collapse" occurring there obviously corresponds to the information transition appearing in every stochastic test, such as turning a card or throwing a dice; only on the macroscopic level does the wave function recover its pure objectivistic signification, as exemplified for instance in photon or electron diffraction patterns.[14] But these statements which are in line with remarks by early authors on probability and frequency, do not tell us *how* the transition from subjectivity to objectivity occurs. The *technical* answer is, of course, that the expectation value of the frequency is precisely the probability;[15] but, by pointing out that after all the probability is *not* the frequency (only its expected value), this straightforward application of first principles stresses that to step outside the enchanted circle of subjective probabilities is by no means a triviality.

At this point Jaynes[2] adduces a remark much in the spirit of earlier ones by Poincaré[16] *plus* an explicit postulate: "*The theory makes definite predictions as to experimental behaviour only when, and to the extent that, it leads to sharp distributions* ... Such sharp distributions for macroscopic quantities can emerge only if it is true that for each of the overwhelming majority of those states to which appreciable weight is assigned we would have the *same* macroscopic behaviour ... It is this principle of 'macroscopic uniformity' which provides the objective content of the calculations, not the probabilities *per se*. Because of it, the predictions of the theory are to a large extent independent of the probability distributions over the microstates."

Interesting as it is this remark (and this form of the necessary postulate) does not settle the matter, because in fact statistical physical theories are operational even when low numbers are at stake, thus producing a direct verification of the *a priori* probability distribution with very little uncertainty; this is true, for instance, in the Wilson-Taylor or Dempster-Batho photon experiments. The argument that this situation is similar to the one in dice or coin-throwing games, thus implying that after all large numbers and "hidden parameters" are at stake, would not in the least solve the problem, but rather "sweep it under the rug".[9,10]

"Consider now the case where the theory makes definite predictions and they are not borne out by experiment. This situation cannot be explained away by concluding that the initial information was not sufficient ... The most reasonable conclusion is that the enumeration of the different *possible* states (i. e., the part of the theory which involves our knowledge of the laws of physics) was not correctly given. *Thus, experimental proof that a definite prediction is incorrect gives evidence of the existence of new laws in physics.* The failures of classical statistical mechanics, and their resolution by the quantum theory, provide several examples of this phenomenon." This pre-eminent statement was made by E. T. Jaynes.[2] But this statement is essentially as old as the probability theory, for it says (1) That the *a priori* probability distribution must conform to the principle of sufficient reason, and (2) That the observed frequencies must reproduce the calculated probabilities. So the umbilical chord is not yet cut ...

The problem which is at stake is the very one underlying all information theory and, before it, the whole of science; that is, the problem of the operational conformity of *statements* with *facts*, or of *representations* with *situations*.[17] Such a problem is radically avoided by the objectivistic school of probability theory, but it exists nevertheless. Generally speaking, it is the problem of the correspondence of mental images with physical states, or in some sense their emergence from physical states. And it is my feeling that information theory, together with statistical physics, has much to say in the future with regard to the solution of this very important problem.

II. Physical Irreversibility: the Link Between the Carnot Principle and the Retarded Waves Principle

I shall now summarize the more detailed examination of this question which may be found elsewhere,[1] and show the connection between the entropy increase principle and the retarded waves principle by rewording the famous von Neumann[18] proof referring to irreversibility in the quantum-mechanical measuring process.

Already in phenomenological thermodynamics it is clear that irreversibility is not *deduced*, but rather *postulated* in the very form of Carnot's assumptions:

(1) When two heat baths are interconnected, the heat flows from the high temperature to the low temperature one (not the opposite way); (2) A monothermic cycle may transform work into heat (not the opposite way).

Very interesting epistemological questions are raised when one tries to reverse the two Carnot assumptions; it turns out that consistent reasoning in the "anti-Carnot" thermodynamics is extremely subtle, because, as Poincaré[19] has shown, the anti-Carnot assumptions are of an anti-causal nature and do not allow physical prediction (only physical retrodiction). These remarks of Poincaré definitely suggest a very close connection between the Carnot principle and the principle of causality — a connection which has been demonstrated in many different ways by various authors in the recent years.

In Caratheodory's[20] axiomatic presentation of phenomenological thermodynamics, irreversibility is introduced by remarking that adiabatic expansion of a gas implies an entropy increase; but then one must ask why an adiabatic expansion rather than contraction is the natural way things go; this is definitely a Popper[21]-like question, and also a statistical argument in disguised from; we shall come back to this later.

Before we turn to the statistical mechanics interpretation of Carnot's principle we must inquire where a time[22] anisotropy[23] occurs in abstract probability theory. Suppose a pack of cards is found "in order": spades, diamonds, clubs, hearts, in four sequels from ace to king (or in any sequence belonging to a specified low populated sub-ensemble). The fact is that, while in predictive problems pertaining to card shuffling, the ordinary combinatorial laws of the probability theory will do very well, they will in fact fail in retrodiction: nobody will easily believe that an ordered pack of cards has been obtained by mere shuffling! The situation is very much the same in physical problems: sandheaps get swept away and nebular gases gather to build up stars, that is, in a given context, more probable situations follow from less probable ones, not the other way.

