THE ARROW OF TIME

Why does time never go backward? The answer apparently lies not in the laws of nature, which hardly distinguish between past and future, but in the conditions prevailing in the early universe

by David Layzer

It seems easy to distinguish the past from the future: memory provides us with a record of the past, but we have no certain knowledge of the future. When events are interpreted according to the most fundamental laws of physics, however, the distinction between past and future all but disappears. Intuitively we perceive the world as being extended in space but "unfolding" in time; at the atomic scale the world is a four-dimensional continuum extended in both space and time. We assign special significance to a particular moment, the present, which we view as the crest of a wave continuously transforming potentiality into actuality and leaving in its wake the dead past. Microscopic physics gives no special status to any moment, and it distinguishes only weakly between the direction of the past and that of the future.

Our intuitive perception of the world as unfolding in time cannot be dismissed as being merely subjective. It has objective counterparts in a variety of processes, including biological, geological and astronomical ones. There is evidence for it in the physiological processes underlying memory, in the growth, development and differentiation of living organisms, and in organic evolution, where random variation and natural selection have generated an immense and constantly increasing variety of ever more highly organized living forms. The earth's crust records the vicissitudes of 4.5 billion years of evolutionary change, and the cratered surfaces of the moon, Mars and Mercury preserve a chronological record of similar duration. Normal stars, red giants, supernovas and white dwarfs represent stages in the evolutionary life cycle of a single star. Finally, the recession of distant galaxies suggests that the entire universe is a product of an evolutionary process: it seems to have sprung from an exceedingly dense, undifferentiated state in the finite past.

All these processes have a quality in common: they generate order, or information; they transform a simpler state into a more complex one. In the phrase of Sir Arthur Eddington, they indicate which way "time's arrow" is pointing; they define what I shall call the "historical" arrow of time.

Paradoxically, the direction of time can also be defined by a class of diametrically opposite processes: those that destroy information and generate disorder. If I drop a lump of sugar into a cup of hot tea and stir, the spatial concentration of the sugar molecules, the organized motion of the tea and the difference in temperature between the tea and its surroundings represent macroscopic information, or order. As the sugar dissolves and the tea comes to rest and begins to cool, that information gradually disappears. The irreversible processes that destroy macroscopic information (in this example molecular diffusion, viscosity and heat conduction) are manifestations of the second law of thermodynamics. This law states that all natural processes generate entropy, a measure of disorder. The irreversible destruction of macroscopic order defines what I shall therefore call the "thermodynamic" arrow of time.

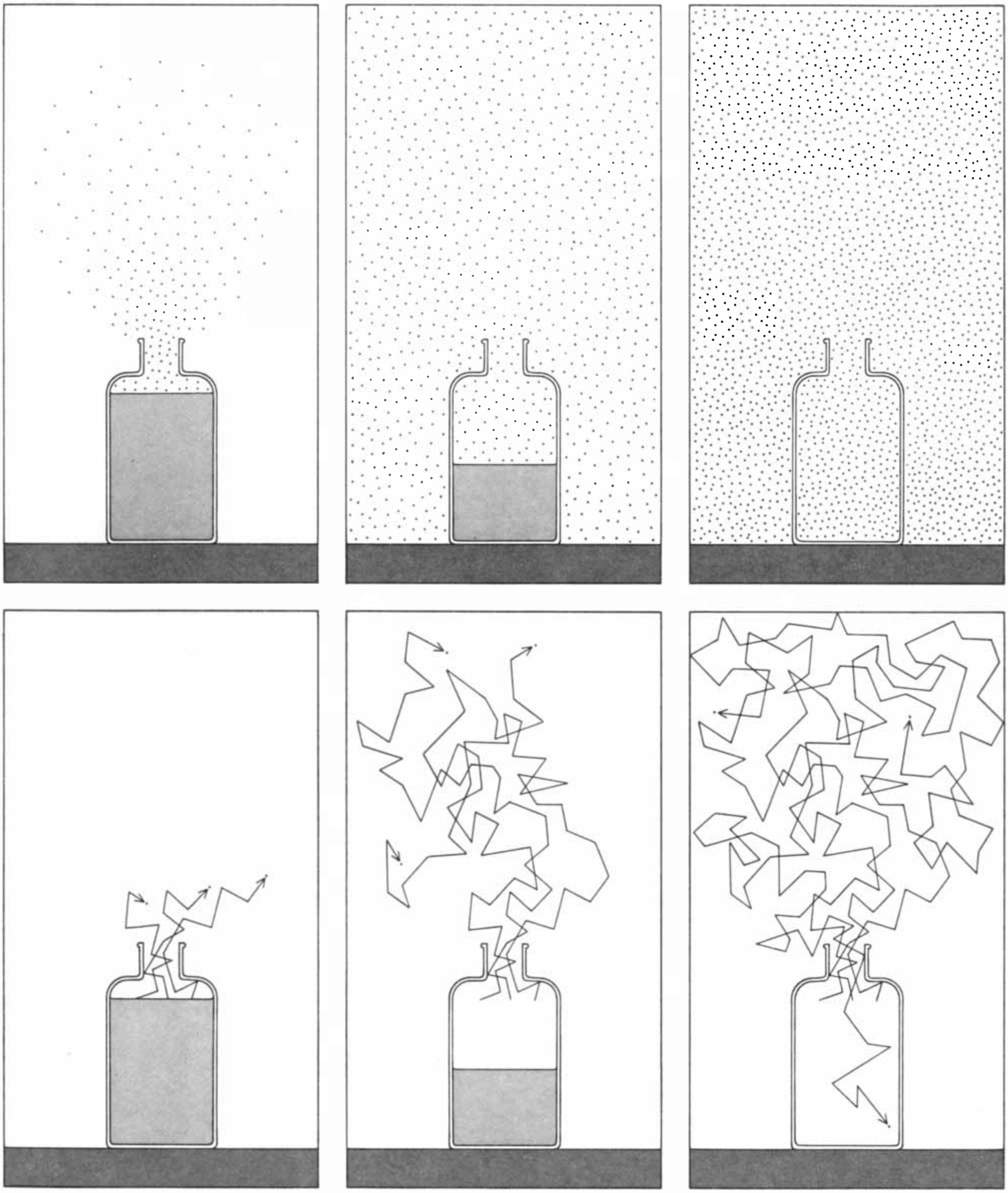
Neither the historical nor the thermodynamic arrow of time can be observed at the microscopic level. The motion of a single sugar or tea molecule generates neither information nor entropy. "Order" is a macroscopic concept, a property of systems made up of many particles; it has no meaning when it is applied to individual atoms or molecules. In the physics of elementary particles the world changes but does not evolve.

I shall argue that neither the macroscopic view of the world as a system degenerating toward complete disorder nor the microscopic view of the world as a changing but nonevolving system of interacting particles and fields is required by fundamental physical laws. I submit that both views derive instead from auxiliary assumptions about the nature and origin of the universe. I propose to replace those assumptions with others that I believe are simpler and equally consistent with observation. The resulting model of the universe, although different from the one accepted by most physicists, resolves the apparent contradiction between the historical and the thermodynamic arrows of time, and it reconciles both with the almost time-symmetric character of physical laws at the microscopic level. The theory implies that the universe is unfolding in time but not unraveling; on the contrary, it is becoming constantly more complex and richer in information.

Irreversibility

The historical and the thermodynamic arrows of time both derive from processes that always have the same direction; they are defined by events that cannot be undone. What makes these processes irreversible? All phenomena can ultimately be described as the interactions of elementary particles; if the laws that govern those interactions do not distinguish between the past and the future, what is the source of the irreversibility we observe in the macroscopic world?

