30. See L. Rosenfeld, "The Velocity of Light and the Evolution of Electrodynamics," *Nuovo Cimento*, supplement to vol. IV, series X, no. 5 (September 1956): 1664.

31. *Enzyklopädie der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen*, vol. V, part 2 (Leipzig: B. G. Teubner, 1904–1922), article 13: "Maxwells elektromagnetische Theorie," by H. A. Lorentz, §§44–45.

32. Ibid., p. 144.

33. Ibid., p. 143.

34. See the references given in note 13 above; also Stein, "'Subtler Forms of Matter,'" pp. 320–321 and 329–330.

35. H. A. Lorentz, *The Theory of Electrons and Its Applications to the Phenomena of Light and Radiant Heat* (reprint of second edition, 1915; New York: Dover, 1952); see preface and pp. 223–230; also (added in second edition) notes 72*–74, 75*–76 (pp. 321–327, 328–334). In particular, compare with the passage earlier cited from Poincaré about the chance agreement of formulas the following remarks of Lorentz's.
- on the last page of the main text of 1909: that Einstein "may certainly take credit for making us see in the negative result of experiments like those of Michelson, Rayleigh and Brace, not a fortuitous compensation of opposing effects, but the manifestation of a general and fundamental principle" (p. 230)
- in a note added in 1915: "If I had to write the last chapter now, I should certainly have given a more prominent place to Einstein's theory of relativity … by which the theory of electromagnetic phenomena in moving systems gains a simplicity that I had not been able to attain. The chief cause of my failure was my clinging to the idea that the variable *t* only can be considered as the true time and that my local time *t'* must be regarded as no more than an auxiliary mathematical quantity." (p. 321)

36. See a series of articles reprinted in Lorentz, *Collected Papers*, vol. V, pp. 229–245, 246–313, 330–355, 363–382.

37. Planck, "Die Kaufmannsche Messungen der Ablenkbarkeit der β-Strahlen in ihrer Bedeutung für die Dynamik der Elektronen," *Physikalische Zeitschrift* 7 (1906): 753–761.

38. Ibid., p. 761.

39. Note the following passage in Sommerfeld's book *Atomic Structure and Spectral Lines*, tr. (from third edition, 1922) H. L. Brose (New York: Dutton, 1923), p. 531: "In this chapter we have seen how the theory of relativity, just as it has remodelled all our physical thought and ideas, has also been able to help forward spectroscopy in a decisive manner. Conversely, we note that, in return, spectroscopy is in a position to lend support to one of the main pillars of the theory of relativity and to decide in its favor the question of the variability of mass of the electron."

40. L. Laudan, "A Confutation of Convergent Realism," *Philosophy of Science* 48 (1981): 19–49; see in particular p. 24 and p. 21, n. 1.

41. Lorentz, "Die Maxwellsche Theorie und die Elektronentheorie," in *Physik*, ed. E. Warburg (Berlin: B. G. Teubner, 1915), pp. 311–333; the quotation is from the closing sentences of the article.

Abner Shimony

The Methodology of Synthesis: Parts and Wholes in Low-Energy Physics

1   Aspects of the Problem

One of the most pervasive features of the natural world is the existence of reasonably stable systems composed of well-defined parts which are to a large extent unchanged by entering into composition or leaving it. The problem of parts and wholes is to understand with the greatest possible generality the relation between the components and the composite system.

The parts-wholes problem has an ontological aspect, which concerns the properties of the components and the composite system without explicit consideration of how knowledge of them is obtained. Among the ontological questions are the following: Is there an ultimate set of entities which cannot be subdivided and which are therefore "atomic" in the etymological sense? If the properties of the components are fully specified, together with the laws governing their interactions, are the properties of the composite system then fully determined? In particular, are those properties of composite systems which are radically different from those of the components, and which might properly be characterized as "emergent," also definable in terms of the latter? Do composite systems belong, always or for the most part, to "natural kinds"? Is the existence of natural kinds explicable in terms of the laws governing the components? Are both the possible taxonomy and the actual taxonomy of natural kinds thus explicable? Is there a hierarchy of "levels of description"—i.e., microscopic, macroscopic, and possibly intermediate—such that laws can be formulated concerning a coarser level without explicit reference to the properties at a finer level of description?

The parts-wholes problem also has an epistemological aspect. Suppose that the most precise and best-confirmed laws turn out to govern relatively simple systems—as is indeed mostly the case in physics—but that the systems of interest are enormously complicated combinations of simple components. Then there will be insuperable experimental difficulties in gathering knowledge about all the initial conditions of the parts, and insuperable mathematical difficulties in deducing from the basic laws the properties of the composite system.

To what extent can the composite system be said to be understood in terms of the laws governing its parts? And if there is independent phenomenological knowledge of laws on a coarse level of description, how do we know that these are in principle derivable from the laws on a finer level?

Aside from the ontological-epistemological distinction, there is a subdivision of the parts-wholes problem according to domains of investigation. There is the domain of inorganic systems (the physical sciences), that of organic systems (biology), that of systems endowed with minds (psychology), and that of groupings of human beings (social sciences). There is no *a priori* reason to believe that the questions listed above should be answered in the same way in all these domains, and it may not even be appropriate to pose the same questions in all of them. Of course, even the separateness of these domains is an aspect of the parts-wholes question. If the properties of an organism can in principle be defined in terms of those of the constituent molecules, and biological laws can in principle be derived from those of physics, then the domain boundary between the inorganic and the organic loses its fundamental significance and is maintained only for reasons of professional specialization. Likewise for the other domain boundaries.

The parts-wholes problem is so vast and ramified that no apology is needed for restricting attention to a small part of it. I shall consider almost nothing outside of physics, and within physics almost exclusively those branches in which the energies involved are typically of the order of an electron volt per particle or less: atomic physics, molecular physics, solid-state physics, and in general the physics of bulk matter. Low-energy physics is roughly Democritean. Electrons and stable nuclei have properties which can be studied in isolation, and to a very good approximation they keep these properties unchanged when they enter atoms, molecules, fluids, and solids. Hence electrons and nuclei behave as "building blocks," even though high-energy investigations reveal a rich internal structure of the latter. Democritus's characterization of his atoms as infinitely hard, indivisible, and immortal certainly does not apply without drastic reservations to these building-blocks. Nevertheless, they preserve their integrity in low-energy interactions to such an extent that Democritus's vision of explaining the full variety of the natural world in terms of the combinations of a very few kinds of unchanging bodies is realized to a remarkable extent. By contrast, high-energy physics—according to which particles are created and annihilated and the nucleon components (quarks) are removed from confinement either with great difficulty or not at all—is radically non-Democritean.

This essay consists largely of summaries of the treatment of parts and wholes in five areas of low-energy physics, concerning atoms, molecules, fluids, infinite Coulomb systems, and spin systems. Roughly Democritean answers are given in each area to most of the ontological questions concerning parts and wholes (if the divisibility of nuclei is set aside). I wish to call the attention of philosophers of science to the great variety and subtlety, and the often surprising nature, of the derivations of properties of composite systems from those of the components. Ontologically, the theory of the wholes is reduced to the theory of the parts, but the pejorative overtones of "reductionism"—a suggestion of flattening and loss of distinctive features—is certainly inappropriate. Epistemologically, the reduction of the theory of the wholes to that the parts has to be understood with serious qualifications. In the derivations that will be mentioned, the first principles of low-energy physics are usually supplemented with secondary principles, which can reasonably be regarded as consistent with the former but rarely rigorously inferrable from them. The methodology of the physical understanding of complex systems turns out to have intricacies that are seldom discussed and are difficult to formulate clearly.

This essay does not pretend to be authoritative, for the simple reason that I am not an expert in any of the branches of physics from which my examples are drawn. I nevertheless hope to convey my enthusiasm for inquiry in the territory that is reconnoitred.

