XII*—THE LANGUAGE OF NATURE IS MATHEMATICS—BUT WHICH MATHEMATICS? AND WHAT NATURE?

by Lawrence Sklar

ABSTRACT In theoretical physics the physical states of systems are represented by components of mathematical structures. This paper explores three ways in which the representation of states by mathematics can give rise to foundational problems, sometimes on the side of the mathematics and sometimes on the side of understanding what the physical states are that the mathematics represents, that is on the side of interpreting the theory. Examples are given from classical mechanics, quantum mechanics and statistical mechanics.

I

When philosophers of science explore the source and ground of our fundamental physical theories they often focus on the questions about how laws of nature are discovered, tested and justifiably added into the corpus of accepted scientific generalizations. There is good reason for paying so much attention to laws of nature, for it is these generalizations that provide the scientific resources needed for the prediction and for the explanation of the phenomena within the domain of a theory.

But there is another task the scientists faces that has received somewhat less scrutiny from methodologists, a task arguably even more fundamental and more important than searching for the acceptable laws of nature. This is the task of finding out how to represent the various states of nature within our theories, the states of nature whose correlations with one another are to be given by the laws.

In one sense finding the right way to represent these states of nature must be preliminary to the task of finding the generalizations. After all, how can we even express a general correlation among states until we know how to say what the states are? In another sense, however, the two tasks of finding the appropriate representation for states and of finding the correct

*Meeting of the Aristotelian Society, held in Senate House, University of London, on Monday, 11th May, 1998 at 8.15 p.m.

generalizations expressing correlations among states must usually be undertaken simultaneously. This is because it is often the case that we can only find appropriate guidance into where to look for the right representatives of the states of nature as part of a project correlative to the project of seeking for the correct generalizations. Indeed, as we shall see in some cases to be looked at shortly, there are even situations where the solution to the problem of finding the right generalizations carries with it the solution to the problem of finding the right representatives for the states. When this happens important philosophical issues arise in trying to understand what nature must be like in order that the representatives found appropriately express it.

In our fundamental physical theory it was realized very early that the appropriate language for representing both the states of nature and their correlations is to be found in mathematics. Saying exactly what that means would be no easy task, for we have no clear idea about just what kinds of expressive language we will count as being in the domain of mathematics. But, surely, anything having to do with quantity or magnitude and its formal laws or with the formal characterization of the basic geometric structures of the world is traditionally counted as mathematics. It is in these theories of magnitude and geometric structure that we find, in Greek astronomy in particular and to a lesser degree in ancient physics, the first clear successes in representing both states of the world (for example the apparent positions of the heavenly bodies) and of the correlations among these states (for example in the geometric laws governing the changes and correlations among these apparent positions). Later, in the great Scientific Revolution that built on the work of the later Medieval dynamical theorists and the heliocentric astronomy of Copernicus, Galileo enunciated directly the theme ever after presupposed as a given by all the following astronomers and physicists: that the language of nature is mathematics.

But exactly how mathematical concepts are to be employed in representing nature, how the right mathematical language is to be found and justified, and what nature must be like in order that it be so representable, is an under explored area of philosophy of science rich with methodological, epistemological and metaphysical questions. It is to the outlining of a few examples of these that I will now turn.

II

When we ask where our mathematical representation of states of the world comes from, the first thought will be that we start with some intuitive, possibly vague, concepts for describing the states. These arise, perhaps, from our primordial, evolutionarily given conceptualizations and, perhaps, from our pragmatic attempts at describing and classifying the world of our direct experience. We then refine these already given concepts by means of the formal resources of mathematics. That is certainly the way things sometimes work.

How our basic concepts for describing the spatial and temporal features of the world of our everyday experience arose is something beyond our comprehension at the present time. The same is true with our basic notions of counting and numbering, and it is true of our crude vocabulary for describing comparative magnitude, say with regard to spatial or temporal extent of things and processes. But it is clear that such concepts were available to those who first initiated the mathematical descriptions of the world that constitute formal geometry and the basic kinematics needed for describing place and the change of place that is motion.