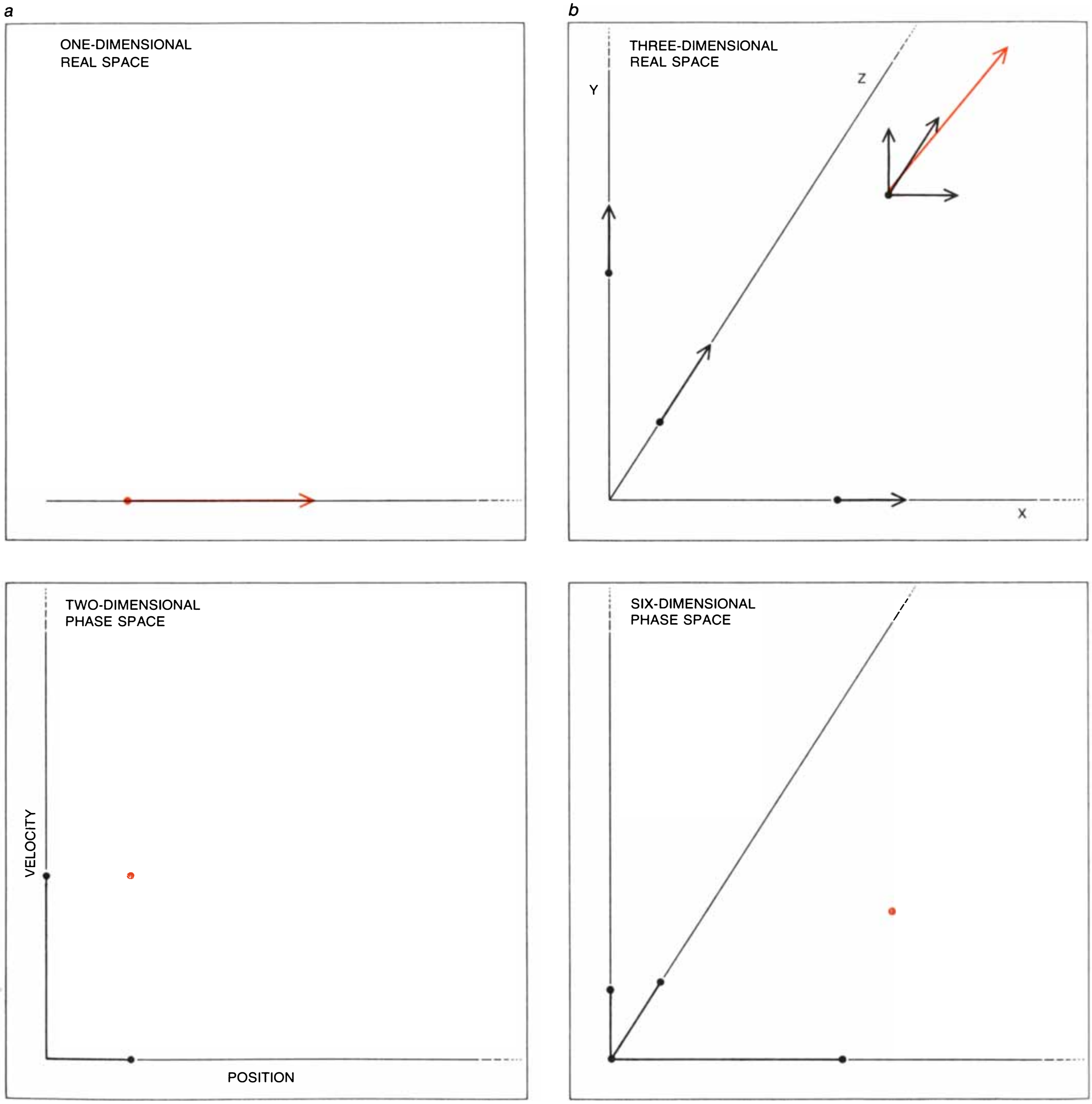
One possibility is that the underlying microscopic laws are not in fact perfectly time-symmetric. Evidence that a temporal asymmetry exists at the level of subatomic particles can be found in the decay of the neutral *K* meson. One of



THOUGHT EXPERIMENT concerning the diffusion of perfume reveals an apparent paradox: the process as a whole always proceeds in the same direction, but it is defined by microscopic events each of which is freely reversible. The bottle of perfume is opened in a hypothetical sealed room that cannot communicate with the outside world. The top series of drawings, when read from left to right, shows molecules beginning to escape the surface of the liquid and gradually filling the room, until eventually all the perfume has evaporated. When the drawings are read in the opposite direction, they represent a process never seen in nature: all the molecules spontaneously reassemble and condense in the bottle. In the bottom drawings the same experiment is depicted in microscopic detail. Individual molecules escape the surface and follow complicated zigzag trajectories that take them to all parts of the room. This sequence of events could well proceed in reverse order, since if every molecule reversed its direction, they would all retrace their paths and return to the bottle. A molecule following a reversed trajectory would obey all the laws of physics, and indeed it would be impossible to determine by examining the path of a single molecule whether it was part of a forward experiment or a reversed one.

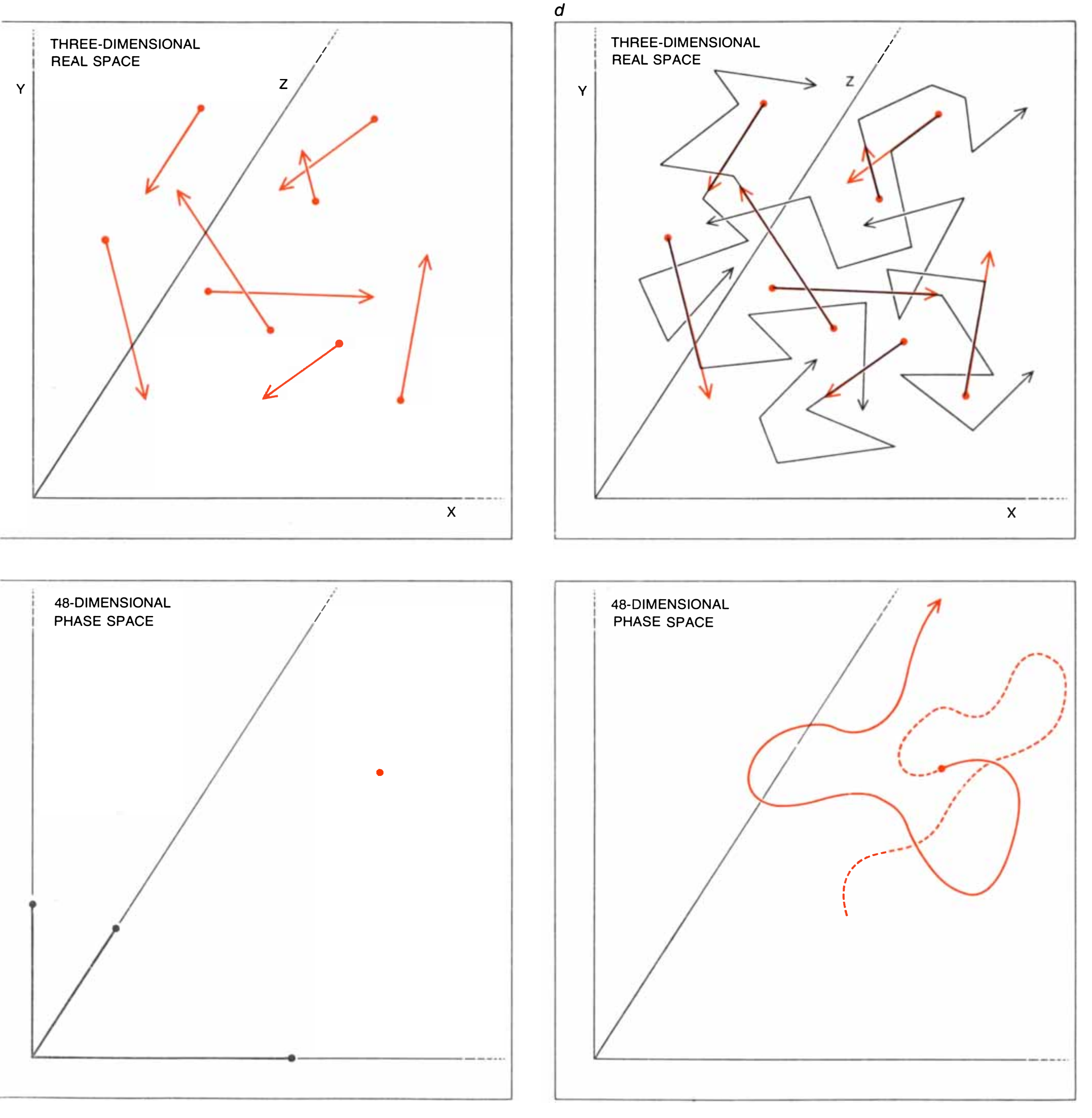
several possible decay modes of that particle seems to violate some symmetry of nature, and the usual interpretation of the event is that the violated symmetry is time-reversal symmetry [see "Experiments in Time Reversal," by Oliver E. Overseth; SCIENTIFIC AMERICAN, October, 1969]. The apparent violation, however, is quite weak: it is observed less than 1 percent of the time. Moreover, *K* mesons are found only in experiments in high-energy physics; they are not constituents of ordinary matter, and they play no role in the macroscopic processes that define the historical and the thermodynamic arrows.

