Quantum Mechanics, Thermodynamics, and the Strong Cosmological Principle

David Layzer

I. Introduction

Historically, dynamics and thermodynamics developed independently. Boltzmann and his successors tried to bridge the gap between the two disciplines by constructing statistical dynamical theories of many-particle systems, but although such theories have enjoyed considerable practical success, their foundations are still far from secure. As Tisza has emphasized,[1] dynamical and thermodynamical descriptions are *complementary*. Dynamical descriptions are *deterministic*, *reversible*, and require (in Tisza's phrase) an *exhaustive* specification of initial data; thermodynamical descriptions are *stochastic*, *irreversible*, and require a *selective* specification of initial data.

It is widely held that dynamical descriptions are more fundamental than thermodynamical descriptions because one arrives at a thermodynamical description by selectively ignoring certain kinds of information about a complex system. But irreversibility and macroscopic order—two features of thermodynamical descriptions that have no counterparts in a dynamical description—seem to be objective properties of the world rather than artifacts of a particular mode of description. Tisza has argued persuasively that thermodynamical descriptions are in fact no less fundamental than dynamical descriptions and may even be more fundamental. Although he believes that a "reasonably unified logical structure" for dynamics and thermodynamics may be found, he stresses the importance of keeping an open mind about the form it will take.

Prigogine[2] has recently proposed to unify dynamics and thermodynamics through a fundamental change in the foundations of dynamics. He suggests that "the basic concepts of classical (or quantum) mechanics, such as trajectories or wave functions, correspond to unobservable

idealizations." In Prigogine's proposed generalization of dynamics, the classical microscopic (Gibbs) entropy is not a functional of the phase density but an operator. Its quantal analogue is not a functional of the density operator but a supperoperator. Thus the classical phase density does not have a definite value unless it happens to be an eigenfunction of the microscopic entropy operator. In this way dynamical descriptions—classical as well as quantal—acquire an irreducibly stochastic character.

A stochastic description, however, does not necessarily single out a preferred direction in time. In Prigogine's theory, irreversibility is built into the microscopic dynamical laws. Prigogine postulates that the (classical) Liouville operator L and the newly defined entropy operator M satisfy a commutation relation

$$-i(LM - ML) = D \leqslant 0.$$

D represents the rate of "microscopic entropy production"—a concept peculiar to Prigogine's theory.

Prigogine's account of *macroscopic* irreversibility differs radically from those offered by Boltzmann and his successors. Boltzman derived his H theorem from the postulate of molecular chaos. Modern kinetic theories of the kind pioneered by Bogoliubov and van Hove postulate that certain kinds of microscopic information (e.g., correlations) are absent at some initial moment or vanish in the limit $t \to -\infty$ (instead of being permanently absent, as Boltzmann assumed). Such assumptions introduce a preferred direction in time, namely, the direction away from the initial state. Prigogine's theory, by contrast, dispenses with assumptions about randomness: "*Both* randomness *and* irreversibility are consequences of the structure of the equations of motion."[3]

In this essay I shall describe an alternative theoretical framework that unites quantum mechanics, statistical thermodynamics, and modern kinetic theories of irreversible processes, and that reconciles Tisza's "two kinds of causality." The key to this reconciliation is the Strong Cosmological Principle, which asserts that in a complete description of the physical universe all points in space are indistinguishable, as are all spatial directions at a given point (section II). Supplemented by the assumption that little or no information was present in the initial state, this postulate provides a *historical* framework within which Tisza's notion of "selective specification of the initial state" (of a macroscopic system) acquires a concrete and explicit meaning. I shall argue that a *complete* description of an actual macroscopic system usually contains far less information than would be contained in a complete microscopic description of the same system (considered ahistorically). What kind of information about a given system actually is present depends on the system's history. Thus the question whether the initial state of a microscopic system satisfies the assumptions of a given kinetic theory of irreversible processes can in principle be settled by a suitable evolutionary calculation. For many processes, for example, heat conduction and molecular diffusion, in most natural contexts, very crude evolutionary considerations suffice to make a convincing case that the initial conditions guarantee the growth of entropy. On the other hand, the initial states of certain specially prepared systems that exhibit "antientropic" behavior (the spin-echo experiment is the classic example) contain enough information to invalidate the premises of relevant irreversibility theorems.

According to the theory to be described, certain kinds of information about physical systems are objectively present, at any given moment, and other kinds are objectively absent. Modern kinetic theories have yielded insight into the microscopic processes that *degrade* information in many-particle systems. But where does the information come from in the first place? I shall argue in section V that it is not present initially. In section VI I discuss the hierarchy of order-generating processes. I shall argue that distinct classes of order-generating processes are defined by distinct classes of initial conditions. Order-generating processes are hierarchically related because each one generates the initial conditions that define its successor. This scheme includes Prigogine's hierarchy of orderly configurations resulting from fluctuations in dissipative systems far from equilibrium, but is much more general. In particular, it includes processes that generate order in astronomical and biological contexts.

The present theory, unlike Prigogine's leaves the mathematical structure of quantum mechanics unchanged. It does, however, supply a new physical interpretation of quantal indeterminacy. The Strong Cosmological Principle implies that certain kinds of *positional information* are unobtainable in principle. I have previously[4] stated that this spatial indeterminacy is distinct from quantal indeterminacy—though its *existence* depends on the discreteness of quantal descriptions (section III). In section VII I argue that positional and quantal indeterminacy are actually facets of a single kind of indeterminacy. By virtue of the Strong Cosmological Principle, the universe may be viewed as the realization of a Gibbs ensemble. Thus, as Einstein believed must be possible, quantal statements about probabilities and averages can be translated into statements about spatial frequencies and averages. (On the other hand, a complete description of microscopic processes requires probability amplitudes, not just probabilities.)

Finally, I shall argue (section VIII) that the present theory resolves certain long-standing difficulties in the quantal theory of measurement. An ideal microscopic measurement creates a correlation between the eigenstates of the observable being measured and macroscopically distinguishable states ("pointer readings") of the measuring apparatus, without disturbing the state of the system on which the measurement is performed. And it leaves the system in one of these eigenstates. The theory described below gives a consistent and self-contained description of such processes.

