Simulating Many-Body Models in Physics: Rigorous Results, 'Benchmarks', and Cross-Model Justification*

Axel Gelfert
Collegium Budapest
Institute for Advanced Study
Szentháromság u. 2
H-1014 Budapest, Hungary
email: axel@gelfert.net

June 11, 2006

Abstract

This paper argues that, for a prospective philosophical analysis of models and simulations to be successful, it must accommodate an account of mathematically rigorous results. Such rigorous results are best thought of as genuinely model-specific contributions, which can neither be deduced from fundamental theory nor inferred from empirical data. Rigorous results often provide new indirect ways of assessing the success of computer simulations of individual models. This is most obvious in cases where rigorous results map different models on to one another. Not only does this allow for the transfer of warrant across different models, it also puts constraints on the extent to which performance in specific empirical contexts may be regarded as the main touchstone of success in scientific modelling. Rigorous results and relations can thus come to be seen as giving cohesion and stability to actual practices of scientific modelling.

*Pre-conference draft for the LSE/CNRS conference 'Models and Simulations', Paris, 12-13 June 2006.

1 Introduction

This paper argues that the philosophical analysis of models and simulations can be advanced by reconsidering the role of mathematics in the process of modelling. It focuses on a class of models intended to describe the physical behaviour of systems that consist of a large number of interacting particles. Such *many-body models* are typically employed in order to account for a range of complex phenomena such as magnetism, superconductivity, and other phase transitions. Because of the dual role of many-body models as models *of physical systems* (with specific empirical phenomena as their explananda) and as *mathematical structures*, they form a sub-class of scientific models, from which one can hope to draw general conclusions about the role of mathematical relationships in constructing and assessing models. Since many-body models lend themselves to computational evaluation via a range of techniques (e.g., Monte Carlo simulations, Green's function techniques, etc.), they are of special significance when it comes to analysing the relation between mathematical rigour and computer simulations.

The structure of the paper is as follows: First, an attempt is made to clarify the relation between models and simulations by drawing on recent work in the philosophy of models. While models have rightly earned their place in philosophical analyses of science, philosophical work on simulations is still in its infancy. It is argued (section 2) that a proper appreciation of simulations in science requires a shift in focus from questions of representation to analyses of how science generates results. The question of how models can be used to generate specific results has sometimes been overlooked; in the case of mathematical models, it has often been regarded as unproblematic. This calls for an analysis of the relation between mathematics and computation (section 3). Section 4 argues that, in addition to numerical results, there also exist mathematically rigorous results and relations, which may play an important role as benchmarks, even when they lack an empirical interpretation. Section 5 argues that it is via the existence of rigorous results and relations that seemingly unrelated models may mutually support each other; it is this phenomenon of cross-model justification which, I suggest in the Conclusion, is essential to the cohesion of the practice of modelling and simulation in science.

2 Simulating Models: From Representation to Results

While scientific models have enjoyed a long history as objects of philosophical inquiry, simulations have only recently begun to garner serious interest from philosophers of science. (See, for example, refs. (22), (12) and works quoted therein.) As a result, when it comes to distinguishing different kinds of simulations, or different philosophical perspectives on simulations, the philosophical terminology is much less stable than in the case of different views of models. Not that anyone is to blame for this: on the one hand, it is an expression of the liveliness and comparative novelty of the debate, on the other hand it reflects the fact that scientist themselves – those who routinely implement simulations, often by using computing power of various kinds – tend to employ the term quite loosely. This also applies to the relation between models and simulations. Scientists frequently speak of 'simulating a model' (in the sense of running a computer simulation of the dynamic behaviour of a model given a certain input, e.g. consisting in parameter values and boundary conditions), but they also profess to engage in 'modelling a simulation': that is, using numerical techniques to construct a *computational* model whose only raison d'être is its ability to generate sets of simulated data. Whether a 'model' is being 'simulated', or a 'simulation' is being 'modelled', is – at least in actual scientific usage – often not so much a question of logical order as of temporal order: new simulation techniques may be applied to well-established models, in which case an old model (say, the Ising model, first proposed in 1925, see ref. (13)) may be simulated in new ways, whereas new computational models may be developed simply because they are particularly suitable for the evaluation by means of certain easy-to-implement, reliable numerical techniques. The situation is not made any easier by a certain ambiguity of the term 'simulation', which can be understood as referring either to the process of applying a set of (usually numerical) techniques, or to the output generated by such a procedure. Given that actual scientific usage offers no clear verdict on the matter, a first task of any prospective philosophical analysis of simulation will be to clarify the logical order and the conceptual relationship between models and simulations; this is the task of the present section.