2 First Principles

The first principles of low-energy physics fall into three groups: those of non-relativistic quantum mechanics, those concerning the dominant forces among the building blocks, and the principles of statistical mechanics.

A. Quantum mechanics is a framework theory that at present is commonly believed to apply to every physical system from elementary particles to the whole of the cosmos. My skepticism about the universal validity of quantum mechanics is largely irrelevant to the present discussion, since there is little doubt that quantum mechanics holds with great accuracy within the domain of low-energy physics. As a framework theory, quantum mechanics has to be supplemented by detailed information about the constitution of any particular system to which it is applied before inferences can be made about the behavior of the system and before experimental results can be predicted. Nevertheless, the framework which quantum mechanics provides is rich.

(i) It prescribes the general characteristics of the space of states of any system, most commonly summed up by saying that there is a one-to-one correspondence between the pure states and the rays of a Hilbert space. Implicit in this characterization of the space of states is the superposition principle.

It prescribes the general structure of the class of observables (as they are commonly called, though a less anthropocentric name such as "dynamical variables" would be preferable).

(iii) It prescribes the rule for calculating the probability distribution of possible values of an observable, contingent upon its actualization (by measurement or possibly by other means), when the state is known.

(iv) Nonrelativistic quantum mechanics offers a general law for the temporal evolution of the pure state of an isolated system, provided that a sufficient characterization of the system is given (namely, its Hamiltonian operator).

(v) Quantum mechanics itself has a remarkable principle of composition. If A and B are two systems, associated respectively with the Hilbert spaces H_A and H_B, then the Hilbert space associated with the composite system $A + B$ (provided that their dynamical interaction does not change them internally) is the tensor-product space $H_A \otimes H_B$. Only in special cases is it possible to write a vector ϕ belonging to $H_A \otimes H_B$ in the simple-product form $u \otimes v$, where u belongs to H_A and v to H_B, and only in these special cases is it correct to say that the pure state represented by ϕ is equivalent to the attribution of pure states to each of the components A and B. In general, A and B have an "entangled" state (to use Schrödinger's expression), so that there is no complete characterization of A without reference to B and vice versa. It is evident that quantum-mechanical entanglement requires a partial retrenchment from the Democritean conception of the parts-wholes relationship, but it leaves open the possibility of realizing other aspects of this conception. The quantum-mechanical ground state of a many-particle system is determined by its Hamiltonian operator, which consists of kinetic-energy contributions for each particle and potential-energy contributions for each pair of particles. Such a Hamiltonian is in the spirit of Democritus—*mutatis mutandis*—since it contains no contributions from the whole that cannot be traced explicitly to the parts. If the quantum-mechanical ground state then turns out to account for the physical properties of the system at a temperature of absolute zero (assumed for the present in order to avoid complications due to thermal excitations), then it is reasonable to say that a Democritean description of the many-particle system has been realized, even though the ground state is entangled.

(vi) Finally, there are symmetrization principles; i.e., if there are n identical bosons (integer-spin particles) in the system, then any vector in the Hilbert space representing a physically allowable state must be invariant under the interchange of two of these bosons, and if there are n identical fermions (particles with half-integer spin), then any vector representing a physically allowable state must change sign under exchange of two of these fermions. The Pauli principle, which prohibits two fermions to occupy the same single-particle state, is a direct consequence of this anti-symmetrization. Both the symmetrization rule and the anti-symmetrization rule are modifications of the tensor-product principle, since each restricts the space of states of a composite system to an appropriate subspace of the full tensor-product space. .

B. One of the great simplifications of the restriction to low-energy physics is that the dominant forces among the constituents are electromagnetic. Furthermore, these forces are treated essentially classically, by potentials, rather than by postulating a quantized electromagnetic field which interacts with matter

fields. Nevertheless, much complexity remains, for electromagnetic forces are exhibited in electron-nucleus interaction, electron-electron repulsion, nucleus-nucleus repulsion, orbit-orbit interaction, spin-spin interaction, electron spin–electron orbit coupling, electron orbit–nuclear spin coupling, and interaction with external electromagnetic fields. The potential terms in the total Hamiltonian can be of any of these types. In addition, of course, there are kinetic-energy terms. When a specified set of nuclei and a specified number of electrons are given, usually together with certain constraints upon the configuration, then the electromagnetic-force laws and the general principles of quantum mechanics yield a definite Hamiltonian. A large set of physical questions (What are the stationary states of the system? What is the ground state? What is the energy of the ground state? At what energy above the ground state does dissociation occur? What are the degeneracies of the various allowable energies? What are the geometrical properties of the system in the ground state? and so on) then become answerable in principle.

C. Statistical mechanics is a science of systematically extracting a relatively small amount of reliable physical knowledge of statistical matters from an ocean of ignorance. It treats mainly of two classes of systems. Systems of the first class consist of very many parts, typically falling into a small number of types (e.g., a homogeneous or a heterogeneous gas, or an alloy). Because of the very large number of parts, it is practically unfeasible to gather exact information about the initial conditions of the system. And even if, *per impossible*, a complete knowledge of the initial conditions were given, it would be humanly impossible to solve the equations of motion (classical or quantum mechanical) in order to infer exactly the state at a later time. The other class of systems considered by statistical mechanics consists of open systems, interacting with an environment the exact constitution of which is not known. (If the constitution were known in detail, then the system plus the environment might constitute a single system of many parts that would fall in the first class just mentioned.) The reliable information that one wishes to extract from the immense background of ignorance consists of probability distributions of especially interesting quantities concerning the system. When the system consists of a large number of parts, then typically the probability distributions are very sharply peaked, so that interesting quantities concerning the system as a whole, such as energy, state of condensation, and magnetization, can be predicted with virtual certainty. Statistical mechanics thus provides a powerful instrument for making inferences about wholes from the properties of parts, though it is an instrument that usually requires supplementation by secondary principles. The enterprise of statistical mechanics is greatly expedited if the environment with which the system exchanges energy is negligibly affected by the exchange, so that it can be considered a reservoir. When the system is in equilibrium with the reservoir—which is a concept that can be defined either phenomenologically or probabilistically—then the probability distribution over

the space of states can reasonably be shown to be the canonical distribution,[1] which can be written as follows for classical systems:

$$\rho(x) = \frac{1}{Z(\beta)} e^{-\beta E(x)},$$

where x represents the state in an appropriate space of states, $E(x)$ is the energy of the system when it is in the state represented by x, β is the inverse of the product of Boltzmann's constant k and the absolute temperature T, and $Z(\beta)$ is a normalizing factor defined in such a way that the integral of $\rho(x)$ over the space of states is 1. Explicitly,

$$Z(\beta) = \int e^{-\beta E(x)} \, dx.$$

If $Z(\beta)$ is considered as a function of β, it is called the partition function; some indication of its importance and utility will be conveyed below. There is also a quantum-mechanical version of the canonical distribution, which will not be needed explicitly in this essay even though it is implicit in much of the discussion of low-energy physics.

