FREE WILL AS A SCIENTIFIC PROBLEM

David Layzer

I. Philosophy and the "basic facts"

According to John Searle[1], "the overriding question in contemporary philosophy" is "How do we fit in?" How can we reconcile "a conception of ourselves as conscious, intentionalistic, rational, social, institutional, political, speech-act performing, ethical and free will possessing agents" with "the basic facts," our present "reasonably well-established conception of the basic structure of the universe." As an example of the tension between our self-conception and the basic facts, Searle cites the problem of freedom and determinism. He argues that free will is a fact of conscious experience and that consciousness is a biological phenomenon, a "higher-level biological feature of the brain" (p. 48). If free will is not an illusion, physical laws and the antecedent state of a deliberator's brain cannot determine the outcome of a deliberative process. It follows that "consciousness is a feature of nature that manifests indeterminism." Up to this point I agree with the argument. But now Searle appeals to what he considers to be one of the "basic facts": "[W]e know that quantum indeterminism is the only form of indeterminism that is indisputably established as a fact of nature" (p. 74). So, Searle concludes, "consciousness manifests quantum indeterminism" (p. 75).

Although quantum indeterminism is indeed a fact of nature – a feature of our present scientific description of the world that future advances seem highly unlikely to change – its origin remains controversial. There is no settled view about how the world that classical macrophysics describes – a refined version of the world of experience – is related to quantum microphysics. This paper sketches a novel scientific approach to this issue and discusses its implications for determinism, the nature of free will, and the relations between physics and biology and between biology and consciousness. One of my conclusions will be that for reasons that have little to do with quantum indeterminism we have the capacity to shape the future through our choices, plans, and actions.

II. What is quantum indeterminism?

Quantum indeterminism is sometimes thought to imply that processes governed by the laws of quantum mechanics have indeterminate outcomes. In fact, quantum mechanics' law of change, Schrödinger's equation, is deterministic in exactly the same sense as Newton's second law of motion: both laws determine how isolated systems – idealized physical systems that do not interact with the outside world – change with time. Although quantum mechanics and classical mechanics describe physical states differently, they agree that the state of an isolated system at any given moment determines its state at any subsequent moment. Yet quantum mechanics also *predicts* that certain processes, such as radioactive decay, have indeterminate outcomes. How does quantum mechanics reconcile the deterministic character of its law of change with its prediction that the lifetime of an unstable atomic nucleus is unpredictable in principle?

The quantum description of the history of an isolated radium nucleus is indeed deterministic. But the decay of an isolated radium nucleus – an event in the history of an isolated radium nucleus – is unobservable. What *is* observable is the result of an interaction between one of the decay products (in this example, a radon nucleus and a helium nucleus, or alpha particle) and a macroscopic detector whose construction has been specified in the language of classical physics. Quantum mechanics predicts that the *measured* lifetime of a radium nucleus has different values on different occasions, and it predicts the probability distribution of measured lifetimes. Experiments confirm these predictions.

Quantum indeterminism manifests itself only in processes of this kind: processes in which a microscopic system initially in a definite quantum state interacts with a classical measuring apparatus designed to measure one of the microscopic system's properties. A quantum measurement leaves the measured system in one of several possible quantum states and leaves the apparatus in a correlated classical state. For example, each possible post-measurement state of the quantum system might be correlated with the position of a pointer. Quantum mechanics predicts both the possible outcomes of such a measurement and their probabilities.

Could measurement-like processes in the brain mediate exercises of free will? Free will, as defended by Immanuel Kant, William James, Robert Kane, and John Searle, is more than the ability to make unforced choices between given alternatives. It is the capacity to shape future events. Our present understanding of neurophysiology offers little support for the view that measurement-like processes in the brain underlie choices that are neither forced nor random. Must we then conclude that our felt capacity to shape the future through our free choices is an illusion? This dilemma invites a closer look at how physicists describe quantum indeterminism.

III. The standard account of quantum indeterminism

Physical theories are self-contained mathematical structures linked to the results of possible measurements by auxiliary rules. Newtonian mechanics represents a physical system's measurable properties by real variables (i.e., mathematical objects whose possible values are real numbers). The auxiliary rule that links these variables to the outcomes of possible measurements identifies the value of each variable either with the result of an idealized measurement or with the result of a calculation that expresses the result of the measurement in terms of measurements of other variables such as position and elapsed time. In this respect, Newtonian descriptions are refined versions of descriptions in ordinary language.

The mathematical structure of quantum mechanics is less closely tied to experience. Quantum mechanics represents an isolated system's possible physical states by "state vectors," vectors of unit length in an abstract multidimensional vector space. It represents the system's measurable properties by operators, mathematical objects that act on vectors in this space. The standard formulation of quantum mechanics[2] contains a rule that links the mathematical description of an isolated physical system to the results of possible measurements of the system's properties. This rule, the *measurement postulate*, equates the average of "a large number" of measurements of a given property to a quantity that depends on the operator that represents the measured property and the state vector that represents the state of the measured system. From this rule and quantum mechanics' mathematical formalism one deduces: (a) that a single measurement does not in general have a definite outcome; (b) the set of possible outcomes; and (c) the probability associated with each possible outcome (or range of possible outcomes). Thus quantum indeterminacy is a consequence not of the mathematical formalism alone but of the formalism plus a rule that links the mathematical formalism to the results of (ideal) measurements.

Experiments leave no room for doubt that the standard formulation of quantum mechanics is correct. As an instrument for making predictions about the outcomes of measurements, the standard formulation leaves nothing to be desired. But many physicists have sought a deeper account of the linkage between quantum physics and classical physics than that provided by the measurement postulate. This account would not just postulate that a composite system consisting of a quantum system coupled to a classical measuring apparatus evolves indeterministically. It would explain why. The account that emerges from the following discussion implies that contrary to conventional opinion, quantum indeterminism is not the only form of indeterminism. A variety of macroscopic processes, I will argue, have indeterminate outcomes; chance is endemic in the macroscopic domain.

IV. Quantum mechanics and classical physics

Quantum mechanics and classical mechanics are closely related. Not only do classical properties like position, momentum, and energy have quantum counterparts. At a deep formal level the laws that govern classical properties and their rates of change are identical with the laws that govern their quantum counterparts. Moreover, the domains in which classical mechanics and quantum mechanics are valid overlap. For example, the classical description of an electron circling a proton approximates the quantum description of an electron in a hydrogen atom when the electron’s angular momentum (the product of its momentum and the radius of its circular orbit) greatly exceeds Planck’s constant. The formal similarity between quantum and classical mechanics and the overlap between their domains have motivated many attempts over many years to formulate quantum mechanics in a way that includes classical mechanics as a limiting case. That these attempts have not yet been completely successful is due largely to a central feature of quantum mechanics: the principle of superposition.

To illustrate the principle, consider a free electron. Because its position coordinates are represented by operators in an abstract vector space, they do not have definite values at a given moment. The electron could, however, be in a state in which its *measured* position was almost certain to lie inside a small sphere S. It could also be in a different state, in which its measured position was almost certain to lie inside a different, non-overlapping sphere S′ arbitrarily distant from S. Call the state vectors that represent these two states *U* and *V*. The principle of superposition says that any linear combination of these state vectors, $aU + bV$, represents a possible state, where *a* and *b* are complex numbers whose squared magnitudes add up to 1. States represented by such state vectors have no classical analogue; a point-like particle cannot be in two non-overlapping regions at the same time. Of course, quantum mechanics (which includes the measurement postulate) does not claim that they can be. It predicts (and experiments confirm) that when an electron is in a superposition of the states *U* and *V* its measured position will *either* be a point in S *or* a point in S′. It also predicts the probabilities of these outcomes: they are the squared magnitudes of the coefficients *a* and *b*.

The domain of quantum mechanics has no clearly defined boundary. Nothing in the mathematical formalism indicates that it applies only to small or simple systems. Let us therefore assume, as most physicists do and as I will do, that the domain of quantum mechanics includes macroscopic systems. In other words, let us assume that an isolated macroscopic system has a set of possible quantum states and that it is in one of these states. Now assume that the measuring apparatus in a quantum measurement, which is necessarily a macroscopic system, is initially in a definite quantum state and that the combined system (measured system + measuring apparatus) remains isolated during the measurement. Then, as John von Neumann, showed in 1932, an ideal measurement would cause the state vector of the combined system to evolve into a superposition of state vectors each of which represents one of the measurement's possible outcomes as given by the measurement postulate; and the probability of each outcome, as given by the measurement postulate, would coincide with the squared magnitude of the coefficient of the corresponding state vector in the superposition.

But a quantum measurement does not, as this account predicts, produce a superposition of outcome states. It produces just one of them. To bring his account into agreement with the measurement postulate (and experiment), von Neumann postulated that the predicted superposition of outcome states no sooner forms than it collapses unpredictably onto one of them, with a probability given by the squared magnitude of the coefficient of the corresponding state vector in the superposition. (This hypothetical process is called the collapse, or reduction, of the state vector.)

