GREGG JAEGGER AND SAHOTRA SARKAR

COHERENCE, ENTANGLEMENT, AND REDUCTIONIST EXPLANATION IN QUANTUM PHYSICS

Abstract. The scope and nature of reductionist explanation in physics is analyzed, with special attention being paid to the situation in quantum mechanics. Five different senses of "reduction" are identified. The strongest of these, called "strong reduction," is the one that purports to capture the relations between macroscopic and microscopic physics. It is shown that the criteria for strong reduction are violated by explanations in quantum mechanics which involve "entangled states." The notion of "coherence" in physical systems is also defined. It is shown that, contrary to many current views, the invocation of coherence does not necessarily lead to the violation of strong reduction. However, entangled systems also exhibit coherence. Therefore, the subclass of coherent systems that are entangled presents problems for strong reduction.

INTRODUCTION

In recent years, the problems surrounding the role of reductionism in science have been extensively studied, mainly in the contexts of biology and psychology. Philosophers seem to have generally assumed that (i) instances of reduction in the physical sciences are ubiquitous; and (ii) these instances are straightforward in the sense that they are trivially captured by at least one of the alternative models of reduction that have been developed by philosophers of science. These assumptions seem to be based on intuitive analyses of putative instances of reduction such as the following:

(i) there is a straightforward sense in which Newtonian mechanics can be obtained from the special theory of relativity (for instance, by taking the $c \to \infty$ limit, where $c$ is the speed of light in vacuum). Thus, Newtonian mechanics is reduced to special relativity;
(ii) similarly, the so-called "Newtonian" limit of general relativity yields Newton’s theory of gravitation. Once again, the latter is thus reduced to the former;
(iii) geometrical optics is reducible to physical optics through Maxwell’s electromagnetic theory in the sense that the behavior of light waves, as predicted by Maxwell’s theory, shows why geometrical optics is correct to a good approximation;
(iv) classical thermodynamics can be derived from, or reduced to, statistical mechanics by the construction of kinetic models of gases. This putative reduction is particularly fascinating because mechanics, the laws of which are invariant under time reversal, is supposed to give rise to the second law of thermodynamics which embodies a direction of time;

A. Ashtekar et al. (eds.), Revisiting the Foundations of Relativistic Physics, 523–542.
(v) quantum mechanics reduces classical chemical bonding theory, and accounts for
the various valency rules of chemistry that have been used since at least the early 19th
century

It is quite possible that a half-century ago most, though not all, physicists and
chemists would have accepted these cases as successful reductions. Most of them
would also have endorsed a positive and unproblematic assessment of the status of
reductionism in the physical sciences. The developments in physics during the last
fifty years, however, have called these judgments into question. Among the most star-
tring have been scaling theory and renormalization “group” techniques in condensed
matter physics. A perusal of these developments has even led Leggett (1987) to sug-
gest that the appropriate relationship between macroscopic and microscopic physics
is only one of consistency: microscopic physics has no significant role in the explana-
tion of macroscopic phenomena.3 Leggett’s views are by no means idiosyncratic. For
instance, Fisher (1988), one of the founders of scaling theory, has argued that there
are aspects of condensed matter physics for which the underlying microphysics (viz.,
quantum mechanics) is explanatorily irrelevant.

Meanwhile, Fröhlich (e.g. 1968, 1969) initiated a research program that aims to
prove that in several instances biological entities must be considered as quantum
mechanical systems that exhibit “coherence” or phase correlations in their dynamics.
Fröhlich (1973) has claimed that biological explanations that invoke such phenomena
are not “mechanistic,” that is, reductionist in the usual sense of that term (see section
2). This claim has also been endorsed by Ho (1989).4 What is particularly philosoph-
ically interesting about Fröhlich’s conjecture is that, if it turns out to be correct, these
attempts at physical explanation in biology would fail to satisfy the strictures of
reduction not because of any peculiarity of biological systems but because of a failure
of reduction (e.g., as it will be construed in section 2) in physics (more specifically,
quantum mechanics) itself.

Finally, it has become apparent that at least some of the ostensibly straightforward
cases of reduction, such as those mentioned above, are not quite as unproblematic as
they customarily have been taken to be. Returning to the cases mentioned above:

(i) Though, from a strictly mathematical point of view, Newtonian mechanics does
descend from special relativity in the $c \to \infty$ limit, that limit is counterfactual: the
speed of light is finite;
(ii) For the Newtonian limit of general relativity to exist and yield Newton’s theory of
gravitation, constraints have to be imposed on general relativity (Ehlers 1981; Mal-
ament 1986). In particular, the spatial part of space-time must be flat;
(iii) A series of approximations is required to obtain geometrical optics from Max-
well’s laws.
(iv) The relation of thermodynamics to statistical mechanics has turned out to be even
more complicated. It should consist of the derivation of the thermodynamic laws
from statistical mechanics in the so-called “thermodynamic limit,” the one in which
the number of particles ($N$) and the volume ($V$) of the system both go to $\infty$ while
the density ($N/V$) remains constant. All thermodynamic parameters must approach
well-defined values in this limit. In the classical (non-quantum) realm, the existence
of the thermodynamic limit so far has been rigorously proved for only very contrived and simple systems (Thompson 1972). For classical systems with only electrostatic interactions, it is even easy to show that the thermodynamic limit does not exist. If quantum mechanics is invoked, however, the limit is defined (Ruelle 1969, 60-68). Perhaps there is even some unexpected insight to be gleaned from this situation: that classical thermodynamics is reducible to quantum statistical mechanics, but not to classical statistical mechanics;

(v) To obtain the classical rules of valency from quantum mechanics, approximations about the nature of the wave-functions, about the Hamiltonian for complex atoms, and assumptions about the convergence of solutions must all be brought into play (Pauling 1960).