The technical means for handling this time dissymmetric situation is provided by a temporal application of Bayes' formula (which is, formally, a combination of the additivity of partial probabilities and the multiplicativity of independent probabilities): in retrodictive problems unphysical conclusions would be reached if one did not multiply the intrinsic probabilities of the dynamical problem by extrinsic *a priori* probabilities — the so-called Bayes coefficients, of which the theory says nothing, except that they must not be equal, because, then, the retrodictive problem would become time-symmetric to the predictive one, and thus belong to the body of laws of the paradoxical anti-Carnot world.

Very characteristically, the retrodictive application of Bayes' formula has often been called the "principle of probability of *causes*",[24] this showing that the connection of what may be called the "Bayes' principle" with the principle of causality was intuitively felt by the classic scientists. As for the interpretation of the Bayes' coefficients, it will become clear in the following that they are expressing the *interaction* out of which the stochastic system under study has been generated; but of course the question is thus raised of why an interaction limited in time produces, in each subsystem, after-effects and not "before-effects".

Turning now towards the statistical interpretation of Carnot's principle, either in the form of mechanical models (Clausius, Maxwell, Boltzmann, Gibbs and their followers), or in the formalistic, abstract[25] and concise form available in "thermostatics",[12,5] it is immediately clear that the statistical interpetation of Carnot's principle is merely a specification of the above stated Bayes' principle; this was first stated clearly by Van der Waals[26] and is also implicit in a much-quoted passage by Gibbs;[27] an analogous though less technical statement is found in the writings of Mehlberg,[28] Grünbaum,[29] and many others, in the form that physical irreversibility is definitely of a "fact-like rather than law-like" character; that is, that it is expressed as a *boundary condition* rather than as a dynamical law.

Perhaps a typical example will help to make this point clear. Poincaré,[30] discussing the problem of the uniform distribution of the little planets on their common orbit, simplifies it by taking the orbit to be a circle and replacing the planets by a fictitious gas.[31] Denoting ω the (constant) angular velocity of a molecule, ϕ its initial longitude, t the time, $f(\omega, \phi)$ the distribution function, Poincaré remarks that the "characteristic function"

$$C(t) = \iint \exp\left(\mathrm{i}(\omega t + \phi)\right) f(\alpha\phi)\, \mathrm{d}\alpha\, \mathrm{d}\phi$$

goes to zero when t goes to ∞, whatever the function f, provided simply that it is continuous in ω. Thus, whatever the initial distribution $f(\alpha, \phi)$, the final one will be uniform. The trouble is that a similar conclusion is drawn if $t \to -\infty$, which would certainly be unphysical. There is no other issue than cutting the Gordian knot, that is stating that the use of Poincaré's formula is allowed in prediction and forbidden in retrodiction. This is, of course, a statement fundamentally similar to the one expressed by the temporal application of Bayes' principle. Incidentally, this Bayesian kind of paradox belongs to the same family as the early Loschmidt[32] and Zermelo[33] arguments, but it is much more general in that no specific dynamics is implied in it. Also, it is of course akin to the well-known Ehrenfest[34] and Smoluchowski[35] consideration pertaining to the temporal isotropy and homogeneity in the evolution of isolated statistical mechanical systems.

Finally, temporal anisotropy is truly postulated at the very basis of any form of statistical mechanics, in the form of an initial boundary condition, macroscopic in its nature, and stating, as Watanabe[36] puts it, that blind statistical prediction is allowed and blind statistical retrodiction forbidden; in other words, it is a specification of the temporal application of Bayes' principle.

Now I will show that the latter is also true for the principle of retarded waves.

That some sort of physical connection must exist between the principle of entropy increase and the principle of retarded waves is obvious in many examples. When a stone is thrown in a pond, retarded volume and surface waves are generated, and it is through them that the dissipation of the energy first concentrated in the stone takes place. Also, the slowing down and consumption of a meteorite entering the earth's atmosphere occurs through retarded ballistic waves. Still more specifically, if, between times t_1 and t_2, a physicist moves a piston in the wall of a vessel containing a gas in equilibrium, Maxwell's velocity distribution law is alterated after time t_2, not before time t_1, that is, the alteration is propagated in the gas as a retarded rather than advanced wave. Commentaries on the paradoxes of blind retrodiction and the corresponding anti-causality, in these examples, are left to the reader.

In this context one must recall what has been qualified as "an inconclusive but illuminating discussion carried on by Ritz and Einstein in 1909, where Ritz treats the limitation to retarded potentials as one of the foundations of the second law of thermodynamics, while Einstein believes that the irreversibility of radiation depends exclusively on considerations of probability".[37] There is little chance to get at the heart of the problem when speaking of non-quantized waves — and all the more since all waves in material media are now believed to be quantized. My personal feeling (which I intend to justify) is that Ritz and Einstein were equally right, and that the only circumstance preventing them from recognizing that they were looking at the same thing from two opposite directions was simply that, at the time they wrote, the undulatory aspect of mechanics was not yet discovered — even if the corpuscular aspect of light was. Had Ritz and Einstein known, in 1909, that every scattering process, in the sense of statistical mechanics, is *also* a scattering of waves, and *vice versa*, then certainly both of them would have recognized that their opposite positions were in fact *reciprocal*, that is, mutually exchangeable.