II. The Strong Cosmological Principle

The indistinguishability of points in space and of directions at any given point are symmetry principles satisfied by all existing physical theories.

The *Strong Cosmological Principle* asserts that *in a complete description of the physical universe all points in space, and all spatial directions at a given point, are indistinguishable.*

This postulate implies that there is a unique decomposition of space-time into space plus time, and a preferred cosmic reference frame. (To an excellent approximation, this reference frame is the one in which the cosmic microwave background is everywhere isotropic.)

A weak form of the preceding assumption was put forward late in the sixteenth century by Giordano Bruno, and, in the following century, by Huygens and Newton, who postulated an infinite universe uniformly sprinkled with stars. Einstein used the assumption, which he called "the cosmological principle," in his 1917 paper on relativistic cosmology, as did Friedmann in his classic paper of 1922. Einstein and Friedmann considered simplified cosmological models characterized by a spatially uniform density and an everywhere isotropic velocity field. Observational cosmology rests on the somewhat stronger assumption, called by Hubble the "principle of uniformity," that the statistical properties of astronomical systems are everywhere the same at a given instant of cosmic time.

The Strong Cosmological Principle goes beyond these familiar cosmological assumptions. It asserts that a *complete* description of the universe contains no information that distinguishes between points in space or, at any given point, between spatial directions. In other words, the strong cosmological principle asserts that spatial homogeneity and isotropy are exact symmetries of the universe itself, as they are of the physical laws that govern the universe.

Is such an assumption reasonable? Not if one thinks of the universe as just another macroscopic system. For as Wigner has emphasized,[5] one expects the underlying regularities of physical phenomena to be expressed by laws and symmetry principles, not by auxiliary (initial, boundary, and symmetry) conditions, which serve to define the specific contexts in which the laws operate. It is this division between laws and auxiliary conditions that allows phenomena of unlimited diversity to be governed by a few simple and exact laws. If we assume that the auxiliary conditions defining a particular physical system—for example, a star—are simple, it is merely for the sake of mathematical convenience; it would be absurd to postulate that any real star has exact spherical symmetry.

But these considerations do not apply to the universe as a whole. It is true that separating the auxiliary conditions from the laws enables us to construct many different models of the universe. But only one of these can describe the actual universe. (Analogously, we are at liberty to test different physical laws. If we wished, we could assert that laws that do not hold in our own universe hold in some other universe.) Because we have access to only one universe, the auxiliary conditions that define it are no less unique than the physical laws that govern its behavior. And because the laws do not define a preferred position or direction in space, it seems reasonable to assume that neither do the auxiliary conditions.[6]

Without the Strong Cosmological Principle or some similar assumption, Einstein's theory of gravitation is, in an important sense, incomplete, for it does not satisfy Mach's requirement that the gravitational field be wholly determined by the distribution of mass. If we add the Strong Cosmological Principle to Einstein's theory, local inertial coordinate systems are determined, to a first approximation, by the large-scale distribution of mass, and local gravitational fields are determined by local nonuniformities in the distribution of mass, in accordance with Mach's principle.[7] Of course, Einstein's theory also enables one to calculate departures of the local inertial coordinate system from the coordinate system defined by the large-scale distribution of mass (dragging of the local inertial frame).[8]

III. Incompatibility between the Strong Cosmological Principle and a Classical Description

Consider a snapshot of a euclidean universe filled with classical particles. A complete description of the positional information in such a snapshot clearly distinguishes between positions in space. For example, the ratio AB/BC between the separations of any three particles A, B, C is a real number. There is zero probability that any other such ratio has the same value, because there is only a denumerable infinity of such ratios in an infinite universe of finite average number density. Thus the value of the ratio AB/BC is almost certain to specify a particular set of three particles, and hence to define a preferred position in space.

It follows that there must be preferred positions and directions in a universe that admits a classical description.

In addition, a statistical description of a classical universe is necessarily incomplete. The mutual separations of particles in the snapshot are specified by an infinite list of real numbers. The number of distinguishable lists of this kind having identical statistical properties (for example, lists describing a Poisson distribution of particles with a given mean number density) is nondenumerably infinite. Hence a statistical description of a classical universe can never be a complete description. There always exist microscopically distinguishable realizations of even the simplest statistical descriptions. Indeed the set of distinct realizations of a given statistical description is nondenumerable.

Thus the Strong Cosmological Principle cannot hold in a universe that admits a classical description.

IV. One-Dimensional Model Universes Satisfying the Strong Cosmological Principle

Consider a straight line divided into numbered cells of equal length. Suppose that each cell contains a particle in the spin state $|\psi\rangle = a|+\rangle + b|-\rangle$, where $|+\rangle$ and $|-\rangle$ are spin-up and spin-down state vectors, respectively, and $|a|^2 + |b|^2 = 1$. Assume that there are no interactions among the particles, so that the single-particle states are mutually independent. The state vector associated with the $2N + 1$ particle occupying the cells labeled $-N, -N + 1, \ldots, N$ is then

$$|\Psi_N\rangle = |\psi\rangle_{-N} \cdots |\psi\rangle_N = \sum_{k=0}^{2N+1} a^{2N+1-k} b^k \binom{2N+1}{k}^{1/2} |\Psi_N^k\rangle, \qquad (1)$$

where

$$\binom{2N+1}{k}^{1/2} |\,\Psi_N^k\rangle = \sum |+\rangle_{s_1} |+\rangle_{s_2} \cdots |+\rangle_{s_{2N+1-k}} |-\rangle_{t_1} \cdots |-\rangle_{t_k}. \tag{2}$$

In (2) the sum runs over all ordered sequences $t_1, \ldots, t_k$ of length k, where the t_i are integers in the range $(-N, N)$. Since there are $\binom{2N+1}{k}$ terms in the sum, the state vector $|\psi_N^k\rangle$ is normalized if the single-particle state vectors $|+\rangle, |-\rangle$ are.