The development of both the rich conceptual vocabulary needed for characterizing the kinematics and the dynamics of systems as described in what we now call classical mechanics was slow and fitful. The need to invoke such subtle notions as the change of the magnitude of change of place, the magnitude of acceleration, and the need to go from our intuitive notions of the push and pull of things (presumably first available to us from kinaesthetic muscular sensations) to the refined concepts of force and the even more refined notions such as energy and action, was only recognized in the long process that simultaneously developed the full vocabulary needed for the science of dynamics and correlatively developed such lawlike notions as inertia and the uniform acceleration of falling bodies. But in this development there is a clear path from the pre-scientifically existing conceptualization of states of things present in common vocabulary to the refined and extended conceptualization of states needed in kinematics and dynamics.

Many of the philosophically interesting problems about the mathematical representation of states that is needed in classical mechanics arise out of the need to continually extend one’s original mathematics in order to develop a rich enough conceptual apparatus to handle the nature of space and time as continua. At each extension questions about the legitimacy of the mathematics invoked arise. This leads to a number of different programs for dealing with allegations of mathematical illegitimacy launched against these extensions. Such programs sometimes consist in trying to justify the extensions by fitting them into a rigorous mathematical scheme and sometimes in trying to avoid the problems incurred by showing that one could do all the physics you need done with a weaker version of the mathematical representation.

In early geometry such problems already arose with the proof that the diagonal of a square whose sides had rational length could not have a length given by any fractional value, leading to the first suspicion of the existence of irrational numbers. In the great developments of classical dynamics it was the need to invoke mathematical quantities representing instantaneous velocities and accelerations, that is to say the need to represent states of the world by the values of derivatives at a point, that forced the expansion of mathematics into realms whose logical consistency and rigour were suspect.

A vast portion of the program for formalizing and rigorizing mathematics that led to the great discoveries of the logical and set-theoretical means of presenting all of mathematics in a single unified framework arose out of the recurrent process of looking for a general scheme in which one could assuage the doubts of those who found the rich but informal mathematics of physics suspiciously vague and potentially rife with contradiction. Even in this century the process continues that starts with dubious mathematics in physics, and then looks for an appropriate rigorous reformulation to do it justice. It can be seen, for example in the programs needed to rationalize Heaviside's operator calculus and in the rigorization of Dirac's mathematics of quantum mechanics by means either of von Neumann's Hilbert space methods or by means of Schwartz's theory of distributions.

Although the last century's work in the foundations of mathematics certainly alleviated many of the initial doubts that a rigorous theory of the continuum and of instantaneous values could be constructed, it eventually led to the deeply problematic aspects of the full theory of sets with its initial inconsistency and later reconstruction on an axiomatic basis that itself was rife with conceptual puzzles, puzzles such as which axioms to adopt once the naive basis of set needed to be rejected. These problems in turn led to the numerous attempts at reconstructing mathematics on a basis that, being weaker in many ways than the full apparatus of standard set theory, might avoid some of its controversial aspects. The programs of intuitionism and various strict finitisms had their inspirations here.

If one wants to reform mathematics by so limiting it, however, one is confronted by the immediate objection that all of traditional mathematics is needed if one is to adequately represent the states of nature needed for classical physics. It is in this context that we see an example of the issues noted in the title of this paper. There exist a number of programs designed to assure us that giving up full set theory in favour of some weaker, say finitist, mathematics will not leave us bereft of any of the mathematical apparatus we need to represent all of the states of nature that we really must posit in order to do classical physics. To convince us of that the programs engage in a systematic effort to reconstruct classical physics so as to eliminate the apparent need for traditional mathematics in its full richness. Typically, for instance, one might argue from the inexactness of all measurement to the desirability of eliminating from one's physics any reference to the exact real valued quantities that demand the full continuum of mathematics for their representation.

So here we have one way in which questions about the relation between the states of nature and their appropriate mathematical representatives can arise. One starts with an intuitive, pre-scientific, notion of the states and refines it using a mathematical structure whose very development is motivated by the desire to have a mathematical formalism adapted to expressing the refined intuitive concepts of state. The mathematics itself gives rise to issues of conceptual legitimacy that then feed back into various programs suggesting reconstructions of the traditional notions of the refined intuitive states. The reconstructions aim at finding reconstructed states that can simultaneously play all the roles needed in physics and yet be representable in a mathematics weakened so as to avoid alleged problematic features of the now traditional mathematics.

III

When one explores the development of the mathematical representation of states in the history of quantum mechanics, however, quite a different picture emerges. Here we find an intricate and tangled history. Curiously it is one in which the relationship of mathematical representative to the state it represents becomes philosophically interesting not so much because of questions about a variety of mathematical proposals introduced to represent intuitively given states, but, instead, because of the difficulties encountered in trying to figure out just what kinds of states of nature the uniformly accepted mathematical structure could possibly be taken to represent.