These observations should indicate that the issues surrounding reduction are far from settled in contemporary physics. Our immediate purpose in this paper is to examine whether explanations invoking the concept of “coherence” (including the claims of Fröhlich) really do violate reductionism. We conclude that the use of “coherence” does not necessarily preclude an explanation from being reductionist. Nevertheless, there exists a class of quantum-mechanical systems exhibiting coherence, those having “entangled states,” for which explanations are no longer clearly reductionist. This class of systems has long been of philosophical interest because one of its subclasses, that of entangled two-particle systems each with an associated two-dimensional Hilbert space, includes the systems that violate Bell’s inequality. It appears likely that the results establishing the relationship between entanglement and Bell-inequality violations can be generalized to the result that all entangled systems are capable of violating Bell-type inequalities. Furthermore, it turns out that though the type of model that Fröhlich has invoked involves coherence, any failure of reduction will be due to entanglement rather than coherence. In the long run, we hope that our analysis will help reintroduce detailed discussions of the place of reductionist explanation in physics.

In section 2 we discuss what we mean by a “reductionist explanation.” In section 3 we give several examples of classical and quantum systems that exhibit “coherence.” We then attempt a general definition of that and related concepts, which is necessary because a sufficiently general definition has proved to be elusive in the past. We are not fully satisfied with our definition; we hope, however, that it will spur further discussion by others. In section 4 we show that all entangled (quantum) states exhibit coherence, and that the coherent states invoked by Fröhlich are entangled. We observe that the quantum states that violate Bell’s inequality are also entangled. In both sections 3 and 4 our more technical conclusions are presented as relatively precise theorems. We give proofs of these theorems in those cases in which they are not explicitly available in the extant literature. In section 5 we show that, while the use of coherence does not necessarily lead to an explanation failing to be reductionist, entanglement leads to such a failure. We note some of the implications of our analysis in the concluding section (section 6).
2. ON REDUCTION

Systematic analysis of the concept of reduction in the natural sciences began with the pioneering efforts of Nagel (1949, 1961) and Woodger (1952), who viewed reduction as a type of inter-theoretic explanation. This approach has since been significantly extended, especially by Schaffner (1967, 1994). All these approaches view explanation as deductive-nomological. Alternative accounts that view reduction as a relation between theories, though not necessarily one of explanation, have been developed by Kemeny and Oppenheim (1956), Suppes (1957) and, more recently, Balzer and Dawe (1986, 1986a). Meanwhile analyses of reduction that view it as a form of explanation but not necessarily as a relation between theories have also been developed (e.g., Kauffman 1971; Wimsatt 1976; Sarkar 1989, 1992, 1998). These analyses have been motivated primarily by the situation in molecular biology, where reductionist explanation seems to be rampant, but the explanations refer to a variety of mechanisms rather than to theories. Using a set of distinctions introduced by Mayr (1982), Sarkar (1992) classified these models of reduction into three categories: (i) theory reductionism which consists of those models that view reduction as a relation between theories; (ii) explanatory reductionism which consists of those models that view it as explanation, but not as a relation between theories; and (iii) constitutive reductionism which consists of those models such as the various types of supervenience that eschew both theories and explanation.

Accounts in all three categories make both epistemological and ontological claims. In particular, all models of reduction share the rather innocuous ontological claim that what happens at the level of the reduced entities (theories or not) is not novel in the sense of being inconsistent with what happens at the level of the reducing entities — otherwise these would not be models of reduction. Throughout our discussion we will continue to make this assumption. However, we will ignore other ontological issues that have been controversial: whether the terms invoked in a reduction refer to “natural kinds”; whether reductions establish relations between “types” at the two levels or between “types” and “tokens”, etc. These issues, though often regarded as philosophically important, are orthogonal to our present purpose. We will also ignore those formal epistemological issues that have persistently been the focus of dispute: (i) whether the factors involved in a reductionist explanation are codified into theories; (ii) whether the structure of the explanation is basically deductive-nomological (Schaffner 1994) or statistical (Wimsatt 1976), etc. Thus our present analysis will be consistent with any of the models of reduction that view it as a type of explanation. Meanwhile, we will focus on three substantive claims that have often been implicit in models of reduction but have rarely been discussed in detail. Our analysis will reveal some rather surprising subtleties about reductionist explanation.

We assume that what we have at hand is an explanation, i.e., it satisfies whatever strictures that one chooses to put on “explanation.” Our problem is solely to specify additional criteria by which we can decide whether that explanation is reductionist. The reasons for this move are to avoid disputes about the explication of “explanation” (on which there is no consensus), and to focus attention precisely on those factors that make an explanation reductionist.
We suggest that, at the substantive level, the three most important criteria are:

(i) **Fundamentalism**: the factors invoked in an explanation are warranted by what is known, either from theoretical considerations (characteristically involving approximations) or only from experiments entirely at the level of the reducing theories or mechanisms which are more “fundamental” than those at the level of the reduced entities in the sense that their presumed domain of applicability is greater.\(^9\) Satisfaction of this criterion is a matter of degree. If the demonstration of such a warrant involves only theoretical derivation, from first principles, with no approximation, and so on, then its satisfaction is most complete. Approximations, especially counterfactual or mathematically questionable approximations, hurt the degree to which this criterion is satisfied;

(ii) **Abstract hierarchy**: the complex entity whose behavior is being explained is represented as having a hierarchical structure (with identifiable levels and an ancestral relation between levels) in which only the properties of entities at lower levels (of the hierarchy) are used to explain the behavior of the complex entity.\(^10\) Such an abstract hierarchical representation can involve any space, not necessarily physical space, that is used to model a system. For instance, it can be a hierarchy in any configuration or phase space in (classical) analytical mechanics or a Hilbert or Fock space in quantum mechanics;

(iii) **Spatial hierarchy**: the hierarchical structure of the entity (that is invoked in the explanation) must be realized in physical space, that is, entities at lower levels of the hierarchy must be spatial parts of entities at higher levels of organization.