The scalar product of (1) with its hermitian conjugate gives the identity

$$1 = \langle \Psi_N | \Psi_N \rangle = \sum_{k=0}^{2N+1} |a|^{2(2N+1-k)} |b|^{2k} \binom{2N+1}{k}. \tag{3}$$

For large N the dominant contribution to the sum in (3) comes from terms with k close to $|b|^2(2N+1)$, and as N increases the range of values of k that contribute appreciably to the sum gets smaller and smaller. In the limit $N \to \infty$

$$|a|^{2(N-k^*)} |b|^{2k^*} \binom{N}{k^*} \to 1, \qquad k^* \to N|b|^2, \tag{4}$$

where k^* denotes the value of k that makes the largest contribution to the sum in (3). This result (a weak form of the law of large numbers) was first derived, in a mathematically identical context, by James Bernoulli. From it we conclude that

$$|\Psi_N\rangle \to |\Psi^{[|b|^2(2N+1)]}\rangle, \tag{5}$$

where $[x]$ denotes the integer nearest x.[9]

The generalization of (5) to the case when ψ is a superposition of any finite (or denumerably infinite) set of mutually orthogonal states is straightforward. Thus if

$$\psi = \sum_k |k\rangle \langle k | \psi \rangle, \tag{6}$$

then

$$|\Psi_N\rangle \to |\Psi_N^{s_1 s_2 \cdots}\rangle, \tag{7}$$

where

$$s_k = [(2N+1)|\langle k|\psi\rangle|^2] \tag{8}$$

and

$$\binom{2N+1}{s_1 s_2 \cdots}^{1/2} |\psi_N^{s_1 s_2 \cdots}\rangle$$

is the sum of all products of single-particle state vectors in which s_1 are of type $|1\rangle$, s_2 of type $|2\rangle$, and so on.

Up to this point the argument has covered familiar ground. Let us now consider what happens to the state vector $|\Psi_N\rangle$ in the "limit" $N = \infty$. If instead of (1) we had considered states of the type

$$|\Psi'_N\rangle = |\Psi\rangle_1 \cdots |\Psi\rangle_N, \tag{9}$$

passage to the limit $N = \infty$ would be a continuous process.[10] With increasing N, the state vectors $|\Psi'_N\rangle$ would differ less and less in all of their properties from the limiting state vector $|\Psi'_\infty\rangle$.

But the state vector $|\Psi_\infty\rangle$, the "limit" of (1) as $N \to \infty$, refers to an *unbounded and statistically uniform* aggregate. This aggregate and the corresponding state vector have a property that is not shared by their finite counterparts, or even by their infinite but bounded counterparts (the aggregate of cells 1, 2, ... and the corresponding state vector): *The product state vectors that make up* $|\Psi_N^{s_1 s_2 \cdots}\rangle$ *become indistinguishable in the limit* $N = \infty$.

I have previously discussed this property of unbounded, statistically uniform aggregates in connection with the problem of thermodynamic irreversibility.[4] The state vectors that make up $|\Psi_N\rangle$ are indistinguishable in the limit $N = \infty$ because (a) in this limit they have identical statistical properties [see equations (7) and (8)], and (b) in a doubly unbounded sequence we can replace each label i by $i + n$, where n is any fixed integer. That is, the product state vectors composing $|\Psi_\infty\rangle$ are invariant under translations of the cell indices. These two properties ensure that it is impossible to exhibit any difference between doubly unbounded products descended from the distinct finite products that compose $|\Psi_N^{s_1 s_2 \cdots}\rangle$. In short, $|\Psi_\infty\rangle$ is a single product, which may be written

$$|\Psi\rangle_\infty = \cdots |k_{i-1}\rangle |k_i\rangle |k_{i+1}\rangle \cdots. \tag{10}$$

Here k_i denotes the ith label in some list $k_1, k_2, \ldots$, but the cell to which this label applies is not, and cannot be, specified. Thus $|k_i\rangle$ does not mean the same thing as $|k_i\rangle_i$, the state vector of the ith cell.

The state represented by (10) differs qualitatively from the state of any bounded or semibounded system. Because the state vector is a simple product of single-particle state vectors, every particle may be said to be in a definite state. But because the number attached to any given cell is purely arbitrary, we cannot know the location of a particle in a given state. Conversely, if we focus attention on a particular cell, we cannot know the state of its occupant.

V. Randomness and the Growth of Disorder in the Universe

The state vector space for any spatially bounded system has a discrete basis. So the preceding discussion of a one-dimensional universe applies to an infinite three-dimensional universe composed of statistically independent three-dimensional cells with identical statistical properties. Such cells can certainly be constructed if, in addition to the Strong Cosmological Principle, we assume that all correlation distances are bounded. Since the rate at which a correlation distance can grow cannot exceed the speed of light and the time during which it can grow cannot exceed the age of the universe, all correlation distances will be finite if the cosmic expansion began from a uniform state or from a nonuniform state in which all correlation distances (measured in mass units) were finite. Astronomical observations indicate that the present scale of local irregularities does not exceed about 3×10^8 light-years.

The assumption that the universe is spatially infinite is also consistent with current astronomical evidence. Relativistic cosmology relates the mean spatial curvature to the mean density of matter. The mean spatial curvature is positive, and space is finite, if the mean density of matter exceeds a certain critical value proportional to the square of the cosmic velocity/distance ratio (the Hubble constant). Indirect arguments suggest that the cosmic mass density is less than this critical value, implying that the mean spatial curvature is negative or zero and hence that space is infinite.

The mean spatial curvature is also related to the time that has elapsed since the beginning of the cosmic expansion (= the age of the universe). With increasingly large negative values of the spatial curvature, the age of the universe approaches the reciprocal of the velocity/distance ratio. This limiting age is currently estimated at 10^{10} years. If the universe has zero spatial curvature, its age is two thirds this value, and if the curvature is positive, the age is still smaller. The most reliable age estimates, based on relative abundances of radioactive isotopes, suggest that the age of the universe exceeds the value corresponding to zero spatial curvature, and thus imply that space is infinite.