If the root of irreversibility is not to be found in the laws that govern microscopic events, it must derive from constraints on how those events take place. Laws and constraints are complementary as-

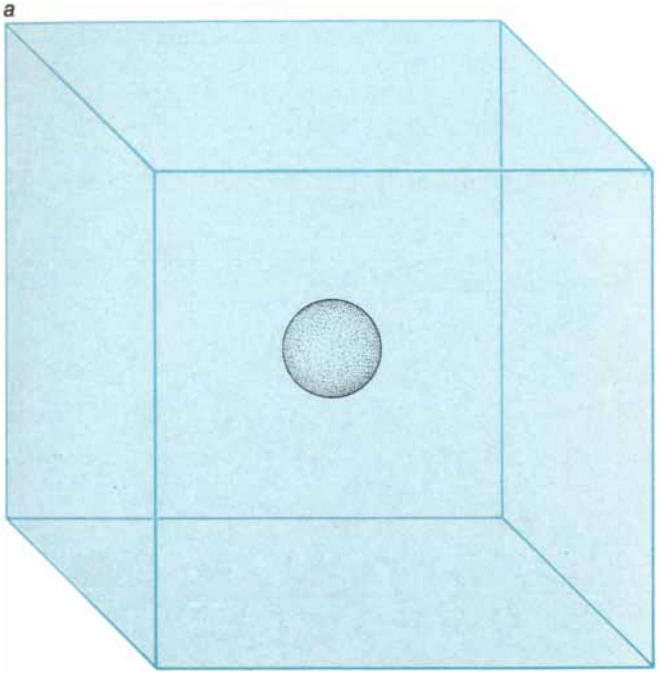


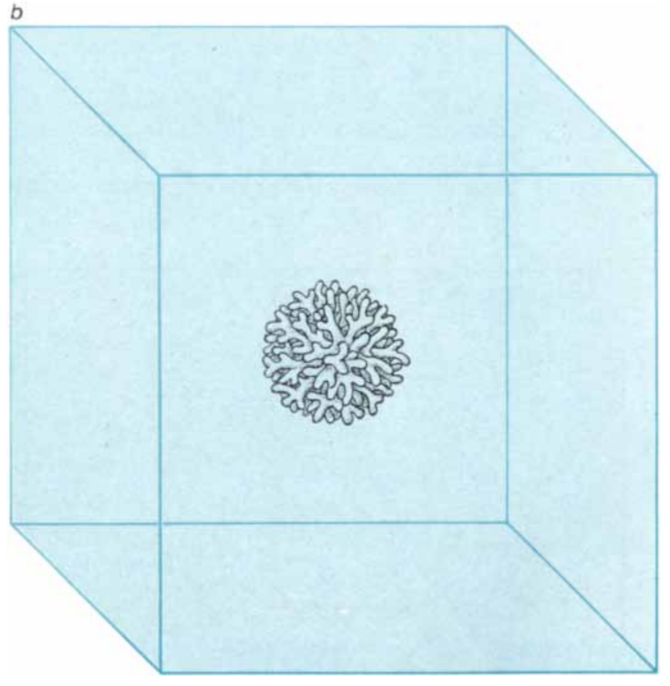
CONCEPT OF PHASE SPACE is employed to represent the dynamical state of a system of particles. Any particle can be described by a vector (*colored arrows*), which defines its position and velocity. For a particle in a one-dimensional universe (*a*) two numbers suffice to specify these quantities, and the state of the particle can be represented in a phase space having two dimensions. Every possible state of the particle corresponds to the location of some point in the phase space. A particle with freedom of movement in three dimensions (*b*) requires six numbers for the specification of its state, since both position and velocity have components on three axes. The corresponding phase space must therefore have six dimensions. Since it is not possible to construct a real space with more than three dimensions, the phase space is represented here by a three-dimensional "slice" of the six-dimensional space. In a system made up of many particles six numbers are required to specify the state of each particle, so that the corresponding phase pects of the physicist's description of nature. Laws describe the regularities underlying phenomena; they are few in number and each applies over a wide domain. Constraints serve to select from the set of all events governed by a given law the particular phenomenon of interest. The laws define what is possible, the constraints what is actual or relevant. The constraints can take the form of initial conditions, boundary conditions or symmetry conditions.

As an illustration of how laws and constraints jointly shape phenomena, consider the motions of the planets in the solar system. From Newton's law of gravitation one could calculate all the past and future positions of the planets, given their positions and velocities at some initial moment. Newton's law explains why each planet moves in an elliptical orbit



space must have a number of dimensions equal to six multiplied by the number of particles. For example, a system of eight particles (*c*) could be represented by a point in a phase space having 48 dimensions. All the information needed to specify the state is embodied in the location of that single point, and every possible state corresponds to a unique point in the 48-dimensional space. The axes of the three-dimensional slice shown are chosen arbitrarily from among the 48. As a system of particles evolves (*d*), its dynamical state changes, and the change is reflected in the movement of its representative point in phase space. The path of the point in both the past and the future is entirely determined by its initial position, and the dynamical history of the system of particles can therefore be predicted in complete detail. Moreover, the point in phase space can follow the same path in either direction, so that the motions of the particles are fully reversible. Once again three dimensions have been selected arbitrarily from among 48.





PROBABILISTIC REPRESENTATION of a system composed of many particles gives a more realistic portrayal of the system's behavior. Probability is represented by a fluid in phase space; the mass of fluid in any region corresponds to the probability that the point representing the state of the system will be found in that region. In the experiment with perfume diffusion all the probability fluid is initially confined to a small volume, since all the molecules of perfume are confined to the bottle. The form of the fluid is actually a hypersphere having $6n$ dimensions, where n is the number of perfume molecules, but it is represented here by a sphere having

with the sun at one focus, why a line connecting the sun to a planet sweeps out equal areas in equal times and why the squares of the planets' orbital periods are proportional to the cubes of their orbital diameters. It could explain those observations for any planetary system. On the other hand, the law of gravitation does not explain why the orbits of the planets are nearly circular, why the orbital planes nearly coincide or why all the planets revolve around the sun in the same direction. As Newton himself recognized, these regularities must arise from initial conditions.

To explain the regularities we would need a theory of the formation of planets. Such a theory could not supply the detailed initial conditions of the solar system, but it would specify certain statistical properties of the primordial systems from which planetary systems, including our own, have evolved. This theory would itself proceed from particular initial conditions, which would in turn exhibit statistical regularities inviting theoretical explanation at a deeper level. In that way we would be led to formulate a series of increasingly general cosmogonic problems, whose solutions would yield increasingly general explanations of the statistical regularities of the astronomical universe. This hypothetical chain of cosmogonic theories must ultimately terminate in a set of constraints, including initial conditions, for the universe as a whole. It is in those cosmological constraints that we can expect to find the seed of irreversibility.

Information and Entropy

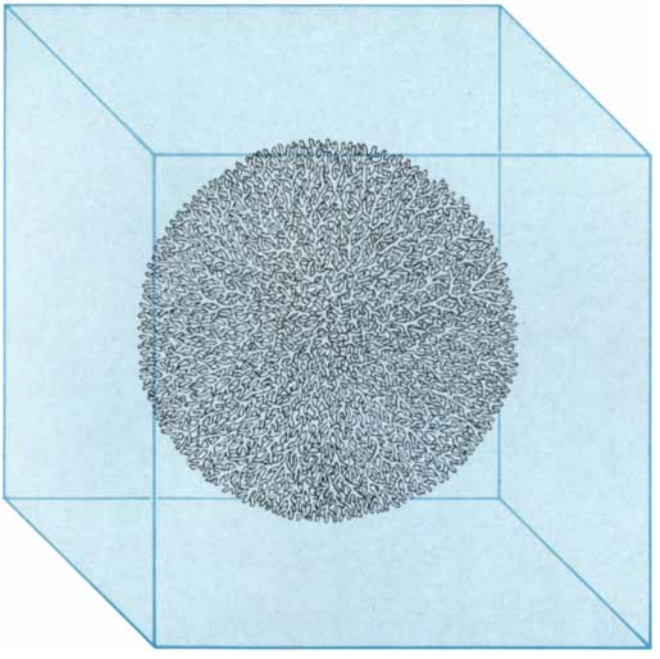
The processes that define the historical and the thermodynamic arrows of time generate information and entropy respectively. As Claude E. Shannon of the Massachusetts Institute of Technology showed in 1946, information is a property of statistical descriptions of physical systems. It is measured in bits, or binary digits; one bit is the quantity of information needed to decide between two equally likely possibilities. Information can also be regarded as a property of physical systems themselves, a measure of how highly organized they are. A fundamental theorem proved by Shannon shows that the information content of a system is the minimum number of bits needed to encode a complete statistical description of the system.

The concept of entropy is closely related to the concept of information. Entropy was first defined (by Rudolf Clausius and Lord Kelvin) in the context of thermodynamics, and it measures the displacement of a system from thermodynamic equilibrium; at equilibrium the entropy assumes its maximum value for given values of temperature and density.