These criteria are not all independent of one another: (iii) can only be satisfied provided that (ii) is. Moreover, if (i) is not satisfied at all, it is doubtful (at least) that an explanation should be considered a reduction. It will be assumed here that, for all reductionist explanations, (i) is at least approximately satisfied. What will distinguish the different types of reduction are the questions whether (i) is fully satisfied and which of the other two criteria, if either, is also satisfied. With this in mind, five different senses of “reduction” based on these criteria can be distinguished. For each of these senses, several illuminative biological examples exist—these are discussed in Sarkar (1996). Here, we only mention putative examples from physics:

(a) criterion (i) is (fully) satisfied while none of the others are: examples include the reduction of Newtonian mechanics to special relativity, and of Newtonian gravity to general relativity. More controversially, the reduction of geometrical optics to physical optics is a reduction of this sort.\(^11\)

(b) criterion (ii) is fully satisfied while (i) is approximately satisfied: this is clearly a very weak sense of reduction. Little more than a hierarchical structure is assumed. In the study of critical phenomena, explanations involving renormalization in parameter space are of this type: the system (say, a ferromagnet) is given a hierarchical representation in parameter space (though not in physical space) but the interactions posited at each stage as the number of units of which the system is composed (in our example, magnetic spins) is iteratively decreased have at best only approximate warrants from the underlying mechanisms;\(^12\)
c) criteria (i) and (ii) are fully satisfied but (iii) is not: in these reductions, the organizational hierarchy does not correspond to a hierarchy in the usual physical space. If quarks really are confined, in the sense that free quarks do not exist, explaining the properties of hadrons on the basis of properties of quarks would be an explanation of this type. There is a hierarchical structure: hadrons consist of quarks but this is clearly not a hierarchy in physical space. Examples of this sort abound in particle physics;
(d) criteria (ii) and (iii) are (fully) satisfied while (i) is approximately satisfied. Once again, in the study of critical phenomena, real-space renormalization, that is, renormalization with a representation in physical space involves a reduction of this type;
(e) all three criteria are satisfied: this is obviously the strongest sense of reduction. It is the one that is invoked in the putative reduction of thermodynamics to statistical mechanics. It is also the one that will be most relevant to the discussion in section 5. We will refer to this sense as "strong reduction."

Criteria (ii) and (iii) together capture the intuition behind those types of reduction which refer to a spatial whole being made up of identifiable constituent parts. When (i) is also satisfied, the explanatory force in a reduction, which comes from the properties of the parts, is supposed to provide a deeper explanation entirely from the lower level. In general, this is the one which is usually assumed to capture the relation between macroscopic and microscopic physics. In the particular case of explanations involving the notion of coherence, this is the intuition that we will explore in this paper. Of course, if what we have already said about explanations involving renormalization theory in condensed matter physics is correct, there are other reasons for doubting that strong reduction describes the relation between microscopic and macroscopic physics—we leave a full discussion of renormalization for another occasion.

3. COHERENCE

We turn, now, to the concept(s) of coherence. Any attempt at an explication of this concept faces a peculiar quandary: though it is quite routinely used in a variety of areas within physics, it is almost never defined outside optics, where it has become associated with various measures of correlation between field variables at two space-time points (cf. Mandel and Wolf 1995, chs. 4-8, 11). Usage in physics and biophysics is so varied that it is open to question whether there are non-trivial criteria that all uses share. This is so despite 75 years of very gradual development and several attempts to broaden the definition beyond quantum optics, where there is a fairly consistent pattern of use (Mandel and Wolf 1970, 1995). Because of this, we start with some remarks in the quantum optics literature regarding this definition. We then attempt to generalize it to be generally applicable across physics. But we are less sure that the result is non-trivial. So we proceed to avoid triviality by relativizing our definition to an independently characterized class of physical systems. At the very least, we hope that our attempt will provoke not just criticism but other attempts at providing a general definition of coherence.
Section 3.1 discusses examples of coherence across physics. Section 3.2 presents some illuminating comments by others regarding usage of the term, as well as our explication of one trivial and one non-trivial concept of coherence; we demonstrate the triviality of the former by an explicit (and itself somewhat trivial) theorem.

3.1 Examples

Skepticism about the possibility of reductionist explanation in physics has often been based on the belief that the components of composite systems exhibit collective behavior that cannot be accounted for by a straightforward examination of the properties of these parts taken in isolation. In physics, descriptions of such behavior often invoke the notion of “coherence.” As we noted above, this notion is almost never given a definition when it is used outside optics. To motivate our suggested definitions to be given at in section 3.2, consider the following four examples. The first is from classical physics. The other three involve quantum concepts:

(i) Coupled harmonic oscillators: Consider a pair of identical oscillators, having the same spring constant $k$ and mass $m$, coupled by a third spring of spring constant $k'$. Such a pair of oscillators are capable of moving in a “high-frequency normal mode” of vibration in which the motion of each oscillator is of constant amplitude and the same frequency $\omega$ higher than their common “natural frequency,” $\omega_0 = (k/m)^{1/2}$. The equations of motion for the two oscillators are:

\[
x_1 = A \cos \omega t , \quad x_2 = A \cos \omega t ,
\]

where $x_1$ and $x_2$ are the displacements of each of the oscillators from its equilibrium position, $t$ is the time, and $A$ and $B$ are the (constant) amplitudes of oscillation. There is a correlation in the position of the pair of oscillators because both oscillators vibrate at the same frequency. Note that, given $x_2$ and knowing $\omega = \sqrt{\omega_0^2 + k'/m}$, we can infer $x_1$ from $x_2$ and vice versa;