Now, superpositions do not collapse in other physical contexts. Every quantum state of a composite system AB made up of interacting quantum systems A and B is a superposition of states in each of which A is in a definite quantum state and B is in a correlated quantum state; and such superpositions never collapse. They are a ubiquitous (as well as distinctive) feature of quantum descriptions. Their existence has been experimentally confirmed on innumerable occasions. What, then, distinguishes quantum measurements from other physical processes governed by quantum laws? Here are three answers that have been offered by physicists:

1. Eugene Wigner[3] once argued that what distinguishes quantum measurements from other physical processes is that a quantum measurement necessarily involves consciousness, for it is completed only when an observer becomes aware of its outcome. And it is at that point that our knowledge of the state of the combined system (measured system + measuring apparatus) changes in a way not governed by Schrödinger's deterministic law of change:

> In other words, the impression which one gains at an interaction [between an observer and a physical system], called also *the result of an observation*, modifies the wave function of the system. The modified wave function is, furthermore, in general unpredictable before the impression gained at the interaction has entered our consciousness: it is the entering of an impression into our consciousness which alters the wave function ... . It is at this point that consciousness enters the theory unavoidably and unalterably (pp. 175-6).

Wigner regarded the argument summarized in the preceding passage as the weaker of two arguments "support[ing] the existence of an influence of … consciousness on the physical world" (p. 181)." The stronger argument "is based on the observation that we do not know of any phenomenon in which one subject is influence by another without exerting an influence thereupon." I will return to the question of how consciousness fits into a scientific description of the world later in this essay.

2. Wigner[4] later defended an instrumental interpretation of quantum mechanics.

> It appears that the statistical nature of the outcome of a measurement is a basic postulate, that the function of quantum mechanics is not to describe some "reality," whatever this term means, but only to furnish statistical correlations between [an observation and] subsequent observations. This assessment reduces the state vector to a calculational tool, an important and useful tool, but not a representation of "reality."

Many if not most contemporary physicists agree with this view. It provides a philosophical justification for the standard formulation, preempting the question "Why does the state vector collapse?"

3. Hugh Everett III[5] postulated that the domain of quantum mechanics includes all physical systems up to and including the physical universe. He argued that we should identify the measuring apparatus in a quantum measurement with the universe minus the measured system. Von Neumann's account then predicts that a quantum measurement creates a superposition of quantum states of the universe. But since quantum mechanics is universally valid, this superposition never collapses. Its components represent macroscopically distinguishable, coexistent states of the universe, all equally real.

Several contemporary cosmological theories incorporate Everett's assumption that quantum mechanics is universally valid. All of these theories raise Searle's question "How do we fit in?" with a vengeance. For in that question "we" now means not just "we humans" but more broadly "the world that classical physics describes." So far as I know, that question has not yet been satisfactorily answered. We lack a physical theory that postulates the universal validity of quantum mechanics and contains classical physics and general relativity, the classical (and strongly confirmed) theory of the physical universe and of the structure of space-time, as limiting cases.

V. Microphysics and macrophysics: statistical mechanics

The measurement postulate and von Neumann's collapse postulate serve in different ways to bridge the gap between microphysics and macrophysics. The gap itself predates quantum mechanics by almost two and a half centuries. In the *Principia* Newton, a convinced atomist, proposed a molecular model of air to account for Robert Boyle's empirical law relating the pressure and the volume of an enclosed sample of air. He attributed the fact that an isolated sample of air expands to fill any enclosure, no matter how capacious, to a hypothetical repulsive force between neighboring air molecules.

In 1738 Daniel Bernoulli proposed a much simpler molecular model. He assumed that the hypothetical air molecules travel freely between relatively short-lived collisions. When a molecule bounces off a wall it transfers momentum to the wall, in accordance with Newton's second and third laws of motion. Averaged over many molecular impacts and over a macroscopic time interval, the result is a steady pressure. This model, like Newton's, predicted that the pressure a gas sample exerts on the walls of its container is inversely proportional to its volume. But it does more. If one identifies the average kinetic energy of a gas molecule with the gas temperature, as was done a century later, Bernoulli's formula for gas pressure includes the remaining empirical gas laws as well as Avogadro's hypothesis.

These successes were gratifying – but also, in one respect, surprising. A macroscopic sample of air has a vast number of microscopic degrees of freedom – six for every molecule in the sample. Yet it seems to have only two macroscopic degrees of freedom: experiments show that the values of two macroscopic properties of the sample, such as its temperature and its pressure, determine all of its measurable properties. In two papers, published in 1860 and 1866, James Clerk Maxwell explained why. He argued that molecular collisions, governed by Newton's laws of motion, cause the distribution of molecular velocities in an isolated gas sample to evolve toward a unique equilibrium distribution that depends on a single parameter, the average molecular kinetic energy. Rudolf Clausius and others had earlier identified this quantity with the gas temperature. Experiments show that a gas sample that is well insulated from its surroundings does indeed quickly relax into a state whose macroscopic properties are averages of appropriate molecular properties over the Maxwell distribution of molecular velocities.

In 1872 Ludwig Boltzmann extended Maxwell's theory. Maxwell had studied the distribution of molecular velocity in a uniform gas sample; Boltzmann studied the joint distribution of molecular position and momentum in (possibly) nonuniform gas samples, and derived a mathematical law – his transport equation – that governs changes in this distribution produced by molecular collisions. Boltzmann also discovered a statistical counterpart of entropy[6] and proved that in a uniform gas sample it increases monotonically until it reaches the largest value that is consistent with the combined energy of the molecules in the sample. The distribution of molecular velocities then becomes the Maxwell distribution. This counterpart to the law of entropy growth is known as Boltzmann's *H* theorem. Its proof depends in part on the fact that a Newtonian description of an encounter between two molecules doesn't change when one reverses the direction of the time axis.

The Maxwell-Boltzmann theory applies to samples of an ideal gas. It predicts that molecular collisions cause an isolated sample an ideal gas to relax into a state in which its molecules are uniformly distributed within the enclosure and have a Maxwell velocity distribution. This state is the counterpart of thermodynamic equilibrium. Relations between appropriate molecular properties, averaged over the distribution of molecular position and velocity, mirror relations between thermodynamic quantities that prevail in thermodynamic equilibrium.

In 1901 Josiah Willard Gibbs published a more general statistical theory of thermodynamic equilibrium. It applies to any system of N particles whose motions and interactions are governed by Newton's laws. Gibbs characterized the macrostates of such a system by probability distributions of its microstates, each specified by a set of $6N$ real numbers, the three position coordinates and three momentum components of each of the system's N particles. As Maxwell and Boltzmann had done, he identified the system's macroscopic properties with mean values of appropriate microscopic properties. In particular, he identified entropy with statistical entropy as given by Boltzmann's formula. He proved that the statistical entropy of his "canonical distribution" – a generalization of the Maxwell molecular-velocity distribution – exceeds that of any other distribution with the same mean micro-energy. He also proved that the statistical entropy of *any N*-particle distribution of microstates of an isolated system is constant in time.

Gibbs's theory, which he called statistical mechanics, requires only small *formal* changes[7] when one uses quantum mechanics rather than classical mechanics to describe microstates. So modified, it reproduces all the laws of equilibrium thermodynamics and goes far beyond it. Like the standard formulation of quantum mechanics it leaves nothing to be desired as an instrument for making predictions about measurement outcomes. Also like the standard formulation of quantum mechanics it raises questions that fall in an area where physics and philosophy overlap. One of these questions concerns the interpretation of probability.

The probability distributions that figure in Maxwell's and Boltzmann's statistical theories represent relative frequencies: we can identify the probability that the position coordinates and momentum components of a gas molecule lie in given ranges with the fraction, or proportion, of gas molecules in a macroscopic sample whose position coordinates and momentum components lie in these ranges at a given moment. Since a macroscopic sample contains a vast number of molecules, this interpretation is virtually exact. But how are we to interpret the probability distributions of microstates that characterize macrostates in Gibbs's theory?

Gibbs imagined a large or infinite collection of replicas of the macroscopic system, each in a definite microstate. He identified the probability associated with a given range of microstates with the fraction of the imaginary replicas whose microstates lie in that range. Recognizing that relative frequencies in an imaginary collection are just as abstract as the probabilities they represent, Gibbs referred to his statistical descriptions as *analogues* of thermodynamics. But because, as I have mentioned, quantum statistical mechanics not only duplicates the predictions of thermodynamics but also goes well beyond them, it is presumably the more fundamental theory. And if that is the case, the probability distributions that represent macrostates in quantum statistical mechanics need a physical interpretation.

The most obvious possibility is to suppose that the isolated macroscopic system whose equilibrium macrostates are characterized by a probability distribution of microstates is actually in one of these microstates. The probability distribution then represents an observer's limited knowledge of the system's microstate:

> [M]acroscopic observers, such as we are, are under no circumstances capable of observing, let alone measuring, the microscopic dynamic state of a system which involves the determination of an enormous number of parameters, of the order of 10^{23}. ... [Thus] a whole *ensemble* of possible dynamical states corresponds to the *same* macroscopic state, compatible with our knowledge.[8]

E. T. Jaynes carried this interpretation a step further. He argued that "statistical mechanics [is] a form of statistical inference rather than a physical theory." Its "computational rules are an immediate consequence of the maximum-entropy principle," which yields "the best estimates that could have been made on the basis of the information available"[9] On this view, the statistical entropy of a probability distribution that represents an isolated system's equilibrium macrostate represents physicists' lack of information about the system's microstate; and Gibbs's theorem that the statistical entropy of the canonical distribution exceeds that of any other distribution subject to the same constraints exemplifies the principle of maximum-entropy inference.