Employing a formula first derived by Ludwig Boltzmann and J. Willard Gibbs, Shannon defined an entropy of information theory, which measures the uncertainty associated with statistical descriptions of a system. The thermodynamic entropy of Kelvin and Clausius and the statistical entropy of Boltzmann, Gibbs and Shannon have identical mathematical properties: they are aspects of a single concept.

Entropy and information are related by a simple conservation law, which states that the sum of the information and the entropy is constant and equal to the system's maximum attainable information or entropy under given conditions. Expressed mathematically, the law states: $H + I = \text{constant} = H_{\max} = I_{\max}$, where H (the Greek letter eta) and I represent the actual values of entropy and information and $H_{\max}$ and $I_{\max}$ are the maximum possible values. Thus a gain of information is always compensated for by an equal loss of entropy.

Suppose some physical system has eight (or 2^3) possible states; in binary notation they could be labeled 000, 001, 010, 011, 100, 101, 110 and 111. The specification of a particular state, for example the one labeled 101, requires three binary digits, which is the quantity of



three dimensions (a). As the system evolves, the probability fluid must migrate to more distant regions of the phase space, but because the trajectories of all the particles are determined the fluid is incompressible; it cannot expand as a gas does. Instead it puts out "fingers" (b), which become more elongated and more numerous (c) as the number of possible states of the system increases. Eventually the entire hypervolume is filled with small branches of fluid, although the total volume of fluid remains constant. From a macroscopic point of view the distribution of fluid now seems uniform, although it is far from uniform when it is examined closely (d).

information associated with the description "The system is definitely in the state 101." The uncertainty or entropy associated with this description is evidently zero. At the other extreme, if we had no information about the state of the system, we would be compelled to assign equal probabilities to each of the eight possible states. In this case the information is zero. Since the sum of the entropy and the information in the system is constant, the entropy must now be three bits. In general, if a system has 2^r possible states, where r is an integer, the maximum quantity information or entropy is equal to the logarithm to the base 2 of 2^r, or r.

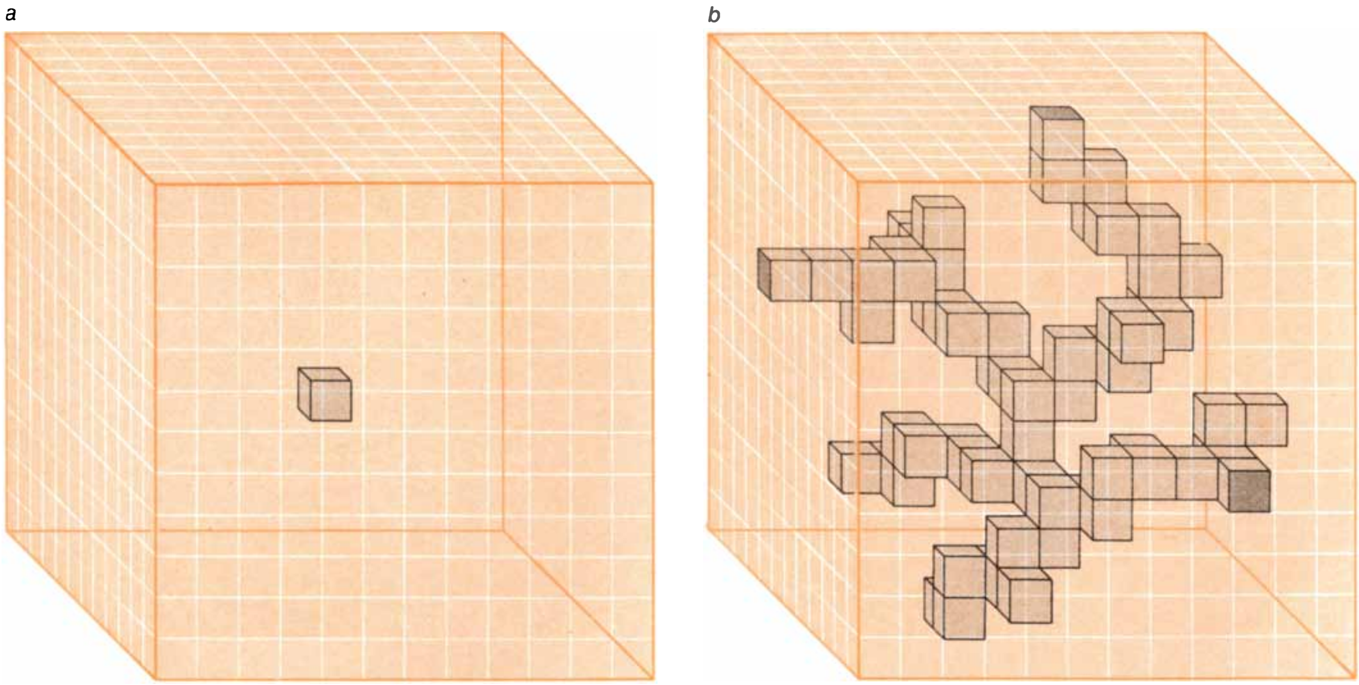
A Thought Experiment

For real systems the number of possible states can be very large, but it is not infinite. The number of states, and therefore the maximum quantity of information, is limited by the uncertainty principle formulated by Werner Heisenberg. The principle states that there is an irreducible uncertainty in our knowledge of the position and momentum of any real particle; the state of the particle cannot be specified with greater precision than our uncertainty allows. As a consequence of the uncertainty principle any finite physical system can be completely described with a finite quantity of information.

In order to understand the connection between cosmology, entropy and information let us conduct a simple thought experiment. In one corner of a room in which the air is perfectly still I open a bottle of perfume. A little later my partner in the experiment, standing in the opposite corner, reports that he can smell perfume. Molecules of perfume have evidently escaped from the surface of the liquid and, colliding with other molecules and following complicated zigzag paths, have made their way across the room. If we wait long enough, all the perfume will evaporate and perfume molecules will be found distributed uniformly throughout the room [*see illustration on page 57*].

Experience and the second law of thermodynamics tell us that the process is irreversible: no matter how long we wait, the perfume molecules will not spontaneously reassemble in the bottle. In principle, however, such an event is not impossible. Imagine that the entire experiment could be recorded on film in microscopic detail, so that we could follow the motions of each perfume molecule individually. We might see a particular molecule in the liquid accelerated by a collision so that it is able to escape from the surface. It then collides with thousands of other particles in the air, and with the sides of the container and the walls of the room, each time changing its speed and direction. At length we find it, still in motion, in a distant part of the room. If such a film were viewed in reverse, we would see the perfume molecules retracing their complicated trajectories, converging on the perfume bottle and there uniting to form a liquid. If we singled out a particular molecule, we would find that its trajectory obeyed all the laws of physics, since the laws that govern molecular motions are symmetric with respect to time reversal. Nothing about the trajectories of individual molecules would enable us to distinguish between the actual film and the reversed one. Why, then, would we be reluctant to accept the reversed film as a record of a real event?

The obvious and conventional answer is that in the reversed film the initial conditions are exceedingly special. At the beginning of the reversed film each one of an enormous number of molecules is on a trajectory that will eventually lead it to converge on a certain small volume of space (the perfume bottle) to the exclusion of all other similar volumes (the rest of the room). Such an initial state is extremely improbable, and that has often been taken as sufficient explanation of the irreversibility of thermody-



INFORMATION THEORY **provides a quantitative interpretation of the distribution of probability fluid. The phase space is divided into small cells of equal hypervolume, and in the initial state all the fluid is assumed to be confined to one cell (*a*). The information required to specify that distribution is equal to the logarithm to the base 2 of the number of cells. As the system of particles evolves, probability fluid occupies progressively more cells (*b*), and eventually the distribution becomes uniform; each cell contains an equal volume of fluid (*c*). The state of the system is then indeterminate, and no information is required to specify it. When the**

namic processes. It is possible to pursue the matter further, however, and ask what makes those initial conditions so unlikely.