Nevertheless, the possibility that space is finite cannot yet be definitely excluded. A finite universe obviously cannot satisfy the Strong Cosmological Principle. For example, it would contain a largest galaxy, whose center of mass would define a preferred position in space. To save the Strong Cosmological Principle, we would have to suppose that the universe is a particular realization of a Gibbs ensemble of universes whose properties are described by a theory that accords equal status to each member of the ensemble. Because a statistically uniform and isotropic universe of positive spatial curvature is finite in time as well as in space (it expands to a state of finite minimum density, then contracts to a final singular state of infinite density), we may imagine the members of the (discrete) ensemble of possible realizations as being contiguous in time. Of course, all realizations other than the one we inhabit are in principle unobservable. They are also causally disjoint, for there can be no causal connection between events separated by a true cosmological singularity. In these respects, distinct realizations of a finite universe are analogous to the finite, causally disjoint "observable universes" into which we may decompose any infinite universe. (Each "observable universe" is bounded by the event horizon of an observer at its center.) Thus the present theory is not tied to the assumption that space is infinite.

Let us now consider in a little more detail the link between a cosmological description satisfying the Strong Cosmological Principle and kinetic theories of irreversible processes in macroscopic systems. Such theories proceed from the assumption that certain kinds of information about a particular kind of macroscopic system—for example, information represented by the off-diagonal elements of a density matrix—are absent at some initial moment. Suppose we had a complete description of the universe itself at some initial moment. The laws of physics would then enable us (in principle) to find a complete description of the universe at any later moment of cosmic time. Such a description would tell us whether the conditions postulated by specific theories of irreversible behavior are satisfied by specific kinds of natural systems. What kind of information is present in a given class of natural systems will depend, of course, on what kind of initial conditions we postulate for the universe as a whole.

The argument up to this point enables us to attach objective meaning to the "selective specification of initial data" for a macroscopic system. But it does not yet tell us whether the data "normally" present in macroscopic systems are consistent with the practical success of chemical thermodynamics. For according to the present interpretation of thermodynamics, the validity of the second law reflects the universal absence of certain kinds of information (e.g., correlations) about the initial states of naturally occurring systems. Equally, it implies the *presence* of certain kinds of *macroscopic* information (e.g., temperature and concentration gradients, chemical disequilibrium). So we shall have succeeded in linking dynamics to statistical thermodynamics, via cosmology, if we can explain why certain kinds of information are regularly absent from the initial states of natural systems and why other kinds of information are regularly present.

The key to this question lies in the cosmic expansion. The rate of the cosmic expansion is proportional to the square root of the cosmic mass density. The rates per particle of particle interactions are proportional to at least the first power of the mass density. (In all cosmological models the temperature increases with increasing density.) Hence at sufficiently early times the expansion is much *slower* than any given particle reaction. Particle reactions, therefore, have plenty of time to degrade microscopic information, i.e., to convert "high-grade" information (expressed by one- and two-particle distribution functions) into "low-grade" information (expressed by higher-order distribution functions). It follows that *any microscopic information that might have been present initially would have been strongly degraded early in the cosmic expansion.*

Since the initial presence of microscopic information would have unobservable consequences, we may as well assume that it was not there to begin with. That is, we may reasonably postulate that the initial state admitted a conventional thermodynamical description in terms of macroscopic variables (mean density, temperature, lepton-baryon ratio, etc.). Notice that such an assumption makes sense *only* for the initial state (or, more precisely, in the limit $t \to 0$), because at later times microscopic information generated by entropic processes acting on macroscopic information (either present initially or produced by the cosmic expansion, as discussed in the next section) *must* be present.

The simplest assumption is that the initial state contains no information whatever—that it was a state of global thermodynamical equilibrium at zero temperature.[11]

The initial absence of microscopic information does not, of course, imply that microscopic information is absent in all macroscopic systems, or even in systems "untouched by human hands." For example, the single-particle distributions of photons and atoms in interstellar space are far from equilibrium. But the microscopic order that *is* present in macroscopic systems "derives" from macroscopic order. It is the kind of microscopic order that characterizes the distribution of perfume molecules in a room immediately after the bottle from which they emerged has been opened, rather than the kind of microscopic order that would allow the perfume

molecules to make their way back into the bottle. Such general statements are necessarily vague. But that is true of all historical generalizations. And, according to the present theory, the validity of the second law hinges on a historical generalization.

VI. Growth of Macroscopic Order

If you ask a physicist to name an irreversible natural process, he is likely to say "friction" or "molecular diffusion" or "heat conduction." To the same question a biologist is likely to reply "evolution" or "development" or "learning." To the physicist, "irreversibility" means the growth of entropy; to the biologist it is more likely to mean the growth of biological order. Present-day physics assigns a central place to the growth of entropy. The second law of thermodynamics asserts that all natural processes generate entropy. There are no analogous laws governing the growth of biological order and the growth of astronomical order. These processes are usually discussed piecemeal, under highly specific assumptions about initial conditions. In short, the growth of entropy seems to be a necessary feature of the world while the growth of order is merely contingent.

In the present theoretical framework, the growth of entropy and the growth of macroscopic order are complementary processes at essentially the same level of generality. According to the preceding discussion, the second law is an inference from (a) more fundamental laws that do not define a preferred direction in time (except for the breakdown of time reversal symmetry in certain weak interactions), (b) the Strong Cosmological Principle, and (c) an assumption about the cosmic initial state. The emergence and growth of macroscopic order are consequences of the same assumptions. Moreover, there exist general rules, analogous to the second law, governing the growth of information under certain broad kinds of initial conditions.

I define the information I associated with a statistical description by the formula

$$I = H_{max} - H, \tag{11}$$

where H is the entropy and H_{max} is the maximum value of H subject to given constraints. Some authors (e.g., Brillouin) define information simply as negentropy $-H$. The difference between this definition and the present one is important because H_{max} is not in general constant during information-generating processes. The quantity H_{max} represents *potential* information. It increases or decreases when the number of states accessible to a given system increases or decreases. For example, in biological evolution, the process of genetic variation may increase or decrease the quantity of potentional biological information present in a given population. Differential reproduction (natural selection), on the other hand, always diminishes the entropy H and hence always generates information. This is an example of a "general rule governing the growth of information." To make it precise, I would need to define the initial conditions more carefully and, more important, I would need to define "biological information." I shall not pursue these questions here, except to remark that the definition of information always depends on the theoretical context. In the context of evolutionary theory, biological information is information that serves to distinguish between genotypes that confer differing expectations of reproductive success.

Let us consider some physical examples of information growth.