Werner Heisenberg interpreted statistical mechanics in much the same way. He also linked physicists' incomplete knowledge of the microstructure of macroscopic systems (and, more generally, of the world) to their inability to predict the outcomes of quantum measurements:

> [The interaction between a measured quantum system and a measuring device] introduces a new element of uncertainty, since the measuring device is necessarily described in terms of classical physics; such a description contains all the uncertainties concerning the microscopic structure of the device which we know

from thermodynamics, and since the device is connected with the rest of the world, it contains in fact the uncertainties of the microscopic structure of the whole world. These uncertainties may be called objective in so far as they are simply a consequence of the description in terms of classical physics and do not depend on any observer. They may be called subjective in so far as they refer to our incomplete knowledge of the world.[10]

Another version of the epistemic interpretation of probability distributions in equilibrium statistical mechanics begins with the remark that every macroscopic system interacts weakly with its surroundings. This interaction causes the system to visit a range of microstates with nearly the same energy. One then identifies the probability that the system's actual (but unknown) microstate lies in a given range of microstates with the relative frequency with which the system's microstate visits this range; see, for example, Schrödinger,[11] Landau and Lifshitz,[12] and Feynman,[13]

In contrast with these epistemic interpretations, I will argue that probability distributions of microstates, viewed in a particular cosmological context, characterize macrostates *completely*. This view does not assume that macroscopic systems are "really" in definite – however short-lived – microstates. Thus it is a variety of what Lawrence Sklar[14] in his insightful account of the foundations of equilibrium and non-equilibrium statistical mechanics calls "tychism."

VI. Microphysics and macrophysics: time's arrow

Whereas equilibrium statistical mechanics fully reproduce the mathematical laws of equilibrium thermodynamics, statistical theories that describe how systems initially in non-equilibrium states relax into equilibrium run into two problems. These are exemplified by Boltzmann's transport equation, which governs changes in the joint probability distribution of molecular position and momentum resulting from molecular collisions in an isolated gas sample. Boltzmann proved that these changes have a one-way character. They cause the statistical entropy of this probability distribution to increase monotonically toward the largest value that is consistent with the sample's mean energy per molecule; and this value characterizes the equilibrium distribution (Boltzmann's *H* theorem). Two questions now arise. 1. How can molecular interactions governed by Newton's time-reversal-invariant laws of motion give rise to one-way macroscopic behavior? 2. How can the growth of molecular statistical entropy be reconciled with Gibbs's proof that the statistical entropy of the joint *N*- particle probability distribution (for an isolated system of *N* particles) is constant in time?

The source of directionality in Boltzmann's derivation of his transport equation isn't hard to spot. Following Maxwell, Boltzmann assumed that the *incoming* velocities of colliding molecules are statistically uncorrelated. That is, he assumed that the joint probability distribution of the incoming velocities of the collision partners is the product of the individual probability distributions. Now, Newton's laws of motion imply that the combined energy and the combined momentum of colliding molecules have the same values before and after a collision. So if the incoming velocities of the collision partners are uncorrelated, their outgoing velocities must be correlated. Boltzmann's derivation assumes, however, that molecular correlations are *permanently* absent. This assumption cannot be true for an isolated gas sample. Even if molecular correlations were absent initially, they would subsequently be produced by molecular collisions.

Yet Boltzmann's equation enjoys strong experimental support. Not only does it include as special cases the phenomenological laws that govern such irreversible processes as heat flow, molecular diffusion, and viscous dissipation of relative fluid motions, it also enables one to express the coefficients that figure in these laws in terms of quantities that characterize molecular properties, molecular motions, and molecular interactions.[15] Predictions based on Boltzmann's equation have passed all experimental tests with flying colors.

Boltzmann's statistical theory belongs to a large class of statistical theories that describe irreversible processes and that contain counterparts to his *H* theorem. Van Kampen[16] has pointed out that these theories all depend on "repeated randomness assumptions," analogous to Boltzmann's assumption that molecular correlations are permanently absent; Sklar, in the book cited in Note 14, calls them "rerandomization posits." What justifies them?

Prigogine has argued that subjective justifications are implausible:

> [In Boltzmann's theory] irreversibility comes from supplementary phenomenological or subjectivist assumptions, from 'mistakes.' But how can we account for the wealth of important results and concepts that derive from the second law? In a sense living things, we ourselves, are then 'mistakes.' [17]

Prigogine and his collaborators have argued that macroscopic irreversibility must be rooted in irreversible microscopic laws that underlie quantum mechanics in its present form.

A more modest suggestion justifies repeated-randomness assumptions by an appeal to environmental interactions. Experimental physicists can effectively prevent an enclosed gas sample from exchanging matter or energy with the outside world. But they cannot prevent the leakage of information associated with molecular correlations. For if the molecules that make up the walls of the enclosure have a maximally random probability distribution (i.e., a probability distribution whose statistical entropy is as large as possible), collisions between gas molecules and wall molecules create statistical correlations between wall molecules and gas molecules and attenuate correlations between gas molecules, thus preventing them from building up to a point where they invalidate the assumption that the incoming velocities of colliding molecules are statistically uncorrelated.

The assumptions that enclosed gas samples initially lack molecular correlations and are embedded in random environments, which wick away correlation information, exemplify an approach that J. M. Blatt[18] and others have called "interventionism." Blatt constructed a mathematical model that allowed him to estimate the rate at which random interactions between an enclosed gas sample and the walls of its container destroys correlation information. He noted that interventionism had not been a popular approach:

> There is a common feeling that it should not be necessary to introduce the wall of the system in so explicit a fashion. ... Furthermore, it is considered unacceptable philosophically, and somewhat "unsporting," to introduce an *explicit* source of

> randomness and stochastic behavior directly into the basic equations. Statistical mechanics is felt to be a part of mechanics, and as such one should be able to start from purely causal behavior [p. 747].

Sklar, in the book cited in Note 14, gave a more detailed critique of interventionism. Shenker[19] responded to this critique and offered a qualified defense of interventionism.

The *physical* problems broached by interventionism are (a) to supply objective definitions of "randomness" and "correlation information" and (b) to justify the assumption that macroscopic systems are initially deficient in correlation information and are embedded in random environments. I will address these problems in due course.

Interventionist theories of irreversibility in statistical mechanics are analogous to decoherence theories of quantum measurement.[20] As emphasized by Niels Bohr, the measuring apparatus in a quantum measurement is necessarily a macroscopic system, and the registration of a measurement outcome is an irreversible macroscopic process. Decoherence calculations show how interaction between the combined system in a quantum measurement and a random environment, such as a dilute gas or a radiation field, effectively randomizes the relative phases of the coefficients in the superposition predicted by von Neumann's account of an ideal measurement. Decoherence calculations explain why the superposition of macroscopically distinguishable quantum states predicted by that account cannot exhibit effects analogous to interference between light waves in diffraction experiments.[21] But as Erich Joos and Hans-Dieter Zeh emphasized in a classic paper[22] on decoherence and quantum measurement, decoherence alone does not explain why quantum measurements have definite outcomes.

Interventionism resolves the apparent contradiction between Gibbs's theorem that the statistical entropy of the N-particle distribution is constant in time and Boltzmann's theorem that the statistical entropy of an initially non-equilibrium one-particle distribution increases monotonically with time:

> If (and only if) molecular correlations are absent, the N-particle distribution reduces to a product of identical one-particle distributions, and the statistical entropy of the N-

particle distribution becomes *N* times the statistical entropy of the one-particle distribution.

It is easy to prove that if molecular correlations are present (so that the *N*-particle distribution does not reduce to a product of identical one-particle distributions), the statistical entropy of the *N*-particle distribution (call it S_N) is less than *N* times the statistical entropy S_1 of the one-particle distribution. The difference represents *correlation information*: $I_{correlation} = NS_1 - S_N$. Since S_N is constant in time, the growth of correlation information requires that the one-particle statistical entropy S_1 to increase with time, as Boltzmann inferred from his transport equation.

If correlation information is initially absent, it is created by molecular encounters, and the one-particle statistical entropy increases. This state of affairs continues to prevail if the gas sample under consideration interacts weakly with a random environment that wicks away correlation information and disperses it to the wider universe.

VII. The relevance of cosmology

Much of physics treats systems whose interaction with the rest of the universe is either negligible or can be described in a simple way. As Heisenberg reminded his readers in the essay I have been quoting,

> [I]t is important to remember that in natural science we are not interested in the universe as a whole, including ourselves, but we direct our attention to some part of the universe and make that the object of our study [p. 52]

But as Zeh[23] has pointed out, we cannot consistently assume that macroscopic systems are truly isolated; we must allow for their interaction with their surroundings. And once we do that, we find ourselves on a slippery slope.

