Emergence, Reduction, and Theoretical Principles: Rethinking Fundamentalism

Margaret Morrison†

Many of the arguments against reductionism and fundamental theory as a method for explaining physical phenomena focus on the role of models as the appropriate vehicle for this task. While models can certainly provide us with a good deal of explanatory detail, problems arise when attempting to derive exact results from approximations. In addition, models typically fail to explain much of the stability and universality associated with critical point phenomena and phase transitions, phenomena sometimes referred to as "emergent." The paper examines the connection between theoretical principles like spontaneous symmetry breaking and emergent phenomena and argues that new ways of thinking about emergence and fundamentalism are required in order to account for the behavior of many phenomena in condensed matter and other areas of physics.

1. Introduction. Fundamentalism comes in many guises. Philosophers typically use the term to refer to the view that fundamental theories like Newton's theory or quantum mechanics and their accompanying laws are the source of explanation of the basic features of the physical world. Typically fundamentalists advocate reduction in one form or another while antireductionists are fond of pointing out that the world has different levels of complexity that cannot be fully explained or understood by reducing them to one set of theories, laws, or basic entities. As is frequently the case with philosophical debates, the views of each side are not without merit. Some forms of reduction have been very successful in producing an understanding of the world that could not have been achieved in other ways; Maxwell's electrodynamics and its reduction of optics and electromagnetism is a case in point. Yet many physical systems, entities, and properties seem to defy reductionist explanations. This paper explores the ways in which some examples of emergence in physics direct

†To contact the author, please write to: Department of Philosophy, University of Toronto, 215 Huron Street, Toronto ON, Canada, M5S 1A2; e-mail: mmorris@chass.utoronto.ca.

Philosophy of Science, 73 (December 2006): 876–887. 0031-8248/2006/7305-0036$10.00
Copyright 2006 by the Philosophy of Science Association. All rights reserved.

https://doi.org/10.1086/518746 Published online by Cambridge University Press

us toward a new way of thinking about fundamentalism, one that can be understood in terms of theoretical principles like symmetry breaking and localization. My claim is that these principles are fundamental because they are responsible for predicting and explaining the behavior of large numbers of phenomena both in condensed matter physics and in the high energy domain. But they speak against traditional forms of reductionism because the stable, generic behavior they produce is immune from changes in the microphysical base. Consequently they provide some interesting insights into how we might go about rethinking issues associated with fundamentalism, reduction, and emergence.

2. Reductionism Redux[2]**.** Let me begin with a quick overview of some of definitions and points of debate. Ontological reductionism is typically associated with the idea that entities, properties, or processes at one level are "nothing more than" a manifestation of entities, properties, and so on that occur at a lower level. For example, the macro properties of a solid are thought to simply consist of its micro level atomic properties where the relation between the two domains can be understood as one of supervenience, that is, the macro properties arise out of or are dependent on the micro properties. The other way of thinking about ontological reduction is in terms of identity where the macro properties are simply *identified with* the micro properties as in the case where heat is identified with kinetic molecular energy. One of the ways of thinking about epistemological reduction concerns the manner in which we express or account for these ontological relations; it deals with the relations between concepts and/or theories. Much of the literature on reduction examines the possibility of reducing one theory to another, as in the much discussed case of thermodynamics and statistical mechanics. However, in order for this kind of reduction to work we need to be able to identify the terms or concepts of the reduced theory with those in the reducing theory, something that is generally considered to be lacking in the case of statistical mechanics (see especially Sklar 1999).

A related notion of epistemological reduction involves the derivation of a law or set of laws from other more basic ones, for example, the derivation of Kepler's and Galileo's laws from Newtonian mechanics. Although philosophers have long been at pains to point out that this deduction is not straightforward in the way that a logical deduction is, it nevertheless embodies a kind of explanatory relation whereby the more inclusive theory (Newton's) explains why the other two hold in the circumstances where they do. Steven Weinberg (1987) discusses a similar situation that he terms "objective reductionism." The idea involves the "convergence of arrows of explanation" to a common source where more basic or fundamental scientific generalizations explain others, thereby giv-

https://doi.org/10.1086/518746 Published online by Cambridge University Press

ing us a "sense of direction in science." Because particle physics is on a level closer to the "source" of the arrows of explanation than any other areas of physics it is more fundamental (437). This framework is objective because it is a fact about nature itself and not about scientific programs (436). He distinguishes this "grand reductionism" from "petty reductionism," which is the view that mere knowledge of the constituents of a complex system would be sufficient to understand it. Grand reductionism claims that scientific principles can, in principle, all be traced down to a small body of simple universal laws (79).