Recent philosophical interest in the use of models in science is in large part due to an approach, pioneered by Margaret Morrison and Mary Morgan, which views models as 'mediating instruments' (17, p. 10). As I shall argue later, the 'models as mediators' view, while by no means the only well-developed view of scientific models, has special affinities with the topic of simulation. Therefore, in the present context, it will be used to set the agenda for the discussion of scientific models. According to the mediator approach, models are to be regarded as more than mere unavoidable intermediary steps in applying our best scientific theories to specific situations. Rather, as 'mediators' between our theories and the world, models inform the interpretation of our theories just as much as they allow for the application of these theories to empirical phenomena. Models, it is claimed, 'are *not* situated in the middle of an hierarchical structure between theory and the world', but operate outside the hierarchical 'theory-world axis'. (17, p. 17f.) Traditionally, unless their role was seen as merely heuristic, models were to be judged by how well they fit with the fundamental theory and the empirical data, or, more specifically, how well they explain the data *by the standards of the fundamental theory*: ideally, a model should display a tight fit both with theory and with empirical data. Indeed, on certain accounts of the formal relation between theory and data, any application of theory to empirical phenomena – that is, any use of a theory that goes beyond the mere deduction of further theoretical statements – must necessarily happen via models embedded in the semantic structure of a theory. Without taking a stance on this issue, it seems obvious that such a view places a rather heavy theoretical load on the concept of 'model' – more, perhaps, than the notion of a scientific model (as derived from scientific practice) can bear.

The 'models as mediators' approach, by contrast, insists that any scientific account of specific processes and phenomena necessarily depends on factors that are extraneous to fundamental theory. Generalising from a number of case studies across the natural and social sciences (see ref. (16)), Morrison and Morgan argue that models 'are made up from a *mixture* of elements, including those from outside the domain of investigation' (17, p. 23); it is this partial independence from both original theory and empirical data that allows models to play an autonomous role in scientific inquiry. In this respect, Morrison and Morgan argue, the role of scientific models is similar to that of tools and scientific instruments; indeed, it is part and parcel of the mediator view that model building involves an element of creativity and skill – it is 'not only a craft but also an art, and thus not susceptible to rules' (17, p. 12). By focusing more on the process of model construction than on the logical relations into which models can enter with other abstract structures, such as theories, one might argue that the mediator view already displays a natural affinity towards questions of simulation. However, differences and divergencies remain, and the rest of this section will spell out some of these.

Before turning to differences between models and simulations, I want to comment briefly on one possible way one might try to reconcile the two debates, namely by regarding both models and simulations as equal constituents of the same scientific activity, 'scientific modelling'. As mentioned earlier, scientists often view models and simulations as, in effect, on a par with one another. However, while such a focus on the activity of scientists may be a useful perspective, for example in the context of sociology of science, for the purpose of discussing the epistemic status and justificatory role of models and simulations, it is useful to not conflate the two. Even if it is not easy to always draw a clear line between them, there appears to be sufficient continuity in order to distinguish at least paradigm cases where the difference between models and simulations is clear. As mentioned earlier, in abstract model theory different definitions of models exist, and in science, in addition to the straightforward use of mechanical or analogical models, certain much-discussed mathematical models, such as the Ising model, pre-date the computational means required for simulating the corresponding physical systems. In this regard, one can successfully distinguish models from simulations.