Phase Space

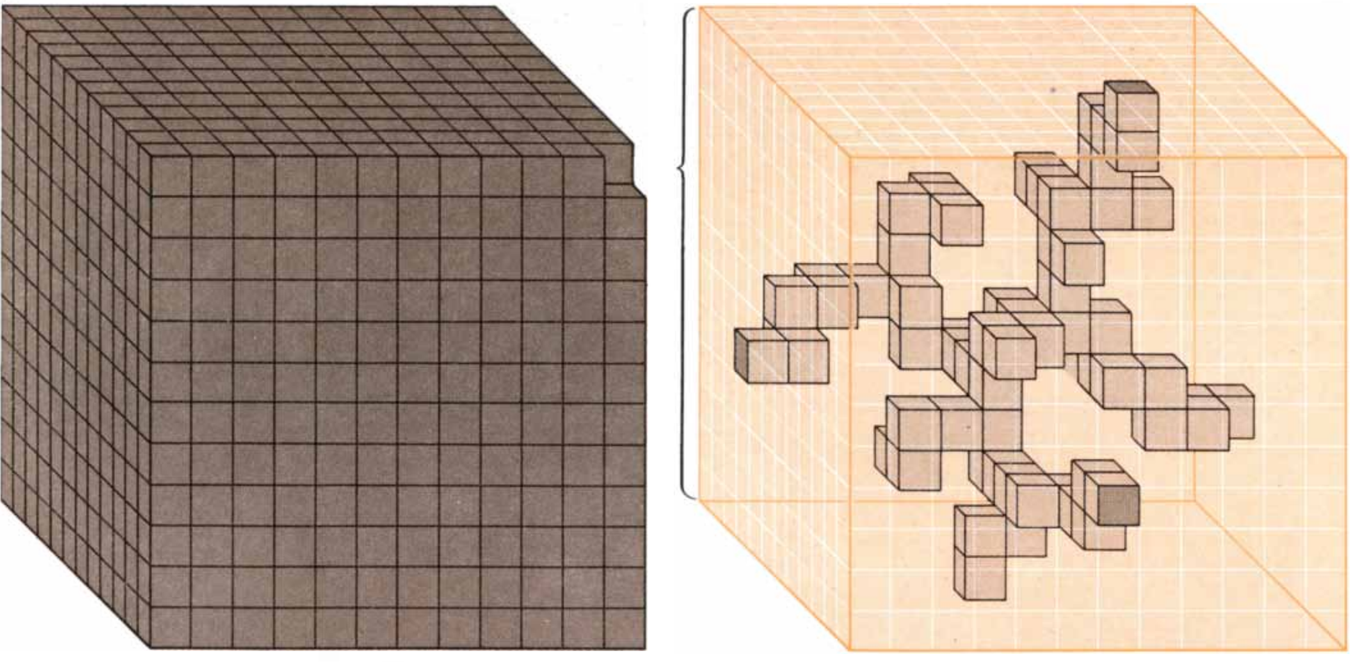
In order to discuss the question we need a convenient means of representing the changing dynamical state of a system containing a vast number of particles. The concept of phase space meets this need.

The dynamical state of a single particle is completely described by its position and its velocity. To express those quantities we need six numbers: three coordinates for the position and three components of the velocity. In Cartesian coordinates the numbers correspond to position and velocity along the x, y and z axes. We may think of the six numbers as the six position coordinates of a point in a space having six dimensions, the particle's phase space. To every point in the phase space there corresponds a unique dynamical state of the particle in real space, and as the particle moves in real space its representative point traces a curve in phase space. If we know the particle's position and velocity at any one moment, we can predict all its subsequent motion with arbitrary precision; in other words, the dynamical history of the particle is completely determined by its initial conditions. Similarly, in phase space the entire curve is determined by its starting point. Moreover, the path of the point in phase space cannot intersect itself or form branches (although it can describe a closed curve). If it did intersect itself, there would be a state of the particle (the state at the point of intersection) with more than one possible subsequent state, and the dynamical history of the particle would not be uniquely determined.

We can use the same technique to describe a closed system of many interacting particles. The dynamical state of a system of n particles is specified by $6n$ numbers: the three position coordinates and the three velocity components of each of the n particles. We may think of these numbers as the coordinates of a point in a space having $6n$ dimensions; to describe the system we must specify the location of a single point in that phase space. Again, the dynamical history of the system is represented by a curve in phase space that is completely and uniquely determined by its starting point. Because of collisions and other interactions between the particles the curve in phase space may have a complicated or irregular shape but it cannot branch or intersect itself.

The diffusion of perfume molecules in our thought experiment is represented by a unique trajectory in a phase space having $6n$ dimensions, where n is the number of perfume molecules. (Typically n is very large; if there is just a gram of perfume, it is about 6×10^{20}.) The trajectory links the points that represent the initial and final states of the experiment, but if we were to examine those points, we could make no qualitative distinction between them. Each point follows as a consequence of the other, and the description of the motions of the molecules is fully reversible.

The analysis of our thought experiment in phase space seems to abolish the arrow of time and to imply an absolute determinism that leaves no room in the future for novelty. That description, however, is unrealistically precise. It assumes the existence of much more information about the system of perfume molecules and air molecules than we could possibly acquire. We do not know the precise initial positions and velocities of the 6×10^{20} perfume molecules, even within the limitations of the uncertainty principle. We know only that all of them are initially confined to a certain small volume, the bottle. As a result we cannot specify the precise coordinates of the system's representative point in phase space; all we can say is that it must lie somewhere within a certain small volc d



phase space is examined at finer scale, however, it is found that the distribution of fluid is far from uniform. By dividing each cell into many smaller cells (*d*) one can show that the amount of information needed to specify the state of the entire system remains unchanged; it is still equal to the logarithm to the base 2 of the number of cells. Information in the initial condition about the macroscopic state of the system has been converted into information about the microscopic state. It can be proved that if microscopic information is initially absent, all the macroscopic information in the system will be converted into microscopic information.

ume, or "hypervolume," of $6n$-dimensional space.

To represent this information let us replace the point in phase space by a blob of imaginary fluid that uniformly fills the small hypervolume corresponding to our actual knowledge of the initial state. The imaginary fluid represents probability, and the mass of fluid in any region of the phase space represents the probability that the dynamical state of the system corresponds to a point within that region [*see illustration on pages 60 and 61*].

How does the probability fluid spread in phase space as the perfume molecules diffuse in physical space? One might guess that it would simply expand in all directions, just as the perfume does, and ultimately fill all the available hypervolume more or less uniformly. Actually it behaves quite differently.

Because the motions of the perfume molecules are completely determined by their initial state (even if we do not know that state), the probability fluid must remain a single, continuous blob. If it were to break up into two or more separate blobs, there would be a dynamical history represented by a branching trajectory, which we have seen to be impossible. Moreover, the volume of the blob cannot change, because the volume is proportional to the number of distinguishable states allowed by the uncertainty principle, and that number cannot change as long as each state defines a unique dynamical history.

From these considerations we can conclude that the probability fluid is continuous and incompressible; it behaves more like a liquid than like a gas. It expands into the hypervolume not by changing its density, as a gas would, but by sending out "fingers" that grow longer and narrower and more numerous as the system evolves. Gibbs compared the process to the manner in which India ink slowly spreads in still water.

As the probability fluid extends fingers at smaller and smaller scales, the total hypervolume occupied by the fluid remains constant but the shape of the occupied region grows steadily more complex. After enough time has passed the fluid will appear to be distributed uniformly throughout the entire hypervolume; when the fluid is examined at a very small scale, however, it will be found that the distribution is still far from uniform. In this description of our thought experiment we have found a striking difference between the initial state and the final state. At the outset the probability fluid is confined to a small region of phase space, which it occupies uniformly; the rest of the hypervolume is empty. In the final state the fluid occupies the entire hypervolume. From a macroscopic point of view it appears to be distributed uniformly, but at microscopic scale its distribution is highly nonuniform.

Flow of Information

The distinction between a uniform and a nonuniform distribution of probability fluid represents a qualitative difference in the information content of the system. In order to measure the information we must divide the accessible region of phase space into small cells of equal hypervolume; for convenience we shall divide it into 2^r cells, where r is an integer [*see illustration on these two pages*]. Initially all the probability fluid is confined to one of the cells. The information required to specify that state is simply the number of binary digits needed to designate a particular cell. The required number of bits is the logarithm to the base 2 of the number of cells, or $\log_2 2^r = r$. Thus the initial state of the thought experiment can be represented by r bits of information.