Quantum mechanics had its origin from two simultaneous but quite distinct research programs, wave mechanics and matrix mechanics. The representation problem from the wave mechanical perspective appears at first to parallel the development of mathematical representation in classical mechanics. One starts with an intuitive physical notion of a state, the notion of a wave-field taken over from the classical theories of sound, light and electromagnetic wave-fields. Then one looks for a mathematical representation of these that is also derived from the familiar and well-developed mathematical representations of classical wave-fields. Some formal variation from the traditional mathematics is needed, but initially this looks like merely a matter of technical variance from the earlier theories.

But then the problem of reconciling the inevitable spatial dispersion of wave-packets, no matter how spatially compact they may be at some initial time, with the observationally detected permanent spatial locality, perhaps even point-like locality, of the elementary particles the waves are supposed to constitute, leads to a need to radically re-understand the nature of the physical states that the mathematical wave-functions are supposed to represent.

In the case of matrix mechanics the historical development of the understanding of the representation problem is different, but equally puzzling. Here the program starts with Heisenberg’s desire to eschew altogether the physical explanation of the observables by reference to unobservable causes. Since the classical physical models ran into insuperable obstacles, his suggestion was the positivist one of trying to do without the inferred, unobservable explanatory elements of a theory. Instead a theory was to be constructed that generated the appropriate correlations among the observables, initially in this case the frequencies, intensities and polarizations of the spectral lines emitted by atoms, and that made no reference to unobservable explanatory causes.

The mathematical technique he used was to radically generalize classical Fourier series into a ‘two-dimensional Fourier analysis’, and to rely on analogy with the way in which the components of mathematical spectral analysis were associated with observable quantities in classical physics to search out the predictions derivable from this mathematical generalization of the older spectral theory. The primary motivation for all of this was the empirical fact that spectral series of lines emitted by atoms needed two integers to characterize their elements instead of the single numbers characterizing the harmonics needed according to classical theory.

Since both the wave and matrix theories predicted the same frequencies and intensities for the spectra of atoms, it seemed clear that the theories must be closely related. Here Schrödinger’s demonstration of a mathematical isomorphism between the theories and the subsequent abstract presentations of the theory by Dirac and von Neumann clarified the story greatly. But problematic questions remained, and indeed still remain. In particular it is far from clear what the states of nature are supposed to be like, the very states that the mathematical apparatus of the theory is to be taken to represent. This is the notorious interpretation problem of quantum mechanics.

Starting with a proposal of Born, it first became clear that the mathematical apparatus was to be read as providing two distinct kinds of quantities. First the theory provided a list of all of the magnitudes that represented possible outcomes when a specific observable quantity was measured on a system. Second the theory provided magnitudes that represented the probabilities that any given one of the possible outcomes would, in fact, be the value obtained in any specific measurement. But what did the existence of such a representation of the possible outcomes of a measurement and of their probabilities of being obtained tell us about the nature of the actual states of the world, the states represented in the theory by wave-functions that, in combination with operators corresponding to the observable quantities, give the predicted values that can be obtained for observable quantities and the probabilities of obtaining these values upon measurement?

This remains an issue of the greatest controversy. Seven decades after the mathematical formulation of the theory there is still no agreed upon understanding of what kind of a world the theory should be taken to describe. There are many reasons for this. Some can be traced back to the fact that although the theory provides representations of probabilities, these probabilities behave in ways utterly unlike the probabilities familiar from earlier physics where interpretations in terms of generalized frequencies or in terms of partial degrees of belief could do at least initial justice to understanding what probabilities were in the theory. In the quantum case one has the peculiar facts that the wave functions that generate the probabilities interfere with one another as if they were ordinary physical wave fields, but have their values spontaneously jump upon measurement as if they represented states of knowledge about proportions in selected ensembles.

Worse yet is the appearance in quantum mechanics of rules for applying the theory that make a discrimination between ordinary physical evolutions of systems and the transitions to be attributed to them as the result of 'measurements'. The place of a distinctive measurement process in the theory remains unclear. Proposals to treat measurement as an ordinary physical evolution describable within the theory vie with those that propose wholly novel physical processes to account for the measurement process and with strange accounts that invoke processes outside the realm of the descriptive capacity of any kind of physics to explain the distinction between ordinary dynamic evolution and the measurement transition.