Zeh's argument is straightforward. An isolated system is in a definite quantum state. But the possible quantum states of a macroscopic system are so closely spaced in energy that they must be "entangled" with the quantum states of the part of the environment with which the system interacts.[24] The same argument applies to every bounded part of the environment. So – this is the slippery slope – if quantum mechanics applies on all scales, the quantum states of any macroscopic system must be entangled with quantum states of the rest of the universe. This conclusion forms the starting point of many-worlds interpretations of quantum mechanics, beginning with Everett's "relative state" interpretation.

But if we accept the conclusion that the universe is in a definite quantum state, we face the (unsolved) problem of explaining how the world of classical physics, which includes the world of experience, fits in. We also face the problem of explaining how Einstein's theory of space, time, and gravitation – general relativity – fits in. A theory that included quantum mechanics and general relativity as limiting cases would, of course, solve that problem; but such a theory doesn't yet exist. Nevertheless, many physicists postulate that quantum mechanics does apply at all scales and that the universe is in a definite quantum state.

The relation between quantum mechanics and classical physics is problematic in another respect. Quantum mechanics and general relativity enjoy overwhelming observational and experimental support in their respective domains. But their domains overlap. And this poses a problem. General relativity is a classical, deterministic theory; but quantum measurements can produce unpredictable macroscopic changes in the structure of space-time. How can these apparently contradictory features of our two most fundamental theories be reconciled?

As I have discussed, many physicists postulate that quantum mechanics is universally valid. They hope and expect that general relativity will one day be shown to be a limiting case of a quantum theory of gravity. Einstein, by contrast, hoped that quantum mechanics would one day be found to be a limiting case of a deeper deterministic field theory – a hope shared by few contemporary physicists.

I will suggest a third way. We do not yet have a unified set of mathematical laws that includes the laws of quantum mechanics and the field equations of general relativity as limiting cases. But I will sketch a theory of initial and boundary conditions that makes quantum mechanics and general relativity compatible in their shared domain. At the same time it offers solutions to the quantum measurement problem and the problem of time's arrow. The framework of the proposed theory of initial and boundary conditions is a version of relativistic cosmology. It unites but does not unify quantum mechanics and general relativity, showing that they coexist peaceably in their common domain.

VIII. Relativistic cosmology: the cosmological principle

Newton speculated that the stars were distant suns uniformly distributed throughout an infinite Euclidean space. But because his theory of gravitation does not apply to an infinite, unbounded distribution of mass, he was unable to formulate a mathematical theory based on this idea. Before formulating his generalization of Newton's theory, general relativity, Einstein had hoped that it would fill this gap, and soon after completing the theory, in 1915, he tried to apply it to an idealized model of the universe: a uniform, unbounded, pressure-free, static medium. He found that his field equations had no solution that satisfied these conditions, and in 1917 he suggested a modification of his 1915 field equations that allowed them to have a static solution. Five years later Alexander Friedmann showed that while the original field equations do not have *static* solutions, they do have non-static solutions, in which space and its contents undergo a uniform expansion from (or towards) a singular state of infinite mass density.

In 1929 Edwin Hubble announced that the most distant galaxies whose distances and line-of-sight velocities could then be measured were systematically[25] receding from Earth at speeds proportional to their distances; they were taking part in a uniform expansion, just as Friedmann, unbeknownst to Hubble, had predicted. Modern observations of galaxies and of the cosmic microwave background, discovered in 1965, support this conclusion.

To estimate the distances of distant galaxies, Hubble assumed that they and their stellar populations have the same statistical properties as nearby galaxies and stellar populations. (This enabled him to use distance criteria calibrated on objects close enough to have measurable parallaxes.) The *cosmological principle*, the starting point for conventional cosmological theories, is a generalization of Hubble's assumption (which he called the principle of uniformity). It says that there is a system of spacetime coordinates relative to which no statistical property of the universe at a given moment serves to define a preferred position or direction in space.

Physicists usually study idealized models of real physical systems. They take it for granted that the initial and boundary conditions that characterize these models hold only approximately. Galileo assumed that the effects of air resistance on the motions of falling and sliding objects masked simple and exact mathematical laws, and in his experiments he took pains to minimize these effects. Astrophysicists know that stars rotate and are chemically inhomogeneous; but they begin by idealizing them as chemically homogeneous, nonrotating gas spheres. In the same spirit one might – and many physicists do – regard the cosmological principle as characterizing a class of simplified models of the universe. By contrast, the following considerations take as their starting point the assumption that the cosmological principle is an exact symmetry of the initial conditions that characterize the universe (as it is of all our present physical laws). I assume further that a statistical description that enjoys this symmetry cannot be augmented by nonstatistical information. In this sense a statistical description that comports with the assumption (which I will refer to as the strong cosmological principle) is complete.

A Newtonian universe cannot satisfy the strong cosmological principle. Consider, for example, a statistically uniform distribution of free particles. A complete Newtonian description of such a distribution at a given moment would specify the distance between every particle and its nearest neighbor. Thus it would assign every particle a real number – the instantaneous distance of its nearest neighboring particle. But the number of particles (and pairs of nearest neighbors) is at most countably infinite. So if the distribution is random, there is zero probability that two of these real numbers coincide. Every particle is uniquely situated with respect to its neighbors.

In contrast, it follows from Heisenberg's indeterminacy principle that quantum mechanics assigns any bounded region of a uniform distribution of free particles a finite number of quantum states, provided the particles' momenta (or energies) are also bounded. Suppose the distribution is infinitely extended, as comparisons between astronomical observations and refined versions of Friedmann's cosmological models indicate, and that all its statistical properties are uniform. Then with probability one, any given bounded region will have infinitely many replicas in the same quantum state. It follows that any two realizations of the same (uniform) statistical description are *finitely indistinguishable* in the sense that any bounded region of one realization has infinitely many exact matches in any other realization. This conclusion and its supporting argument can easily be extended to *any* statistical description of an infinite universe that satisfies the cosmological principle. *In effect, a statistical description of an infinite universe that does not privilege any point or direction in space has a single realization.*

IX. The growth of order and the growth of entropy

Extrapolating the present state of the (observable) universe backward in time, one arrives at an era when, at each moment, the cosmic medium closely approximates a mixture of free particles in thermal and chemical equilibrium.[26] The relative concentrations of particle kinds in thermodynamic equilibrium depend on the mass density and the temperature. As the medium expands, its mass density and its temperature decrease. At sufficiently early times the rates of equilibrium-maintaining particle reactions greatly exceed the rate at which the mass density and the temperature are changing, so particle reactions are able to maintain equilibrium at the instantaneous values of the mass density and the temperature. Now, particle reaction rates and the rate at which space is expanding both decrease with decreasing mass density; but the expansion rate decreases more slowly. Eventually the particle reactions tasked with maintaining the relative concentrations of particle kinds appropriate to chemical equilibrium at the instantaneous mass density and temperature become unable to do so, and the relative abundances of helium and some other light elements become frozen in.[27]

If we define the statistical information of a probability distribution as the amount by which the distribution's statistical entropy falls short of its largest allowed value, then the process just described – nucleogenesis – creates information – specifically, chemical information. Much later, thermonuclear reactions in the core of the Sun degrade some of this information when they burn hydrogen into helium. Some of the energy released by these reactions is converted into sunlight, which drives the biological processes that sustain life on Earth.

The expansion also creates structural information. Self-gravitating astronomical systems could not have existed at the high mass densities that prevailed when helium and light nuclei were formed. They must have come into being later in the cosmic expansion. There is no consensus about how this happened. On one scenario[28] an initially cold cosmic medium solidifies as metallic hydrogen.[29] As the expansion continues, the medium breaks up into fragments whose cohesion energies are approximately equal to their gravitational binding energies. These fragments, the first self-gravitating systems, are less massive by one or two orders of magnitude than the giant planets. At this stage the cosmic medium is a cold "gas" whose "particles" are solid-hydrogen fragments. Because the "particles" are randomly distributed, the gas's internal energy contains a negative contribution associated with the fluctuating part of their local gravitational interactions as well as a positive contribution due to the particle motions relative to the expanding background produced by the fluctuating local gravitational field. Initially these contributions are equal, but the expansion attenuates the positive contribution faster than the negative contribution, so that eventually small self-gravitating clusters of "particles" separate out as self-gravitating systems. These newly formed self-gravitating systems now take over the role of particles, and the process – gravitational clustering – continues, giving rise to self-gravitating systems on progressively larger scales.

This scenario predicts that the initial binding energy per unit mass of a self-gravitating system is proportional to the one-third power of the system's mass. Astronomical measurements are consistent with the predicted relation over a range of masses that extends from giant planets and their satellites to rich galaxy clusters – eighteen powers of ten.

The assumption that the early universe was cold, first suggested by Zel'dovich in 1962, conflicts with the standard interpretation of the cosmic microwave background as a relic of a primordial radiation-dominated phase of cosmic evolution, the hot big bang. The standard interpretation accounts for some observed features of the cosmic microwave background. The cold-universe scenario, in contrast, interprets the cosmic microwave background as thermalized radiation from an early generation of supermassive stars.[30]

Self-gravitating systems evolve toward states of dynamical equilibrium, in which the cohesive effect of gravity balance the disruptive effect of internal motions. But these states differ radically from states of *thermodynamic* equilibrium. Consider, for example, a self-gravitating gas cloud of nearly uniform temperature. As the cloud loses energy by radiation, it contracts *and its temperature increases*. Thus a self-gravitating gas cloud in dynamical equilibrium has negative heat capacity. (By contrast, a system in thermodynamic equilibrium necessarily has positive heat capacity.) As the cloud evolves it departs progressively further from the featureless state of thermodynamic equilibrium: a radial temperature gradient develops and heat flows outward from the center. If the core temperature becomes high enough, thermonuclear reactions produce a radial gradient of chemical composition.