(ii) Rabi oscillators: Consider an electron capable of being in one of two coupled states $|\phi_1\rangle$ and $|\phi_2\rangle$ with energies $E_1$ and $E_2$ (respectively), of an atom, ion, or molecule, so that its quantum state, which can be written

\[
|\psi(t)\rangle = a_1(t)|\phi_1\rangle + a_2(t)|\phi_2\rangle ,
\]

where $a_i(t)$ $(i = 1, 2)$ are state probability amplitudes and $t$ is the time, obeys the standard time-dependent Schrödinger equation. It is convenient to decompose $|\psi(t)\rangle$ using the eigenvectors, $|\psi_+\rangle$ and $|\psi_-\rangle$, of the Hamiltonian, $H = H_0 + W_{12}$, where $W_{12}$ is the operator representing the “energy” of the “force” coupling the two states. Let $|\psi_+\rangle$ and $|\psi_-\rangle$ have the energies $E_+, E_-$ (respectively). Taking $|\psi(0)\rangle = |\phi_1\rangle$, we have $|\psi(t)\rangle = \lambda e^{-iE_+ t/\hbar}|\psi_+\rangle + \mu e^{-iE_- t/\hbar}|\psi_-\rangle$, where $\lambda$ and $\mu$ are constants. The probability of the electron being in either of the two states (or in the other) varies periodically with a frequency that depends on the strength of the
coupling $|W_{12}|$ between them: assuming $W_{12}$ to be purely non-diagonal, the probabilities $P_2(t)$ and $P_1(t)$ of respectively being in states $|\psi_2\rangle$ and $|\psi_1\rangle$ are

$$P_2(t) = \sin^2\theta\sin^2\left(\frac{E_+ - E_-}{2h}t\right)$$  \hspace{1cm} (3.3a)

$$P_1(t) = 1 - P_2(t) = -\sin^2\theta\sin^2\left(\frac{E_+ - E_-}{2h}t\right)$$  \hspace{1cm} (3.3b)

where $\theta = \left(\frac{2|W_{12}|}{E_1 - E_2}\right)$ is the oscillation period and $h$ is $\hbar/(2\pi)$ (with $\hbar$ being Planck's constant).

There is a strict correlation between the probabilities of the electron occupying the two states, thus of the atom, ion or molecule having the feature of one energy level or its alternative energy level occupied and the probabilities can be inferred from one another using eq. (3.3b). Note that this discussion can be extended beyond two-state systems to cases with several states and different associated energy levels. The correlations are then less trivial;

(iii) Fröhlich systems: (Fröhlich 1968, 1975) claims that quantum states of a macroscopic system described by "macro wave functions" are needed to explain a wide range of biological phenomena. These wave functions are supposed to exist when there is "off-diagonal long-range order" (ODLRO) in a system. Yang's (1962) definition of ODLRO, which is a property of a "reduced density matrix," will now be used to show how coherence is present in such a system. 14

Let a large system $\Sigma$ containing a subsystem $\sigma$ be described by the density matrix $\rho$. The "reduced density matrix" representing subsystem $\sigma$ is first obtained by taking a weighted average of the parameters specifying the portion of $\Sigma$ not including $\sigma$. This averaging is achieved by "tracing out" the parameters from the density matrix $\rho$ representing $\Sigma$, yielding a reduced density matrix $\rho_{\sigma}$ for $\sigma$. ODLRO is a property of $\rho_{\sigma}$, illustrated here by the following two examples.

The simplest and most interesting example in our context is the one-particle reduced density matrix $\rho_1$ describing single particles. This matrix has the elements

$$\langle i | \rho_1 | i \rangle = \text{Tr}(a_j^\dagger \rho a_i),$$ \hspace{1cm} (3.4)
\[ \langle k | p_2 | i j \rangle = Tr (a_k a_d p a_d^\dagger a_i^\dagger), \]  

where the new labels \( k \) and \( l \) refer to particle states just as \( i \) and \( j \) do. (Note that elements of this reduced density matrix are labelled by four indices.) Other reduced density matrices for yet larger subsystems of particles, \( \rho_n \), for \( n \)-particle states can be similarly defined. The pertinent subsystems of many-particle systems, which we will call “Fröhlich states (F-states)” (following Pokorny 1982), are those whose single-particle reduced density matrices exhibit ODLRO. These consist of a large number, \( N \), of bosons (either simple bosons or bosonic quasiparticles formed from fermions) that are said to exhibit ODLRO when they can be represented by single-particle reduced density matrices of the form

\[ \rho_1 (r', r^\prime) = \alpha N \Phi (r') \Phi^* (r^\prime) + \chi (r', r^\prime), \]

where \( \Phi (r) \) is the quantum “macro wavefunction” attributed to the subsystem, \( \Phi^* (r) \) is its complex conjugate, \( \chi (r', r^\prime) \) is a positive operator, \( 0 \leq \alpha \leq 1 \) and \( r' \) and \( r^\prime \) represent two (spatial) positions of the subsystem.