The first clear hint of a link between the two principles of entropy increase and (quantized) waves retardation is found in Planck's definition[38] of the entropy of a monochromatic light beam: this entropy increases in a scattering process. That *scattering* as associated with *entropy* *increase* is of course no surprise. The point is that scattering, or diffusion, is a *retarded waves* process. Should advanced waves be used, then a paradoxical phase coherent "con-fusion" instead of "dif-fusion" would occur, and there would correspond to it an entropy decrease.

In Planck's formula for the entropy of a monochromatic light beam the h constant enters; that is, the quantization of light waves (according to Bose statistics in this case) is essential in defining an entropy of the light beam. In other words, *quantization is essential for stressing the link between the Carnot (or Bayes') principle and the principle of retarded waves.*

There are of course as many ways for stating such a link as there are ways of handling probability; I will choose as specially fundamental the one using von Neumann's[18] micro-entropy[39] in quantum-mechanical systems; the above stated link is made clear simply by rewording von Neumann's proof pertaining to irreversibility in the quantum-mechanical measuring process; and this I will do using, for brevity, a simplified form of the argument.

Ψ_i denoting the projection operators associated with orthonormalized wave functions ψ_i and

$$D = \sum_i p_i \Psi_i \quad (0 \leqq p_i \leqq 1, \sum p_i = 1)$$

the density operator, the probability q_i of finding the pure state ϕ_i in a future measurement is

$$q_i = \text{trace } D\Phi_i = \langle \psi_i' | D | \psi_i' \rangle = \sum_j p_j < \psi_i' | \Phi_j | \psi_i' \rangle \,.$$

Introducing the expansion coefficients of ϕ_i in the ψ_j system, majoring all p's by their upper bound P, and using the identity

$$\sum_j C_{ij}^+ C_{ij} = 1 \,,$$

one obtains

$$q_i = \sum_j C_{ij}^+ C_{ij} p_j \leqq P :$$

none of the q's expected in a future measurement can be larger than the largest of the initial p's. This describes a levelling or equalizing procedure typical of the Carnot family.

The point is that the deduction has being carried out as a "blind prediction" and that retarded solutions of the Schrödinger equation, starting from the initial state, have been implicitly used. Had a symmetrical "blind retrodiction" been performed, this would have used advanced solutions of the Schrödinger equation — and would have yielded an entropy decrease. The one to one connection between the use of retarded (light or matter) waves and application of the statistical Bayesian postulate "blind retrodiction forbidden" is thus established on a very fundamental level.

A shorter and more intuitive way for saying the same thing is that, as retarded and advanced waves are the mathematical devices respectively associated with statistical prediction and retrodiction in quantum mechanics, the Bayesian principle forbidding blind retrodiction is just another name for the principle excluding advanced waves on the macroscopic scale.

III. The Einstein-Podolsky-Rosen Paradox and Information Transfer

The *intrinsic* time symmetry of prediction and retrodiction problems is fairly obvious in the probability formalism of the quantum theory: retarded and advanced waves are generated, *via* a resolution of the Cauchy problem, respectively towards the future and the past; this, in the relativistic case, implies the use of the Jordan-Pauli propagator D, that is, of a probability amplitude yielding symmetric probabilities in prediction and blind retrodiction. It is thus clear that the physical dissymmetry between prediction and retrodiction is *not* intrinsic in the quantum dynamics; that is, it is of macroscopic nature, and generated through the consideration of ensembles and use of Bayes' formula, as previously explained.

At first sight the situation seems less clear in the case of transition amplitudes calculated by Feynman's rules. This is because an explicit time dissymmetry, manifesting itself in higher order processes,[40] is introduced there by the use of the so-called "causal" D_C Stueckelberg-Feynman propagator for describing virtual particles; this propagator has a time symmetric twin, which may be called the "anti-causal" propagator D_F. But a closer examination reveals the reason for the trouble.

Jauch and Rohrlich[40] have explained how the use of the Stueckelberg-Feynman D_C describing virtual particles automatically yields (1) the Lamb shift of energy levels in atoms (2) the exponential decay from higher to lower energy states through emission of retarded radiation. A similar calculation then shows that use of the anticausal D_F would yield (1) the Lamb shift (unchanged) and (2) a paradoxical exponential build up of higher energy levels from lower ones through a paradoxical absorption of advanced radiation, while use of the time symmetrized $\bar{D}$ would yield only the Lamb shift and no preference with respect to high or low energy levels, neither for retarded nor advanced radiation.[41]

Now, the statistical origin of exponential decay laws is nothing new: it is found in radioactivity, and, even before, in the theory of chemical reactions. The point is that an exponential decay follows from blind statistical prediction (while an exponential build-up would follow from blind statistical retrodiction). So we find once again the association of retarded waves with blind statistical prediction, and, through the Jauch-Rohrlich demonstration, even something more: that the description of virtual particles by the time asymmetric D_C (rather than the time-symmetric $\bar{D}$) has an extrinsic origin: *it is in a one to one correspondence with the principle of exclusion of advanced waves, which principle is essentially macroscopic in its nature.*

The most striking operational proof of the complete time symmetry in the transmission of information at the elementary quantum level is certainly provided by the so-called Einstein-Podolsky-Rosen[42] (non-relativistic) paradox, or by its relativistic twin considered by Schrödinger.[43] The problem essentially refers to quantum statistics of an elementary transition, and so we are not dealing with frequencies; that is, we are dealing with the *information* gained in a stochastic test.