3   Atoms

The structure of atoms is one of the best-understood kinds of composition of wholes from parts in all of physics. Some important features of this composition can be understood quantitatively from first principles, but other features are described only by semi-empirical rules, for which there is some explanation but by no means a rigorous general derivation.[2]

Well established is the possibility of describing the ground state of any atom by a configuration, which tells how many electrons are in single-electron states of given principal quantum number n and given orbital angular momentum quantum number ℓ. The ascribability of definite configurations to each species of atom is the essence of the shell model of the atom. For example, for aluminum the configuration is $1s^2 2s^2 2p^6 3s^2 3p$, where the letters s, p, and d conventionally stand for the values 0, 1, and 2, respectively for the quantum number ℓ; the number preceding a letter prescribes the value of n; and the superscript after the letter gives the number of electrons with the specified values of n and ℓ. The wave function ψ_{al} (a vector in the Hilbert space of square-integrable functions), representing the ground state of aluminum, is a superposition of a number of vectors, each of which is a product of single-particle wave functions, of which two have $n = 1$, $\ell = 0$; two have $n = 2$, $\ell = 0$; six have $n = 2$, $\ell = 1$; two have $n = 3$, $\ell = 0$; and one has $n = 3$, $\ell = 1$. In each term in the superposition, spin states for each of the 13 electrons are given. Furthermore, the superposition is so contrived that ψ_{al} is antisymmetric under exchange of any two electrons, as required by principle A(vi) above. This is already an

enormous amount of information, and it is derived almost rigorously from first principles, together with the specification of the charges of the electrons and nuclei and the spin $\frac{1}{2}$ of the electron. The value $\frac{1}{2}$ for the spin not only implies that the electron is a fermion but also fixes the dimensionality of the spin space associated with each electron to be 2. Consequently, when a fixed one-particle spatial wave function is given—characterized by n and ℓ together with one other quantum number, the magnetic quantum number m—there is a further option of making the spin either up or down along a specified axis, thereby choosing between two orthogonal directions in the two-dimensional spin space. Hence, a pair of electrons can be characterized by n, ℓ, and m without violating the exclusion principle. Another ingredient in the derivation of the configuration is the approximate spherical symmetry of the effective potential which each electron feels, due to the small size of the nucleus in comparison with the atomic radius and due also to the effective "smearing" of the charge distributions of all the other electrons. This treatment of the electronic charge distribution is an instance of a mean-field approximation, which will recur below. Because the effective potential is nearly spherically symmetric, the angular momentum of the electron is a conserved quantity and hence ℓ is a good quantum number. Further exploitation of spherical symmetry yields rigorous information about angular momentum: for example, that for fixed n a given value ℓ of the angular momentum quantum number is compatible with exactly $2\ell + 1$ values of the component of the angular momentum along any specified axis, and hence $2\ell + 1$ values of the magnetic quantum number m (i.e., $m = -\ell, \ldots, \ell$). That fact was implicit in the configuration stated above for aluminum—e.g., there are six electrons in the one-electron state 2p, because the p is an abbreviation for $\ell = 1$, which permits three possible values of m (1, 0, -1), for each of which there are two possible spin orientations. That is why the closed shell 2p has six electrons. To be sure, the multiplicity of 3p is only 1 in aluminum, but that is because the 3p shell is incomplete; in all the atoms from argon onward there are six 3p electrons.

Although the configuration provides much information about the ground state, there is more to be determined, and almost everything else is much harder to extract from first principles. One wishes to know, for example, the radial dependence of the single-particle wave functions, which in turn determines such important properties as the average size of the atom. To find the radial dependence, one must know the effective potential felt by a single electron and use it to solve the differential equation in the radial variable r which is obtained from the time-independent Schrödinger equation. But the potential depends upon the wave functions of all the electrons. How does one emerge from a maze in which the potential depends upon the wave functions and the wave functions depend upon the potentials—even if one sets aside all the mathematical difficulties of solving the differential equations? A way out is provided by the self-consistent field method of Hartree and Fock. A sequence

of successive approximations is applied in practice, and it is reasonable to expect rapid convergence to a situation in which wave functions and potentials fit each other self-consistently.[3]

Except in the simplest atoms (through beryllium), the configuration does not determine how the orbital angular momenta ℓ_i and the spins s_i of the individual electrons are combined to yield a total orbital angular momentum quantum number L for the whole set of electrons, a total spin S, and also a total angular momentum J. In principle the possible combinations could be checked to see which yields the lowest-lying energy, but that requires detailed knowledge of the wave functions, which—as just noted—is a formidable obstacle. There are, however, Hund's rules based on spectroscopic evidence, which are quite reliable. The first two of Hund's three rules are the following:

1. The LS with the largest S compatible with the configuration has the lowest energy.
2. In the case where the largest S is associated with several different allowable values of L, the largest L has the lowest energy.

There is no generally valid proof for these rules, though they have been confirmed for a number of atoms by detailed calculations. But much plausibility can be given to them, especially to the first, by general arguments from first principles. The larger S is, the more parallel are the spins. When the spins are parallel, they make a symmetrical contribution to the complete wave function, and hence the antisymmetrization principle for electrons implies that the spatial part of the wave function is antisymmetric. But that has the general effect of keeping the electrons, on the average, farther from one another than would be the case with symmetrical spatial wave functions, and therefore the potential energy of electrostatic repulsion among the electrons is diminished, thus lowering the energy.[4]

The mathematical complications of determining the ground-state wave function of an atomic species should not make us lose sight of a fundamental Aristotelian fact: that the species is a natural kind. All atoms of a given species (i.e., a given number of neutrons and protons in the nucleus, and as many electrons as protons for electric neutrality) have the same ground state when there are no external perturbations, and they have the same array of excited states. The *identity* of all atoms of a given species ensures that they all emit and absorb electromagnetic radiation at the same frequencies and have the same size, shape, and internal motion. Weisskopf (1979, p. 87) emphasizes two other features of atoms: stability ("The atoms keep their specific properties in spite of heavy collisions and other perturbations to which they are subjected") and regeneration ("If an atom is distorted and its electron orbits are forced to change by high pressure or by close neighboring atoms, it regains its exact original shape and orbits when the cause of distortion is removed"). Atoms thus exhibit form in Aristotle's sense, and even have the tendency to maintain

this form, which phenomenologically is like his final cause. But the Aristotelian form is achieved by Democritean means—by interactions among the electrons and the nucleus, which leave these building blocks intact. Of course, one additional element is required for the achievement of form: the principles of quantum mechanics. Classical physics did not have the resources to explain natural kinds of composite systems, even when natural kinds of the indivisible building blocks were postulated, whereas Aristotelian science accounted for natural kinds uneconomically, by postulating an irreducible principle for each. In contrast to both, quantum mechanics ensures that the formal properties of very many kinds are implied by the properties of a small number of kinds of components.

4   Molecules

On the whole, molecular structure is more difficult to understand than atomic structure, primarily because molecules lack the spherical symmetry of the single-nucleus atom. Nevertheless, there is one principle of molecular composition that was extracted from laboratory experience and is simple enough to be taught in elementary chemistry: the principle of valence.[5] The valence of an element is defined phenomenologically as the number of atoms of hydrogen that one atom of the element can combine with or take the place of in forming compounds. From H_2O, HCl, NH_3, and CH_4 we infer that O, Cl, N, and C, respectively, have valences 2, 1, 3, and 4. Very soon, however, one finds that valence is not an intrinsic characteristic of an element, since in CO carbon behaves as if it has valence 2, and iron would have to be assigned valence 2 because of FeO and valence 3 because of Fe_2O_3. Clearly, valence is a crude and complicated principle of composition—and there are wonderful monsters like xenon-fluoride to increase the complication.