The mediator view, too, while not predominantly concerned with contrasting models and simulations, holds on to several of these distinctions, though it gives them a slightly different twist. One of its fundamental tenets, quoted above, is the thesis that models constitute a '*mixture* of elements', some derived from theory, other originating from extra-theoretical considerations: 'model construction involves a complex activity of integration' (15, p. 44). More often than not, this integration is neither perfect nor complete. When certain elements of a model are incompatible, the integration cannot be perfect. This, for example, is the case in the Bohr model's conflicting demands that the electrons in an atom should be conceived of as orbiting the nucleus on circular paths without losing energy, while at the same time viewing them as objects of classical electrodynamics. Integration may also remain incomplete for the simple reason that not all features of a system are eventually reflected in the model. As Daniela Bailer-Jones argues, 'selection of aspects for the purposes of modelling is an accepted and well-practised creative strategy' (1, p. 66). Which aspects are deemed relevant may depend on a range of criteria, including such factors as computational accessibility or explanatory interest, which themselves are determined less by theoretical first principles than by contingent facts of scientific practice. This conception of scientific models has sometimes been characterised as being fuelled by anti-theoretical sentiments. The most outspoken 'anti-theory' theorist has been Nancy Cartwright, and her conviction that 'theories in physics do not generally represent what happens in the world – only models represent in this way' (4, p. 180), can be seen as emblematic for this view of models. Morrison seconds Cartwright's view when she writes that 'the proof or legitimacy of the representation arises as a result of the model's performance in experimental, engineering and other kinds of interventionist contexts' – not by reference to theory. (14, p. 81) It is noteworthy, though, that neither Cartwright nor Morrison call into question the overall epistemic goal of all modelling and theorising: namely, to *represent*. Indeed, their claim is not that representation itself is unattainable, but rather that it can only be attained by means of models not theories.

However, there remains a gap between, on the one hand, the aspirations of the mediator view to solve the problem of scientific representation and, on the other hand, the way it assesses the success of scientific models. Merely asserting that models are instruments for intervening in the world, and that their representational success is to be assessed by their performance in 'interventionist contexts' leaves open how we derive knowledge from their application. Bailer-Jones recognises this: 'If one chooses to interpret "representation" in the way Morrison does, then there still remains a gap between good performance "in experimental, engineering and other kinds of interventionist contexts" and "giving useful information".' (1, p. 67) The gap is not made any smaller by insisting that models are 'inherently intended for specific phenomena' (19, p. 75), and that models are superior to theories because 'they provide the kinds of details about specific mechanisms that allow us to intervene in the world' (14, p. 83).

If it is indeed the case that models derive their justification exclusively from instrumental success in specific empirical phenomena, then what is needed is a measure for empirical success, which typically will hinge on comparison of measurements with the model's predictions, both at the quantitative level of numerical results and at the qualitative level of system behaviour. An uninterpreted model does not in and of itself, without numerical evaluation, deliver quantitative or qualitative predictions about specific empirical phenomena. This is why it has been said that, at least across much of the so-called 'hard' sciences, 'the proper object of epistemic evaluation is a model in conjunction with a numerical method' (7, p. 743). Scientifically important questions of accuracy and prediction are not exhausted by a philosophical analysis of whether or not a model stands in a representational relationship to certain aspects of reality. In addition to an epistemology focused on representation, which has long been at the heart of the philosophical debate about models, what is needed is an 'epistemology of results', as it were. It is at this level that simulation gains significance: often, especially in the case of complex models, it is via the use of simulation techniques that specific numerical results and predictions are being derived from models. Somewhat similar to the way observation and measurement techniques furnish *empirical* data, simulation techniques generate specific instances of *simulated* data. Lest it be blind to this analogy, the philosophy of models, with its emphasis on representation, needs to be complemented by a philosophy of simulation, which takes due account of the non-trivial nature of generating results from models. As Eric Winsberg puts it, 'we need an epistemology of simulation because simulation modeling is a set of scientific techniques that produces results.' (21, p. 275)

3 Mathematical and Computational Models

Mathematical models can take different forms and fulfill different purposes. They may be limiting cases of a more fundamental, analytically intractable theory, for example in the case of modelling planetary orbits as if planets were independent mass-points revolving around an infinitely massive sun. Sometimes, models connect different theoretical domains, as is the case in hydrodynamics, where Prandtl's boundary layer model interpolates between the frictionless 'classical' domain and the Navier-Stokes domain of viscous flows. (See ref. (15).) In both cases, models allow for good quantitative predictions despite the intractability of the full theory. Even where a fundamental theory is lacking, mathematical models may be constructed, for example by fitting certain dynamical equations to empirically observed causal regularities (as in population cycles of predator-prey systems in ecology) or by analysing statistical correlations (as in models of stock-market behaviour).