What is significant here is that the reduction is carried out in terms of principles, not entities. This is important for a number of reasons. First, no physicist, nor anyone else for that matter, would deny that a good deal of scientific activity involves breaking down entities into their fundamental constituents. Physics, biology, and chemistry all involve this kind of analysis. And while that analysis provides a good deal of information about the physical world, we also need to know the ways in which the constituent parts behave, together with the laws that govern them, when they form complex systems. Similarly, we want to know the relation of those laws to the ones that govern the constituents when they act in isolation. Weinberg's reductionism assumes that relatively few laws will give us all the information we need, about both the constituents and the systems they comprise.

Several years earlier Philip Anderson (1972) presented a view that acknowledged the importance of reduction while advocating emergence as necessary for understanding many physical phenomena. Anderson claims that the reductionist view that all animate and inanimate matter obeys the same fundamental laws is accepted without question within science. But that in no way entails the truth of what he calls the "constructionist" hypothesis, which states that "the ability to reduce everything to simple fundamental laws does not imply the ability to start from those laws and reconstruct the universe" (1972, 393). The problem facing the constructionist hypothesis is that the behavior of large and complex aggregates of elementary particles cannot be understood simply by referring to the properties and laws that govern small numbers of elementary particles taken in isolation. As Anderson points out, at each level of complexity "entirely new properties appear and the understanding of new behaviours requires research which is . . . as fundamental in its nature as any other" (1972, 393). In fact, he claims that the fundamental laws investigated by elementary particle theorists are not especially relevant to real problems in the rest of science.

What Anderson is objecting to here is the idea that is implicit if not explicit in the reductionist program, namely, that one can explain the macro world using only the laws that explain the micro world. This links

https://doi.org/10.1086/518746 Published online by Cambridge University Press

up with the debate about fundamental physics via the notion of explanation. Because the micro world of elementary particles, and the laws that govern them, form the foundation for many of our explanations they are considered more "fundamental" than the phenomena they explain. Of course there is an important sense in which this is right; we do understand the macro world in terms of its constituent parts, and a good deal of explanation of macro phenomena is in terms of micro structures and processes. But how far does this kind of explanation extend? In other words, are these reductive explanations as powerful as we think?

The nonrelativistic Schrodinger equation presents a nice picture of the kind of reduction Weinberg might classify as "fundamental." It describes in fairly accurate terms the everyday world and can be completely specified by a small number of known quantities: the charge and mass of the electron, the charges and masses of the atomic nuclei, and Planck's constant. Although there are things not described by this equation, such as nuclear fission and planetary motion, what is missing is not significantly relevant to the large scale phenomena that we encounter daily. Moreover, the equation can be solved accurately for small numbers of particles (isolated atoms and small molecules) and agrees in minute detail with experiment. However, it can't be solved accurately when the number of particles exceeds around ten. But this is not due to a lack of calculational power, rather it is a catastrophe of dimension (Laughlin and Pines 2000). It is possible to perform *approximate* calculations for larger systems, and it is through such calculations that we know why atoms have the size they do, why chemical bonds have the length and strength they do, why solid matter has the elastic properties it does, and so on. More experimental detail enables predictions of atomic conformations of small molecules, simple chemical reaction rates, structural phase transitions, ferromagnetism, and sometimes superconducting transition temperatures. However, the schemes for approximating are not first principles deductions but instead require experimental input and local details. Hence, we have a breakdown not only of the reductionist picture but also of what Anderson calls the "constructionist" scenario.