The hope of physical chemists is to put the valence concept on a firm basis, with all appropriate qualifications, by using the principles of composition of low-energy physics, especially quantum mechanics and the electromagnetic forces. Even before the new quantum mechanics, G. N. Lewis proposed to interpret chemical bonding in terms of the sharing of a pair of electrons by each of two atoms in a molecule, and the valence of an atom was interpreted as the number of electrons available for sharing. W. Heitler and F. London later gave a quantum-mechanical treatment of this idea in which the valence electrons are those that are not paired in the atom's configuration, where pairing means having the same spatial wave function (or "orbital") but opposite spins. In the bonding of two atoms in a molecule, a valence electron from one of the atoms is free to combine with a valence electron of the other in such a way that their combined spin state is the singlet state

$$|\uparrow\rangle_1|\downarrow\rangle_2 - |\downarrow\rangle_1|\uparrow\rangle_2.$$

This is a vector in the two-electron spin space which obviously changes sign under the exchange of electrons 1 and 2. Conseqently, the overall antisymmetrization of electrons 1 and 2 is achieved by a symmetric spatial wave function for the two electrons. The effect of the symmetry of the spatial wave function can best be seen if one considers a coordinate r along an axis passing through the nuclei of the two atoms, utilizing the fact that the nuclei are more massive than the electrons and hence more precisely localizable without violating the uncertainty principle. An antisymmetric spatial wave function will have an amplitude close to zero halfway between the nuclei, and a symmetric spatial wave function will usually have a large amplitude and even be peaked near the midpoint. Therefore the symmetric spatial wave function makes the two electrons, on the average, closer to each other than does a comparable antisymmetric one, and hence the contribution of the repulsive electrostatic potential of the two electrons (a positive energy) is larger in the symmetric case than in the antisymmetric; but this positive contribution is more than compensated by the fact that, on the average, the electrons spend more time close to the nuclei in the symmetric case, thereby increasing the attractive contribution. Hence the singlet spin state is favored for the achievement of the lowest possible energy. This is an excellent argument, except for some hand-waving concerning the relative magnitude of the competing repulsive and attractive contributions. In the case of the H_2 molecule, Heitler and London were able to perform a quite accurate quantitative calculation from first principles to confirm the hand-waving argument. The extrapolation to more complex molecules is always somewhat risky, however.

Without any consideration of more detailed calculations it is evident that some modification of the foregoing account of valence is needed. The ground-state configuration of carbon is $1s^2 2s^2 2p^2$, which provides two unpaired electrons in the 2p shell (unpaired because there are three mutually orthogonal spatial wave functions available in this shell—$m = 1$, $m = 0$, and $m = 1$). Usually, however, carbon exhibits a valence of 4, though it was noted that in CO it exhibits valence 2. Pauling (1931) offered a resolution of this and similar difficulties by arguing that in molecular bonding one must consider not only the atomic ground-state configuration but also the lower-lying excited configurations. The reason is that the energy expended to raise some of the electrons to excited levels is more than compensated for by the availability of additional unpaired electrons. This suggestion is a fruitful new secondary principle of composition in molecular physics, known as "the hybridization of molecular orbitals." It is not merely an *ad hoc* device to save the Heitler-London theory of spin pairing, since good calculations of molecular geometry can be made by means of hybridization theory. Hybridization has to be used with tact, and sometimes must be refined by additional secondary principles. For example, Pauling and Keaveny (1973, p. 93) answer as follows an argument that hybridization theory breaks down for iron-group elements: "It has been recognised that the orbital occupied by a bonding electron may be either expanded or contracted considerably with respect to the orbital in the isolated atom. This expansion or contraction is accompanied by a contraction or expansion of other orbitals occupied by unshared electrons. Expansion or contraction of one orbital decreases or increases its shielding of nuclear charge for another orbital." Machine calculations and empirical data seem to support this refinement of hybridization theory.

Machine calculations of wave functions are not imperturbable, objective judgments independent of all appeal to secondary principles, because the choice of initial guesses of wave functions in an iteration procedure is crucial for the rate at which converges occurs. If a principle such as hybridization of molecular orbitals is a good approximation to the truth, then the initial wave function that it suggests will lead to rapid convergence. If not, the convergence will be too slow to be useful.

5 Coulomb Systems in the Thermodynamic Limit

A vast range of phenomena can be described quite well by the formalism of thermodynamics. Essential to this formalism is the possibility of characterizing a system at a macroscopic level in terms of a small number of extensive variables: V, U, and $N_1, \ldots, N_k$, where V is the volume, U is the total energy, and N_i is the mole number of the ith chemical constituent. By *extensive* is meant that when the system is spatially subdivided into subsystems (not too small) the value of the extensive variable of the system as a whole is the sum of the values of that variable in all the subsystems. Furthermore, it is assumed that there is an entropy S, which is a linear homogeneous function of the extensive variables and which therefore is extensive too. The maximization of S relative to constraints determines the thermodynamic state of the system.[6]

Extensiveness as a constructive principle is so often taken for granted in ordinary applications (for example, the energy necessary to heat a kilogram of milk one degree is 1,000 times the energy needed to heat a gram of it one degree, largely independent of the shape of the container) that it takes some reflection to see that it is thoroughly nontrivial. Extensiveness depends, for example, on the negligibility of surface forces, which of course is not always the case. In order to avoid the complications of the surface, the standard procedure is to idealize by going to the thermodynamic limit, in which there is infinite volume and infinitely many particles, but to take the limit in such a way that

$$\lim(N/V) = \text{density}$$

exists. It remains to be established, however, that U and S are extensive variables in the thermodynamic limit, so that $U = Vu$, where u is the internal energy per unit volume, and $S = Vs$, where s is the entropy per unit volume.

There is no *a priori* assurance that the constructive principle of thermodynamics is consistent with the constructive principles of atomic and molecular physics, in which the properties of the composite are determined by the properties of component particles and their interactions. One might properly worry, for example, that if the fundamental forces among particles are long-range, then putting together more and more particles will not determine an internal energy proportional to volume, even if the density is fixed.

These questions turn out to be remarkably difficult. Using classical mechanics, one can prove the existence of the thermodynamic limit, provided that the intermolecular forces are repulsive at short distances and attractive at long distances but that the attractiveness falls off very rapidly as the distance increases.[7] Coulomb forces, which are by far the most important of electromagnetic forces at low velocities, do not have these characteristics, and the physically important question is whether the thermodynamic limit holds for a neutral Coulomb system (equal densities of positive and negative charges). The answer is no if classical mechanics rather than quantum mechanics is taken as the framework theory. But, most remarkable, the answer is yes if quantum mechanics is used. The proof is extraordinarily difficult, but hinges upon two simple facts: that the exclusion principle keeps fermions from coming too close to one another, thereby providing an effective repulsive core if the negatively charged particles are electrons (or, more generally, if either the positively charged particles or the negatively charged particles are fermions of one kind); and that electrostatic shielding in a neutral Coulomb system counteracts the long-range character of the Coulomb force.[8] The very existence and stability of bulk matter with ordinary macroscopic properties depends crucially upon an intimate quantum-mechanical principle that underlies the first of these two facts. In 1519–1522 Magellan proved that the Pacific Ocean existed by sailing across it, but 450 years had to elapse before Dyson, Lenard, Thirring, Lebowitz, and Lieb proved by quantum mechanics that something like the Pacific Ocean *could* exist. Thus science progresses.

6   Normal Fluids

In a very large, spatially extended system, it is trivial to define macroscopic physical variables by summing or averaging over a very large number of microscopic variables. The local matter density $\sigma(\mathbf{x})$ of a fluid can be defined classically by summing the masses of the particles in a small region $R_\mathbf{x}$ centered about x, which is large enough to contain on the average many particles but is small in comparison with the macroscopic dimensions of the system, and then dividing by the volume of $R_\mathbf{x}$. An analogous definition can be given quantum-mechanically, but it makes use of expectation values of positions of particles. Likewise, a local mean velocity $\mathbf{v}(\mathbf{x})$ of the fluid can be defined quantum-mechanically as well as classically. And similarly for other macroscopic quantities such as pressure and temperature, which also may be local quantities, varying with $\mathbf{x}$. Enough has been said to make it clear that *local* can be *macroscopic*. What is nontrivial, and clearly not simply a matter of definition, is whether there are any laws of nature that can be formulated solely in terms of a set of macroscopic quantities. If there are such laws, then the microscopic level is not the only level of description at which laws are exhibited. Actually, we are quite confident that there are laws at a macroscopic level of description, because of the great empirical success of such disciplines as fluid dynamics. The difficult problem is to explain why this is so.