Let us discuss the "paradox" in the slightly different form considered first by Einstein[44] and later by Renninger.[45] During some time interval Δt, *one single* quantum corpuscle goes through a hole of section Δs and is later absorbed at a point on some photographic plate; the *a priori* impact probabilities are of course calculated, as classical optical intensities would be, using the retarded wave emitted by the hole. The point is that, if an observer O_1 registers the impact in, say, A_1 at time T_1, he is immediately *certain* that there is no impact of this same particle at a distant point-instant (A_2, T_2) separated from (A_1, T_1) by a space-like interval; he *knows* that another observer O_2 operating at (A_2, T_2) cannot observe the particle. Conversely, if O_1 knows that the particle has been emitted, and he has not observed it in (A_1, T_1) after the necessary time has elapsed, he is *certain* that an impact of that particle may have occurred at some other (A_2, T_2) point-instant with a space-like separation from his (A_1, T_1) – provided of course a receptor was there.

The "paradox" consists in this "instantaneous" or "space-like" information transfer; such prominent thinkers as Einstein, Schrödinger, de Broglie, felt that the only resolution of the paradox is through some kind of hidden determinism, which would make the statistics purely subjectivistic. But such a solution would entail other difficulties, such as those discussed by Landé.[9] So I believe that the right answer must be more subtle.

Through which channel is information transferred in the hidden variables version? Positive information (answer *yes*, or 1; that is, impact) is transferred along the alleged trajectory, while inferred negative information (answer *no*, or *O*; that is, no impact) is also transferred along a trajectory, but this time an unoccupied one. To solve the paradox we have simply to retain this scheme, but without making the dubious assumption that probabilities are purely subjectivistic in an elementary quantum test. *So, the space-like information transfer from (A_1, T_1) to (A_2, T_2) [or from (A_2, T_2) to (A_1, T_1)] is in effect the combination of two time-like information transfers along a Feynman zigzag joining, in space-time, (A_1, T_1) and (A_2, T_2) with its angle in the past $\Delta s\, \Delta t$ (time-like) 3-surface where the hole is open.*

In other words, it is my feeling that the various correlation paradoxes of the Einstein and Schrödinger family (the *technical* solution of which is straightforward through the quantum formalism) are a *direct operational proof of the time symmetry of information inference at the elementary quantum level.* The principle of wave retardation is a macroscopic principle, and it does not hold at the elementary quantum level, where a time-symmetric wave propagation principle replaces it. This is obvious in the *formalism*; that it is also *operationally* true in the form of the Einstein and Schrödinger correlation "paradoxes" is just one more curiosity in the realm of quantum mechanics.

In a static situation and in terms of Euclidean 3-dimensional space, nobody finds any paradox in the fact that if one single ball is found in one definite closed box belonging to a set, it may not be found in any of the others. The surprise here is that a similarly rigid correlation of informations exists in Minkowski's *space time*; and this of course emphasizes, on the quantum level, what Minkowski,[46] Weyl,[47] Einstein[48], Feynman,[49] and many others, have said about the epistemological status of the realistic space-time; it is, according to them, very similar to the one of the former Euclidean space: neither less nor more realistic. The surprise here is that the proof comes in the form of a correlation of informations; this may well be inherent to the fact that, with quantum mechanics, we are experimenting with statistics on truly elementary physical phenomena.

IV. Bayes' Principle and Causality Principle

Let us make it clear now that probability or entropy increasing[50] evolutions conform to the principle of causality, while paradoxical probability decreasing evolutions would conform to the time reversed principle of finality. This is obvious in terms of waves, because the former are described by retarded waves and the latter by advanced waves. So let us make the point also in terms of probabilities.

Poincaré[19] remarks, as Grünbaum[51] puts it, that, in an anti-Carnot world, it would be dangerous to get into a lukewarm bathtub, because one could not foretell which end is going to boil and which to freeze. It would also be dangerous to bowl, if friction were an accelerating rather than damping process. This is because a Carnot world is such that blind statistical prediction is possible in it, while an anti-Carnot world would be one in which blind statistical retrodiction would work. Also, in an anti-Carnot world, converging waves would build up in quiet ponds, concentrate upon a stone, and precisely throw it into the hand of a passer-by. Or shuffling cards would produce at will an ordered pack. That is, improbable situations would come out of probable ones.