Here, then, we have quite a different situation from that encountered in the mathematical representation of classical mechanics. To be sure there exists special problems in the quantum context concerning the legitimacy of the mathematics employed in the theory. As in the case of classical mechanics these have been attacked both by the methods of rigorization that seek to legitimate the application of concepts whose mathematical rigour is in doubt, and by the methods that reconstruct the original theory in ways that will allow one to dispense with some components of the mathematical rigorization that, for one reason or another, seem either damaging or superfluous to the needed physical theory. This latter approach plays a particularly important role in the quantum theory of fields.

But in the classical case we start with a conceptualization of states that is given to us from our pre-scientific experience and world-view. Our mathematical representations within the scientific theory function to make these initial conceptions precise and rigorous and, at the same time, to extend them in ways unimagined in the pre-scientific conceptualization. But the grounding of the appropriateness of the representation on the intuitive concept of a state of the world always remains. In the quantum mechanical case, though, it is the mathematical representation, derived by analogy and extension from the mathematical representations used in classical mechanics, that comes first in the matrix mechanical version of the theory. In the wave mechanical version of the theory one starts as in classical mechanics with an intuitive physical conceptualization that is then mathematically represented. But when the theory is better understood in its legitimate application, it is then realized that the mathematical representation cannot be taken to represent the original intuitive physical model. In both cases, then, the situation becomes one where the mathematical representation and its role for predictive purposes becomes quite clear, but where it remains mysterious exactly what kind of physical state of nature the mathematics is purporting to represent.

The interpretation of some of the parameters of the mathematical representation as picturing probabilities of outcomes already introduces complexity into our understanding of the kind of world the mathematical formalism is intended to represent. This is because there are famous difficulties in trying to say just what it is in the world (frequencies? limits of frequencies? proportions? idealizations of these? dispositions or propensities? states of partial knowledge of the world?) physical probabilities are supposed to be. But in the quantum case the peculiar behaviour of the probabilities (their interference) and the peculiarity of the rules for the application of the theory (its distinction between dynamic evolution and the measurement process) make the job of trying to understand just what kind of world this mathematical theory represents a task of enormous difficulty. This is the interpretive problem for quantum mechanics that has exercised the foundations of the theory for three-quarters of a century.

IV

The interpretive problems of quantum mechanics are notorious. It is less commonly realized that other fundamental branches of physics offer their own deep problems about the nature of the physical states and their relation to their mathematical representatives in the theory,

Thermodynamics describes the fundamental lawlike behaviour of energy in its manifold forms when it is not merely the conservation of energy that comes into play, but also the intrinsic tendency of systems to spontaneously evolve, even when isolated, toward certain unchanging 'goal' states, the states of equilibrium. It also deals with the limitations on the transformation of heat energy and work imposed by the fundamental irreversibility of such transformations summed up in the Second Law of Thermodynamics as well. In order to capture these fundamental facts about the behaviour of systems, thermodynamics introduces its own conceptual vocabulary of equilibrium, temperature and entropy, and other concepts derivable from them. And it posits a new range of fundamental properties of systems to which these concepts refer. Here, already, fascinating problems of interpretation are lurking.

But the current view of thermodynamics, originating from work in the middle of the nineteenth century, is one that denies a fully autonomous status to that theory. The atomistic and mechanistic theory of matter brought with it the idea that there should be some way in which the thermodynamic concepts and laws could be 'grounded' in the concepts and laws that one used to describe the constitution of matter out of its microscopic components and the dynamical behaviour of those components. This idea gave rise to the kinetic theory of matter and later to full-fledged statistical mechanics.

In its original form the theory seemed to propose that the thermodynamic quantities attributed to a system could simply be constructed out of the dynamic properties of its constituents and that the thermodynamic laws could be derived from the laws governing the dynamics of these constituents. In response to critical attacks on the earliest versions of the theory, however, probabilistic elements were added to the theory. The thermodynamic quantities were now spoken of as 'mean values' or 'most probable values' of quantities dynamically defined and the laws governing the thermodynamic evolution of systems were held to hold only 'most probably' or 'on the average'. In its usual format due to Maxwell, Boltzmann and Gibbs, the theory defines the thermodynamic quantities in terms of 'ensembles' and takes the evolution of individual systems over time to be modelled by the time evolution of such collections of systems.