An outstanding example is the Navier-Stokes equation, which is the dynamical law for a viscous fluid in nonturbulent motion:

$$\sigma(\mathbf{x})[\dot{\mathbf{v}}(\mathbf{x}) + (\mathbf{v}(\mathbf{x})\,\mathrm{grad})\mathbf{v}(\mathbf{x}) = -\mathrm{grad}\,p(\mathbf{x}) - \eta_1\,\mathrm{curl}\,\mathrm{curl}\,\mathbf{v}(\mathbf{x}) + \eta_2\,\mathrm{grad}\,\mathrm{div}\,\mathbf{v}(\mathbf{x}).$$

The matter density $\sigma(\mathbf{x})$ is the mass per unit volume in the neighborhood of $\mathbf{x}$; $\mathbf{v}(\mathbf{x})$ is the local velocity; $p(\mathbf{x})$ is the local pressure; and η_1 and η_2 are friction constants. There have been many derivations of this equation, but I find that of Fröhlich (1973) impressive for its generality and lucidity. The derivation begins with the exact time-dependent Schrödinger equation for a system consisting of N particles of specified mass, but restates this equation in terms of the von Neumann density operator Ω instead of the wave function, as is appropriate where there is imperfect knowledge of the true quantum state. In other words, Ω is the quantum-mechanical analogue of the distribution function of classical statistical mechanics. Ω contains information about the correlated behavior of all the N particles, but for practical purposes most of these correlations are of no interest. It therefore is useful to define reduced density operators, Ω_1, Ω_2, Ω_3, etc., concerning (respectively) single particles, pairs of particles, triples of particles, etc. From the original dynamical equation for Ω, a hierarchy of equations for Ω_1, Ω_2, Ω_3, etc. follows easily. The complication of this hierarchy lies in the fact that the equation in Ω_1 contains a term dependent on Ω_2, the equation for Ω_2 contains a term dependent on Ω_3, and so on. The macroscopic quantities in the Navier-Stokes equation depend only upon Ω_1 and Ω_2, and with appropriate additional assumptions the equation itself can be derived. First, it is assumed that in equilibrium there is rotational and translational invariance, so that the pair-correlation function $P(\mathbf{x}, \mathbf{y})$, which is essentially the probability of jointly finding particles located at $\mathbf{x}$ and $\mathbf{y}$, becomes at equilibrium a function $P_e(|\mathbf{x} - \mathbf{y}|)$ which depends only on the distance between the two points. Second, it is assumed that the interaction between particles is short-range, so that large changes of σ, $\mathbf{v}$, and p do not occur over distances in which the interaction is non-negligible. And third, attention is restricted to situations in which the deviation of $P(\mathbf{x}, \mathbf{y})$ from the equilibrium pair-correlation function $P_e(|\mathbf{x} - \mathbf{y}|)$ is a function of $\mathbf{x}$ only via the distance $|\mathbf{x} - \mathbf{y}|$ and via linear dependence upon the macroscopic fields $\sigma(\mathbf{x})$ and $\mathbf{v}(\mathbf{x})$. Since these assumptions are quite mild, Fröhlich's derivation has

great generality. In principle, values of η_1 and η_2 can actually be calculated from Ω_2, though to do so the hierarchy of equations for Ω_i must be terminated at $i = 2$ by an approximation, and the resulting dynamical equation for Ω_2 must be solved. For practical purposes it usually is sufficient to know that the form of the Navier-Stokes equation is legitimate and to depend upon measurement to supply the friction constants.

7   Phase Transitions in Spin Systems

Most of the discussion of composition so far has drawn only upon the first two classes of principles summarized in section 2, those of quantum mechanics and of electromagnetism. These are the principles that one would expect to dominate at low temperature, when thermal fluctuations are negligible. At elevated temperatures, however, the thermal fluctuations become more and more important as a disordering influence. The derivative principles of composition that one expects to find at finite temperatures are likely to be the result of competition between the ordering tendencies implicit in the rules of quantum mechanics and the disordering tendencies due to thermal excitations. Splendid instances of such competition are provided by phase transitions, such as from a solid to a fluid phase, from a superfluid to a normal fluid, or from a ferromagnetic to a nonferromagnetic arrangement of spins in a crystal.

Below 1,043°K (the Curie temperature), a single iron crystal exhibits net magnetization in the absence of an external magnetic field; above this temperature it does not do so. This phenomenon is interpreted as the preferential alignment of the unpaired electronic spins (each of which behaves as a little magnetic dipole) of the atoms below the Curie temperature, whereas no preferred direction survives above it. Qualitatively, this interpretation is in accordance with the foregoing remark that elevation of temperature has a disordering effect. It is nevertheless remarkable that there is a sharp transition from the ferromagnetic phase (in which magnetization occurs in the absence of an external field) to the nonferromagnetic phase, and that discontinuities occur in the specific heat and in other quantities.

A more detailed explanation requires a statistical-mechanical calculation.[9] First a reasonable expression is written for the energy of a specified configuration of spins:

$$E[\mu_1, \ldots, \mu_N] = \sum_{1 \leqslant i < j \leqslant N} \phi(|\mathbf{r}_i - \mathbf{r}_j|)\mu_i\mu_j - H\sum_{i=1}^{N} \mu_i,$$

where each μ_i can be either $+1$ or -1, according as the ith spin is up or down with respect to a given direction (ordinarily that of the external magnetic field, if it is nonvanishing), H is the intensity of the field, and ϕ is a function that depends upon the distance between the locations of the ith and jth spins. In

principle the calculation of the partition function

$$Z(\beta, H) = \sum_{\text{all configurations}} e^{-\beta E[\mu_1, \ldots, \mu_N]}$$

can be carried out when the function ϕ is given, and from Z the magnetization, the specific heat, and other macroscopic quantities of interest can be calculated by differentiations and other mathematical operations. For almost all choices of ϕ the calculation of Z is extremely difficult. A great simplification is the mean-field assumption, according to which each spin effectively interacts with the mean field produced by all the other spins together, with the result that

$$\sum_{1 \leqslant i < j \leqslant N} \phi(|\mathbf{r}_i - \mathbf{r}_j|)\mu_i\mu_j \cong \frac{\text{const}}{N} \sum_{1 \leqslant i < j \leqslant N} \mu_i\mu_j.$$

This calculation can be performed quite easily, and it exhibits a phase transition: the spontaneous magnetization per spin, as a function of the absolute temperature T, has the following form:

$$m_0 = [3(1 - T/T_c)]^{1/2} \quad \text{for } T \text{ close to but less than } T_c,$$

$$m_0 = 0 \qquad\qquad \text{for } T > T_c.$$

This derivation of a phase transition is an incomplete triumph. In the first place, the behavior of the magnetization, the specific heat, and other quantities near T_c does not agree with experiment. Furthermore, the mean-field assumption is entirely implausible physically, since a crystal can be enormously large in comparison with a single atom, and it makes no sense that the far-off spins should contribute as much to the total field felt by a given spin as the ones closer by.