Of course, the local macroscopic processes that take place in a self-gravitating gas cloud – the transfer of heat from the cloud to its cooler surroundings, the flow of heat down the steepening radial temperature gradient, the thermonuclear reactions in the cloud's core – all generate entropy. But these entropy-generating processes drive the cloud and its surroundings away from global thermodynamic equilibrium.

Thus in bounded self-gravitating systems, as in the expanding cosmic medium, gravity opposes the macroscopic processes that seek to establish thermodynamic equilibrium.

X. Entropy and the law of entropy growth

If thermodynamic equilibrium prevails locally in a self-gravitating system, one can define the system's entropy as the sum of the entropies of its infinitesimal parts. The law of entropy growth then applies to an expanding universe composed of self-gravitating systems and radiation. But one cannot infer from this extended law of entropy growth that the universe is tending toward the unchanging, featureless state of global thermodynamic equilibrium ("heat death") envisioned by Clausius and Kelvin in the midnineteenth century, because as discussed above, the expansion of the cosmic medium and the contraction of bounded self-gravitating systems drive local conditions away from thermodynamic equilibrium.

Clausius extrapolated the law of energy conservation and the law of entropy growth (the first and second laws of thermodynamics) from macroscopic systems and processes to the universe as a whole. Neither extrapolation is valid.

The law of energy conservation does not apply in the uniformly expanding space predicted by Friedmann's cosmological solutions to Einstein's field equations. For example, if the cosmic medium is a uniform ideal gas, the theory predicts that every particle slows down relative to its local standard of rest; its momentum and its kinetic energy both decrease as the medium expands. The energy of an ideal-gas sample likewise decreases with time, though it does no work on its surroundings.

The thermodynamic law of entropy growth applies in a much narrower domain than the law of energy conservation: Clausius's definition of entropy applies only to systems in local thermodynamic equilibrium. Boltzmann's definition of statistical entropy is far more general. Statistical entropy is a property of the probability distribution of microstates that characterizes a macrostate. And if such probability distributions have an objective character, as I argue below, statistical entropy is just as objective as thermodynamic entropy. Yet Boltzmann's *H* theorem cannot be viewed as an instance of a universal law, because its derivation depends on the assumption that information associated with molecular correlations is permanently absent. This assumption in turn follows from an initial condition (that correlation information is absent), a boundary condition (that nominally isolated gas samples actually interact with random environments), and a plausible but not rigorous physical argument (that correlation information flows from a sample to its random environment). As mentioned above, analogues of Boltzmann's *H* theorem for other macroscopic systems rely on analogous initial and boundary conditions and an analogous argument about the role of the environment.

Because experiments confirm the predictions of Boltzmann's *H* theorem and its analogues, we can infer that the initial and boundary conditions on which the derivations of these theorems rest are *ordinarily* satisfied: the probability distributions that characterize macrostates of newly formed – or newly prepared – macroscopic systems ordinarily lack correlation information; and these systems ordinarily have random surroundings. These initial and boundary conditions have a quasi-universal character. They are ordinarily but not *necessarily* satisfied.

E. L. Hahn's spin echo experiment[31] shows that when appropriate kinds of correlation information are present initially and are sufficiently resistant to degradation by random interactions, a random distribution of microstates can evolve into a highly nonrandom distribution. The microstates in question are orientations of magnetic moments of nuclei in a macroscopic liquid sample. The sample is in an applied magnetic field whose direction is the same throughout the sample but whose magnitude has a small position-dependent random component. The state of the collection of magnetic moments is characterized by the joint distribution of their orientations and positions. The collection is prepared in a state in which the magnetic moments are all accurately parallel (or antiparallel) to a direction perpendicular to the direction of the applied magnetic field. The joint distribution of orientations and positions then contains a large quantity of orientation information and virtually no information associated with orientation-position correlations. The magnetic field exerts a torque on each magnetic moment, causing it to rotate in a plane perpendicular to the direction of the field. Owing to the random component of the applied field, the magnetic moments rotate at slightly different rates, gradually getting out of alignment. During this part of the experiment orientation information is converted into correlation information. Eventually the orientations of the magnetic moments are randomly distributed in directions perpendicular to the direction of the applied magnetic field; the orientation information that was present initially has all been converted into correlation information. In a gas sample, correlation information produced by the decay of single-particle information is quickly dispersed by molecular interactions. In the spin echo experiment it remains localized and is amenable to experimental manipulation. An ingenious experimental intervention now reverses the flow of information, converting the correlation information back into orientation information. The process just described – the conversion of orientation information into correlation information and back again into orientation information – is accompanied by the ordinary entropic decay of single-particle information through particle-particle interactions, but on a time scale significantly longer than that of the "anti-entropic" process.

To sum up, I have argued that the thermodynamic law of entropy growth does not apply beyond its original domain: isolated macroscopic systems in local (or global) equilibrium. In particular, it does not apply to self-gravitating systems. Boltzmann's transport equation, his *H* theorem, and their generalizations (master equations, generalized *H* theorems) apply to macroscopic systems that are not in local thermodynamic equilibrium, but they, too, are not laws. They rest on initial and boundary conditions that are ordinarily, but not necessarily, satisfied by both natural and prepared systems. This, I will now argue, is a consequence of the simplest account of the structure and evolution of the universe that is consistent with our most fundamental and most highly confirmed physical laws.

XI. Initial and boundary conditions; the prevalence of chance

The initial and boundary conditions that characterize physical systems are products of historical processes. We can think of these processes as episodes in a history of the physical universe. Of course, we are not yet able to construct, or even sketch, a complete history of the physical universe. The fragmentary history I propose rests on two assumptions: the strong cosmological principle; and the assumption that at some early time the cosmic medium closely approximated a uniform, uniformly expanding distribution of free particles in local thermodynamic equilibrium.

A full history would ground the second assumption in antecedent initial conditions and in physical laws that contain fewer unexplained constants than our current laws and cosmological models. But if both assumptions should turn out to be correct, a full account would preserve the distinctive features of the present account:

1. The classical variables that figure in Einstein's description of the structure and contents of spacetime are to be interpreted as random variables – mathematical objects

characterized not by a definite value at each point of space-time but by a set of possible values and corresponding probabilities. We can interpret these probabilities as relative frequencies, or proportions, in infinite samples whose members are randomly distributed throughout space. For example, the probability that the mass density at a point lies in a given range of values is the fraction of points in a uniformly and randomly distributed sample of points at which the mass density lies in that range; the joint probability that the mass densities at two points with a given separation lie in given ranges is the fraction of a sample of pairs of points that have the given separation in which the mass densities at the two points lie in the given ranges, and so on.

As discussed below, this interpretation of Einstein's description of spacetime and its contents resolves the *prima facie* conflict between the deterministic character of Einstein's field equations and the fact that quantum measurements alter the macroscopic structure of spacetime unpredictably.

2. The probability distributions of microstates that characterize early states of the universe contain little or no statistical information per unit mass. As the universe expands, macroscopic processes create information or change its qualitative character (through processes that always generate statistical entropy). But the quantity of information per unit mass remains far smaller than its largest allowed value. Thus randomness prevails.

3. The initial and boundary conditions that characterize macroscopic systems and processes are expressed by probability distributions of microstates, which in turn are determined by their history.

4. Such histories usually determine the values of macroscopic mechanical and thermodynamic variables but do not usually create information associated with persistent micro-level information (though as the spin echo experiment illustrates, they *can* do so). Theories that describe irreversible macroscopic processes rest on instances of this generalization. This remark explains why the arrow of time defined by varied macroscopic processes in nominally isolated macroscopic systems coincides with the arrow defined by the cosmic expansion.

5. As discussed below, many macroscopic processes other than quantum measurements have indeterminate outcomes.

The present account of chance resembles in important ways an account given a century ago by Henri Poincaré[32] in a popular essay. Poincaré asked why the outcomes of certain deterministic processes seem to be correctly predicted by "the laws of chance." As his first example Poincaré considered an ideal cone initially balanced on its tip. Imprecision in its initial positioning and tiny uncontrollable external disturbances cause the cone to topple in an unpredictable direction. But if the experiment is repeated many times, the final azimuth of the cone's axis will be smoothly (though not necessarily uniformly) distributed between 0 and 2π radians. In this example a deterministic law maps small differences between initial values of the azimuth of the cone's axis onto large differences between its final values. The smooth distribution of final azimuths requires only that the initial azimuths be smoothly distributed over a narrow subrange of their possible values.

In the 1880s Poincaré discovered the phenomenon now called deterministic chaos. The outcomes of chaotic processes depend sensitively on their initial conditions. In the discovery context small differences between the initial conditions of test particles in a gravitating system may cause their orbits to diverge at an exponential rate. If the initial values of the parameters that define an orbit are smoothly distributed over a small subrange of their possible values, the possible values of these parameters at a later time will be smoothly distributed over the entire range. Examples of chaotic processes are legion, ranging from meteorology to biology.