It is thus clear that the temporal application of Bayes' principle, which excludes blind retrodiction, is the *statistical* expression of both Carnot's principle, and of the principle saying that interactions limited in time develop after-effects and not before-effects. It should incidentally be noted that such interactions, though certainly increasing the entropy of the total system, may very well decrease at first the entropy of one subsystem. If, for instance, (Schlick's example) one sees footprints on a sandbeach, one concludes that a man has passed by, and not that a man will step in the footprints and thus erase them. In fact, all registering apparatus are of that kind, that is, subsystems which, before coupling, are in a high entropy state; a formalization of the process is given by J. M. Oudin.[53]

A thorough investigation of these coupling and decoupling processes is found in Reichenbach's[54] and Grünbaum's[29] works under the name of "theory of branch systems". The idea that Carnot's principle and the physical causality principle are in fact the same one under two different names, besides having been strongly emphasized by Reichenbach, Grünbaum, and myself,[1] is also more or less implied in the writings of such physicists as Watanabe,[36] E. N. Adams,[55] J. Mc Lennan,[56] Wu and Rivier,[57] Penrose and Percival;[58] it is explicit in a paper by Terletsky.[59]

An interesting question is that of the relation between the (subjective) principle of sufficient reason and the (objective) principle of causality. The classical answer would perhaps be that the former is merely the subjectivistic duplication of the latter. Nevertheless, the preceding analyses have shown, as Jaynes[2] puts it, that the two concepts of subjective and objective probability refuse to be united; so the expression of the connexion between the two principles of sufficient reason and of causality will require more subtlety.

Before tackling this point, I wish to show how easily the postulate of a one way time arrow creeps in when one makes use of either the principle of sufficient reason or the principle of causality. E. Borel, in his book *Le Hasard*,[60] adopts the objectivistic or frequency theory of probabilities, and writes: "A coin is thrown and bets are laid as to which side will show up when it has fallen. This is the simplest of probability problems, if one adds the hypothesis that chances are equal for heads and tails. Regarding the equality of chances, we won't get involved in any philosophival discussion, and will merely take it as an experimental fact, or, if you like, as the very *definition* that the coin is *good* ...". On the other hand P. Lévy, in his *Calcul des Probabilités*,[61] is an avowed proponent of the subjective probability concept; regarding this very same heads or tails game, he writes: "When two possible cases are equally probable, that is, when we have no reason to expect the one more than the other, we are *ipso facto* expecting that, if the test is carried out a sufficiently large number of times, the two cases will show up with almost the same frequency."

The point I wanted to make is that, in both arguments, the principle of probability increase by testing — that is, the temporal application of Bayes' formula — has been postulated.

V. Information and Entropy: Brillouin's "Generalized Carnot Principle"

It will be especially interesting to apply the general conclusion reached in recent years by many authors — that statistical irreversibility is of a fact-like rather than law-like character — to the two way transition

$$\textit{negentropy} \rightleftarrows \textit{information}$$

discovered by cybernetics.

In this context a preliminary question will be: is it true, as G. N. Lewis,[13] S. Watanabe[36] and H. Mehlberg[28] would have it, that physical evolutions are completely time symmetrical, and that "Gain in entropy always means loss of information and nothing else? Is it a subjective concept?"[13] This reviewer feels that it would be difficult to go so far, and that there is certainly *something objective* in the *fact* that sandheaps get swept away or that nebular gases gather to build up stars. That the sun radiates would remain true even if no earthly physicists were there to observe the phenomenon and develop its thermodynamics. But, on the other hand, such observations and the corresponding theories may help to devise procedures and build devices to partly recapture the dissipated energy: so there is after all *something* true in Lewis', Watanabe's and Mehlberg's views — and this is precisely the concern of cybernetics.

In the present context, the fundamental discovery of cybernetics is,[62] as Gabor puts it, that "one cannot get anything for nothing, not even an observation". He, and Brillouin,[63] have given very interesting examples showing that the information ΔI gained in any physical measurement must be paid for by an (at least) equal loss of the surrounding negentropy[64] ΔN_1:

$$\Delta N_1 \geqq \Delta I;$$

or, if information is measured in natural binary units and entropy in practical thermodynamic units,

$$\Delta N_1 \geqq k \ln 2 \, \Delta I$$

where k denotes Boltzmann's constant. This immediately places every physicist, and even every layman, very much in the situation of a gambler. That this general law of Nature had escaped recognition for so long is of course due to the smallness of the Boltzmann's constant in practical units; and this occurrence reminds one of those of the quantum and relativity discoveries, where the universal constants h and $1/c$ are also very small when expressed in practical units.

Now, the second fundamental discovery of cybernetics is that the possession of a certain amount of information allows its owner to restore an (at most) equal amount ΔN_2 of (coarse grained) negentropy in the surroundings:

$$\Delta I \geqq \Delta N_2 .$$

For instance Maxwell's demon, being well informed on the microstate of affairs, is able to convert the corresponding high value of the fine grained negentropy into a coarse grained one. So, cybernetics take the *acting transition*

information → *negentropy*

as the reciprocal of the *learning transition*

negentropy → *information* .