The reduced density matrix for the subsystem has the spectral resolution

\[ \rho_1 (r', r^\prime) = \sum_{i=1}^{\infty} \mu_i \xi_i (r') \xi_i^* (r^\prime), \]

where the \( \xi_i (r) \) are energy eigenstates and \( \mu_i \) are weights. Most of the \( N \) particles in our system lie in the same state, \( \xi_n (r) = \Phi (r) \), for which \( \mu_n = \alpha N \) (for some value \( n \) of \( i \) and \( \alpha \) is a constant in \([0,1] \) near 1) and the weights \( \mu_i \ll 1 \) for all \( i \neq n \). Thus

\[ \rho_1 (r', r^\prime) = \alpha N \Phi (r') \Phi^* (r^\prime) + \sum_{i=1}^{\infty} \mu_i \xi_i (r') \xi_i^* (r^\prime), \]

where \( \sum_{i=1}^{\infty} \mu_i = (1-\alpha)N \). With \( \sum_{i=1}^{\infty} \mu_i \xi_i (r') \xi_i^* (r^\prime) = \chi (r', r^\prime) \), which describes that portion of \( \rho_1 \) describing individual particles of our subsystem not in the state \( \Phi \), we have

\[ \rho_1 (r', r^\prime) = \alpha N \Phi (r') \Phi^* (r^\prime) + \chi (r', r^\prime), \]

where \( \chi (r', r^\prime) \) is small compared to \( \alpha N \) except when \( r' = r^\prime \). As a result of the presence of the first term, \( \rho_1 (r', r^\prime) \neq 0 \) even as \( |r' - r^\prime| \to \infty \). Thus, for F-states, an identity of quantum state—and, therefore, perfect correlation—will persist over
large spatial distances. This is the sense in which ODLRO is a form of long-range order.

\[ |\Psi\rangle = \frac{1}{\sqrt{2}} [|+\rangle_e |-\rangle_p - |\rangle_e |+\rangle_p]. \]  

(3.10)

where, in each term, the subscripts refer to the Hilbert space of the corresponding particle, \(|+\rangle\) is the single-particle state “spin up,” and \(|\rangle\) is the single-particle state “spin down.”

The system is correlated in two ways: (a) there is a correlation of both quantum phases, the angles \(\theta\) of the complex exponential part of each complex probability amplitude, for the combined system (as explained in theorem 3.1 below); and (b) the spins are strictly anticorrelated, that is the spin states of the electron and positron are such that if one has spin “up” then the other has spin “down”; knowing the spin state of either particle allows one to infer the spin state of the other. The example will be relevant to section 4.

Such a system is traditionally referred to as being in a “coherent” superposition of states. Here this involves two quantum-mechanical states: one in which the electron has the relevant component of its spin in the “up” state and the positron has the same spin component “down” (i.e. it is described by \(|+\rangle_e |-\rangle_p\), and one in which the electron has spin “down” and the positron has spin “up” (described by \(|-\rangle_e |+\rangle_p\).

3.2 Tentative Explication
'coherent' is also widely used in physics to indicate correlation between two or more functions of either space or time, (such as its use in the description of two light beams obtained by the splitting of a single beam), although the functions themselves may have some random properties.

Klauder and Sudarshan (1968, 56) echo these sentiments in their own attempt to characterize "coherence" broadly:

... we must come to terms with the vague concept of 'coherence.' Being purposely general, let us say that a 'coherent feature' of a statistical ensemble is an observable aspect held in common by each member of the ensemble. Different ensembles will in general have different coherent features depending on what collection of quantities is deemed 'observable.' This suggests that we call a 'relative coherent feature' one which fulfills the criteria for a coherent feature for a subset of the observables. That is, the criteria for a relative coherent feature are necessary but not sufficient for a coherent feature. ... 'Full coherence' can be said to exist if the members of the ensemble are identical in all their observable aspects.

This definition, which is explicitly intended as a general one (though motivated by Klauder and Sudarshan's interest in optics), is unnecessarily restricted to situations where a statistical ensemble exists. One should be able to speak of "coherence" in systems that are not normally regarded as ensembles. This point has been long recognized (e.g. by Hopkins 1952, 263; and Senitzky 1962, 2865). It is also suggested by our examples, all of which are systems of this sort. Furthermore even in the context of ensembles, it seems unreasonable to require that all properties of a system be correlated for a system to be called "fully coherent." Similarly, in one of the most highly regarded definitions requires that a fully coherent field be "defined as one whose correlation functions satisfy an infinite succession of stated conditions" (Glauber 1963), which also seems unreasonably strict.

However, Klauder and Sudarshan are correct to indicate that even a context-dependent definition of coherence can invoke nothing more than the existence of a correlation. Ho and Popp (1993) also come to the same conclusion. 15 We might, therefore, attempt modify and extend such a putative definition in the following way: a system is coherent if and only if, for at least two of its features, A and B, there exists a correlation between their values. 16 Then, either the value of B can be estimated from that of A or the value of A from that of B.

The trouble with this definition of coherence is its weakness: almost all systems exhibit coherence. 17 For the purposes of discussion we will instead call this property "trivial coherence." In classical physics almost all systems exhibit such coherence because of the fully general laws such as Newton's third law, which routinely provide the framework in which all models are constructed; each such model will have coherence automatically built into it. In the case of Newton's third law, \( F_{ij} = -F_{ji} \) (where \( F_{ij} \) is the force of one object, i, on another, j). each time two objects interact, the forces involved will correlate their motions in accordance with Newton's second law, \( F_{ij} = m \dot{a}_{ij} \), where \( \dot{a}_{ij} \) is the contribution to the acceleration of each object due to the force \( F_{ij} \) of one object, i, on the other j. In quantum mechanics, for all pure states a similar situation arises because all systems are represented by a ray in an associated Hilbert space. The following trivial theorem can be proved:
Theorem 3.1: All isolated nonstatistical (pure) quantum states exhibit trivial coherence.