It is in comparison with this diversity of examples of scientific models, I want to suggest, that several characteristic features of *mathematical models* can be singled out. The first such feature concerns the medium of expression, which for mathematical models is, naturally, the formulaic language of mathematics. It would, however, be misguided to simply regard a model as a set of (uninterpreted) mathematical equations, theorems and definitions, as this would deprive models of their representational potential: a set of equations cannot properly be said to 'model' anything, neither a specific phenomenon nor a class of phenomena, unless one gives some of the variables an interpretation that connects them with (some aspects of) observable phenomena. After all, one of the key motivations for constructing a model, at least in cases where a 'full' theory is presumed to hold 'in principle', is the recognition that theories are about abstract objects (e.g., 'mass points') rather than real objects (e.g., planets).[1]

[1] Cf. Ronald Giere, who argues that there are good reasons to regard 'Newton's laws as *defining* idealized abstract objects rather than as *describing* real objects' (8, p. 52)

Irrespective of one's stance towards the dispute over the primacy of fundamental theory, it is important to acknowledge that mathematical models cannot merely be uninterpreted mathematical equations if they are to function as mediators of any sort; that is, if they are to *model* a case that, for whatever reason, cannot be calculated or described in terms of theoretical first principles.

The fact that mathematical models, like other kinds of models, require background assumptions for their interpretation, of course, does not rule out that in each case there may be a core set of mathematical relationships that model *users* regard as definitive of the mathematical model in question. In fact, where these mathematical features are not merely 'inherited' from an underlying fundamental theory, they may provide a mathematical model with precisely the autonomy and independence (from theory and data) that its role as mediator requires. An extreme example of independence from theory and data, though perhaps not one that is representative of the majority of mathematical models used in the natural and social sciences, can be found in *numerical modelling*, where this involves fitting a set of – sometimes quite arbitrary – equations to empirical data. In cases with considerable uncertainty about the causal processes and dynamics laws governing a system (e.g., in analyses of the stock market), such mathematically informed 'curve-fitting' may be the theoretician's last resort. It does, however, lie at the extreme end of possible ways of constructing mathematical models and, importantly, differs radically from simulation. Whereas 'curve-fitting' typically accommodates existing (past) data to an, often crude, mathematical model, simulation is essentially about the generation of new 'data-like' material – that is, of *simulated data* that were not antecedently available, neither via empirical observation nor via theoretical derivation.

While it may be true that, as Giere puts it, '[m]uch mathematical modeling proceeds in the absence of general principles to be used in constructing models' (8, p. 52), there are good terminological reasons to speak of a mathematical model *of* a phenomenon (or a class of phenomena) only if the mathematics employed (i.e., the *kind* of mathematical techniques and concepts) is in some way sensitive to the *kind* of phenomenon in question. For example, while it may be possible, if only retrospectively, to approximate the stochastic trajectory of a Brownian particle by a highly complex deterministic function, for example a Fourier series of perfectly periodic functions, this would hardly count as a good mathematical model: There is something about the phenomenon, namely its stochasticity, that would not be adequately reflected by a set of deterministic equations; such a set of equations would quite simply not be a mathematical model *of Brownian motion*.

In addition to the requirement that the core mathematical techniques and concepts be sensitive to the kind of phenomenon that is being modelled, a further condition can be imposed on what should count as a *mathematical* model. Loosely speaking, the mathematics of the model should do some work in integrating the model's various other elements; after all, it follows from the discussion in the previous section that, for a mathematical construct to count as a *model* of a phenomenon or process, it must extend beyond its formal, theoretical, or mathematical representation as a set of uninterpreted equations.