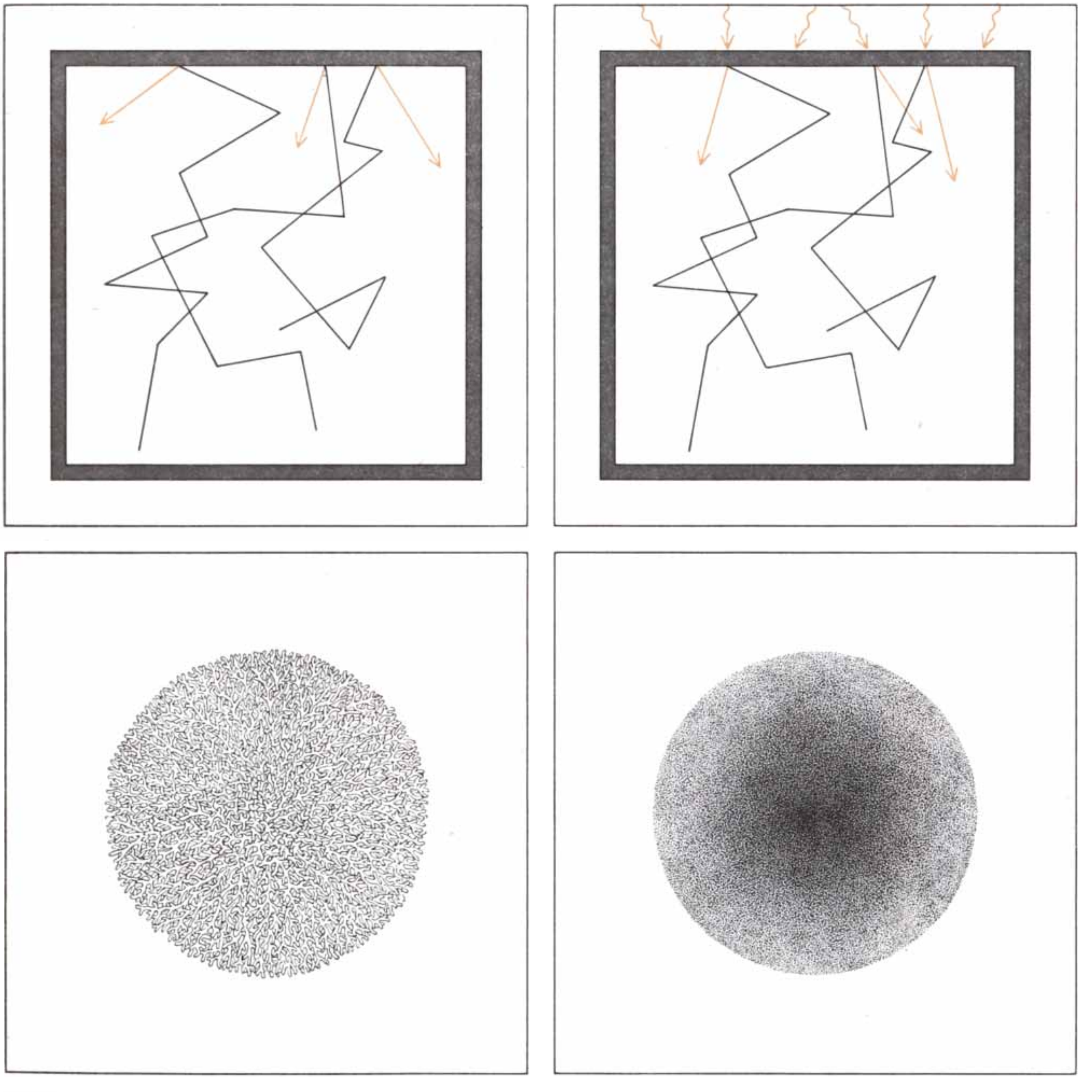
In the final state, when the fluid is distributed uniformly among the cells, each of the 2^r cells contains the same volume of probability fluid. At that level of description the final state is completely indeterminate and no information is needed to specify it. In the evolution of the system all the information contained in the initial state seems to have disappeared.

If we examine the distribution of the fluid on a finer scale, however, we can discover where the information has gone. If each cell contains an equal volume of fluid and the total volume of fluid has not changed, then within each cell the probability fluid must occupy only $1/2^r$ of the cell's volume. Although the density of the fluid has not changed, the shape of the occupied region is now very complex. By dividing the cell into sufficiently small "microcells" it can be shown that the information required to specify the distribution of the fluid in the entire region of phase space is again $\log_2 2^r = r$. The macroscopic information present in the initial state has not disappeared; it has merely been converted into microscopic information in the final state.

This conclusion can be made completely general and precise. No matter how we choose to partition the phase space into "macrocells," we can define macroscopic information as the information needed to specify the set of probabilities associated with these macrocells; the information needed to specify

a

b



RANDOM PERTURBATIONS from outside a system of particles tend to dissipate microscopic information. In a system that cannot communicate with its environment (*a*) the paths of all particles are forever determined, and the probability fluid in the system's phase space is incompressible. No real system, however, is completely isolated. Heat is communicated through the walls of any container, and particles inside a container can interact gravitationally with distant matter. As a result random disturbances destroy all information about the microscopic state of the system. Because the future state of the system can no longer be predicted from its present state the probability fluid is no longer incompressible (*b*); it expands like a puff of smoke to fill the entire region of phase space.

the distribution of fluid within the macrocells we define as microscopic. As the closed system of molecules evolves, the total quantity of information needed to specify the distribution of probability fluid in the system's phase space remains constant, but macroscopic information can be converted into microscopic information and vice versa.

What do these two kinds of information represent? We can identify macroscopic information with our knowledge of the statistical properties of the system, and microscopic information with our detailed knowledge of the individual molecules. In particular, microscopic information represents our knowledge of correlations between the velocities of particles. In our thought experiment microscopic information was initially absent, because in the initial state there were no correlations between molecular velocities; knowledge of the velocity of one molecule would not have enabled us to predict the velocities of any other molecules. As the system evolved, collisions created correlations between particle velocities, and all the macroscopic information present was ultimately converted into the microscopic information represented by those correlations.

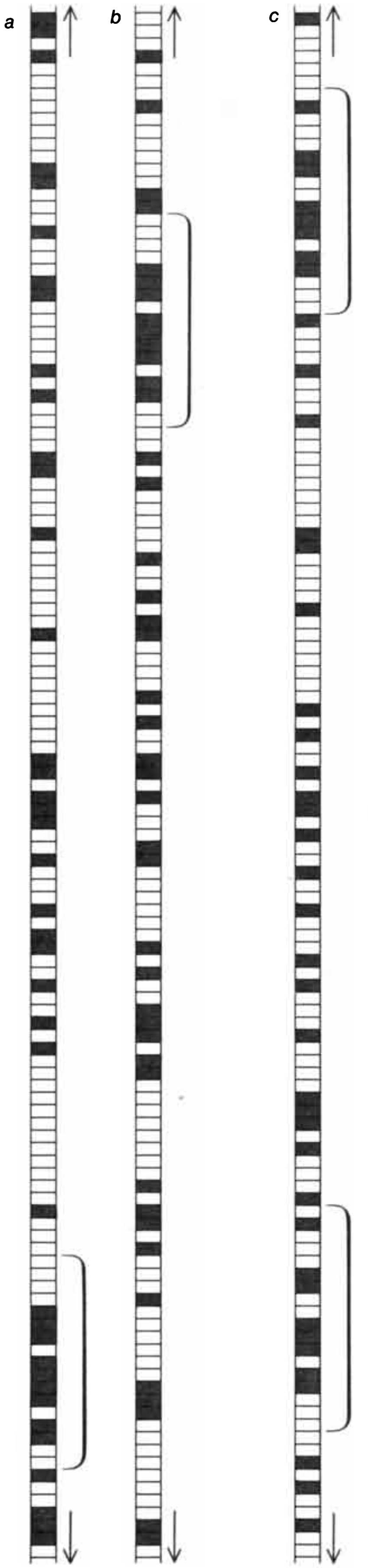
For certain kinds of physical systems and under certain initial conditions it can be shown that this process is inevitable. If microscopic information is initially absent from a closed system composed of many interacting particles, then the information needed to specify the system's macroscopic state must decrease steadily until it has all been converted into microscopic information. Since 1946 theorems with this general form have been established (for particular classes of physical systems and for particular definitions of microscopic information) by Nikolai Bogolyubov, Leon C. P. van Hove, Ilya Prigogine, Radu Balescu, Mark Kac and others.

Because macroscopic information invariably decreases in those situations where thermodynamic entropy increases, it is tempting to define thermodynamic entropy as negative macroscopic information. In fact, such a definition leads directly to the equation presented earlier: $H + I = H_{\max} = I_{\max}$. We now interpret H as thermodynamic entropy and I as macroscopic information. The entropy is then always positive or zero, and if the maximum entropy remains constant, as it must in a closed system, then the entropy must increase as the macroscopic information decreases. We have now traced the origin of the thermodynamic arrow of time to a property of the initial states of closed systems: The entropy of a closed system will increase only if macroscopic information is initially present in the system and microscopic information is initially absent.

Random Perturbations

Those special initial conditions may provide an explanation of the thermodynamic arrow of time, but it is hardly a satisfying one. Why are those particular initial conditions regularly satisfied in nature? Microscopic information seems easy enough to generate. Why does it appear only in the final states of natural systems and not in their initial states? What does it mean to say that microscopic information is absent from a certain state? One can always acquire such information by expending enough energy. Finally, what significance can we attach to the distinction between the macroscopic and the microscopic level of description? A plausible way of dealing with these questions was described in 1912 by the French mathematician Émile Borel. In recent years Borel's argument has been rediscovered and elaborated by John M. Blatt, by Peter G. Bergmann and Joel L. Lebowitz and by Philip Morrison.