On the subjectivistic side, *cognizance awareness* (where the representation follows in time the physical situation) is present in the latter case, and *willing awareness* (where the representation precedes in time the situation) in the former one. And it certainly is a good thing that cybernetics seems at first sight able to take care of human work and labour, and not only of speculation and contemplation. Also, it may be noted that cybernetics has rediscovered, without having searched for it, the connection between the two senses of the old Aristotelian information concept: *gain of knowledge* and *planning power* (the second of which had almost been forgotten).

It thus seems plausible that cognizance awareness and willing awareness should respectively emerge in regressing and growing up fluctuations. Now, according to Brillouin and others, the "generalized Carnot principle" is written as

$$\Delta N_1 \geqq \Delta I \geqq \Delta N_2$$

(the original one being simply

$$\Delta N_1 \geqq \Delta N_2);$$

and of course the kind of irreversibility expressed in the new form of the principle must remain, as in the old one, "of a fact-like rather than law-like character". Thus, that information gained in a learning transition is smaller than the negentropy from which it is borrowed and larger than the one it could restore in a willing transition, expresses a *de facto* situation of ours where observation is easier than action; in the two way transition

negentropy ⇄ *information*

the upper arrow is easier to follow than the lower one. Remembering what was said above of the connection of retarded actions with decreasing entropy, this amounts to saying that causality is largely predominant over finality — a statement which is of statistical rather than absolute character, because of course there are fluctuations, some of them large and long lasting; this was pointed out by Schrödinger[65] in the case of living things.

Incidentally, such a fact-like situation contains a possible explanation of Boltzmann's[66] idea that life is bound to explore the entropy curve in the time direction where entropy is growing: if one postulates (which seems inevitable: one cannot "anti-read" a book from the last to the first sentence, thus erasing one's previous knowledge in the field) that life essentially implies an incoming information flux, then this $\Delta I > 0$ requires that the time axis is followed in the direction so that $\Delta S > 0$, as Boltzmann had postulated.

It is my personal feeling that the *de facto* smallness of Boltzmann's constant k in practical units is in relation with the above stated existential situation (just as the largeness of the c constant in practical units may well be a reflection of the low value of our nervous influx, which relates *for us* small length intervals to large time intervals). Due to the smallness of k, to gain information is very cheap in negentropy terms, while to produce negentropy costs very much in information terms; going to the pre-cybernetical limit $k \to 0$ would render observation costless and action impossible, a version of "epiphenomenal consciousness" which is no longer tenable. In a similar vein, the reason why the causality concept is so obvious to observational awareness while the finality concept is so obvious to willing awareness seems to me very clear: it is because, by their very nature, observation and action are respectively connected with entropy increasing and entropy decreasing processes, that is also, according, to the preceding analyses, with retarded and advanced actions.

VI. Brief Conclusion

As one should not say too many things in one single paper I had to leave out many topics such as the consideration of the Wheeler-Feynman semi-classical radiation theory, more developments on the Reichenbach-Grünbaum theory of branch systems, and the connection which may very well exist between cosmological and statistical irreversibility (Sciama and others).

On the whole, I have the feeling that the already long history of probability, irreversibility, information in general, information in quantum theory, is very far from finished, and that in the future we may learn much from it concerning the way knowing and willing awareness (of men and of animals) are connected with the mechanics of the universe. In his well-known book, Brillouin[63] writes: "Relativity theory seemed, at the beginning, to yield only very small corrections to classical mechanics. New applications to nuclear energy now prove the fundamental importance of the mass-energy relation. We may also hope that the entropy-information connection will, sooner or later, come into the foreground, and that we will discover where to use it to its full value".