But what is an ensemble? The intuitive idea is of a collection of individual systems that share some common constraining feature. But it is clear that an ensemble is not meant to represent some actual collection of systems in the world in any direct way. For one thing the ensembles are defined as nondenumerably infinite collections of systems. In retrospect what one sees happening is that it is possible to obtain the desired definitions for the thermodynamic quantities in terms of the constitutive and dynamical features of the system, and it is possible to find some dynamical surrogate for the usual thermodynamic laws, only if the theory represents the physical states of the world in a highly indirect fashion. Mathematically ensembles are simply probability distributions. The thermodynamic quantities are, then, defined not as ordinary dynamical quantities of the individual systems, but as quantities definable only with respect to a chosen probability distribution. And the laws derived are no longer laws of the behaviour of individual systems in any simple sense. They characterize, rather, the dynamical behaviour of those quantities defined out of the constitution and dynamics of the individual system only with the aid of the chosen probability distribution.

What is the probability a probability of? Basically it is a probability over the possible dynamical micro-states of a system compatible with the macroscopic constraints placed upon it. Given a volume and fixed internal energy for a gas, for example, many possible locations and momenta for the molecules of the gas are possible. The ensemble probability distribution defines probabilities for any range of such micro-state possibilities.

The very fact that the basic concepts of the theory are associated with probabilities already makes the issue of how the theory represents the actual physical states of the world problematic. How are we to understand how theories that generate probabilities are connected to what happens to individual systems in the world?

Here there have been many contending interpretive views about probability. And these alternative understandings of probability in general carry over into alternative understandings of how statistical mechanics is to be understood as representing the world.

Frequentist and proportionalist views take the theory to be referring to proportions of outcomes in actual collections of many systems that have some common constraining feature. But these accounts must resort to intricate ways of trying to tell us how the idealized probabilities of the mathematical theory connect up to the actual finitistic frequencies encountered in real collections of systems. Dispositionalist accounts of probability take probabilities as describing actual states of individual systems, but in the case of statistical mechanics seem to fail to do justice to the underlying physical assumption that exact individual micro-states, and not merely probabilities of them, characterize the individual systems. There are even ‘subjectivist’ versions of statistical mechanics that take its probabilities to refer not to the world of the experimental systems at all, but, rather, to our partial knowledge of their hidden micro-states. Some versions of such theories try to find the source of the probability distribution posited by the theory in a kind of *a priori* logic of probabilities rather than in some empirically discovered feature of the physical world. These last interpretations, not surprisingly, have difficulty in accounting for the full role of the theory in its descriptive and explanatory accounting for the world of observation and experiment.

But the difficulties in understanding how the mathematical constructs of the theory represent the actual states of nature do not end with the general problems of trying to tie probabilistic assertions to the individual occurrences of the world. Statistical mechanics works by the combining of facts about the physical constitution of the systems in question, the dynamical laws governing the behaviour of the constituents of the systems, and fundamental probabilistic posits. It attempts to derive the fundamental features of the world described by thermodynamics. These include the existence of temporally stable equilibrium states, the describability of these states in terms of a small number of physical parameters, lawlike functional relationships between these parameters when the system is in equilibrium, the fundamental facts about the temporally asymmetric tendency of systems when not in equilibrium to approach the equilibrium state, and certain lawlike descriptions about the manner in which this approach to equilibrium takes place.

But in order to obtain the appropriate consequences about the probability distributions in statistical mechanics that are to function as the 'analogies' within the probabilistic theory to the thermodynamic laws, a number of deep idealizations must be made. In deriving the equilibrium properties of systems, for example, it is often necessary to deal with systems idealized as having an infinite number of constituents and having infinite volume, an idealization rationalized by the fact that realistic systems do indeed have vast numbers of constituent molecules. In other equilibrium problems, and in order to deal with the problem of the approach to equilibrium from non-equilibrium, it is often necessary to work in an idealization in which one is concerned with what happens to a system over an infinite period of time. Here the rationale often given (a weak one) utilizes the long period of time macroscopic processes take relative to those that happen at the microscopic level. Other problematic idealizations are used as well, for example idealizing the system as having zero density (the Boltzmann–Grad limit), an idealization used in one important derivation of the laws governing approach to equilibrium. In all of these cases, however, deeply problematic issues arise about the legitimacy of mathematically representing what happens to real, finite systems in finite times by utilizing the kind of radical idealizations necessary to make the theory work.