A more plausible model is the Ising model, in which a given spin is assumed to interact only with its nearest neighbors:

$$E[\mu_1, \ldots, \mu_N] = -J \sum_{\substack{i=1 \\ j = \text{nearest neighbors}}}^{N} \mu_i\mu_j - H\sum_{i=1}^{N} \mu_i.$$

The one-dimensional version of this model was analyzed by E. Ising in 1925 and shown not to exhibit a phase transition. (Ising erroneously argued that the same was true in two dimensions.) In 1944 L. Onsager performed a rigorous calculation for the two-dimensional Ising model and found that

$$m_0 = [1 - (\sinh 2v)^{-4}]^{1/8} \quad \text{for } T < T_c,$$

$$m_0 = 0 \qquad\qquad \text{for } T > T_c,$$

where $v = J/kT$ and $\sinh 2v_c = 1$. An intuitive explanation can be given for the failure of a phase transition in one dimension and its occurrence in two. In one dimension, the occurrence of a single spin flip due to a thermal perturbation interrupts the long-range order; in two dimensions, each spin has four neigh-

bors, and if one is deviant because of a thermal perturbation, the other three are likely to maintain the long-range order, except in the improbable event of two or more becoming deviant. (The political implications of this argument are obvious.) A beautiful variant of this reasoning explains how the periodic structure of a crystal is maintained over lengths many orders of magnitude greater than the lattice spacing, so that deviations from exact periodicity do not become cumulative with distance (Peierls 1979, pp. 85–91).

The spectacular work of the last two decades by K. Wilson, B. Widom, M. Fisher, and L. Kadanoff on phase transitions and critical phenomena[10] has exhibited the power of several further principles of composition. Their main results concern the values of critical exponents and relations among them, these exponents being numbers that characterize the singular behavior of thermodynamic quantities, such as susceptibilities and specific heats, in the vicinity of singular points in thermodynamic space (e.g., critical points and lines of phase transition). These critical exponents turn out to depend crucially upon a small number of parameters, notably the dimension d of the system and the "spin dimensionality" D or its analogue.[11] What is striking is that the values of the critical exponents are independent of many microscopic features that one might intuitively think to be important, such as the geometry of the lattice array and the strengths of the interactions. The explanation for the insensitivity of the critical exponents to these factors is that near the singular points fluctuations are very large. (In some cases the fluctuations can even in a sense be visible; near the critical point of a fluid there is "critical opalescence," in which giant fluctuations cause unusually large scattering of light.) But if the extent of the average fluctuation is much larger than the average spacing between atoms, then the details of the interactions become irrelevant or are swamped out. This is the essence of the hypothesis of universality for each class characterized by a few parameters. It is at first surprising, but then upon reflection quite reasonable, that general statements can be made about whole classes of macroscopic systems in virtue of giant fluctuations, which are disordering factors. The neglect of fluctuations is the chief respect in which the mean-field treatment of phase transitions fails. The rigorous exploitation of the hypothesis of universality requires the machinery of the renormalization group (a group of transformations in each of which a change of scale plays a crucial role); in a few cases, quantitative calculations of critical exponents have actually been achieved by this machinery.

Although the universality results just mentioned concern only phase transitions and critical phenomena, they are probably indications of a principle of much greater generality. This is the principle enunciated by P. Anderson (1984, p. 85) and called by him "the principle of continuation." Anderson writes: "In a very deep and complete sense we can refer back whole classes of problems involving interacting fermions to far simpler and fewer problems involving only noninteracting particles. While there can be great or small quantitative differ-

ences between a real metal, for instance, and a gas of noninteracting electrons, the essentials of their qualitative behavior—specific heat, T dependence of susceptibility, T dependence of various transport coefficients, etc.—are the same." Later, Anderson states the conjecture that the renormalization group "affords at least one way of putting mathematical teeth into the basic concept of continuation . . . , and perhaps it is even *the* most fundamental way of doing so" (p. 167).

8   Ontological Comments

On the basis of the foregoing summaries a number of propositions can be asserted with some confidence concerning the ontological aspects of the parts-wholes relationship in physics. All the ontological questions listed in section 1 will be addressed, though the first only briefly. Furthermore, a few additional matters not anticipated in the list of questions will be touched upon.

A. It is an open question in elementary-particle physics, which has been bypassed in this essay, whether there is an ultimate level of noncomposite particles. Electrons and other leptons may have no internal structure, since they exhibit pointlike behavior in scattering experiments;[12] the same may be true of quarks, which are the components of protons and other heavy particles. Whatever may be the answer to this question, the Democritean equating of noncompositeness with immortality is surely abandoned. Electron-positron pairs can be created and annihilated, even though each member of a pair is created or annihilated as a whole rather than by fabrication and dismantling.

B. The Democritean picture is transformed by the fact that, from the point of view of quantum field theory (the best framework of fundamental physics that we have), a particle is a quantum or excitation of a field, and a field in some sense is a holistic entity, given over all of space-time. Although this statement seems inconsistent with a Democritean answer to the parts-wholes problem, this *prima facie* judgment must be carefully qualified by locality considerations. The propagation of a field is presumably governed by relativity theory, which forbids direct causal connection between two points with spacelike separation. As a result, there is complete freedom to specify a classical field over a spacelike surface, constrained only by smoothness and good behavior at infinity. Except for these restrictions, a classical relativistic field on the whole spacelike surface is determined by the field values on an exhaustive set of disjoint subregions of the spacelike surface, and hence a residuum of the Democritean point of view concerning parts and wholes is found in classical field theory. Analogous (but suitably modified) statements can be made about a quantized relativistic field.

C. From the standpoint of fundamental physics, the low-energy domain to which we restricted attention in sections 3–7 is derivative. The elementarity of electrons may be an open question, but there is no doubt that all nuclei other

than that of ^{1}H are composite, even if one does not probe the quark structure of the protons and neutrons. Nevertheless, there are several crucial physical facts which ensure that the level of low-energy physics has a certain autonomy, allowing it to be investigated with little attention to deeper and finer levels, and has a quite firmly Democritean character. One fact is the scale of energies in nature. Typically, the energy required to excite a nucleus is of the order of 100,000 electron volts, whereas the excitation energy of an atom is of the order of one electron volt and that of a molecule even less; excitations in solids can be much less. Consequently, in energy transactions typical of low-energy physics, the nuclei are not disrupted or even excited (though there are unstable or nearly unstable nuclei for which this statement must be amended). A second fact is that the lightest of the particles with nonvanishing rest mass, the electron, has (according to the relativistic equation $E = mc^2$) an energy equivalent of about 500,000 electron volts, and therefore the creation of an electron-positron pair has an energy threshold far beyond the domain of low-energy physics. A third fact is that even though there is no such hindrance to the production of low-energy photons, since photons have zero rest mass, their existence can largely be neglected in the low-energy domain, where the semiclassical radiation theory makes few predictions that differ from those of quantum electrodynamics.[13]

D. The properties of atoms, molecules, normal fluids, infinite Coulomb systems, and spin systems are derivable in principle from the properties of their components and the physical laws governing the interactions of the components. It is not strictly correct, however, to say that the laws underlying these derivations are none other than the principles of low-energy physics summarized in section 2. One reason for caution is that corrections from relativistic quantum mechanics and gravitational theory are required in order to obtain in full detail the properties of the systems surveyed; for instance, relativistic effects are already found in the hydrogen atom, and they become increasingly important with increasing atomic number. Furthermore, the rigorous derivations of phase transitions in spin systems did not assume electromagnetic interactions among the elementary spins, but rather mathematically concocted interactions that hold only between pairs of spins of a restricted class (e.g., nearest neighbors or next-nearest neighbors). The results of Onsager and his successors are not to be depreciated for this reason. What that work shows is that long-range order can be accounted for in terms of interactions with strictly short range, which is a more severe restriction than the falling off with distance of electromagnetic forces. Consequently, the motivation for postulating some kind of hitherto unobserved long-range interaction for the purpose of accounting for long-range order is removed. (Similarly, the success of the Bardeen-Cooper-Schrieffer theory of superconductivity showed that electromagnetic forces in the framework of quantum mechanics could account for that kind of long-range order, without the postulation of a new fundamental long-range force.)

E. Many of the properties of composite systems which are implied by the laws governing the components are emergent, in the sense that they are qualitatively radically different from the properties of the components. The long-range order of a ferromagnet is an example of an emergent property, as are the rigidity of a crystal and the viscosity of a fluid. The term *emergent* is a sometimes used to mean "underivable in principle from the properties of the components," but this is not the meaning that I adopt, nor would it be semantically economical for anyone maintaining a roughly Democritean point of view to do so.