Extra-theoretical considerations as well as background assumptions that do not lend themselves to formalisation must all be in place for a model to be a tool of scientific inquiry. A bare mathematical structure alone does not lend itself to application to individual cases. The perhaps vague demand that the mathematical aspects of a model should contribute to the integration of all, or at least *a wide range*, of the model's elements, can be given a concrete interpretation by way of example. If, say, a mathematical model employs the calculus of partial differential equations, then it should also indicate which (classes of) initial and boundary conditions need to be distinguished. Through specifying dynamic equations and their initial and boundary conditions, mathematical models can efficiently subsume different domains under the same basic structure. As an example consider Prandtl's boundary-layer model of fluid dynamics, which in this way succeeds in integrating not only different spatial domains (the boundary layer surrounding an object, and the infinite flow into which it is immersed), but also different domains of dynamic behaviour (laminar versus turbulent flow), as well as various background assumptions (Bernoulli's 'no-slip' condition, Helmholtz's principles etc.). Michael Heidelberger, in his detailed study of the development of Prandtl's model, attributes the model's success to precisely this capacity of mathematical models to integrate different elements: '[I]f unification is taken to mean a close relationship among the elements used – which one could call structural coherence – then "unification" would indeed be the right expression to characterize Prandtl's advance over the rational mathematicians and especially over his predecessor Helmholtz.' (10, p. 58) Other authors have referred to this capacity of mathematical models to successfully integrate different elements, or different aspects of the same phenomenon, as 'mathematical moulding':

> Mathematical moulding is shaping the ingredients in such a mathematical form that integration is possible, and contains two dominant elements. The first element is moulding the ingredient of mathematical formalism in such a way that it allows the other elements to be integrated. The second element is calibration, the choice of the parameter values, again for the purpose of integrating all the ingredients. (3, p. 90)

Calibration is essential to the function and functioning of models. However, as will be argued in the next section, it need not be understood narrowly as fixing the parameter values of a given model; rather, calibration may also take place across different models, by inquiring into their quantitative and qualitative behaviour as well as into the non-empirical relationships that hold between them.

Before moving on to a concrete class of mathematical models – models of physical systems that consist of many interacting particles – which illustrate the complex interplay between modelling and simulation, as well as certain general features of mathematical models, it is useful to consider an intermediary stage between mathematical models (in our sense) and simulations, namely *computational* models. In discussions of artificial intelligence, network design and the theory of computation, where the term 'computational model' appears to originate, it appears to have a rather more specialised meaning than in most scientific contexts. It would be wrong, however, to assimilate its meaning in science entirely to that of the term 'computer simulation', as indeed other authors have pointed out in various contexts. (5), (21), (10) Computational models are typically implemented in the form of an algorithm, either on a computer or on a network of computers. Their main structural and computational features are determined by such factors as network topology, numerical methods and algorithms used, computing power etc. In this regard, they differ from mathematical models, which are typically represented in an analytically closed form by a set of equations and whose structural characteristics are determined by mathematical constraints, not by constraints of realising a technological implementation. While computational models are often a crucial step in the actual implementation of computer simulations, it makes sense not to conflate them with simulations either. As R.I.G. Hughes has urged, one ought to distinguish between 'the use of computer techniques to perform calculations, on the one hand, and computer simulation, on the other'. (11, p. 128) Since computational models may be used both for mere 'number-crunching' and for 'genuine' simulation, it would be unwise to attempt to assimilate them to the latter.

4 Rigorous Results as Benchmarks for Simulations

The present section aims to apply the general framework outlined above to a particular class of mathematical models intended to describe and explain the physical behaviour of systems that consist of a large number of interacting particles. Such models, usually characterised by a specific Hamiltonian (energy operator), are frequently employed in condensed matter physics in order to account for phenomena such as magnetism, superconductivity, and phase transitions. Many-body models are particularly suitable as an example in the present context, since they form a class of models that, on the one hand, picks out a wide, yet well-defined range of physical phenomena as their explananda and, on the other hand, can be characterised mathematically by a narrow range of representational techniques (e.g., the formalism of second quantization). Many-body systems are also among the systems most widely studied using computer simulation, and it is the use of certain mathematical features of many-body models as benchmarks for the simulation of many-body systems, which will serve as a tool by which to analyse the interplay between models and simulations for the case of mathematical models more generally.