Proof. Consider the case when the Hilbert space of the quantum-mechanical system in question is countable. Every isolated such a system can be written as a pure state of the following form: $|\psi\rangle = \sum_{j=1}^{n} \lambda_j |\psi_j\rangle$, where $\lambda_j = r e^{i \xi_j}$, $\xi_j = -\frac{1}{\hbar} E_j t$ ($E_j$ being the energy of the system in state $|\psi_j\rangle$, $t$ the time, $r_j$ real constants) and the eigenvectors $|\psi_j\rangle$ form an orthonormal basis for the Hilbert space of the system. A sustained correlation will exist between the properties corresponding to the eigenvectors $|\psi_j\rangle$ since the quantity $\Delta_{kl} = \frac{1}{t} (\xi_k - \xi_l) = -\frac{1}{\hbar} (E_k - E_l)$ is constant for any $k$ and $l$, for all times $t$; since, for isolated systems, the energies $E_j$ are constant in time, a correlation exists, represented by the energy difference $\Delta_{kl}$ between the values of $\xi_k$ and $\xi_l$ at any time $t$. This correlation is observable in the form of quantum interference.

(The proof in the case of uncountably infinite Hilbert spaces is identical but with the sums replaced by integrals and discrete indices replaced continuous ones.)

Every physical system will usually exhibit some such trivial coherence which makes it, therefore, not a very satisfying notion. We will, therefore, introduce the notion of (non-trivial) coherence and relativize it to a class $K$ of physical systems which forms the background against which new correlations become interesting. A system is (non-trivially) coherent with respect to $K$ if and only if it exhibits a type of non-trivial coherence with respect to $K$. This definition is context-dependent since it is relativized to the class $K$. Nevertheless it remains, in a sense, largely context independent since it is applicable to a wide variety of situations. Non-trivial coherence is most interesting when class $K$ is almost universal, for instance, the class of all Newtonian systems ($N$) or the class of all quantum-mechanical systems ($Q$). Returning to our examples, it is obvious that the coupled oscillators exhibit non-trivial coherence with respect to $N$ and the other three cases exhibit non-trivial coherence with respect to $Q$.

4. ENTANGLEMENT AND COHERENCE IN QUANTUM SYSTEMS

In the quantum mechanics of composite systems, entangled states are those states that cannot be expressed as a product of states of its individual subsystems. For simplicity, consider a system involving just two subsystems, each representable by a finite-dimensional Hilbert space. In this case, each possible state of the composite system is represented by a vector in a Hilbert space $H_1 \otimes H_2$, where $H_1$ and $H_2$ are the Hilbert spaces of each of the two subsystems in isolation. The product states of such a system are those represented by vectors $|\psi\rangle_{1+2} = |\psi\rangle_1 \otimes |\psi\rangle_2$, for a pairs of states, $|\psi\rangle_1 \in H_1$ and $|\psi\rangle_2 \in H_2$. The entangled
states of such a system are those represented by vectors $\psi_{1,2} \in \mathcal{H}_1 \otimes \mathcal{H}_2$ such that $\psi_{1,2} = \eta_1 \theta_2$ for any of the possible pairs of states $\eta_1 \in \mathcal{H}_1$ and $\theta_2 \in \mathcal{H}_2$. Note that the state vector of any composite system of two objects (each having a finite-dimensional Hilbert space) can be written as a linear combination of unit vectors $|\alpha_i\rangle|\beta_j\rangle$, $i = 1, \ldots, n$ forming an orthonormal basis for $\mathcal{H}_1 \otimes \mathcal{H}_2$. This basis is known as the Schmidt basis and the linear combination is called the “Schmidt decomposition.” This basis makes manifest the character of a quantum state: if only one of these components is non-zero then the state is a product state; otherwise it is an entangled state. (Note that the Schmidt decomposition cannot, in general, be achieved for systems with more than two parts. Nevertheless, in such cases, any $n$-dimensional system can still be rather artificially decomposed into two subsystems, one with dimension $k (0 < k < n)$ and the other with dimension $n - k$.)

It is hard to over-emphasize the importance of entanglement for understanding quantum systems: as Schrödinger (1935, 555), who introduced the concept, put it: “I would not call [entanglement] one but rather the characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought.” However, what are most interesting here are the following results.

**Theorem 4.1 (a):** All isolated composite systems (having finite-dimensional Hilbert spaces) in entangled states are coherent with respect to $Q$; (b): However, not all quantum-mechanical systems that are coherent with respect to $Q$ are entangled.

**Proof (a):** Any entangled state for which the Schmidt decomposition can be performed will exhibit correlations between at least two pairs of properties of its component systems, namely whose eigenvectors form the Schmidt basis. Such an entangled state may be written as $\psi = \sum_{i=1}^{n} \mu_i(t)|\alpha_i\rangle|\beta_i\rangle$, where the $\mu_i(t)$ are complex numbers and $|\alpha_i\rangle$ and $|\beta_i\rangle$ are the basis vectors. For that basis, there are clearly $n$ ordered pairs $(\alpha_i, \beta_i)$ ($i = 1, \ldots, n$, where $n > 1$) of values of the observables $A$ (of system 1) and $B$ (of system 2). These $n$ ordered pairs of properties are perfectly correlated. Such a correlation is not found in all quantum systems. Thus a system in an entangled state will be non-trivially coherent with respect to $Q$;

**Proof (b):** This is shown by exhibiting a counterexample: the system above exhibiting Rabi oscillations (see section 3) exhibits coherence but is not entangled. There are many other counterexamples that could be supplied here as well.

The discussion of entanglement that we have given above refers only to systems with two subsystems. This definition of “entanglement” and theorem 4.1, can be straightforwardly extended to systems with countably infinite subsystems. However, entangled states of two-particle systems have long been of interest to philosophers because, to begin with:
Theorem 4.2 (a): States of two-particle (quantum) systems each having Hilbert spaces of dimension 2 that violate the Bell inequality are entangled. (b): Moreover, states of two-particle (quantum) systems that violate the Bell inequality are non-trivially coherent with respect to Q.