1. *The cosmic expansion generates chemical information.* I define chemical information as information that specifies deviations of the relative abundances of chemical elements from the equilibrium values appropriate to the prevailing temperature and density. As I discussed in the last section, the rates of chemical reactions are proportional to a higher power of the cosmic density than the rate of the cosmic expansion. The cosmic expansion causes the temperature and density of the cosmic medium, and hence the equilibrium values of the chemical concentrations, to change at certain calculable rates. When the density is sufficiently high, the rates of key equilibrium-maintaining reactions are fast enough to maintain equilibrium. Eventually, however, the rates of these reactions must fall below the cosmic expansion rate. The relative concentrations of the reactants are then frozen in. These considerations are perfectly general; they hold for weak and nuclear reactions as well as for ordinary chemical reactions. The chemical composition of the universe is, of course, very far from equilibrium. The chemical disequilibrium of the sun is the source of the free energy on which terrestrial life depends. This illustrates how one kind of order-generating process (the cosmic expansion, leading to chemical disequilibrium) can give rise to initial conditions in which a qualitatively different order-generating process (biological evolution) can occur.

2. *The cosmic expansion generates "morphological" information (the information needed to specify the nonuniformity of the mass density).* The truth of this assertion is obvious under the assumption that the cosmic mass density was uniform in the limit $t \to 0$. This assumption characterizes cosmological theories that postulate a cold initial state.[12] Cosmologies that postulate a hot initial state must also postulate initial nonuniformities. But these presumably[13] contain less information than the cosmic mass distribution that evolves from them.

3. *The evolution of an isolated system composed of a large number of gravitating particles generates information.* In such a system the central density and temperature increase steadily, while the peripheral regions expand and become less dense. Thus a system of this kind evolves *away* from the maximum-entropy state appropriate to its energy, mass, *and radius*. A spherical system of gravitating particles confined by a reflecting spherical wall will evolve toward a stable equilibrium configuration if the ratio of the central density to the surface density in this configuration is less than a certain critical value. If the ratio exceeds this value, the equilibrium configuration is unstable and the core will continue to collapse indefinitely.

Figure 1 illustrates the hierarchical relations among order-generating processes and the structures and initial conditions they generate. Each order-generating process creates a new class of initial conditions and structures, which provide the setting for a new class of order-generating processes. The figure is not complete. For example, it omits the sub-hierarchy of dissipative structures.

VII. Spatial Interpretation of Quantal Indeterminacy

The state of any finite portion of the universe is in general a superposition of eigenstates of a given observable. On the other hand, as we have seen, the state of the universe itself is a direct product of eigenstates, each associated with a cell whose dimensions exceed all correlation scales. This product represents a state in which each eigenstate is realized somewhere; but, by virtue of the Strong Cosmological Principle, it is impossible to say where. Thus the quantal indeterminacy represented by the statement "It is impossible (in general) to specify which eigenstate of a particular observable a given cell is in" is equivalent to the statement "It is impossible to say *where* a given eigenstate is realized." Moreover, the probability that a given eigenstate will be realized in a given cell is equal to the frequency with which that eigenstate is realized in the infinite assembly of cells that constitutes the universe. Transition probabilities also admit a spatial interpretation. Thus the statement that a certain transition has probability p of occurring in a certain time interval means