Recall the idea of 'mathematical moulding' mentioned in the previous section: namely, the capacity of mathematical models to integrate, through the use of certain mathematical techniques, diverse elements – some deriving from fundamental theory, others of non-theoretical origin – and subsume them under one mathematical structure. While this capacity is essential for the applicability of mathematical models to specific scientific problems, it would be quite misleading to regard the mathematical features of a model as merely *auxiliary.* The mathematics of a model does not merely serve the 'sanitary' purpose of integrating already existing elements into a coherent formal structure; it also contributes new elements. By virtue of their mathematical structure, mathematical models possess features and characteristics that extend beyond their function as representations of physical systems. Importantly, they can stand in a formal relation with other mathematical models, even when these are models of different physical systems. Thus, a mathematical model may contribute new elements to the theoretical description of the physical system, or class of systems, under consideration – elements which are not themselves part of the fundamental theory (or, as it were, cannot be 'read off' from it) but which may, in turn, take on an interpretative or otherwise explanatorily valuable role.

One important class of examples of such newly contributed elements are *rigorous results and relations* in statistical physics and many-body physics. Over the years, these have attracted considerable attention and have even given rise to a special branch of theoretical physics which concerns itself with rigorous results. (For a summary of some groundbreaking earlier developments, see Baxter 1982 and Griffiths 1972; for a philosophical case study see Gelfert 2005.) The term 'rigorous result' calls for some clarification. What makes a result 'rigorous' is not the qualitative or numerical accuracy of a theory or model. In fact, the kind of 'result' in question will often have no immediate connection with the empirical phenomenon (or class of phenomena) the model or theory is supposed to explain. (In this regard, the derivation of rigorous results is unlike, say, the simulation-aided generation of [data-like] results, philosophical analysis of which I urged towards the end of section 2.) Rather, it concerns an exact *mathematical* relationship between certain mathematical variables, or certain structural components, of the mathematical model, which may or may not reflect an empirical feature of the system that is being modelled. Examples of rigorous results include, but are not limited to, conditions on the asymptotic behaviour in certain limiting cases (some of which may be 'unphysical' in the sense that they do not correspond to actual physical scenarios – such as the limit of 'infinitely strong' interaction among particles in a system), symmetry requirements for certain mathematical elements of a model, 'impossibility theorems' that rule out certain kinds of macroscopic or dynamic behaviour of a model, and so forth.

The 'active' contribution of the model – that is, its contributing *new elements* rather than merely *integrating* theoretical and experimental (as well as further, external) elements – is not only relevant to interpretative issues, but also has direct consequences for assessing the techniques used to evaluate the model in specific circumstances, either by computing observable quantities or by simulating possible scenarios using a range of techniques. This is particularly salient in the case of the rigorous results mentioned in the preceding paragraph. Rigorous results are exact results that are true *of a model* (or a class of models) rather than of a theory. They often take the form either of exact relations holding between two or more quantities, or of lower and upper bounds to certain observables. If, for example in a model of a magnetic phase transition, the order parameter in question is the magnetization, then rigorous results – within a given model – may obtain, dictating the maximum (or minimum) value of the magnetization or the magnetic susceptibility. Quite often, rigorous results and relations provide a partial mapping of a model's mathematical structure onto relationships between observables. By checking the results of computer simulations against those relationships, one can then hope to find out whether a given simulation technique respects the model's fundamental features.

The partial independence of rigorous results from fundamental theory, and the fact that they are model-specific, makes them interesting 'benchmarks' for the numerical and analytical techniques of calculating observable quantities from the model. R.I.G. Hughes notes this, albeit only in passing, in his case study of one of the first computer simulations of the Ising model: 'In this way the verisimilitude of the simulation could be checked by comparing the performance of the machine against the exactly known behaviour of the Ising model.' (11, p. 123) The significance of 'benchmarks' for the purposes of simulations can hardly be overestimated. As Winsberg emphasises, simulations are often performed to investigate systems for which data are sparse; hence, 'comparison with real data can never be the autonomous criterion by which simulation results can be judged' (21, p. 287). If empirical data are not available, other reliable means of calibration must be found as a substitute. This is where rigorous results play an important role, and indeed may be crucial to the assessment of a simulation's success, given that '[t]he first criterion that a simulation must meet is to be able to reproduce known analytical results' (21, p. 288). Rigorous results thus can be seen to play an essential role in the verification of a simulation, where *verification* 'is taken to mean the testing of the model in relation to existing analytical solutions [...] as a benchmark' (as opposed to a simulation's *validation* against empirical data). (10, p. 59)