Poincaré argued that the initial conditions that characterize the cone balanced on its tip as well as those that characterize chaotic orbits in the solar system are in fact smoothly distributed on very small scales because historical processes have smoothed out irregularities on the smallest scales. The present historical account of initial and boundary conditions suggests a closely related but somewhat simpler explanation: The experimental setup that creates the initial state of Poincaré's cone specifies a probability distribution of initial conditions that does not contain enough information to specify the cone's final azimuth. Similarly, the historically determined probability distribution that characterizes the initial position and velocity of an asteroid in a chaotic orbit does not contain enough information to specify the asteroid's position after a lapse of 4.5 billion years.

We can define a classical microstate of Poincaré's cone, in part, by the azimuth of its axis. We can define the cone's macrostates, in part, by the precision of a given measuring apparatus. Initially the cone is in a macrostate in which the azimuth of its axis doesn't have a definite value, but as the cone's angle of tilt increases, the number of experimentally distinguishable azimuths – and hence the number of distinguishable macrostates – increases. Analogously, the orbit of an asteroid may be sensitive to small changes in its initial position and velocity. A historical account characterizes the initial state by a joint probability distribution of positions and velocities, which evolves into a distribution that characterizes a multitude of observationally distinguishable orbits.

To accommodate such situations we need to modify the rule that links probability distributions of (classical or quantum) microstates to classical macrostates. The standard rule equates the value of a macroscopic variable in a given macrostate to the result of averaging the corresponding microscopic variable over the probability distribution of microstates that represents the given macrostate. We modify it in three ways.

First, we characterize macrostates by experimentally distinguishable ranges (or aggregates) of microstates, as in the above examples. A probability distribution of microstates may then represent two or more experimentally distinguishable macrostates.

Second, we equate the result of averaging a microscopic variable over such a probability distribution to the result of averaging the measured value of the corresponding macroscopic variable over a "large number" of replicas of the measurement.

Finally, to incorporate into our rule the fact that neither physical laws nor initial and boundary conditions that comply with the strong cosmological principle serve to define a particular position, we interpret the set of replicas mentioned in the preceding paragraph as a "cosmological ensemble" – a set of replicas randomly and uniformly distributed throughout an infinite space. (Like Gibbs's ensembles, a cosmological ensemble is made up of imaginary replicas. But each replica in a cosmological ensemble is in a definite *macro*state. And cosmological ensembles have a physical interpretation: they allow us to express the assumption that physics cannot make unconditional predictions about where in the universe given measurement outcomes are realized.)

These rules enable us to calculate the probabilities of experimentally distinguishable measurement outcomes from measurements of mean values:

Following an argument given by Dirac[33] in a related context, let V denote a macroscopic property whose possible values are real numbers. Let the index k label the possible outcomes of a measurement of V and let the index r label replicas in a cosmological ensemble. Let $I(V, k, r)$ be the function of V, k, and r that is equal to 1 if a measurement of V at the rth replica has the kth outcome and is equal to 0 otherwise. The value of $I(V, k, r)$ averaged over the members of a cosmological ensemble is the fraction $f(V, k)$ of replicas for which a measurement of V has the outcome k. We can think of the set $\{k\}$ of outcomes as a sample space and the set of fractions $\{ f(V, k)\}$ as a set of probabilities on this sample space.

XII. Quantum measurement

The preceding rule for linking a probability distribution of (classical or quantum) microstates to the possible outcomes of a measurement and their probabilities applies to quantum measurements. The isolated macroscopic system now consists of a quantum system one of whose properties we wish to measure, a macroscopic measuring apparatus that interacts with the quantum system, and a bounded random environment[34] that interacts with the measuring apparatus. We assume, as in decoherence calculations, that this system has quantum states that evolve in accordance with Schrödinger's equation. But we do not make the customary assumption that the system is initially in one or another of its microstates. We assume instead that it is in a macrostate characterized by a probability distribution of its microstates. Application of the preceding rule then reproduces the measuring postulate of the standard formulation without further ado: it predicts that ideal measurements have definite outcomes given, along with their relative frequencies in a cosmological ensemble, by the measuring postulate.

XIII. QM and GR

As mentioned earlier, general relativity's deterministic description of the evolution of space-time structure clashes with the fact that quantum measurements affect the local structure of space-time in unpredictable ways. The present account dissolves this contradiction. From a macroscopic standpoint the unpredictability of the post-measurement position of a pointer in a quantum measurement is no more problematic than the unpredictability of the final orientation of Poincaré's cone. In both cases macroscopic unpredictability results from an objective absence of information in the probability distribution of microstates that characterizes the initial state of an isolated system. In both cases a deterministic law – Schrödinger's equation in the first case, Newton's laws of motion and gravitation in the second – governs the evolution of the system's microstates.

XIV. The irreducibility of macrophysics and the unity of physics

The holy grail of physics is a Theory of Everything. Such a theory would include as limiting cases our present strongly confirmed laws and would contain far fewer adjustable physical constants than figure in these laws. As I have already emphasized, accounts of physical systems and processes depend on initial and boundary conditions as well as laws; and it has long been understood that laws and initial/boundary conditions are not entirely distinct categories. To derive macrophysical laws such as Boltzmann's *H* theorem and its generalizations from more fundamental microscopic laws one needs to impose appropriate initial and boundary conditions. In this essay I have argued that these conditions are products of a historical process whose description rests on simple cosmological initial conditions and a strong version of the cosmological principle.

The account I have sketched of this historical process knits together our present laws in other ways as well. It shows how initial and boundary conditions link the temporal direction of macroscopic processes to the direction of the cosmic expansion, it offers a simple and direct answer to the question of why quantum measurements have definite but unpredictable outcomes, and it reconciles the unpredictability of quantum measurement outcomes with the deterministic character of Einstein's field equations.

Besides joining these loose ends, a historical account of initial conditions offers a new view of the role of chance in macroscopic processes. Physicists have conventionally held that the outcomes of macroscopic processes other than quantum measurements are predictable in principle. Some, though not all, evolutionary biologists have taken issue with this doctrine, which also seems to be at odds with judgments based on ordinary experience. But physics as conventionally interpreted assures us that to a contemporary version of the omniscient mind posited by Laplace in his essay on chance, nothing except quantum measurement outcomes would be unpredictable. The historical account of initial conditions sketched in this essay supports the contrary view suggested by evolutionary biology and experience: much of what we observe in the world around us is influenced by chance. This generalization applies not just to aspects of our physical environment, like weather. As discussed in a little more detail below, randomness plays an essential role in the biological world.

XV. Is biology a part of physics?

Some physicists consider physical theories to be nothing more than devices for linking measurements to other measurements. Others – realists – regard our present theories as descriptions, perhaps partial or approximate, of a unified mathematical structure behind experience. The second view, which was held by Einstein, draws support from the history of physics. Strongly confirmed theories have not been overturned by their successors. They have remained in place as limiting cases, valid in circumscribed domains, of the successor theories.[35] And as the scope of physical theories and the accuracy of their predictions has increased, the fundamental theories have become fewer, more comprehensive, and more abstract. History thus supports the view that our present physical theories capture, or at least approximate mathematical regularities behind experience and that these regularities belong to a unified mathematical structure.

What characterizes the objects and processes that belong to the world that physics describes and physicists try to understand? Consider atoms. Two and a half millennia ago Leucippus and Democritus tried to link the sizes and shapes of hypothetical atoms to observed properties of bulk matter. Newton in the *Principia* tried to account for Thomas Boyle's empirical law relating the pressure and the volume of an enclosed sample of air by positing air atoms moving and interacting in ways governed by his laws of motion. Half a century later, Daniel Bernoulli introduced a much simpler atomic model of air, and in the nineteenth century Rudolf Clausius and James Clerk Maxwell significantly extended Bernoulli's model. But Maxwell realized that Newtonian physics and his own theory of electricity and magnetism could not explain the observation that atoms always absorb and radiate light at a fixed set of frequencies. Meanwhile Ernst Mach argued that the atomic hypothesis was methodologically unsound because it invoked unseen entities. Physical theories, in his view, should seek to represent, rather than explain, experience.

In his 1905 paper on Brownian motion Einstein invoked a different criterion for physical hypotheses: falsifiability (as Popper later called it). If his predicted relation between the motions of a liquid's hypothetical molecules and the observable motions of microscopic particles suspended in the liquid should be shown to be incorrect, he wrote, this would "provide … a weighty argument … against the molecular-kinetic conception of heat [i.e., the atomic hypothesis]." Experimental confirmation of Einstein's law, which came a few years later, strengthened the case not only for the atomic hypothesis but for a way of doing theoretical physics that relies more heavily on mathematical invention and the testing of theoretical predictions than on the analysis of facts.

At the other end of the size scale, the physical universe crossed the boundary that separates physics from metaphysics in two steps. In 1915 Einstein published a theory of gravitation that applies to an unbounded, statistically uniform distribution of mass, and in the 1920s Edwin Hubble supplied observational evidence that the astronomical universe is indeed unbounded and statistically uniform.