1 O. Costa de Beauregard, 'Irreversibility Problems', in *Proceedings of the 1964 International Congress for Logic, Methodology and Philosophy of Science*, North Holland Publishing Co, Amsterdam, 1965, p. 313—342. Also: *Le Second Principe de la Science du Temps*. Ed. du Seuil, Paris, 1963.
2 E. T. Jaynes, 'Information Theory and Statistical Mechanics', *Phys. Rev.* 1957, *106*, 620—630.
3 P. de Laplace, *Théorie Analytique des Probabilités*, Paris, 1812.
4 T. Bayes, *Essay towards solving a Problem in the Doctrine of Chance*, 1763.
5 Modern writers following this line are, for instance, P. Levy, *Calcul des Probabilités*, Gauthier Villars, Paris, 1925, and R. T. Cox, *The Algebra of Probable Inference*, John Hopkins Press, Baltimore, 1961. Laplace was referring to the "principle of insufficient reason" (for preferring one possibility), while Poincaré (ref.[30] below) referred himself to the "principle of sufficient reason". Keynes has suggested the technical term of "Principle of Indifference".
6 J. Bertrand, *Calcul des Probabilités*, Paris, 1888 — and others have given examples of very tricky problems in *a priori* estimations in applications of probability theory. See for instance E. P. Northrop, *Riddles in Mathematics*, D. Van Nostrand Co, New York, 1944, chap. 8.
7 C. E. Shannon and W. Weaver, *The Mathematical Theory of Communication*, Univ. of Illinois Press, Urbana, 1949.
8 Please notice that the "principle of indifference" is still implied in the estimation of the *a priori* weights p_i.
9 A. Landé, *New Foundations of Quantum Mechanics*, Cambridge Univ. Press, 1965, chap. II.
10 K. R. Popper, *The Logic of Scientific Discovery*, Hutchinson, London, 1956 and 'Indeterminism in Quantum Physics and in Classical Physics', *Brit. J. Phil. Sci.* 1950, *1*, 1.
11 M. von Smoluchowski, 'Gültigkeitsgrenzen des zweiten Hauptsatzes der Wärmetheorie', *Oeuvres*, vol. 2, 1927, p. 361—398.
12 L. Szilard, 'Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen', *Zeitschrift f. Phys.* 1929, 53, 840.
13 G. N. Lewis, 'The Symmetry of Time in Physics', *Science* 1930, *81*, 569—577.
14 Not only with "bosons", but also with "fermions", there is a kind of objectivity of the wave function emerging on the macroscopic level: the one showing up upon Davisson and Germer's photographic plate.
15 M. Tribus, P. T. Shannon and R. B. Evans, 'Why Thermodynamics is a Logical Consequence of Information Theory', *A. I. Ch. Journal*, March 1966, p. 244—248.
16 H. Poincaré, *Calcul des Probabilités*, 2nd ed., Gauthier Villars, Paris, 1912.
17 This is the idea in Einstein's and de Broglie's remarks: "What remains ununderstandable is that the world is understandable" (A. Einstein, 'Physik und Realität' in *Zeitschrift für freie deutsche Forschung*, 1, n. 1 p. 5—19 and n. 2 p. 1—14) and "There is much more obscurity than people believe in the simple fact that some science is possible" (L. de Broglie, 'L'Invention Scientifique' in *Neuvième Semaine Internationale de Synthèse*, Alcan, Paris, 1938, p. 125).
18 J. von Neumann, *Mathematische Grundlagen der Quantenmechanik*, Springer, Berlin, 1932, chap. V.
19 H. Poincaré, *Science et Méthode*, Flammarion, Paris, 1908, chap. IV — *Le Hasard*.
20 C. Caratheodory, 'Untersuchungen über die Grundlagen der Thermodynamik', *Math. Annalen* 1909, 67, 365. See also A. Landé, *Handbuch der Physik* 1926, 9, 281—300 and M. Born, *Natural Philosophy of Cause and Chance*, Oxford Univ. Press, 1949.
21 K. Popper, *Nature* 1956, *177*, 538 and *178*, 381; 1957, *179*, 1296; 1958, *181*, 402. See also *ibid* commentaries by Schlegel, Hill, Grünbaum and Bosworth.
22 Discussions of a possible connection between the PC-violation recently discovered in elementary particle physics and the statistical time arrow are yet merely academic.
23 That Kolmogorov (*Foundations of the Theory of Probability*, Chelsea, New York, 1950) has been able to formulate the probability theory in a completely atemporal fashion does not preclude that *in physics* we are interested in *tests*, and that the concept of time is implied in the very idea of a test.
24 Such an analysis may be found in E. Borel, *Le Hasard*, Alcan, Paris, 1914, 339.
25 It should be noted that present day statistical mechanics has become a highly *formalistic* theory and is