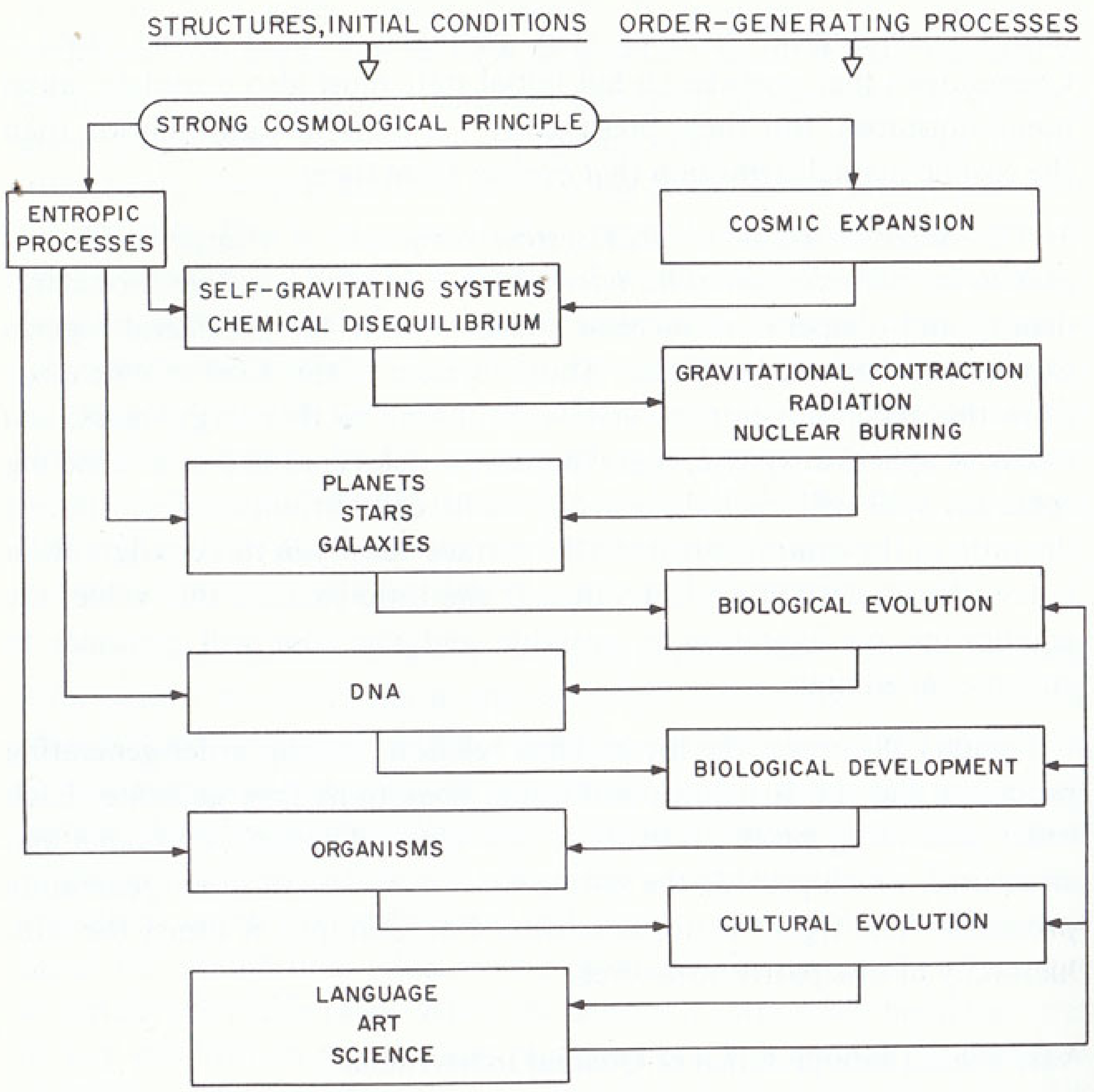

Figure 1

that during this time interval this transition occurs in a fraction p of the cells into which we divide the universe (or in a fraction p of the system's identical copies distributed throughout the universe).

The present interpretation of quantum mechanics may remind some readers of the "many-worlds" interpretation,[14] in which the "universe is constantly splitting into a stupendous number of branches, all resulting from the measurementlike interactions between its myriads of components."[15] In the present interpretation, the many worlds are causally disjoint regions of a single spatiotemporal continuum. In the many-worlds interpretation, each of the many worlds is equipped with its own space-time. The multiplicity of realizations in the present interpretation is not postulated ad hoc, as it is in the many-worlds interpretation, but is a consequence of the Strong Cosmological Principle and the assumption of initially bounded correlations. These metaphysical and methodological distinctions are comparatively unimportant, however. The many-worlds interpretation was put forward as a solution to the problem of measurement. In the next section we shall see that the present interpretation of quantum mechanics offers a radically different solution to that problem.

VIII. Measurement

According to Bohr and to most present-day writers, classical notions must figure in any careful formulation of the basic principles of quantum mechanics. As Landau and Lifshitz remark in their well-known textbook,[16] "Quantum mechanics occupies a very unusual place among physical theories: it contains classical mechanics as a limiting case, yet at the same time it requires this limiting case for its own formulation."

Since classical and quantal notions are mutually inconsistent, it is perhaps not surprising that the quantal theory of measurement, which treats interactions between classical and quantal systems, should still, after more than half a century, be a source of conjecture and controversy.

To fix ideas, let us review the main ideas of the theory, as formulated by von Neumann.[17] Let $|k\rangle$ denote the kth eigenket of an observable Q, and let $|\bar{k}\rangle$ denote the corresponding eigenket of an apparatus designed to measure Q. A measurement of Q must yield one of its eigenvalues, and there is a definite probability p_k that it will yield the kth eigenvalue. Thus the state of the system and the apparatus after a measurement of Q is represented by the density matrix

$$\hat{\rho} = \sum_k |\hat{k}\rangle|k\rangle p_k \langle k|\langle\hat{k}|. \qquad (12)$$

But if the system was in the state ψ before the measurement and if the measurement did not disturb it, the principle of superposition implies that the state of the system and the apparatus after the measurement is represented by

$$\sum_k |k\rangle|\hat{k}\rangle\langle k|\psi\rangle$$

with the corresponding density matrix

$$\rho = \sum_{k,k'} |k'\rangle|\hat{k}'\rangle\langle k'|\psi\rangle\langle\psi|k\rangle\langle\hat{k}|\langle k|, \qquad (13)$$

which is not the same as (12) unless the initial state was an eigenstate of Q.

Four ways of dealing with this contradiction have been widely discussed.

1. Von Neumann[17] postulated that measurements and measurementlike processes are not governed by the laws of quantum mechanics but do in fact convert pure states into mixtures. This postulate represents a kind of supplementary law serving to join classical and quantum mechanics.

2. Wigner[18] suggested that the laws of quantum mechanics apply to all physical processes, including measurement, but not to certain mental processes, namely, those in which an observer becomes aware of the outcome of a measurement. This interpretation preserves the unity of physics but demands a radical separation between physical and mental processes. Wigner has defended this separation on philosophical grounds.[19]

3. The many-worlds interpretation[14] postulates that the "world"—or, more accurately, an ensemble of equally "real" worlds—is described by "a universal wave function" evolving according to the laws of quantum mechanics. According to this view, mixtures do not exist. Thus the contradiction disclosed by the orthodox theory of measurement does not arise. The many-worlds interpretation explains the observation that microscopic measurements have definite (though unpredictable) outcomes by positing that every such measurement splits the world in which it occurs into a number of distinct worlds, corresponding to individual branches of the state vector.

But this "explanation," it seems to me, ignores, rather than resolves, the difficulty. If the "measurementlike" process is reversible, as in Wigner's example of the double Stern-Gerlach experiment,[18] the many-worlds interpretation must be false because the individual branches of the state vector do not then evolve independently. But distinct branches of the state vector are never strictly independent. Hence the evolution of the "universal wave function" can never be accurately represented by a splitting of worlds.

Of course, the *intent* of the interpretation was that only irreversible branching processes should produce a splitting of worlds. But, as Everett himself clearly perceived, there is no room in the many-worlds interpretation for the notion of (macroscopic) irreversibility:

> This type of irreversibility [thermodynamical irreversibility] . . . arises from a failure to separate "macroscopically indistinguishable" states into "true" microscopic states. It has a fundamentally different character from the irreversibility of Process 1 [a "discontinuous change brought about by the observation of a quantity with eigenstates $\phi_1, \phi_2, \ldots$, in which the state ψ will be changed to the state ϕ_j with probability $|(\psi,\phi_j)|^2$"], which applies to micro-states as well and is peculiar to quantum mechanics. Macroscopically irreversible phenomena are common to both classical and quantum mechanics, since they arise from our incomplete information concerning a system, not from any intrinsic behavior of the system.[20]

Since the many-worlds interpretation does not apply to reversible "measurements" like the double Stern-Gerlach experiment, and ordinary thermodynamical irreversibility is not an admissible criterion, there would seem to be no objective way (that does not make a tacit appeal to classical concepts) of deciding when or whether a splitting into multiple worlds has occurred.