In short, a combination of mathematical invention, experiment, and observation shapes a physical realist's conception of the physical world. The question "Is *X* a constituent of the physical world" can be rephrased as "Does *X* figure in a mathematical theory that is tightly linked to fundamental physical theories and is strongly confirmed by experiment or observation?" According to this criterion, quarks are constituents of the physical world, while "dark energy" is not – or at least not yet.

What about living organisms and biological processes? Living organisms are physical systems, because they are made up entirely of atoms and molecules drawn from the nonliving environment; and biological processes are physical processes, because they obey the same physical and chemical laws as nonliving systems. Yet living organisms and biological processes are not *just* physical systems and processes. They have a distinctive character, which they owe entirely to their distinctive initial and boundary conditions.

Like the initial and boundary conditions that characterize nonliving systems, those that characterize living organisms and biological processes were shaped by history. Although the opening chapter of the history of life exists only in rough, competing drafts. the authors of these drafts agree that life arose by chance in a nonliving environment through processes governed by well-understood physical and chemical laws. Can we then conclude that biology is at bottom a branch of physics, like condensed-matter physics and astrophysics? Do biological systems and processes belong to the world that physics describes or could describe? Or, as Ernst Mayr[36] and other biologists have argued, is biology an autonomous science?

Biology has a number of terms that do not appear in the physical sciences, such as *function*, *fitness*, *purpose*, *adaptation*. Can such terms be explained, however clumsily, in the language of physics and chemistry, augmented if necessary by explicit definitions?

Take *function*. Molecular physics and chemistry supply detailed accounts of the physical structure and chemical properties of hemoglobin. For example, they explain its capacity to bind oxygen molecules. But molecular physics and chemistry alone cannot tell us that in vertebrates the biological role of hemoglobin depends on its affinity for oxygen. *Bio*chemistry continues the chemical story. It seeks to understand not only how hemoglobin performs its biological function but also how the molecule and its function have evolved from simpler precursors. Can this continuation of the chemical story be recast in the language of physics and chemistry?

Part of it is already in that language. The chemical processes that involve or depend on hemoglobin belong to the common subject matter of chemistry and biochemistry. The other part of the story involves the notion of fitness. Changes in the structure of hemoglobin that affect its ability to bind and release oxygen molecules under specific environmental conditions affect an animal's prospects for survival and reproduction. Fitness is a measure of these prospects.

And there's the rub. Population geneticists have defined fitness in various ways, but all the definitions are prospective; they all refer to the future.

In a given population, genetic changes that have a significant random component give rise to variants of hemoglobin. *If the population's environment doesn't change for a sufficiently long period of time*, heritable variants whose oxygen affinity is optimal for the given, unchanging conditions will gradually come to dominate the population's gene pool: the possessors of sub-optimal variants will have fewer descendants than the descendants of possessors of optimal variants.

This example illustrates an essential aspect of evolution: the emergence and subsequent fixation of new or modified traits through random genetic variation and natural selection *in an unchanging environment*.

The example of hemoglobin evolution in an unchanging environment resembles the "evolution" of a cone initially balanced, imperfectly, on its tip. The fate of any given molecular variant is predictable; so is the path of a cone whose initial orientation and angular momentum have been specified. The cone's initial state is characterized by a probability distribution of micro-conditions; analogously, one might perhaps be able to assign a probability per unit time to the appearance of each possible variant of the hemoglobin molecule and then go on to predict the relative frequencies of variants after many generations. So the story of hemoglobin evolution in an unchanging environment can perhaps be recast in language familiar to physicists and chemists.

The emergence of evolutionary novelties poses a more formidable challenge to the view that evolutionary stories can be recast as physic-chemical stories. Evolutionary novelties are the results of a creative process:

> Evolution is a creative process, in exactly the same sense in which composing a poem or a symphony, carving a statue, or painting a picture are creative acts. An artwork is novel, unique, and unrepeatable; ... The evolution of every phyletic line

> yields a novelty that never existed before and is a unique, unrepeatable, and irreversible proceeding.[37]

Evolutionary novelties are also unpredictable:

> Evolutionary change in every generation is a two-step process, the production of genetically unique new individuals and the selection of the progenitors of the next generation. The important role of chance at the first step, the production of variability, is universally acknowledged (Mayr 1962), but the second step, natural selection, is on the whole viewed rather deterministically: Selection is a non-chance process. What is usually forgotten is the important role chance plays even during the process of selection. In a group of sibs it is by no means necessarily only those with the most superior genotypes that will reproduce. Predators mostly take weak or sick prey individuals but not exclusively, nor do localized natural catastrophes (storms, avalanches, floods) kill only inferior individuals. Every founder population is largely a chance aggregate of individuals, and the outcome of genetic revolutions, initiating new evolutionary departures, may depend on chance constellations of genetic factors.[36]

According to the central argument of the present essay, the random element in both genetic variability and natural selection is objective and irreducible. In physical contexts the randomness inherent in initial conditions entails that many kinds of macroscopic processes (besides quantum measurement) have objectively unpredictable outcomes; but the *possible* outcomes of such processes are predictable, along with their probabilities; and we may interpret the probability that attaches to a particular outcome or range of outcomes as its relative frequency in a cosmological ensemble. In the context of evolution, the possible outcomes – evolutionary novelties – are themselves unpredictable. Even if these outcomes and their probabilities could be predicted – a feat that would perhaps require the expenditure of more free energy than the Sun could supply in its lifetime – the prediction would be useless, for it would assign an infinitesimal probability to the novelties that have actually been produced in the course of evolution.

To sum up, although living organisms are physical systems and biological processes obey physical laws, life and its history are not part of the world current physical theories describe. Life is a natural phenomenon, firmly anchored in the physical world and its laws; but the initial and boundary conditions that characterize living systems and biological processes ensure that the history of life is creative and hence, in a way that transcends physical indeterminacy, unpredictable.

XVI. The problem of free will

Defenders of libertarian free will usually grant at the outset that events other than the outcomes of quantum measurements are determined by universal physical laws and antecedent conditions. They must then explain how it can be that we are able to shape the future through our choices and decisions. In this essay I have argued that the premise is false: Events in the macroscopic world are not determined by universal physical laws and antecedent conditions; a wide class of macroscopic processes have indeterminate outcomes. And if the processes involved in reflective choice belong to this class, there is no scientific reason why we should not accept the proposition that we shape the future through our choices and decisions.

Biology supplies a strong positive case for libertarian free will.[38] As mentioned earlier, Mayr and other evolutionary biologists have stressed the central role of chance in evolution. Genetic variation has a random component, but if genetic variation were entirely random, complex adaptations could never have evolved. The evolution of complexity requires genetic regulation of the ways in which chance manifests itself in genetic variation.

It is easy to see why. A central and strongly confirmed tenet of evolutionary theory is that complex organs such as eyes evolved from less complex but fully functional predecessors. These predecessors themselves evolved from less complex but also functional predecessors, and so on until we arrive at “a simple light sensor for circadian (daily) and seasonal rhythms around 600 million years ago"[39] At each stage of this multistage process genetic modifications that improved the eye's function emerged through the usual combination of genetic variation and selection. But for the process to work, the genetic variations that occur at each stage must not significantly impair the function achieved at that stage. This means that at each stage the genes and gene combinations that encode the eye's developmental program must be held safe from harmful variation – not from all variation, just harmful variation. And indeed experiments show that genetic variation is suppressed to varying degrees at different genetic loci. For example, proofreading and error-correction processes suppress transcription errors (which when not suppressed are a source of variability); exchanges of genetic material between homologous chromosomes during meiosis are nonrandom in ways that preserve gene combinations whose disruption would lower fitness while allowing others. Molecular mechanisms that regulate and channel genetic variation are themselves products of evolution's two-stage process.

Open behavioral programs enable animals, including single-celled animals, to thrive in environments that change in unpredictable ways. All animals learn from experience. They tend to repeat behaviors for which they have been rewarded in the past and to avoid behaviors for which they have been punished. Some animals also learn from experience in ways that allow for risk-taking, exploratory behavior, and delayed rewards. Economists, students of animal behavior, and cognitive neuroscientists have developed algorithms that seek to mimic such flexible learning strategies, and have constructed hypothetical neural networks that instantiate algorithms of this kind.[40]

In light of what ethologists have learned about the behavior of monkeys and apes, we can plausibly conjecture that much of human learning and decision-making is mediated by neural architecture that embodies complex and sophisticated algorithms of this kind. But humans have an extra, qualitatively different capacity for learning and decision-making: a capacity for *reflective choice*. We are able construct mental representations of scenes and scenarios that are not directly coupled to external stimuli (as in perception) or to movement (as in reflexes). We call on this capacity when we imagine possible courses of action and then go on to imagine the possible consequences of each of these invented candidates for choice. Our ancestors used it when they painted pictures on the walls of their caves and created the first human languages. We use it when we compose an original sentence or a tune or when we try to solve an abstract problem. It allows us to reconstruct the distant past from fragmentary evidence and to envision the distant future. It makes possible the life of the mind.