Yet, as Laughlin and Pines (2000) point out, there is a peculiarity here; the parameters e, h, and m appearing in the Hamiltonian for the Schrodinger equation *can* be accurately measured in laboratory experiments involving large numbers of particles. For example, the magnetic field generated by a superconductor that is mechanically rotated measures e/mc while four point conductance measurements on semiconductors in the quantum Hall regime accurately determine the quantity e^2/h. However, because the Schrodinger equation involves approximations we are unable to derive or predict these values by direct calculation. The natural questions that arise here are what accounts

https://doi.org/10.1086/518746 Published online by Cambridge University Press

for the accuracy and stability of the results. The answer involves an appeal to theoretical principles like symmetry breaking and localization, principles that sometimes figure in theory construction but are also responsible for different kinds of phenomena whose behavior appears independent of the content of microphysical theories. The Josephson effect is exact because of the principle of continuous symmetry breaking and the quantum Hall effect because of localization.[1]

3. Emergent Phenomena and Theoretical Principles. Phenomena occurring as a result of processes like symmetry breaking and localization are also classified as "emergent," meaning that they arise from but are not dependent on lower level processes and entities. Philosophical discussions of emergent phenomena are most prevalent in the philosophy of mind literature (e.g., Kim 1999) where the goal is to connect mental properties to some physical base while avoiding the pitfalls of reductionism. The formation of a crystal lattice shows how the connection between theoretical principles and emergent phenomena plays out in physics. Crystallization is manifested only by a sufficiently large and complex system, exactly the sort of complex aggregates Anderson was referring to in his discussion of the constructionist hypothesis. The electrons and nuclei that make up a crystal lattice do not have rigidity, regularity, elasticity—all characteristic properties of the solid. These are only manifest when we get "enough" particles together and cool them to a low "enough" temperature. In fact some particles—atoms of either isotope of helium, or electrons in a metal—simply do not stack at all and remain fluid right down to absolute zero. This illustrates one of the most important facts about broken symmetry: quantum mechanical as well as thermal fluctuations are inimical to it. The symmetric crystalline state is an example of an emergent property of broken symmetry because the crystal has less symmetry than the atoms of the fluid from which it crystallized.

Superconductivity provides a nice example of how theoretical principles like broken symmetry can function in accounting for the stability and predictability of certain phenomena/effects. Standard textbook treatments tell us that superconductors exhibit startling phenomena that can be predicted with extraordinary accuracy, as in the case of very low electrical resistance where currents can circulate for years without perceptible decay. But, in deriving these results the textbooks generally use models that are, for the most part, little more than reasonably good approximations. There

1. Laughlin and Pines (2000) call these organizing principles, but that seems to suggest that they are somehow tools for classification. I have referred to them as theoretical principles in order to capture the idea that there is a dynamical process associated with these principles responsible for producing certain kinds of behavior.

https://doi.org/10.1086/518746 Published online by Cambridge University Press

are macroscopic models like Ginzburg-Landau and the microscopic model of the Bardeen-Cooper-Schriffer (BCS) theory where electrons are assumed to interact only by single phonon exchange. So here again the question arises: How can one use such approximations to derive predictions about superconducting phenomena that are of essentially unlimited accuracy? The answer is that although high precision predictions about superconductors follow from the models, more generally they follow from the fact that the models exhibit a spontaneous breakdown of electromagnetic gauge invariance in a superconductor (Weinberg 1986, 1996). One needs the detailed models like BCS to explain the mechanism for spontaneous symmetry breaking (SSB), at what temperature superconductivity is produced and as a basis for approximate quantitative calculations, but not to derive the most important exact consequences of this breakdown such as infinite conductivity.[2]

The notion of emergence relates to superconductivity in the following way: In the $N \to \infty$ limit of large systems (the macroscopic scale) matter will undergo mathematically sharp, singular phase transitions to states where the microscopic symmetries and equations of motion are in a sense violated. The symmetry that has been broken "leaves behind as its expression," to use Anderson's words, only certain characteristic behaviors such as long wave length vibrations (sound waves), the conduction phenomena of the superconductor, or the very rigidity of crystal lattices and most solid matter. Hence in the case of this kind of broken symmetry in the properties of macroscopic material bodies we see that the whole becomes "not only more than but very different from the sum of its parts" (Anderson 1972, 395).