5 Cross-Model Justification and the Coherence of Simulational Practice

Practices of modelling and simulating physical systems raise a number of justificational questions. Are the methods that are being used reliable? Does the outcome successfully describe reality? Do models and simulations enhance our understanding of the phenomena that are being studied? As discussed above, such questions have typically been discussed in terms of whether or not the model in question is a faithful representation of the physical system. However, it is by no means obvious how, in practice, the representational relationship between a model and reality could be assessed globally. At best, one can hope to probe this relationship locally and test the model's performance in specific circumstances. The mediator view of models argues that it is a model's performance in specific 'interventionist contexts', in connection with 'specific phenomena', which is the main source of justification and determines the model's validity. On this account, the specific outcomes in different instances of employing a model determine its instrumental value which, in turn, is considered a measure of the model's justification. However, in the present section I want to suggest that such a purely 'outcome-based' perspective does not exhaust the range of actual sources of justification in our modelling and simulation practices.

Before turning to an entirely different source of justification, it should be noted that the outcome-based approach is not limited to comparison with empirical data. As the more and more widespread use of simulation techniques suggests, the data-like results of computer simulations can take on a similar role for the purposes of validation. From a purely descriptive perspective, it is by no means clear that the main activity of researchers is to assess the model's performance in experimental or other empirical contexts. An at least equal amount of work goes into comparing and calibrating different methods of numerical evaluation against each other. That is, the calibration often takes place not between models and empirical data, but amongst different methods of evaluation, irrespective of *their* empirical accuracy. This can be made particularly salient in the case of the many-body systems referred to earlier. Even in cases where quasi-exact numerical results are obtainable for physical observables (for example by Quantum Monte Carlo simulations), these will often be compared not to empirical data but instead to other relations derived at by other numerical methods. When it comes to the use of many-body models in solid-state and condensed matter physics, it is not uncommon to come across whole papers on, say, the problem of 'magnetism in the Hubbard model', which do not make a single reference to empirical data. (As an example, see Tusch, Szczech, Logan 1996.) Rather than adjust the parameters of the model to see whether the behaviour of a specific physical system can be modelled with empirical accuracy, the parameters will be held fixed to allow for better comparison of the different evaluative techniques with one another, often singling out one set of results (e.g., those calculated by Monte Carlo simulations) as authoritative.

However, beyond the narrow focus on empirical or simulated outcomes, there is a quite different source of justification, based on the rigorous results and relations discussed in the previous section. The results of models, as well as the application of simulation techniques, may be vindicated not only by referring to empirical performance, but also by exploiting certain rigorous relations between different mathematical models, especially where these take the form of (mathematically exact) mappings of one model on to another. Such mappings may connect different mathematical models in quite unexpected ways, thereby also allowing for cross-checks between evaluative methods and simulation techniques that were originally intended for very different domains. Such connections can neither be readily deduced from fundamental theory, since the rigorous results do not hold *generally* but only for certain mathematical models, which themselves are, as it were, *permitted* but not *entailed* by theory; nor can rigorous results be justifiably inferred from empirical data, as they may concern features of the model that lack an empirical interpretation. As an example consider again the theory of magnetic phase transitions, which aims to explain magnetism both in systems with fixed spins and in systems with itinerant electrons. The physical interpretations of the two cases are very different: In the former case, spins – the 'elementary magnets', so to speak – are spatially located at atoms in a crystal lattice, whereas in the latter case they are associated with freely moving, delocalized electrons. Yet mathematically, the two models are intimately related. For example, it can been shown rigorously (e.g., ref. (6)) that, under certain conditions (at half filling – i.e., when half of the quantum states in the conduction band are occupied – and in the strong-coupling limit, when the parameter representing the relative strength of the electron-electron interaction goes to infinity) the Hubbard model can be mapped on to one version (namely, the spin-1/2 antiferromagnetic) Heisenberg model. Under the specified conditions, the two models, despite their different physical interpretations, are *de facto* isomorphic and display the same mathematical behaviour. Of course, the Hubbard model with *infinitely strong* electron-electron interaction cannot claim to describe an actual physical system, where the interaction is necessarily finite, but to the extent that various mathematical und numerical techniques can nonetheless be applied in the strong-coupling limit, it provides a test for the adequacy of the Hubbard model by comparing it with the numerically and analytically more accessible antiferromagnetic Heisenberg model. In light of the fact that the conditions under which the mapping holds are empirically unattainable, it would be quite meaningless to ask for an experimental validation of these results, or for their instrumental usefulness in intervening in actual physical systems.