Our conclusion that the microscopic information of a system increases as the macroscopic information decreases is valid only for closed systems, that is, systems that do not communicate with their surroundings. Borel showed that no finite physical system can be considered closed. For example, consider the room in which we conducted our thought experiment on the diffusion of perfume. Even if the room has no door or windows, and even if the walls are insulated



"TOY UNIVERSE" consists of a straight line extending infinitely in both directions and divided into small domains that are either occupied (*dark squares*) or empty (*open squares*). If the distribution of squares obeys certain statistical properties, it can be shown that the toy universe contains no microscopic information. The detailed properties of particular sequences of squares, for example, have no meaning. Such properties cannot distinguish one representation of the universe from another (*a*, *b*), since any sequence of any finite length can be found somewhere in all infinite representations. Nor can a particular sequence define a unique position in a single representation, since the same sequence is certain to be found elsewhere (*c*). The argument can be extended to the real universe, which satisfies the required statistical conditions.

and made very thick, the system of molecules cannot be isolated from the rest of the universe. Perfume and air molecules must collide with the walls, which are also in contact with the outside world. More important, it is impossible in principle to shield the molecules from gravitational interactions with distant matter. The effects of such interactions are exceedingly small, but they are not trivial. Borel calculated that the change in gravitational potential caused by displacing one gram of matter by one centimeter at the distance of the star Sirius would, in the course of one microsecond, substantially alter the microscopic state of a macroscopic volume of gas.

The unavoidable interaction of a nominally closed system with the rest of the universe operates as a small random perturbation that destroys correlations between the velocities of particles. The perturbation therefore dissipates microscopic information, and it perpetually re-creates the initial condition needed to ensure the decay of macroscopic information and the growth of thermodynamic entropy. Because the system is no longer isolated, its dynamical history is no longer completely determined. The probability fluid in phase space is no longer incompressible; it expands like a puff of smoke to fill the available hypervolume [*see illustration on page 64*]. In the real world, then, macroscopic information decays into microscopic information, but the microscopic information is dissipated by random perturbations.

The Cosmological Principle

Borel's argument hinges on the assumed randomness of interactions of supposedly closed systems with the rest of the universe. If the positions and velocities of all perturbing particles were known, we could expand the definition of a closed system to include the perturbing particles. That larger system, however, would itself be subject to external perturbations. Ultimately we would need to include the entire universe in our description. Given a complete microscopic description of the universe (within the limitation imposed by the uncertainty principle) there could be no qualitative distinction between the two directions of time, for such a description would be symmetric with respect to time reversal. Is it possible, however, even in principle, to compile such a description?

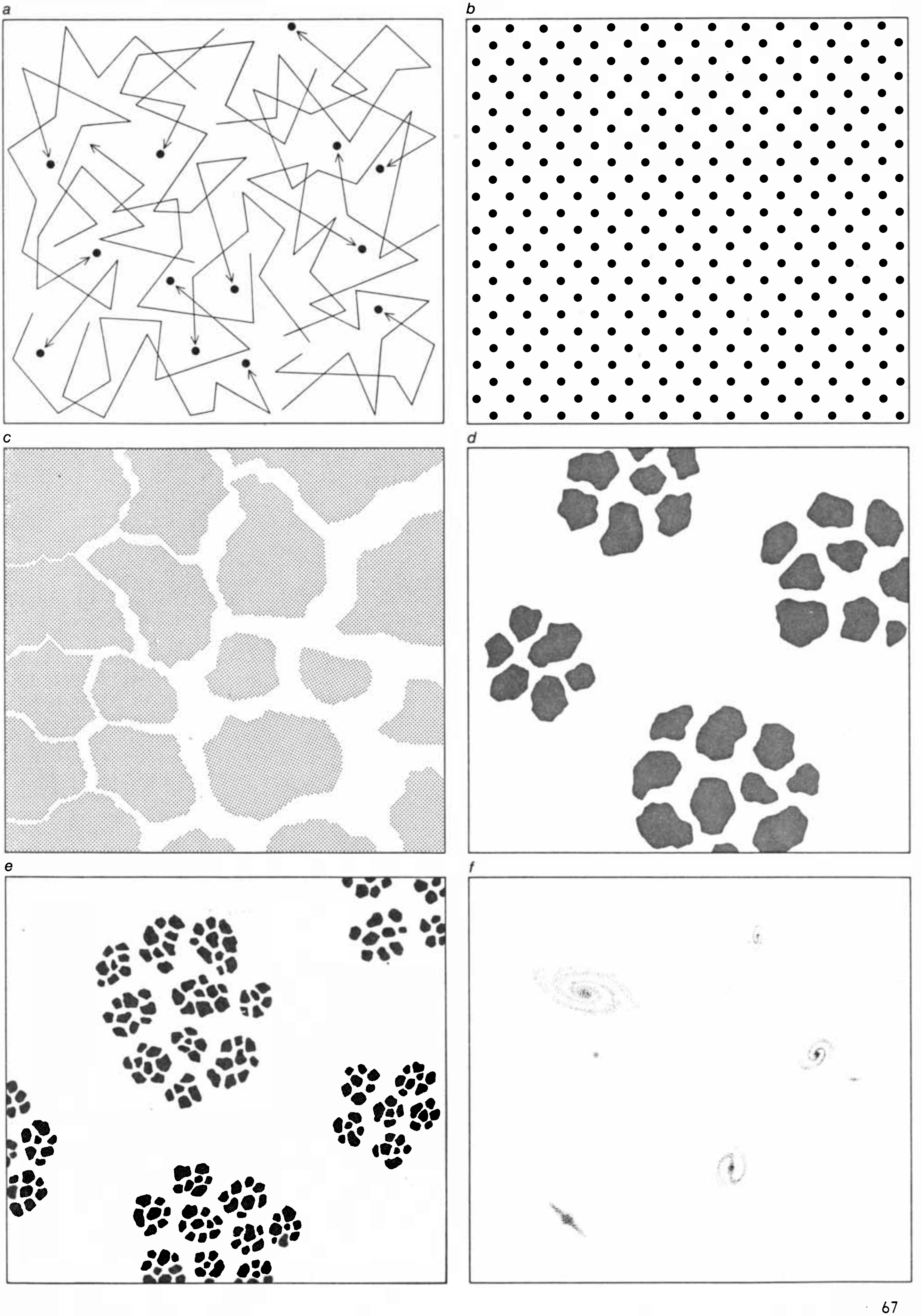
Every finite physical system admits of a complete microscopic description containing a finite quantity of information, and it may seem that the universe as a whole could also be described comprehensively. Whether such a description would contain a finite quantity or an infinite quantity of information would depend on whether the volume of the universe is finite or infinite. (Relativistic cosmology admits of both possibilities.) The universe, however, has certain distinctive properties not shared by its subsystems. In particular every finite subsystem of the universe is bounded, but the universe itself, whether it is finite or infinite, is assumed to be unbounded. Moreover, it seems to conform to what I shall call the strong cosmological principle, which states that no statistical property of the universe defines a preferred position or direction in space. The (ordinary) cosmological principle, so named by Albert Einstein in 1916, states that the spatial distribution of matter and motion in the universe is homogeneous and isotropic apart from local irregularities; the strengthened version adds that the local irregularities themselves are statistically homogeneous and isotropic. The strong cosmological principle is a direct descendant of Copernicus' thesis that our planet does not occupy a privileged position in the cosmos. It is the simplest and most comprehensive symmetry postulate one can make without contradicting observational evidence or established physical laws. It has an unexpected consequence directly related to our search for the origin of the thermodynamic arrow of time. I shall argue that the strong cosmological principle implies that microscopic information about the universe is objectively absent: it cannot be acquired or specified. This limitation of our knowledge represents a kind of cosmic indeterminacy, related to but distinct from the indeterminacy demanded by Heisenberg's uncertainty principle. It is a property of the universe as a whole but not of bounded subsystems, for which microscopic information can be freely specified or acquired.

The notion of cosmic indeterminacy can be illustrated by considering a "toy universe" of pointlike particles distributed randomly but with uniform average density along an infinite straight line. We can estimate the statistical properties of this one-dimensional universe with arbitrary precision; for example, we can estimate the mean number of points per unit length to any desired accuracy by taking averages over ever longer line segments. Can we specify any nonstatistical, or microscopic, properties of the toy universe? What would constitute a microscopic property? Suppose we are given two representations of the toy universe, identical in all their statistical properties. I assert that we define a microscopic property if we find some way to distinguish between the representations, since the only kind of information on which such a distinction could be based is nonstatistical, and hence microscopic, information.