4. The fourth and now perhaps most widely held view is that measurements and measurementlike processes do not, as von Neumann postulated, transform pure states of the system plus apparatus into mixtures, but that they are *irreversible* (in the ordinary thermodynamical sense).[21] Supporters of this interpretation argue that thermodynamically irreversible processes give rise to pure states that are *practically* indistinguishable from mixtures: The branches into which the state vector is split by an irreversible process do not subsequently interfere, at least in practice. For example, the time required to bring about measurable interference between macrosocopically separated branches of the state vector corresponding to distinct "pointer readings" may exceed the age of the universe. I shall call this the reformed view, to contrast it with the orthodox view.

According to the reformed view, the laws of quantum mechanics apply to measurement as they do to all other physical processes, macroscopic as well as microscopic. What, then, distinguishes irreversible from reversible processes?

Within the conventional framework of quantum mechanics, the answer to this question must be that given by Peres and Rosen: "As long as experiments can be performed in which interference effects may show up, then [the state vector] is a superposition. It becomes a mixture beginning from the stage at which such experiments become inconceivable. The striking feature of this approach is that the determination of the nature of [the state vector] . . . has a certain subjective aspect: A poorly equipped physicist may interpret it as a mixture, while a better endowed one might still be able to display interference effects."[21]

The key idea in the reformed theory of measurement is to replace the *postulate* that measurements *destroy* information about an isolated physical system by the *definition* of measurement as an *irreversible* process that does not disturb the state of the system being measured and creates a correlation between eigenstates of the observable being measured and macroscopically distinguishable states of the apparatus. Irreversible processes do not contravene the laws of quantum mechanics, which imply that the microscopic entropy of an isolated system is constant in time, because they merely *redistribute* information. The theory sketched in the preceding sections of this paper gives an objective meaning, compatible with the laws of quantum mechanics, to such thermodynamical notions as irreversibility and the redistribution of information. According to this theory, the state vector of a complex system that has not been specially prepared is in general a true mixture. The fact that a poorly equipped physicist can acquire less information about the state vector of a given system than his better-equipped colleague has no more to do with the quality and quantity of the information objectively present in the system than the ability of a chemist to analyze a given sample has to do with the sample's chemical composition.

It may lend concreteness to the preceding remarks to consider in a little more detail the connection between modern quantal theories of thermodynamic irreversibility and the Strong Cosmological Principle. Let ρ denote the density matrix of a dynamical system. The Gibbs entropy of the system is

$$H_G = -k\,\mathrm{Tr}\{\rho \ln \rho\}, \tag{14}$$

where k denotes Boltzmann's constant and $\mathrm{Tr}\{\ \}$ denotes the diagonal sum (which is the same in all representations). In the diagonal representation

$$\rho = \sum_n |n\rangle p_n \langle n|, \tag{15}$$

so that

$$H_G = -k \sum_n p_n \ln p_n. \tag{16}$$

In another representation, labeled by the quantum numbers α, the diagonal elements of ρ still represent occupation probabilities,

$$\langle\alpha|\rho|\alpha\rangle = \sum_n |\langle\alpha|n\rangle|^2 p_n = p_\alpha, \tag{17}$$

but the Gibbs entropy [whose value is still given by (16)] depends also on the off-diagonal matrix elements of ρ and on the matrix elements of $\ln \rho$:

$$H_G = -k\,\mathrm{Tr}\{\rho \ln \rho\} = -k \sum_{\alpha,\alpha'} \langle\alpha|\rho|\alpha'\rangle\langle\alpha'|\ln \rho|\alpha\rangle. \tag{18}$$

We may, however, associate a different kind of entropy,

$$H_B = -k \sum p_\alpha \ln p_\alpha, \tag{19}$$

with the diagonal part of the density matrix in the α representation. H_B is a generalization of the Boltzmann entropy and, under appropriate conditions, to be discussed, satisfies an H theorem. The Gibbs entropy, on the other hand, is constant in time.

Because the function $-x \ln x$ is convex, it follows from (17) and a well-known inequality for convex functions that

$$H_B \geqslant H_G, \tag{20}$$

with equality only if the density matrix is diagonal in the α representation. We may therefore write

$$H_G = H_B - I_\phi, \tag{21}$$

where I_ϕ is a positive quantity. It represents the information associated with the relative phases of the amplitudes $\langle\alpha|n\rangle$.

We define the information associated with the probability distributions $\{p_n\}$ and $\{p_\alpha\}$ in the usual way:

$$I = H_{max} - H, \tag{22}$$

where H_{max} denotes the maximum value of the entropy consistent with given constraints. Then (21) becomes

$$I_G = I_B + I_\phi, \tag{23}$$

which expresses the Gibbs (total) information as the sum of information associated with the occupation numbers of the states α and information associated with the relative phases of the amplitudes $\langle\alpha|n\rangle$.

The quantal H theorem, due essentially to van Hove,[22] states that the entropy H_B is a nondecreasing function of the time if (a) $I_\phi = 0$ initially, and (b) the interaction Hamiltonian of the system satisfies certain mathematical conditions. Van Hove showed that condition (a) is necessary and sufficient for the set of probabilities $\{p_\alpha\}$ to be related to their initial values by a generalized (non-Markoffian) master equation.[23] In other words, if information about the relative phases of the basis vectors is initially absent, it is permanently irrelevant to the evolution of the occupation probabilities represented by the diagonal elements of the density matrix.

In the present context it is the first condition, the initial vanishing of the off-diagonal elements of the density matrix, that concerns us. The assumption raises two questions: What dictates the choice of the α representation (in which $\langle\alpha|\rho|\alpha'\rangle = \delta(\alpha,\alpha')p_\alpha$) for a given physical system? And what objective meaning, if any, attaches to the assumption that in this representation information about the relative phases of the basis vectors is initially absent?

The present theory answers these questions. As we saw in section V, a complete description of the universe could in principle contain a very small quantity of information. The subsequent evolution of the universe generates statistical information (section V). At the same time, information is continually flowing (irreversibly) from low-order correlations, where it is present initially, to higher-order correlations.