The significance of the dependence on theoretical principles like SSB for the reductionist hypothesis is that results of the sort mentioned above (e.g., infinite conductivity) would continue to be accurate even in the face of significant changes to our microphysical theories. In other words, we might have an entirely different model from the current one(s) of *how* SSB occurs, but this would in no way affect the predictions derived from the principle itself. But this raises a more difficult philosophical question: How does this seeming independence from the microphysical world affect the reductionist strategy of explaining macro phenomena on the basis of physical laws governing the micro world. There are a number of possi-

2. Again, see Weinberg (1986, 1996) for a derivation of effects like the Meissner effect, flux quantization, and the AC Josephson current from the assumption of spontaneous breakdown of electromagnetic gauge invariance. As Weinberg points out (1996, 336), most textbooks explain superconductivity in terms of dynamical models and rarely mention broken symmetry, but in fact one can *define* a superconductor simply as a material in which electromagnetic gauge invariance is spontaneously broken.

https://doi.org/10.1086/518746 Published online by Cambridge University Press

bilities. First, it suggests that the reductionist goal implicit in the search for a theory of everything cannot explain many fundamental features of our scientific world picture. Or, relatedly, attempts to write down the ultimate equations that will tell us about physical phenomena at both high and low temperatures may no longer be a productive strategy. This of course is not to say that we need to abandon all forms of ontological and epistemological reductionism since a good deal of knowledge of the physical world comes via reductionist strategies. Rather, we need to be aware of its limits and refrain from excusing its failures with promissory notes about future knowledge and ideal theories.

4. Philosophical Matters: Some Problems for Wholes, Parts, and Natural Kinds. What are philosophers of science to make of this "emergent physics"? Not only does it call into question the very idea that an understanding of the fundamental laws that govern the microphysical world can explain macro level phenomena, it also casts doubt on the claim that when the former strategy fails our understanding of physical behavior must be restricted to local models. The relation to higher level theoretical principles like symmetry breaking and localization shows that certain kinds of stable behavior, though not derivable from fundamental theory, can nevertheless be explained in a systematic way, one that doesn't rely on the contingencies of particular situations. Typically those who point out problems with reductionism and fundamental theory for understanding the physical world tend to look to models as the source of explanation for many phenomena that physics deals with. But, as we saw above, this is unsatisfactory not only because they provide no explanation of accuracy and stability but because one cannot derive exact consequences from approximate models.

A denial of the kind of reductionism mentioned above together with the role of theoretical principles points to a hierarchical view of the world where emergent phenomena arise out of their constituent parts but are not dependent on them in the sense of being reduced to, explained by, or predicted from these constituents. In other words, while there is an ontological relation that holds between the different levels, there is no corresponding theoretical relation except, say, the straightforward claim that all ordinary matter obeys basic theories like quantum mechanics and electrodynamics. Unlike cases in the philosophy of mind, the problem of trying to provide some physical "grounding" for emergent phenomena/properties simply isn't present in the physics cases discussed above. There is no question about phenomena like superconductivity and ferromagnetism being "physical" in the robust sense of the word. Still, one would like to be able to say a bit more about the ontological relation that holds

https://doi.org/10.1086/518746 Published online by Cambridge University Press

between the emergent phenomena and the lower level entities that constitute them.

There have been several characterizations of the relation between emergent properties and their constituent parts; two notable ones are Humphreys (1997) and Teller (1992). On Humphreys's account the constituent parts are thought to no longer exist once the emergent property has formed, while for Teller the emergent property is thought to simply "transcend" the parts from which it arises. In keeping with the scientific examples discussed above the common thread in these philosophical characterizations is that emergence depends on a hierarchical view of the world. However, if we look closely at other examples from physics we can see that the part/whole relation is more problematic than the philosophical accounts have suggested.

The most striking feature of emergent physical phenomena is that it is not the constituents that cease to exist once the new phenomenon is produced; instead the emergent phenomenon typically disappears when the system is taken apart. For example, when a large number of atoms condense into a crystal, the phonon, the elementary quantum of sound, becomes a perfectly legitimate particle at low energy scales. It propagates freely, does not decay, interacts by simple rules that can be verified experimentally, carries momentum and energy relative to wavelength and frequency, and mediates the attractive interaction responsible for conventional superconductivity. None of these things depends in detail on the underlying equations of motion; they are simply generic properties of the crystalline state (Laughlin 1999). However, the phonon ceases to have meaning when the crystal is taken apart because sound makes no sense in an isolated atom. Similarly with superfluidity—a low energy collective effect of huge numbers of particles. It cannot be deduced from the microscopic equations of motion in a rigorous way and completely disappears when the system is taken apart. The point here is that the excited states of these systems are particles in the same way that an electron is, yet they do not qualify as elementary insofar as they do not exist outside the stable state of matter that gives rise to them. These states or quantum "protectorates" and their accompanying emergent behavior demonstrate that the underlying microscopic theory can easily have no measurable consequences at low energies.[3]