The brain of every animal contains a model (or "theory") of the world – a hierarchical set of schemata that regulate the animal's behavior. Some of these schemata can be modified by the animal's experience. According to the psychologist Thomas Suddendorf,[41] children exhibit "the ability to entertain and collate offline mental models (e.g., about past, future, or imaginary situations) in addition to the primary reality model" around the age of two. He cites behavioral evidence that great apes, *but not monkeys*, also have this capacity. Around age four, children demonstrate a new capacity: they become able to understand representations *as* representations. Their world models begin to include a "theoretical" component: a set of beliefs about animals, gods, weather, the Sun, the stars, other people, and ourselves. Although great apes have "offline" mental models, Suddendorf writes, they show no behavioral evidence for this second capacity – the capacity to understand that mental representations are products of their own imagining. Creative thought requires the second capacity. We have it; according to Suddendorf, nonhuman primates do not. So free will in the strong or libertarian sense is a distinctively human biological capacity. Like evolution itself, it harnesses chance in the service of creativity.

XVII. Free will, consciousness, and the brain

How does free will fit into a scientific picture of the world? The preceding account of free will rests on psychology and, more broadly, on biology. I have argued that the capacity to invent and evaluate possible courses of action is a distinctively human biological adaptation whose precursors are found in other animals. I have also argued that biology is not only different from physics in the ways that Mayr and other evolutionary biologists have discussed; it is also not reducible to physics. Fitness, for example, is a property of genes or gene combinations that encode particular developmental programs in members of a particular population; but for the reasons I have discussed, it is not a *physical* property. Talk about the *functions* of biomolecules, organs, and behavioral traits, is a necessary (and perfectly scientific) part of biological discourse, but it cannot be paraphrased in the vocabulary of the physical sciences.

Nor can talk about mental models and the conscious aspects of creative decision-making be translated into talk about neural circuitry and neural processes. Neuroscience seeks to understand the biological underpinnings of mental states and of conscious and unconscious mental processes, but as Max Bennett and P.M.S. Hacker[42] have argued in considerable detail and with great clarity, mental states and processes are not reducible to brain states and processes. Just as biology rests on but cannot be reduced to physics, psychology rests on but cannot be reduced to neuroscience. Nevertheless mental states and processes can be studied and at least partially understood by the methods of psychology and comparative psychology.

Do conscious acts of will cause our voluntary actions? From a thorough examination of the evidence bearing on this question the psychologist Daniel Wegner[43] has concluded that the answer is no. "Conscious will arises from processes that are psychologically and anatomically distinct from the processes whereby mind creates action [p. 29]." This conclusion accords well with the arguments and conclusions of the present essay. I have argued, as Henri Bergson did a century ago, that we act most freely when we act most creatively. Whether conscious acts of will are essential features of the extended mental processes involved in reflective decision-making is an empirical question. My own experience, for what it is worth, suggests that conscious acts of will play at most a minor part in reflective decision-making. What seems undeniable is that we believe we can alter the course of events through our plans and projects. This essay has argued that that is not an illusion.

17 August 2011

NOTES

[1] Searle, John R., 2004. *Freedom and Neurobiology*, Columbia University Press, New York

[2] Dirac, P. A. M., 1967. *The Principles of Quantum Mechanics*, fourth edition (revised), Clarendon Press, Oxford

[3] Wigner, E. P., 1962 Wigner, "Remarks on the Mind-Body Question" in *Symmetries and Reflections*, MIT Press, Cambridge

[4] Wigner, E. P., 1976. "Lecture Notes" in J. A. Wheeler and W. H. Zurek, eds., *Quantum Theory and Measurement,* Princeton University Press, Princeton NJ, 1983

[5] Everett, Hugh, III, 1957. *Reviews of Modern Physics* **29**, 454

[6] The statistical entropy of a discrete probability distribution is the mean value of the negative logarithm of the probability. The statistical entropy of a continuous probability distribution is the mean value of the negative logarithm of the probability density.

[7] Notably, classical microstates are continuous, whereas quantum microstates are discrete.

[8] Jancel, R., 1963, *Foundations of Classical and Quantum Statistical Mechanics* Pergamon, Oxford, p. xvii

[9] Jaynes, E.T., 1957, *Phys. Rev.* **106**, 620; **108**, 171

[10] Heisenberg, W., 1958. *Physics and Philosophy*, Allen and Unwin, London, pp. 53-54

[11] Schrödinger, E. 1948. *Statistical Mechanics*, Cambridge University Press, Cambridge, 1948), p. 3

[12] Landau, L. and Lifshitz, E.M., 1980. *Statistical Physic*, *Part 1*, *revised*, translated from the Russian by J.B. Sykes and M.J. Kearsley, Addison-Wesley, Reading, Massachusetts

[13] Feynman, R., 1972. *Statistical Mechanics, a Set of Lectures*, W.A. Benjamin, Reading, Massachusetts

[14] Sklar, L., 1993. *Physics and Chance*: *Philosophical Issues in the Foundations of Statistical* Mechanics, Cambridge University Press, Cambridge [England], New York

[15] Clausius, in 1857, derived these phenomenological laws from Bernoulli's model, supplemented by an additional molecular parameter: the average distance a molecule travels between collisions. And Maxwell, in the two papers mentioned above, improved Clausius's theory by constructing a detailed account of molecular collisions and their effects.

[16] Van Kampen, N.G., 2007. *Stochastic Processes in Physics and Chemistry*, 3rd ed., Elsevier, Amsterdam

[17] Prigogine, I.,1980. *From Being to Becoming*, W. H. Freeman, New York, p. 157

[18] Blatt, J. M. 1959, *Prog. Theor. Phys.* **22**, 745

[19] Shenker, O.R., 2000, "Interventionism in Statistical Mechanics: Some Philosophical Remarks" (preprint)

[20] For a clear and comprehensive review of decoherence theories, see Schlosshauer, M., 2008, *Decoherence and the Quantum-to-Classical Transition*, corrected 2d printing Springer, Berlin

[21] How two light waves with the same wavelength interfere at a given point depends on their relative phases at that point: they interfere constructively if they are in phase destructively if they are out of phase. Interference between state vectors in a superposition depends on the relative phase of their coefficients in the superposition. The interaction between the combined system in a quantum measurement and a random environment, such as a dilute gas or a radiation field, effectively randomizes the relative phases of the coefficients in the superposition predicted by von Neumann's account of an ideal measurement.

[22] Joos, E. and Zeh, H. D., 1985. *Zeitschrift für Physik B* **59**, 223

[23] Zeh, H-D, 1970/ *Foundations of Physics* **1**, 69

[24] Interaction between two systems entangles their quantum states in the same sense as the interaction between a measured system and a measuring apparatus entangles the quantum states of the system and the apparatus, or the interaction between the electrons in a helium atom entangles the quantum states of the individual electrons.

[25] Galaxies also have "peculiar" velocities associated with local deviations from uniformity. Until Hubble's observations, these masked the systematic component associated with the expansion of space.

[26] Weinberg, S., *Gravitation and Cosmology*, 1972, Wiley, New York

[27] Weinberg (footnote 26) gives a detailed account of nucleogenesis in an initially hot universe; Anthony Aguirre (*Astrophysical Journal* *** 20**) discusses nucleogenesis in an initially cold universe.

[28] Layzer, D. *Constructing the Universe,* W. H. Freeman, New York, 1984; *Cosmogenesis*, Oxford University Press, New York. 1990

[29] Layzer, D. and Hively, R., 1973, *Astrophysical Journal* **179**, 361

[30] Aguirre, A., 1999. *Astrophysical Journal* **521**, 17 (1999)

[31] Hahn, E. L., 1950, *Phys. Rev.* **80,** 580

[32] Poincaré, H. "Chance" in in *Science and Method* (New York, Dover, 2003)

[33] Dirac, P.A.M., Reference cited in Note 2, p. 47

[34] The random environment could consist of the microscopic degrees of freedom of the macroscopic measuring apparatus.

[35] Thomas Kuhn, argued in *The Structure of Scientific Revolutions* (1959) that a successor theory overturns its predecessor because *meanings* of scientific terms common to the two theories are incommensurable. But physical theories are mathematical constructs; they are not verbal-conceptual constructs that have been made more precise through the use of mathematical "language." The axioms of Newtonian mechanics are neither inconsistent nor incommensurable with the axioms of relativistic mechanics; in a precise and completely describable sense they approximate those axioms.

[36] Mayr, Ernst, 1988. "How To Carry Out The Adaptationist Program?" in *Toward a New Philosophy* of *Biology*, Cambridge, Harvard University Press, p. 159

[37] Dobzhansky, Theodosius, 1970. *Genetics of the Evolutionary Process*, Columbia University Press, New York

[38] For a more detailed version of the following argument see Layzer, "Naturalizing Free Will" (preprint)

[39] Lamb, T. D. "Evolution of the Eye," *Scientific American*, July 2011

[40] For recent reviews, see *Nature Neuroscience*, Volume 111, Number 4, April 2008, pp. 387-416.

[41] Suddendorf, T., 1998. *The Behavioral and Brain Sciences*, 21, 131

[42] Bennett, M. R. and Hacker, P.M.S.,2003. *Philosophical Foundations of Neuroscience*, Blackwell, Oxford

[43] Wegner, Daniel M., 2002. *The Illusion of Conscious Will*, MIT Press, Cambridge