There is another reason why the issue of the representation of the actual states of individual systems and their lawlike behaviour by the results of statistical mechanics is replete with problems. One frequently gets the result one wants in statistical mechanics only by getting it for the 'wrong' probabilistically defined quantity. For example results are often derived for 'average' values of quantities calculated using the posited probability. But what one wants are results for most probable values instead. After all, a collection of systems can have some average value of a quantity even if not one single individual system in the collection has a value for that quantity even close to that average value. Here again much critical inquiry is needed in order to understand the degree to which one has derived what one sought in the first place. And once again the issue hinges on the appropriateness of some mathematical quantity of the theory to represent the physical quantity whose behaviour one wanted explained all along.

V

At this point I would like to outline an extended example that once again shows us how many difficult problems can become entangled in the issue of finding the appropriate association of physical state and mathematical representative that will give us a conceptual understanding of an important physical theory. This is an example from statistical mechanics.

I noted above that one gets the desired results out of statistical mechanics only by adding to the facts about the constitution of a system and about the dynamics of the system's constituents some probabilistic posit over the possible micro-states of the system compatible with the constraints placed upon the system. One fundamental achievement of statistical mechanics is the discovery of the right versions of these posits. The most profound foundational questions about statistical mechanics are those about the explanatory grounds that could account for these posits. In particular there is the deep question of the extent to which the posits must be taken as additional, irreducible physical assumptions about the world. Must the posits be simply added to the remaining components of the theory, or is there some sense in which they can be derived from the other posits concerning constitution and dynamics of constituents? Actually these problems bifurcate into two separate classes of questions in statistical mechanics, one dealing with the role of the probabilistic posits in what is called equilibrium theory and the other with the posits needed in the theory of non-equilibrium.

In the equilibrium theory one seeks for a probabilistic posit such that quantities calculated with it, say by computing averages over features of the system fixed by its microscopic state, will be related to one another as the macroscopic parameters of the system are related to each other in the equilibrium condition. Such probability distributions were discovered in the earliest days of statistical mechanics. But why should that particular probability distribution be the right one to characterize the equilibrium state?

There is an explanatory account, ergodic theory, that has a curious structure. Using the underlying dynamics of the system one tries to show that the preferred probability distribution is the only such distribution that will remain constant in time as individual systems in a collection have their microscopic states vary according to the dynamical laws. If that could be shown, one would have a kind of 'transcendental deduction' of the standard probability distribution. It would have been shown from constitution and dynamics that this distribution was the unique one qualified to characterize states of the system that were invariant in time, as is, by definition, the equilibrium state.

But even this grounding of the probability distribution in the remaining parts of the theory has its difficulties. Uniqueness cannot really be shown. All that can be shown is that the standard distribution is uniquely stationary among those that agree with it about which sets of conditions get probability zero. This starts a new foundational discussion: What entitles us to ignore probability distributions that give non-zero probability to sets of states normally taken as having probability zero? After all, having probability zero does not mean 'being impossible' in any standard reading of probability over infinite sets. And since it is the very reasonableness of the standard probability that is in question, it is far from clear why sets to which that distribution attributes zero probability ought to be ignored.

In the case of the non-equilibrium theory the issue of rationalizing the choice of a probability distribution is even more problematic. Here the basic problem is to choose a probability distribution corresponding to the initial non-equilibrium state of the system and then to show that the dynamics will drive this distribution toward the one corresponding to equilibrium in an appropriate fashion. There are actually only a limited number of case where it is known how to do this. In the non-equilibrium case, however, even when we know how to choose an appropriate initial probability distribution, we do not have even the limited transcendental deduction for this distribution that is available in the equilibrium case. Distributions are known in some cases that give good predictive results and they have a kind of intuitive plausibility. But one would like a general physical and mathematical rationale for selecting the appropriate distribution. The problem of showing that the initial distributions evolve toward equilibrium as they should is also one replete with difficulties.