There are also rather counterintuitive aspects to the part/whole re-

3. Emergent phenomena/properties are sometimes referred to as "quantum protectorates," meaning that they are protected properties of matter in the sense that their behavior is not subject to changes in the microphysical base. Instead their stable behavior is governed solely by the kinds of theoretical principles discussed above. These and other examples are discussed by Laughlin 1999.

https://doi.org/10.1086/518746 Published online by Cambridge University Press

lation of some emergent phenomena. The fractional quantum Hall effect is a good example. It implies that many electrons, acting in concert with powerful gauge forces between them, can create new particles having a charge smaller than the charge of any constituent electron, a charge that is exactly 1/3, 1/5, 1/7, etc. of an electronic charge, depending on the conditions under which it has been prepared. The idea that objects combine to form smaller ones is rather peculiar, yet it is known "with certainty" that none of the electrons has split up into pieces (Stormer 1999). Quantum numbers, usually integers or half-integers, also turn out to be fractional, such as 2/5, 4/9, and so on. Moreover, bits of magnetic field can attach to each electron creating yet other objects that have properties very different from those of the electrons. For instance, their mass is unrelated to the mass of the original electron but arises solely from interactions with their neighbors. The interesting feature from the point of view of emergence is not just the peculiarity of these phenomena but the fact that quantum numbers and gauge fields, two central components of the standard model that were thought to be fundamental, can now be characterized as emergent phenomena.

In fact the notion of emergence extends well into the domain of quantum field theory with the use of effective field theories that allow for the exploration of phenomena at one level without concern about what happens at higher energies. Similarly, renormalizability, which is usually thought of as a constraint on "fundamental" quantum field theories can be reconceived as an emergent property of matter both at quantum critical points and in stable quantum phases. To see why this is so simply think of the original use of renormalization—to get rid of the divergent terms in the perturbation by redefining the parameters of the original theory, for example, e and m in QED. The upper limit of the integral was replaced by a cut-off Λ which was merely a calculational device, so one takes the limit and lets Λ go to infinity at the end of the calculation. With the introduction of the renormalization group (RG) equations by K. Wilson and others the cut-off can be interpreted realistically because one need only consider the energy scale for the phenomena we are concerned with.[4] The masses and charges, for example, depend on the scale under consideration because as one lowers the energy the microscopic details become less important. The RG equations can then be used to reevalute these parameters at higher or lower scales. The point is that they have no absolute values but are scale dependent. So what started off as a mathematical technique has become reinterpreted, to some extent, as evidence

4. For a summary of this work, see Wilson 1983.

https://doi.org/10.1086/518746 Published online by Cambridge University Press

for the multiplicity of levels required for understanding physical phenomena.

Given the rather tenuous connection between emergent phenomena and the entities/properties that give rise to them, something needs to be said about how to characterize similarities and differences between different *kinds* as opposed to *states* of matter. In other words, at what point do we identify a phenomenon as truly emergent as opposed to just a different form or state of the basic underlying constituents. This of course is an age-old philosophical problem, one that evokes unresolved debates about essences and natural kinds, notions that are often more complex and ambiguous than the problems they are intended to solve. With respect to emergent phenomena in physics the tool for differentiation is the reversible adiabatic map. A change in the underlying equations of motion for a system requires a check at each stage to make sure that the ground state and the low-lying excitations have evolved in a one-to-one way (see Laughlin 1999). If the system contains a small number of particles this check isn't really necessary since the mapping is guaranteed to be one to one by the adiabatic principle. However, if the system contains a thermodynamically large number of particles a small change to the equations of motion can result in a violent rearrangement of the ground state and low-lying excitations, and a corresponding breakdown of the one-to-one mapping. In such cases we have a quantum phase transition which functions as the criterion for individuation. In other words the distinction we are concerned with is different *phases* of matter rather than different types or kinds. Two states are the same phase of matter if they can be slowly transformed into each other without encountering a quantum phase transition and different if they cannot. On this definition metals, insulators, and superconductors are all different while two metals with slightly different electron-electron repulsion strength are the same.