Proof (a): This has long been part of the folklore of foundations of quantum mechanics. (b): This now follows directly from theorem 4.1(a).

Furthermore, these results are very likely extendable to apply to larger systems and systems of higher dimensionality. Finally, returning to F-states, the coherence of which has been a source of disquiet about reductionist explanation in biology, it is easy to prove that:

Theorem 4.3 (a): Some F-states exhibit entanglement; (b): Not all F-states exhibit entanglement.

Proof (a): Yang (1962) has demonstrated that ODLRO is present only in density matrices of systems of bosons or fermion pairs forming bosonic quasiparticles (via pair occupation of single-particle states by fermions). Being composed entirely of bosons, Frohlich systems must have wave functions \( \Psi(1, 2, ..., N) \) that are symmetric under exchange of particle labels. Whenever the particles are not in the same single-particle state, \( \phi \), the overall system state will exhibit entanglement because this symmetry requirement yields a state of the system that is unfactorizable, i.e. \( \Psi(1, 2, ..., N) = \Phi(1)\Phi(2)\ldots\Phi(N) \) where \( \Phi(i) \) are wavefunctions describing individual particles. For example, if one particle is in a state \( \phi \), the state of the other \( N-1 \) particles, \( \Psi \) is be a sum of \( N \) different terms each containing one factor \( \phi(i) \), for exactly one value of \( i \in \mathbb{Z}^+ \), and \( N-1 \) factors \( \phi(j), j \neq i \).

Proof (b): F-states that have reduced density matrices for which \( \chi(r', r'') = 0 \), in which case all particles are in the same single-particle state, are not entangled. For them, \( \rho_1(r', r'') = N\Phi(r')\Phi^*(r'') \) and \( \Psi(1, 2, ..., N) = \prod_{i=1}^{N} \Phi(i) \), where \( \Phi(i) \) is the wavefunction attributed to the \( i^{th} \) particle (for example, in the case of the ground state of the entire system, \( \Psi = \phi_0(1)\phi_0(2)\phi_0(3)\ldots\phi_0(N) \), so that for all \( i, \Phi(i) = \phi_0(i), \) the single particle ground state). Exchanging particle labels in \( \Psi \) leaves the state unchanged in compliance with the symmetry requirement for bosons. From the definition of entanglement, such product states, which are F-states, are not entangled.

5. THE QUESTION OF REDUCTION

In the explanation of the coherent behavior of all the systems that we have discussed, our first criterion of reduction (fundamentalism) is easily satisfied. Thus, none of our examples fail to satisfy the condition for the first sense (a) of reduction. It is important to note that this is a non-trivial claim and should not be confused with the ontological question whether any new process exists at the higher level. Explanations from the
more fundamental assumptions routinely involve making approximations or taking limits and this can be philosophically problematic. The relevant procedures might well be contrived in the sense that the particular way in which a limit is taken or an approximation made may be motivated specifically by the desired result (Sarkar 1996, 1998). In such a situation, the fundamentalist criterion is violated. It is an open question how often this happens in physics.

Turning now to strong reduction, the analysis in section 4 permits the following three conclusions to be drawn:

(i) The existence of coherence does not always imply a failure of strong reduction for either classical or quantum systems. The coherent motions of the (classical) coupled oscillators system can be simply explained by describing each oscillator separately and noting the correlations between their positions. The coherent features of the Rabi oscillations are explained by the coupling between the states, an interaction between parts of a hierarchy that can be spatially instantiated (though, as in most quantum systems, it cannot be easily and accurately visualized);

(ii) Explanations of coherence in quantum systems that involve entangled states violate strong reductionism because they violate the second condition (hierarchical structure) and, ipso facto, also the third condition (spatial instantiation). If a composite system is described by an entangled state, definite states in general cannot be attributed to its individual subsystems. As Schrödinger (1935) pointed out, in such a state the subsystems cannot be in general be individuated, and locutions such as “subsystem A” and “subsystem B” do not refer to any precise entity within the entire state. Consequently, often no hierarchical relation between the subsystems’ states and that of the composite system exists, let alone is instantiated in physical space. This conclusion has some serious consequences. Consider, for example, a hydrogen atom which consists of an electron and a proton interacting with each other. In a fully quantum description of this atom, it would be represented by an entangled state. Therefore it cannot be represented as a hierarchical structure with identifiable individual states for the proton and electron. The situation is the same for atoms other than hydrogen (Jaeger 2000). What this means is that once entanglement becomes involved, the usual hierarchical picture of the composition of matter breaks down in quantum mechanical explanations.

(iii) Such explanations invoking F-states can violate strong reduction. This is trivially a consequence of the last point. Note, however, that by theorem 4.1b the F-state may exhibit no entanglement, in which case all three criteria for strong reduction can be satisfied. However, composite systems that are represented by product states and, therefore, do not pose problems for (at least abstract) reductionist explanation can yet exhibit unusual properties, such as superfluidity and superconductivity: fluid motion with zero viscosity and electrical conduction with zero resistance, respectively. Superfluids or superconductors in their ground states are such systems. Nevertheless, the point is that if there is a the real culprit is entanglement, not coherence. Unusual behavior is no guarantee that the hierarchy criterion is violated,
6. CONCLUSIONS

We only have two major conclusions. **First**, we have shown that the invocation of quantum-mechanical entanglement violates the conditions for strong reduction within physics though coherence alone is not sufficient for that purpose. Entanglement destroys the possibility of strong reduction because an entangled system cannot be described as being hierarchically organized. This does not, of course, imply that this is the only way in which strong reduction can fail in physics. Leggett (1987) and Fisher (1988), for instance, have argued that the fundamentalist assumption itself fails in some types of explanation in condensed matter physics. In our discussion of examples of types (b) and (d) of reduction we, too, have indicated that strong reduction fails in those explanations of critical phenomena in condensed matter physics that involve renormalization. Whatever be the merit of those claims, we hope that our analysis has drawn attention to the fact that not all the issues surrounding the role of reductionist explanation in physics have been resolved.