The mathematical language in which statistical mechanics is framed includes the mathematics appropriate for representing the underlying dynamics of the constituents of the systems, whether that be the language of classical mechanics or the language of quantum mechanics. But it also includes the special mathematical language appropriate for framing its special probabilistic assumptions. This is the language of measure theory, the conceptual framework of the mathematical formulation of probability. The basic posits of this mathematical theory are of extraordinary simplicity: A collection of events is given along with a set of subsets of that collection. A function assigns real numbers normalized to be between zero and one to members of the collection of subsets. When two subsets are disjoint, the number assigned to their union is the sum of the numbers assigned to the two subsets. This additivity postulate is generalized in the case of an infinite collection of events to countable additivity for a countable union of disjoint sets in the subset collection.

As we have noted, understanding what in the world probabilities are to be taken to represent is never a simple matter. In the case of statistical mechanics the probabilities connect up to the world in two ways. First a posited probability distribution is taken to idealize some notion of the proportions with which sets of possible micro-states of the system are realized. Second various quantities calculated by using the probability measure, average values or most probable values of functions of the micro-states of the system, are to be taken as representative of the macroscopic thermodynamic features of the system such as temperature or entropy.

There have been several proposals, however, to the effect that we need an additional mathematical representation of the world in order to supplement to the usual measure theoretic language if we are ever going to be able to fully understand, in an explanatory way, the remarkable predictive success of statistical mechanics. The concepts of the measure theoretic representation are meant to connect, in a subtle and idealized way to be sure, with the notion of the proportion of systems in a collection that have their microscopic condition in a specified range, or, alternatively, with our justifiable expectations about the range of values in which the microscopic condition of an individual system will be found. But these probability distributions do not, in and of themselves, make close contact with such matters as the ways in which nature or experimenters can control the microscopic condition of systems or prepare specified collections of them. Once an initial probability distribution is chosen, then the dynamics of the system determines the probability of the distribution for future times. But nothing in the measure theoretic representation gives us a direct clue about why the initial distributions should be as they are. In the equilibrium case this leads to the problem of justifying the neglect of sets of systems of measure zero. In the non-equilibrium case it leads to the problem of justifying the choice of appropriate measure overall.

Another branch of mathematics, topology, deals with such issues as the continuity of sets of points and their 'closeness' to one another. Without going into the details of how this works, we can simply note that this suggests that topological notions are the appropriate ones for dealing with issues such as how stable physical systems are under slight perturbations or how difficult it might be to differentiate two close initial states of systems from one another by some experimental means. In the case of the spaces appropriate for statistical mechanics, these topological notions are intimately connected with certain notions generalizing distance, so-called metric notions. The topology of the space of micro-conditions of the system is determined by the 'distances' between the points in the space representing the possible initial states of a system. The distances involved here are not distances in physical space. They are, rather, distances between points in a multi-dimensional phase–space. Each point in this phase–space represents the entire exact micro-state of a system, that is the position and momentum of every one of its molecules.

All of this strongly suggests that one might resort to continuity facts about the collection of the states of nature, and to their representation in topology, in order to try and fill some of the gaps left by the measure theoretic representation. There are numerous proposals in this vein. Some rely upon the fact that one can argue that the sets of measure zero in equilibrium theory lack stability, so that it would be impossible for experimenters or nature to contrive the construction of such ensembles of systems. The idea here is that any such attempt at preparation would be immediately frustrated by even the smallest perturbing influence on the systems from their environment.

In non-equilibrium theory another proposal for invoking topology tries to overcome the problem of the seeming lack of justification for picking any one measure as appropriate for representing the system. In measure theoretic non-equilibrium theory one can construct indicators of how randomizing the dynamics of a system is. It can be shown, once again invoking crucial idealizations, that a sufficiently randomizing dynamics will make an initial probability distribution of an appropriate sort approach, in a special coarse-grained sense, an equilibrium distribution. The most important measure of the randomizing effect of the dynamics is the Kolmogorov–Sinai entropy of the dynamical shift. The value of this measure, however, depends crucially upon the basic probability measure chosen for the phase–space, giving rise to questions about how to rationalize the choice of one particular measure over the others.

Looking at a collection of systems topologically, however, one can discern a topological measure of the randomization of the collection as well. This is called the topological entropy of the shift. Crudely, it works by asking how many points are needed to keep each point representing a micro-state in the collection within a specified distance of one of the designated points for a specified period of time and then divides that number by the time elapsed. Then limits are taken as time goes to infinity and as the specified distance goes to zero. The more the system randomizes the higher the value of this topological entropy of the shift. What turns out to be remarkable is that this topological entropy of a shift is provably the maximum over all the measure theoretic entropies for any measure having the desirable property of time invariance. The suggestion here is, then, that the topological measure of the randomization provides a measure less infected than is the measure theoretic notion by seeming arbitrariness.