no more a modelistic one, as it was in its early days, when challenging "energetism". A clear characterization of the relation between formalism and modelism in modern statistical mechanics is found in reference[2].
26 J. D. Van der Waals, 'Über die Erklärung der Naturgesetze auf statistisch-mechanischer Grundlage', *Phys. Zeits.*, 1911, *12*, 547–549.
27 J. W. Gibbs, *Elementary Principles in Statistical Mechanics*, Yale Univ. Press, New Haven, 1914, p. 150.
28 H. Mehlberg, 'Physical Laws and Time's Arrow' in *Current Issues in the Philosophy of Science*, H. Feigl and G. Maxwell eds, Holt, Rinehart and Winston, New York, 1961.
29 A. Grünbaum, *Philosophical Problems of Space and Time*, Knopf, New York, 1963, Part II. Also, many previous articles quoted there.
30 'Le Calcul des Probabilités' in *La Science et l'Hypothèse*, Flammarion, Paris, 1906, chap. XI 2 III.
31 *This* transition from a discrete to a continuous distribution raises no awkward questions.
32 J. Loschmidt, 'Über das Wärmegleichgewichtigkeit eines Systems von Körpern mit Rücksicht auf die Schwere', *Wien Ber.*, 1876, *73*, 128 and 136.
33 E. Zermelo, 'Über einen Satz der Dynamik und der mechanischen Wärmetheorie', *Ann. der Phys.*, 1896, *57*, 485.
34 P. and T. Ehrenfest, *Encycl. Math. Wiss.*, IV, 4, D, Leipzig 1911, 332 p. 43.
35 Ref.[11] p. 361–398.
36 S. Watanabe, 'Symmetry of Physical Laws', III, Prediction and Retrodiction', *Rev. Mod. Phys.*, 1955, *27*, 179–186. See also 'Reversibility of Quantum Electrodynamics', *Phys. Rev.*, 1951, *84*, 1008–1025; 'Le Concept de Temps dans le Principe d'Onsager' in *Transport Processes in Statistical Mechanics*, Interscience Publishers, New York, 1958, p. 285–310 (explicit reference to Bayes' principle); 'Time and the Probabilistic View of the World' in *The Voice of Time*, Braziller, New York, 1966, p. 527–563; 'Réversibilité contre Irréversibilité en Physique Quantique' in *Louis de Broglie, Physicien et Penseur*, A. George ed., Albin Michel, Paris, 1953, p. 385–400. (compare to Lewis, ref.[13] and Mehlberg, ref.[28].)
37 W. Ritz and/or A. Einstein, 'Papers' in *Ann. de Chim. et de Phys.*, 1908, *14*, 145; *Phys. Zeits.*, 1908, *8*, 903; 1909, *10*, 185, 323; 1910, *10*, 817. Incidentally, the fact-like character of retarded waves on the one hand, and of increasing probabilities on the other, were not stressed in this discussion.
38 See M. Planck, *The Theory of Heat*, Macmillan, London 1932, Parts III and IV. See also P. Rosen, 'Entropy of Radiation', *Phys. Rev.* 1954, 96, 537 and A. Ore, 'Entropy of Radiation', *Phys. Rev.* 1955, 98, 887.
39 This, as Landé, ref.[9], chap. V, has stressed, corresponds to the "chemical potential" contribution in the classical definition of entropy.
40 See J. M. Jauch and F. Rohrlich, *The Theory of Photons and Electrons*, Addison Wesley Publishing Co, Cambridge (Mass.) 1955, p. 405–410.
41 Technically, the whole affair rests on the choice of an integration path in the complex plane, as was already the case in the early Weisskopf-Wigner Theory (*Z. Phys.*, 1930, *63*, 54 and *65*, 18). The choice of D_C (resp. D_F) has the effect of precisely equating the decay (resp. build-up) coëfficient with the transition probability per time unit.
42 A. Einstein, B. Podolsky, N. Rosen, 'Can the Quantum Mechanical Description of Physical Reality Be Considered Complete?', *Phys. Rev.*, 1935, *47*, 777–780.
43 E. Schrödinger, *Naturwiss.*, 1935, *23*, 787, 823, 844.
44 A. Einstein in *Rapports du 5ème Conseil Solvay*, Paris, 1928, p. 253.
45 W. Renninger, *Zeits. Phys.*, 1963, *136*, 251.
46 H. Minkowski, in A. Einstein, *The Principle of Relativity*, Methuen, London, 1923, p. 75.
47 H. Weyl, *Philosophy of Mathematics and Natural Science*, Princeton Univ. Press, 1949, p. 116: "The objective world simply *is*, it does not happen".
48 A. Einstein, Dingler's *Polytechn. Journal*, 1930, 345, 122.

49 R. P. Feynman, *Phys. Rev.*, 1949, *76*, 749.
50 With a logarithm basis higher than 1, of course.
51 A. Grünbaum, Ref.[29], p. 227.
52 M. Schlick, *Grundzüge der Naturphilosophie*, Gerold and Co, Wien, 1948, p. 106.
53 J. M. Oudin, *C. R. Acad. Sci.*, 1961, *252*, 3008 (second part of the Note).
54 H. Reichenbach, *The Direction of Time*, University of California Press, 1956.
55 E. N. Adams, 'Irreversible Processes in Isolated Systems', *Phys. Rev.*, 1960, 120, 675–681.
56 J. A. Mc Lennan, 'Statistical Mechanics of Transport in Fluids', *Phys. of Fluids*, 1960, *3*, 493–502.
57 T. Y. Wu and D. Rivier, 'On the Time Arrow and the Theory of Irreversible Processes', *Helv. Phys. Acta*, 1961, *34*, 661–674.
58 O. Penrose and I. C. Perceval, 'The Direction of Time', *Proc. Phys. Soc.*, 1962, *79*, 605–616.
59 P. Terletsky, 'Le Principe de Causalité et le Second Principe de la Thermodynamique', *J. Phys. Rad.*, 1960, *21*, 681–684.
60 E. Borel ref.[24] chap. II, No 9. I have reinserted a footnote into the main text.
61 P. Levy, *Calcul des Probabilités*, Gauthier Villars, Paris, 1925, Part I, chap. II, 10.
62 D. Gabor, *M. I. T. Lectures*, 1951.
63 L. Brillouin, *Science and Information Theory*, Academic Press, New York, 1956.
64 By "negentropy" one means of course the entropy with its sign changed.
65 E. Schrödinger, *What is Life*? Cambridge Univ. Press, 1944, chap. VII.
66 L. Boltzmann, *Vorlesungen über Gastheorie*, Barth, Leipzig, 1896, Bd II, p. 257. Incidentally, this Boltzmann conception implies the idea of an actually laid down time development of things as in relativity theory.