F. The ontological reduction of the properties of the composite system to those of the components, as asserted in paragraph D, does not constitute a renunciation of the entanglement of states of composite systems, which was one of the peculiarities of quantum mechanics emphasized in section 2. Quite the contrary, in many instances the ontological reduction is possible only because of entanglement—e.g., in the role of hybridization in molecular bonding.

G. Composite systems fall into natural kinds, as the discussion of atoms pointed out explicitly and that of molecules tacitly. The same can be said for nuclei and the elementary particles composed of quarks, though these systems lie outside the domain of this essay. The stable configurations permitted by quantum mechanics and the force laws determine in principle the complete taxonomy of possible kinds of stable systems composed from the given components. In the case of relatively simple composite systems—the composite elementary particles, the nuclei, the atoms, and the smaller molecules—the actual taxonomy is also determined, provided that the environmental conditions are suitably specified. (In the case of the heavier nuclei, the requisite environmental conditions are satisfied only in sufficiently hot stars.) The reason is that one can reasonably assume enough trial encounters to have occurred among the components that the space of possible configurations has been well explored in the time interval under consideration. In the case of complex systems, such as macromolecules, there is certainly not enough time within the relevant intervals (e.g., the existence of the earth) to explore the space of possible configurations, and hence the actual taxonomy is in large part the result of the contingencies of evolutionary history. It should be added that within the framework that has been sketched, imperfections in natural kinds (such as flaws in crystals and errors in DNA) are as comprehensible as the near perfections. The Democritean treatment of the central Aristotelian doctrine of natural kinds must be reckoned as one of the great triumphs of modern physics.

H. One of the remarkable features of the physical world is the existence in many situations of several well-defined levels of description for the same

physical system. Three levels were considered in the analysis of normal fluids and spin systems, but can be found in many other types of systems. The deepest level—that of relativistic field theory—was only briefly mentioned above, and its existence is acknowledged but not normally exploited by working condensed-matter physicists. The second level is the nonrelativistic quantum-mechanical many-body treatment of interacting electrons and nuclei. The third and coarsest level is that of macroscopic description, in which the fundamental variables for spin systems are temperature, applied magnetic-field strength, and magnetization, and those for fluids are local density, local velocity, and local pressure (all these local quantities being defined by averages over regions large compared to an atomic volume, and all of them possibly varying with position). The remarkable feature of a level of description is not that there are quantities at that level which are defined by summing or averaging over the quantities at a deeper level of description, but that there are laws which govern the specified level without supplementary information from a deeper level. The Navier-Stokes equation for fluid flow and the equation of state $M = M(H, T)$ for a spin system are formulated entirely in terms of appropriate macroscopic variables. To be sure, there are parameters entering these equations which are not supplied by the macroscopic theory and which must either be measured empirically or derived by resorting to a deeper level, but there are no physical variables in these equations except those of the macroscopic level.[14]

I. As pointed out in the summary of Fröhlich's derivation of the Navier-Stokes equation, the derivation of a macroscopic equation from the laws governing a deeper level is not absolutely general; it holds only in a range of circumstances which fortunately is very wide. Sometimes this failue of generality is cited as an objection against the thesis of reducibility of macroscopic physics to physics at a deeper level of description. The most common version of this objection is that classical thermodynamics is a nonstatistical theory, whereas statistical fluctuations require changes in the thermodynamics which Boltzmann and Gibbs purport to derive from statistical mechanics. My answer to this objection is that excessive rigidity concerning the concept of reducibility can mask the truly wonderful relations between levels of description. From an Olympian point of view, unclouded by the difficulties of mathematical inference, the laws at the deeper level imply, for each set of system constitutions, initial conditions, and boundary conditions, a definite set of values or possibly a definite set of probability distributions of the macroscopic quantities; the extent to which the putative macroscopic laws are satisfied is then made explicit. If the agreement is good (by a reasonable standard), in a set of circumstances that is wide (by a reasonable standard), then a striking relation holds which captures the intuitive and practical sense of "reduction." What constitutes "reasonable standards" obviously cannot be specified at the level of generality of the present discussion, but the treatment of fluctuations in statistical mechanics shows that sensible things can often be said about this question.

9   Epistemological Comments

The epistemological problem that has most preoccupied philosophers of science is the legitimation of scientific principles on the basis of empirical evidence. Some variant of the hypothetico-deductive method is recognized widely to be appropriate for this purpose, though, of course, there are many different opinions about the details of the method and the cognitive claims that can be justifiably made for it. The deductive steps in the hypothetico-deductive method are usually given less attention than the other steps; deduction is supposed "in principle" to be well understood, even though in practice it encounters enormous technical difficulties. In other words, there is a tendency to take an Olympian point of view for granted in discussing the deductive steps of the hypothetico-deductive method.

In the enterprise of reducing macrophysical theories to microphysical theories and, in general, explaining the properties of composite systems in terms of those of the components, the principles governing the components are taken as established, and therefore the problem of deduction emerges from the background and assumes a central role. The special epistemological problems concerning the relation between parts and wholes in physics arise just because humans are debarred by their intellectual limitations from taking an Olympian point of view. If there is not a satisfactory surrogate for Olympian deduction, then human beings are not entitled to assert that macrophysics is reduced to microphysics and that wholes are understood in terms of parts, no matter how matters appear to the deities. The five illustrations discussed in sections 3–7 contain a number of procedures, none of which by itself provides a satisfactory surrogate for Olympian deduction but which work together remarkably well to this end.

One procedure is systematic approximation to the exact solutions of difficult mathematical equations—for example, the Hartree-Fock iteration method mentioned in section 3. Approximations might be thought to be purely mathematical, carried out without reference to the physical problems that posed the difficult equations. In practice, however, the sublimation of the mathematics from the physics is incomplete, and physical considerations direct and supplement purely mathematical ones in several ways. First of all, the starting point of an approximation method is very important, and observational data together with intuitive understanding can suggest a good beginning. Conversely, a qualitatively wrong starting point—e.g., one with a different symmetry from that of the phenomenon of interest—will ensure that the approximation method will never converge to the true solution; a barrier of singularity separates a state with the wrong symmetry from those with the correct symmetry

(Anderson 1984, pp. 27–30). Even if the starting point is not radically wrong in this sense, a clumsy choice will entail convergence so slow that the outstanding features of the true solution will not be recovered by a computation of reasonable duration. Another way in which physical considerations guide and direct mathematics is by indicating that an approximation method has been carried far enough in a situation where no rigorous proof exists of convergence or where no rigorous bounds can be put upon the error of truncation. Example: In the ground state of a system with finitely deep potential wells, the energy is a minimum; if the energy computed with an approximate solution is very little above the experimentally measured energy, then it is reasonable to conclude that the approximation is good and has captured the major features of the true solution.

A second procedure is the performance of simplifying steps in the course of physical analysis of a problem. This procedure is quite different from the construction of a model, in which the simplifications are made initially in characterizing the constituents of the system and the forces among them. The procedure to which I am now referring is exemplified at several points in section 6—e.g., it was assumed that changes of σ, $\mathbf{v}$, and ρ are small over distances of the order of the interaction range between atoms, and that the dependence of the pair-correlation function $P(\mathbf{x} - \mathbf{y})$ on powers of $\mathbf{v}$ higher than the first is negligible. These simplifications are made quite naturally in the course of the analysis, but would be hard to insert into a rigorously treated model. Usually these simplifications are intuitively appealing, and when they seem particularly significant they are often dignified with the name of *Ansatz*.[15] Simplifications are made plausible not only by physical intuition but also by the fact that there is an empirically well-established *terminus ad quem*, a phenomenological equation that is known to work well for normal liquids that are not driven to the point of turbulence. This empirical evidence provides inductive support for the supposition that the simplifications are valid broadly and not merely in highly idealized special cases.