The basic idea here is that the distance between the points representing the micro-states of systems, being a property of the underlying dynamics, is derivative from the fundamental structure of the system and its dynamics in a way in which a measure over sets of initial conditions is not. But new questions then arise concerning the justification for picking one such criterion of ‘distance’ rather than others that remain mathematically legitimate.

Perhaps, though, one can find a more adequate representation of nature by either supplementing the mathematical representation using measure theory by one framed in topological concepts, or even by supplanting the measure theoretic by the topological concepts.

VI

Obviously this is not the place to try and explore any of the technical results nor to enter the complex debate about the degree to which they can be successfully deployed to resolve foundational problems in statistical mechanics. What I want to convey, rather, is the way in which in this case a new set of issues about the nature of the world and the appropriateness of a proposed mathematical representation of it has arisen.

In the case of classical mechanics, we started with a pre-scientific, intuitive notion of a state. Our mathematical representation, designed to fit these intuitions, allowed us to extend and refine the intuitions as well. But it did so at the price of introducing a mathematical realm whose very power and extent led some to think it over-extended in ways that threatened its own legitimacy and to seek more modest representations that might do the job.

In the case of quantum mechanics we find a very different situation. Here a mathematical formalism is discovered that is brilliantly adequate in its descriptive and predictive tasks. But the task of trying to understand what kind of states of nature such a representation can be taken as representing becomes one of enormous difficulty.

The aim of statistical mechanics is the understanding of the nature of the thermodynamic properties of systems and of their lawlike relations to each other. But in trying to comprehend how thermodynamics fits in with the remainder of physics, with the theory of the constitution of matter and with the theory of the dynamics of matter’s constituents, it is discovered that one must think about the world in ways which go beyond the attention to the specific fundamental physical state of individual systems at particular times. One must think also in terms of probabilistic representations corresponding to something that idealizes the notion of the proportions of systems having these individual states within certain specified ranges, and in terms of how those proportions change with time driven by the dynamics governing each individual system. And one must also think in terms of topological representations that correspond to the ways in which distinct systems have their states related to one another in terms of some measure of closeness that allows us to capture our intuitions about just how these states can be experimentally discriminated from one another and what the limitations on such discriminations might be.

The fundamental roles played in the theory by both the measure theoretic representations and the topological representations force us to think deeply, then, about just what aspect of nature the theory is to be taken to represent. Our notion of the state of a system must go beyond a focus on the kind of individual state of a single system that is dealt with in ordinary dynamical theories.

When we first look at thermodynamics it appears to be telling us that we must describe individual systems by properties left out in their dynamical descriptions, properties such as temperature and entropy. Initially these seem to be properties in a normal sense, that is to say particular attributes of individual systems. There are some versions of statistical mechanics that allow us to continue to think that way. But in many interpretations of the theory, these properties are taken to be, fundamentally, not properties of individual systems, but, rather, properties of idealized collections of individual systems that share certain common constraints. Only indirectly, in these accounts, do the thermodynamics properties hold of the individual systems.

A familiar pattern appears in statistical mechanics. We invoke the notion of an ensemble, or, better, a probability distribution, to characterize the system. We represent features of that collection using the language of mathematics. Usually that mathematical language is measure theoretic, but in other cases it is topological. It is often possible, by making suitable assumptions and idealizations, to derive within the chosen mathematical representation results that bear pleasing analogies to the familiar thermodynamic laws. But it then becomes a complex problem of interpretation to understand physically and philosophically in just what sense we are justified in taking these mathematical results as truly explaining the physical facts about the individual systems that the thermodynamic concepts and laws were taken to describe.

What we find, then, are three quite different ways in which our understanding of what we are to think of as the states of the world and what we are to take as their appropriate mathematical representatives can present us with ongoing challenges. In one case, classical mechanics, the main problem is in fully understanding the mathematics. In a second case, quantum mechanics, the major problem is in understanding what we take the fundamental individual states of systems to be. And in the third case, statistical mechanics, the problem is in understanding how such states of nature as those represented by the measure theoretic and topological features of ensembles function over and above the usual individual dynamical states of systems to allow us to describe the world and predict and explain its behaviour.

Department of Philosophy
The University of Michigan
Ann Arbor
MI 48109
USA