This notion of sameness is rather counterintuitive from the point of view of fundamental physics, but that is exactly the point. When thinking in terms of reduction and derivation from first principles we are interested in differentiating matter into specific kinds according to its defining properties in much the same way as we classify elementary particles according to their quantum number, charge, and so on. Elements are differentiated on the basis of their atomic number and substances like water according to their chemical composition and molecular/atomic structure. This process whereby phenomena are thought to have essential features presupposes an underlying microphysical structure that allegedly underwrites and justifies the classification. Moreover, an explanation of these properties can typically be deduced from first principles. But, as we have seen, in the case of emergent phenomena this relation to microphysics is absent. Instead, the behavior is regulated by higher level theoretical principles—

https://doi.org/10.1086/518746 Published online by Cambridge University Press

core physical principles (like symmetry breaking) upon which more detailed dynamical models are based. In order for the reductive program to work these principles would need to be derivable from the underlying microphysical equations, which is not really possible.

These different characteristics of emergent physical phenomena point to the fact that philosophical accounts of emergence need to be restructured if they are going to capture essential features of "emergent" physics and that the model provided by philosophy of mind may not be the best resource on which to draw. Not only do we need to reorient our thinking about the role of natural kinds as a method for differentiating essential features of matter, but new ways of thinking about the part/whole relations involved in defining and describing emergent phenomena are required. Although the part/whole distinction remains ontologically important, it provides no predictive or explanatory information and as such ceases to function in an essential way. And perhaps most importantly, a reevalution of reductionist strategies for defining the relationship between different theoretical levels is crucial for making sense of emergence in physics.

In connection with the "levels" approach characteristic of this picture of emergence, it is important to note the differences between emergence and the fact that phenomena at different scales may obey different fundamental laws. While general relativity is required on the cosmological scale and quantum mechanics on the atomic these differences do not involve emergent behavior in the sense discussed above. With emergent phenomena we have generic, stable behavior that cannot be explained in terms of microphysics and is immune from changes to the equations of motion of the system.

To conclude, then, the existence of emergent phenomena undermines the kind of reductionism that is presupposed in the search for a theory of everything. But, at the same time it does not endorse the claim that the only way to understand nature is by modeling—that models are all we have. Both of these positions overlook the role of theoretical principles and their role in producing certain effects, not only principles that are the basis of "emergent" phenomena but those that figure in our so-called fundamental theories as well. It is really *they* that we ought to think of as the basis of fundamental physics. Consequently fundamental physics need not be equated with fundamental theories and reductionism, and to that extent its domain can extend beyond the world of elementary particles.

REFERENCES

Anderson, P. W. (1972), "More Is Different," *Science* 177: 393–396.
Humphreys, P. (1997), "How Properties Emerge," *Philosophy of Science* 64: 1–17.
Kim, J. (1999), "Making Sense of Emergence," *Philosophical Studies* 95: 3–36.

https://doi.org/10.1086/518746 Published online by Cambridge University Press

Laughlin, R. B. (1999), "Nobel Lecture: Fractional Quantization," *Reviews of Modern Physics* 71: 863–874.
Laughlin, R. B., and Pines, D. (2000), "The Theory of Everything," *Proceedings of the National Academy of Science* 97: 28–31.
Sklar, L. (1999), "The Reduction (?) of Thermodynamics to Statistical Mechanics," *Philosophical Studies* 95: 187–199.
Stormer, H. (1999), "Nobel Lecture: The Fractional Quantum Hall Effect," *Reviews of Modern Physics* 71: 875–889.
Teller, P. (1992), "A Contemporary Look at Emergence," in A. Beckerman, H. Flohr, and J. Kim (eds.), *Emergence or Reduction? Essays on the Prospects of Nonreductive Physicalism*. Berlin: de Gruyter, 139–153.
Weinberg S. (1986), "Superconductivity for Particular Theorists," *Progress of Theoretical Physics Supplement* 86: 43–53.
——— (1987), "Newtonianism, Reductionism and the Art of Congressional Testimony," *Nature* 330: 433–437.
——— (1996), *The Quantum Theory of Fields*. Vol. 2, *Modern Applications*. Cambridge: Cambridge University Press.
Wilson, K. G. (1983), "The Renormalization Group and Critical Phenomena," *Reviews of Modern Physics* 47: 583–600.

https://doi.org/10.1086/518746 Published online by Cambridge University Press