**Second**, physicalist explanations in molecular biology have so far satisfied the conditions for strong reduction (Sarkar 1989, 1992, 1996, 1998). However, if entangled F-states have to be invoked in such explanations, strong reduction will fail because of the failure of the hierarchy assumption. Whether F-states have any explanatory role in molecular biology remains highly controversial. It suffices here simply to note that there is at present no plausible candidate for a biological mechanism involving an F-state. It remains an empirical question whether, eventually, such a mechanism will be discovered, and lead to a failure strong reduction in biology. Should that happen, what will perhaps be most ironic is that it will fail not because of any special property of biological systems (as anti-reductionists have usually held), but because of the nature of physics itself. Thus, if the reduction of biology to "fundamental" physics is a problem, then what is problematic is not the relation of biology to physics, but the relation of macroscopic physics to "fundamental," that is microscopic physics. The full oddness of this situation seems not to have been fully appreciated, despite its having been occasionally pointed out at times before (e.g. Shimony 1978).

*Boston University*

*University of Texas at Austin*

**ACKNOWLEDGEMENTS**

We would like to thank Abner Shimony for initiating our interest in the problems of reductionism in quantum mechanics. Besides Abner, we would like to thank Tian Yu Cao, Don Howard, Simon Saunders, John Stachel and Chuck Willis for comments and criticism of an earlier version of this paper. This research was partly supported by an NIH grant (No. 7 ROI HG00912-02) to SS.
NOTES

1. See Wimsatt (1979), Sarkar (1992, 1998), and Schaffner (1994) for reviews.
2. See, for example, Nagel (1961). However, Specter (1978) and Shimony (1987) are notable exceptions.
4. See also Ho and Popp (1993).
5. For models of supervenience see, e.g., (Davidson 1970; Rosenberg 1984) and, especially, Kim (1993).
6. A detailed discussion of ontological issues can be found in Wimsatt (1995). Arguments against the philosophical importance of these issues are developed by Sarkar (1998).
7. This analysis is not straightforwardly compatible with models of reduction which do not view it as a form of explanation — e.g. Balzer and Dawe (1986, 1986a) and Ramsey (1993). However, it is unclear that these models should at all be regarded as those of "reduction" in any usual philosophical sense, given that this term has almost always been used to refer to a particular mode of explanation.
8. The approach that is summarized here is developed in Sarkar (1998). That work includes a discussion of the epistemological problems posed by approximations, which is glossed over in this paper.
9. The provision of a "warrant" may be weaker than logical deduction or even derivation. A warrant can potentially be a theoretical claim that is made on the basis of experimental facts known at the reducing level (see Sarkar 1989). Note that the reducing level need not be a lower level of any hierarchy as, e.g., in the reduction of Newtonian gravitation to general relativity or in the reduction of Newtonian mechanics to special relativity.
10. The simplest kind of such a hierarchical structure is a (graph-theoretical) tree. More complicated structures would only require a directed graph with an identifiable root and no cycles.
11. That one must distinguish between reductions only satisfying (i) and those that satisfy the other criteria was pointed out by Nickles (1973) though, in his usage, the reductions occurred in the opposite direction than the one discussed here (Wimsatt 1976).
12. For a particularly clear exposition, see Binney, Downick, Fisher and Newman (1993). Renormalization theory in condensed matter physics has not received nearly the kind of philosophical attention that it deserves. (For discussions of renormalization theory in the context of quantum field theory see (Brown 1993).)
13. Two quantum states of the system are said to be "coupled" if the system is subject to a force, with corresponding energy (here $W_{12}$), capable of causing the system to change from one state to the other.
14. See (Yang 1962; Penrose and Onsager 1956; Penrose 1951; and Ginzburg and Landau 1950).
15. Ho (1993, 141) admits that "[c]oherence in ordinary language means correlation, a sticking together, or connectedness; also, a consistency in the system." He eschews any attempt at explicit definition and later (pp. 150-151) only provides two jointly sufficient conditions for coherence, but even these are restricted to a quantum context.
16. We mean "feature" quite generally in the sense that even possible states of a system will be regarded as "features."
17. For instance, a classical system consisting of a free particle with mass $m$, energy $E$, and momentum $\hat{p}$ is coherent because $E$ can be estimated from $\hat{p}$, though $\hat{p}$ cannot be estimated from $E$. Therefore this system will not be fully coherent.
19. This does not, of course, mean that such explanations necessarily satisfy the criteria of strong reduction — that will depend on whether the other explanatory factors involved also satisfy these criteria.
20. T. Y. Cao (personal communication) has argued that the proper response to this situation is to abandon the characterization of "hierarchy" that we use, and adopt a less restrictive one and thus save the usual hierarchical picture of matter. We find this unpalatable for two reasons: (i) the notion of hierarchy we use is the standard one appropriate not only for biology and the social sciences, but also for classical physics. It would be odd philosophical strategy to weaken a generally useful notion simply to include
one special case; (ii) such a strategy would prevent a recognition of yet another way in which quantum mechanics undermines our classical intuitions and this is precisely what we think is of most interest.

21. See, for example, (Cooper 1978; Yushina 1982; Fröhlich 1983; Mishra and Bhowmik 1983; Ho and Popp 1993).

REFERENCES