In order to represent the influence of the uncertainty principle we must divide the one-dimensional universe into cells of equal length, the length representing the precision with which the position of a single particle can be specified. If we then specify the number of particles occupying each cell, the toy universe is represented by an infinite, doubly open-ended sequence of "occupation numbers." Microscopic information is now defined as information that would enable us to distinguish between two such sequences of occupation numbers having the same statistical (or macroscopic) properties. We might try to establish that two sequences are different by matching them up, cell for cell, along their entire length. Since neither sequence has a beginning or an end or any other preferred point, however, there are infinitely many ways to align them. It is impossible in principle to complete an infinite series of tasks, so that by this method we could never prove the impossibility of a match.

Alternatively, we could try to prove that the two sequences are identical. First we would select from one line of occupation numbers a sub-sequence of arbitrary length, then we would search for an identical sub-sequence in the oth-

EVOLUTION OF THE UNIVERSE represents a growth of macroscopic information. In a model devised by the author and his students the initial state of the universe is assumed to be devoid of all information and structure. In the period immediately following the "big bang" (*a*) the universe is in thermodynamic equilibrium, maintained by the rapid interaction of particles and radiation. After expanding for about 15 minutes, the universe crystallizes, or freezes, into an alloy of metallic hydrogen and helium (*b*). Because of the continuing cosmic expansion this solid universe shatters into fragments of approximately planetary mass (*c*), which form a "gas" in the sense that they interact frequently and randomly much like the molecules of an ordinary gas. In the planetary gas density fluctuations eventually develop (*d*) as groups of several fragments adhere; the fluctuations grow at larger and larger scales as groups of fragments themselves aggregate (*e*). Eventually a hierarchy of structures is formed, corresponding to stars, galaxies and clusters of galaxies seen today (*f*).


a
b
c
d
e
f

er line. In an infinite line any sub-sequence of finite length occurs infinitely often. The law of large numbers guarantees that our search must succeed after a finite number of trials. Moreover, we will succeed no matter how long a sub-sequence we choose, provided only that its length is finite. The two infinite sequences are operationally indistinguishable; if they were not, it would be possible to exhibit at least one sub-sequence of one that could not be duplicated in the other. Thus there is only one infinite sequence of digits with the statistical properties that define the toy universe. Two representations of the universe with identical statistical properties are indistinguishable. Since microscopic information, by definition, is what could make a distinction possible, we must conclude that it is objectively absent.

The argument can readily be extended to infinite models of the real, three-dimensional universe that satisfy the strong cosmological principle and the additional requirement that the scale of local structure be finite. The pattern of stars and galaxies visible from the earth is so complex and distinctive that it would seem to define our position in the universe as uniquely as a thumbprint defines its owner. In an infinite, statistically homogeneous and isotropic universe, however, it is certain that the same pattern of stars and galaxies recurs repeatedly. If our universe satisfies the strong cosmological principle, its meaningful properties are all statistical and its microscopic state is completely indeterminate. Since the time of Newton it has been implicit in cosmological thinking that the universe, in principle, admits of a complete microscopic description. We can now see that that need not be so. If there is enough symmetry in the universe, there is no place for microscopic information.

The Origin of Macroscopic Information

We have seen that the thermodynamic arrow of time results from the absence of microscopic information and the presence of macroscopic information in the initial states of closed systems. We found that microscopic information is objectively absent in a universe satisfying the strong cosmological principle; on the other hand, we have found no reason why macroscopic information should not also be lacking. Indeed, the complexity of the astronomical universe seems puzzling. Isolated systems inevitably evolve toward the featureless state of thermodynamic equilibrium. Since the universe is in some sense an isolated system, why has it not settled into equilibrium? One answer, favored by many cosmologists, is that the cosmological trend is in fact toward equilibrium but that too little time has elapsed for the process to have reached completion. Fred Hoyle and J. V. Narlikar have written: "In the 'big bang' cosmology the universe must start with a marked degree of thermodynamic disequilibrium and must eventually run down." I shall argue that this view is fundamentally incorrect. The universe is not running down, and it need not have started with a marked degree of disequilibrium; the initial state may indeed have been wholly lacking in macroscopic as well as microscopic information.

Suppose that at some early moment local thermodynamic equilibrium prevailed in the universe. The entropy of any region would then be as large as possible for the prevailing values of the mean temperature and density. As the universe expanded from that hypothetical state the local values of the mean density and temperature would change, and so would the entropy of the region. For the entropy to remain at its maximum value (and thus for equilibrium to be maintained) the distribution of energies allotted to matter and to radiation must change, and so must the concentrations of the various kinds of particles. The physical processes that mediate these changes proceed at finite rates; if these "equilibration" rates are all much greater than the rate of cosmic expansion, approximate local thermodynamic equilibrium will be maintained; if they are not, the expansion will give rise to significant local departures from equilibrium. These departures represent macroscopic information; the quantity of macroscopic information generated by the expansion is the difference between the actual value of the entropy and the theoretical maximum entropy at the mean temperature and density.

This argument does not depend on the cosmic expansion as such but on the finite rate at which density and temperature can change. The conclusion would be the same if the universe were contracting from a state of equilibrium instead of expanding: if the rate of contraction were greater than the rates of those processes that maintain thermodynamic equilibrium, both macroscopic information and entropy would increase. The result therefore does not fix the direction of the cosmological arrow of time with respect to the direction of the thermodynamic arrow. It does establish that macroscopic information and entropy are generated as the universe evolves away from a hypothetical state of local thermodynamic equilibrium.

Is it reasonable to suppose the universe ever was (or ever will be) in local thermodynamic equilibrium? In order to answer this question we must compare the rates of equilibration processes (those that generate entropy) with the rate of cosmic expansion or contraction. Neither rate is constant. As we proceed backward in time toward the big bang the rate of expansion increases, and at the origin of time—the cosmological singularity—the expansion rate is infinite. The rates of equilibration processes also increase, however, as we approach the cosmological singularity, because encounters between particles become more frequent with increasing density and temperature. In fact, in the period immediately following the singularity the rates of equilibration processes are much higher than the rate of cosmic expansion. As a result the big bang is an exceedingly gentle process; local equilibration processes easily keep pace with the changing macroscopic conditions of temperature and density during the first fraction of a microsecond. It is only for this brief initial phase in the evolution of the universe that local thermodynamic equilibrium can be assumed, but from that assumption it follows that the expansion of the universe has generated both macroscopic information and entropy. Thus the cosmological arrow, the historical arrow and the thermodynamic arrow all emerge as consequences of the strong cosmological principle and the assumption that local thermodynamic equilibrium prevailed at or near the initial singularity. Remarkably, neither of these assumptions refers directly to time or temporal processes.

One final question remains if this formulation is to be considered plausible: Does the cosmic expansion generate the particular kinds of macroscopic information that characterize the universe today? It is possible that some of the information was present from the outset, perhaps in the form of density fluctuations. The question cannot yet be answered with confidence, but it is important to note that there is no theoretical necessity for structure in the initial state. My students and I have developed a model of the evolution of the universe that begins with a state of complete thermodynamic equilibrium at zero temperature [*see illustration on preceding page*]. Hence it is at least possible that the astronomical universe, with all its richness and diversity, has evolved from a state wholly devoid of information and structure. If we postulate the existence of such a primordial state, we can even dispense with the separate assumption of the strong cosmological principle. The statistical homogeneity and isotropy of the universe follow from the fact that all known physical laws are invariant under spatial translation and rotation.

Novelty and Determinism

We have now traced the thermodynamic arrow and the historical arrow to their common source: the initial state of the universe. In that state microscopic information is absent and macroscopic information is either absent or minimal. The expansion from that state has generated entropy as well as macroscopic structure. Microscopic information, on the other hand, is absent from newly formed astronomical systems, and that is why they and their subsystems exhibit the thermodynamic arrow.

This view of the world evolving in time differs radically from the one that has dominated physics and astronomy since the time of Newton, a view that finds its classic expression in the words of Pierre Simon de Laplace: "An intelligence that, at a given instant, was acquainted with all the forces by which nature is animated and with the state of the bodies of which it is composed, would—if it were vast enough to submit these data to analysis—embrace in the same formula the movements of the largest bodies in the Universe and those of the lightest atoms: nothing would be uncertain for such an intelligence, and the future like the past would be present to its eyes."

In Laplace's world there is nothing that corresponds to the passage of time. For Laplace's "intelligence," as for the God of Plato, Galileo and Einstein, the past and the future coexist on equal terms, like the two rays into which an arbitrarily chosen point divides a straight line. If the theories I have presented here are correct, however, not even the ultimate computer—the universe itself—ever contains enough information to completely specify its own future states. The present moment always contains an element of genuine novelty and the future is never wholly predictable. Because biological processes also generate information and because consciousness enables us to experience those processes directly, the intuitive perception of the world as unfolding in time captures one of the most deep-seated properties of the universe.