In the two procedures considered so far, informal inductions are made in order to remedy the shortcoming of human deductive powers. Disagreement with empirical evidence would then constitute *prima facie* evidence that the truncation was premature or the simplification unjustified. Conversely, agreement with empirical evidence provides inductive support for the legitimacy of the truncation or the simplification, though the informality and the lack of rigor must be acknowledged; it is certainly possible that good agreement is achieved, not because of the smallness of the terms neglected relative to those retained, but because of fortuitous cancellations among the neglected terms. Confirmation is always riskier than disconfirmation. As in inductive inference generally, however, when good agreement is found in a variety of experimental situations, and when the physical picture becomes more and more coherent, then confidence in the confirmation justifiably increases.

A third procedure is the construction of rigorously soluble models possessing some of the crucial qualitative features of real systems—for example, the short-range character of the microscopic magnetic moments that compose an actual ferromagnet. As noted in sections 7 and 8, the great significance of these results is the rigorous establishment of emergence; that is, the exhibition of macroscopic properties radically different from those of the constituents. Conceptually, though not technically, it is a larger step to establish the possibility of emergence at all than to show that real composite systems behave qualitatively in the same way as the models.

A fourth procedure is to implement what Anderson calls "the principle of continuation," thereby providing the missing linkage between models and real systems. For this purpose the renormalization group is very powerful, but an exposition of that is beyond the scope of this essay and its author. Evidently, however, Anderson's suggestion that the renormalization group is not merely a technique for studying a special body of phenomena but an instrument of great generality in many-body physics deserves very careful study.

The systematic deployment of these procedures for the purpose of understanding the properties of composite systems in terms of their components deserves the epithet "the methodology of synthesis," which has previously been applied to statistical physics (Toda et al. 1983, p. v). The epistemological significance of the physics of composite systems is summed up in this phrase.

One final epistemological comment is appropriate. Throughout this essay, physical first principles have been assumed to be given. Of course, they are not given, but have to be established by ingenious experimentation and profound analysis of the experimental results. In view of the great mathematical difficulties of determining what experimental prediction a theory makes concerning a many-body system, it is obviously desirable to carry out the critical experiments for assessing first principles upon the simplest possible systems.[16] But a reasonably successful execution of a physics of composite systems within a given framework of first principles, carried out in accordance with the "methodology of synthesis" sketched above, should be regarded as very impressive supporting and reconfirming evidence for those principles. Contrariwise, the failure to execute successfully a physics of composite systems within a proposed framework in spite of much effort (for which we now have a good standard of strenuousness) should be regarded as serious disconfirming evidence. The investigation of the properties of composite systems would usually be classified as "normal science," but it should be obvious from the foregoing discussion that "normal" cannot be equated with "mundane" or "routine."

Acknowledgment

Notes

1. See, for example, Khinchin 1949.

2. See, for example, Bethe and Jackiw 1968.

3. A good treatment is found in Merzbacher 1970, pp. 535–539.

4. A good discussion is given in Baym 1969, p. 458.

5. See, for example, Baym 1969, pp. 486–498; Roby 1973.

6. See, for example, Tisza 1966, pp. 102–193.

7. See, for example, Thompson 1972, pp. 67–71.

8. A survey of the argument is given in Lieb 1976.

9. See, for example, Thompson 1972, chapters 4–6.

10. See, for example, Ma 1976.

11. As already pointed out, the qualitative behavior of the Ising model depends upon whether d is 1 (in which case there is no phase transition) or 2 or 3 (in which case there is a transition). D is 1 if the spin can be either up or down in a fixed direction; it is 2 if the spin can be a vector confined to a plane; it is 3 if it is a vector that can point in any direction in space. There are generalizations of this meaning of spin dimensionality when one is dealing with other types of transitions between order and disorder.

12. See, for example, Bransden et al. 1973, p. 210. For the view that the electron may be composite, see Greenberg 1985.

13. See, for example, Sargent et al. 1974, p. 97.

14. The Ω_i in section 6 could be considered to constitute a hierarchy of levels of description intermediate between the second and third levels listed here, except that the dynamical equation for Ω_i refers to Ω_{i+1}. With suitable assumptions, the hierarchy of equations can be truncated at the kth without serious error (this was in fact done for $k = 2$ in section 6), and then there is an autonomous intermediate level with its own dynamical law. Grad 1967 contains a fine discussion of levels of description.

15. A famous example is L. Boltzmann's *Stosszahlansatz* in the kinetic theory of gases, which says essentially that the velocity distribution of a molecule that has just undergone a collision is the same as that of a randomly selected molecule. As a matter of fact, this *Ansatz* is automatically satisfied by an appropriate model (Grad 1967, p. 56).

16. It may be objected that a many-body system supplied the data that inspired the beginning of quantum theory: the radiation in a black cavity. This objection is not decisive, however, because black-body radiation is in fact extremely simple, consisting of noninteracting excitations.

References

Anderson, P. W. 1984. *Basic Notions of Condensed Matter Physics*. Menlo Park, Calif.: Benjamin/Cummings.

Baym, G. 1969. *Lectures on Quantum Mechanics*. New York: Benjamin.

Bethe, H. A., and R. W. Jackiw. 1968. *Intermediate Quantum Mechanics*, second edition. Reading, Mass.: Benjamin/Cummings.

Bransden, B. H., D. Evans, and J. V. Major. 1973. *The Fundamental Particles*. London: Van Nostrand Reinhold.

Fröhlich, H. 1973. "The Connection between Macro- and Microphysics." *Rivista del Nuovo Cimento* 3: 490–534.

Grad, H. 1967. "Levels of Description in Statistical Mechanics and Thermodynamics." In *Delaware Seminar in the Foundations of Physics*, vol. 1, ed. M. Bunge. New York: Springer-Verlag.

Greenberg, O. W. 1985. "A New Level of Structure." *Physics Today* 38, no. 9: 22–30.

Khinchin, A. 1969. *Mathematical Foundations of Statistical Mechanics*. New York: Dover.

Lieb, E. 1976. "The Stability of Matter." *Reviews of Modern Physics* 48: 553–569.

Ma, Shang-keng. 1976. *Modern Theory of Critical Phenomena*. Reading, Mass.: Benjamin.

Merzbacher, E. 1970. *Quantum Mechanics*, second edition. New York: Wiley.

Pauling, L. 1931. "The Nature of the Chemical Bond. Application of Results Obtained from the Quantum Mechanics and from a Theory of Paramagnetic Susceptibility to the Structure of Molecules." *Journal of the American Chemical Society* 53: 1367–1400.

Pauling, L., and I. Keaveny. 1973. "Hybrid Bond Orbitals." In *Wave Mechanics*, ed. W. Price et al. New York: Wiley.

Peierls, R. 1979. *Surprises in Theoretical Physics*. Princeton University Press.

Roby, K. R. "Mathematical Foundations of a Quantum Theory of Valence Concept." In *Wave Mechanics*, ed. W. Price et al. New York: Wiley.

Sargent, M., M. O. Scully, and W. E. Lamb. 1974. *Laser Physics*. Reading, Mass.: Addison-Wesley.

Thompson, C. 1972. *Mathematical Statistical Mechanics*. New York: Macmillan.

Tisza, L. 1966. *Generalized Thermodynamics*. Cambridge, Mass.: MIT Press.

Toda, M., R. Kubo, and N. Saito. 1983. *Statistical Physics* I. Berlin: Springer-Verlag.

Weisskopf, V. F. 1979. *Knowledge and Wonder*, second edition. Cambridge, Mass.: MIT Press.