Thus it is possible, at least in principle, to calculate the cosmic density matrix at any moment, and hence to decide whether a given set of necessary conditions for irreversible information flow are actually satisfied by a given physical system. Moreover, the absence or presence of specific kinds of information is an objective property of the universe. Whether or not a given system exhibits irreversible behavior does not depend on what we know about it but on its history.

According to the present theory, the universe is in a mixed state, described not by a state vector but by a density matrix that contains only the kinds of information present in the initial data or generated during the course of cosmic evolution. The form of this density matrix exhibits the essential unity of quantum mechanics and thermodynamics. Statistical thermodynamics may be defined as the study of the statistical aspect of the cosmic density matrix.

IX. Quantum Mechanics and Reality

In 1949 Einstein summarized years of thought on the physical interpretation of quantum mechanics in the following words: "One arrives at very implausible theoretical conceptions, if one attempts to maintain the thesis that the statistical quantum theory is in principle capable of producing a complete description of an individual physical system. On the other hand, those difficulties of theoretical interpretation disappear, if one views the quantum-mechanical description as the description of ensembles of systems."[24]

The conclusions of the present paper strongly support this view of quantum mechanics. But the present theory goes further. It offers an explanation of why quantum mechanics cannot give a complete description of individual physical systems: If a complete description of the universe were to contain complete descriptions (in Einstein's sense) of individual systems, it would necessarily define preferred positions in space. A theoretical description that is invariant under spatial translation and rotation is necessarily a quantal description (a classical description would have preferred positions and directions) and necessarily describes ensembles rather than individual systems. The statistical character of the description is a consequence of its symmetry; if God plays dice, it is because He doesn't play favorites.

Notes

1. L. Tisza, *Synthèse* 14:110–131 (1962); *Rev. Mod. Phys.* 35:151–185 (1963). Both papers are reprinted in *Generalized Thermodynamics* (MIT Press, 1966).
2. I. Prigogine, *From Being to Becoming* (Freeman, 1980).
3. *Op. cit.* p. 176.
4. D. Layzer, *Astrophys. J.* 206:559–569 (1976); *Sci. Am.* 233[6]:56–69 (1975).
5. E. Wigner, "Events, laws of nature, and invariance principles," in *Symmetries and Reflections* (MIT Press, 1970).
6. But this argument cuts two ways, as Abner Shimony has kindly pointed out to me. We could equally well postulate that spatial uniformity and isotropy are exact symmetries of *neither* the cosmic auxiliary conditions *nor* the laws of physics. Perhaps, as C. S. Peirce and, more recently, Philip Morrison and others have suggested, the laws of physics are themselves evolving. Then symmetry principles might be analogous to thermodynamic equilibrium: ideal limits that are approached ever more closely but never actually attained.

While this is an attractive hypothesis, it is, in its present stage of development, less simple than the Strong Cosmological Principle, and hence more difficult to refute.

7. D. Layzer, "Cosmogonic processes," in *Astrophysics and General Relativity*, vol. 2 (Gordon and Breach, 1971).

8. Mach's principle is also satisfied in any finite universe, whether or not it satisfies the cosmological principle. On both theoretical and observational grounds, the Strong Cosmological Principle seems more attractive than the assumption that the universe is finite.

9. This result, for the mathematically equivalent case of a time series of independent identical measurements, was stated (but not proved) by H. Everett, *Rev. Mod. Phys.* 29:454–462 (1967). Proofs were subsequently supplied by J. Hartle, *Am. J. Phys.* 36:704 (1968) and R. N. Graham, PhD thesis, U. of N. Carolina (1970), neither of whom mentioned the connection with Bernoulli's law of large numbers.

10. See J. Hartle, reference in note 9 for details.

11. For a discussion of the astronomical consequences of this assumption, see D. Layzer, *Cosmology* (forthcoming).

12. See D. Layzer, "Galaxy clustering: Its description and its interpretation," in *Galaxies and the Universe*, ed. Sandage et al. (The University of Chicago Press, 1975); D. Layzer, "The structure of matter and the structure of the astronomical universe," in *International Journal of Quantum Chemistry* (1977); D. Layzer, *Cosmology* (forthcoming).

13. "Presumably" because a consistent theory for the growth of density fluctuations in an initially hot universe is still lacking.

14. H. Everett, III, *Rev. Mod. Phys.* 29:454 (1957); see also the articles collected in *The Many-Worlds Interpretation of Quantum Mechanics*, ed. B. S. DeWitt and H. Graham (Princeton University Press, 1973).

15. B. S. DeWitt, "Quantum mechanics and reality," reprinted in *The Many-Worlds Interpretation of Quantum Mechanics*, ed. B. S. DeWitt and N. Graham (Princeton University Press, 1973), p. 161, from *Physics Today*, 23[9] (September 1970).

16. L. D. Landau and E. M. Lifshitz, *Quantum Mechanics* (Pergamon, 1958).

17. J. von Neumann, *Mathematical Foundations of Quantum Mechanics*, translated by R. P. Beyer (Princeton University Press, 1955).

18. E. Wigner, *Am. J. Phys.* 31:6 (1963).

19. E. Wigner, "Remarks on the mind-body question" and "Two kinds of reality," in *Symmetries and Reflections: Scientific Essays of Eugene P. Wigner* (MIT Press, 1967), pp. 171–184.

20. H. Everett, III, "Theory of the universal wave function," in *The Many-Worlds Interpretation of Quantum Mechanics*, ed. B. S. DeWitt and N. Graham (Princeton University Press, 1973), p. 99.

21. See, for example, A. Peres and N. Rosen, *Phys. Rev.* 135b:1486 (1964); K. Gottfried, *Quantum Mechanics* (NY, 1966), vol. 1, pp. 165–189; L. N. Cooper and D. van Vechten, *Am. J. Phys.* 37[12]:1212–1220 (December 1969).

22. L. van Hove, *Physica* 21:517 (1955).

23. R. Zwanzig, *J. Chem. Phys.* 33:1330 (1960), has given a simpler and slightly more general derivation of this result, applying projection operators to the quantal analogue of Liouville's equation.

24. A. Einstein, "Reply to criticisms," in *Albert Einstein: Philosopher-Scientist*, vol. 2, ed. P. A. Schilpp (Open Court, 1969), p. 671.