It is important to realise that the mechanism by which rigorous results and relations confer justification is quite different from that of outcome-based numerical comparison. The reasoning behind using rigorous relations as benchmarks is not that numerical conformity is a sign of a model's being a faithful representation of the physical system in question; rather, the fact that a model or simulation, *in conjunction with* a method of numerical evaluation, respects rigorous results and relations is regarded as an indication that our models and simulations are employed in a consistent and mutually supportive way. Rather than conferring justification for isolated predictions of a given model in a specific empirical context, this provides an internal vindication of our practices of modelling, numerical evaluation, and simulation across a range of contexts. It also allows for the application of well-established techniques to new domains of inquiry. For example, in those parameter regions where one model can be mapped on to another, as in the case of the Hubbard and Heisenberg models, techniques that have proved useful for one model may also be employed for the other model. An increase in reliability in the case of one model (as, for example, indicated by a greater overall numerical stability of the methods used to evaluate the model), may well be interpreted as conferring additional justification also to results derived for the other model. Unlike comparisons with empirical data, rigorous results and relations may thus confer cross-model justification.

One might object that, while a good deal of preliminary testing and cross-checking of one's numerical and simulation techniques has to happen before the model's predictions can be compared with empirical data, nonetheless the latter is the ultimate goal. While this may be a consistent, if somewhat narrow interpretation of scientific practice, it should be noted that in many cases this activity of cross-checking and 'bench-marking' is what drives research and makes up the better part of it. At the very least, it must be acknowledged that some of the most heavily researched models typically are not being assessed by their performance in specific empirical contexts. In part, this is because many models never were intended for *specific* phenomena in the first place, but for a qualitative understanding of a range of physical systems. This is true of the Hubbard model, which has been studied in connection with an array of quite diverse physical phenomena, including spontaneous magnetism, electronic properties, high-temperature superconductivity, metal-insulator transitions and many others, and it is particularly obvious in the case of the Ising model, which, even though it has been discredited as an accurate model of magnetism, continues to be applied to problems ranging from soft condensed-matter physics to theoretical biology. In many areas of research, as R.I.G. Hughes points with respect to the physics of critical phenomena, 'a good model acts as an exemplar of a universality class, rather than as a faithful representation of any one of its members' (11, p. 115).

6 Conclusion

In this paper, I have argued for the relevance of mathematically rigorous results and relations to the actual scientific practice of simulating models. Rigorous results are genuinely new contributions of a model; they are neither entailed by theoretical 'first principles', nor can they be inferred from empirical data. To the extent that they help to coordinate and calibrate one's tool box of numerical techniques and approximations, rigorous results are internal to a model, or class of models, and quite independent of both fundamental theory and empirical data. As such, they illustrate the capacity of models to take on roles beyond both fundamental theory and performance in empirical and interventionist contexts. It is such rigorous results, I claim, which guide much of research by providing fixed points for modelling strategies and attempted refinements of evaluative techniques, whether by numerical means, analytical evaluation, or computer simulation. The example of the mapping of models on to each other, for example of the strongly-coupled Hubbard model at half-filling on to the Heisenberg model, is but one example of how rigorous relations can set the agenda for further research. Characteristically, rigorous results provide non-empirical constraints, which may serve as general 'benchmarks' for model construction and computer simulation; it is such constraints and benchmarks which guide the process of model refinement. The existence of rigorous results and relations makes salient that a model's justification need not derive exclusively from empirical considerations: since rigorous results generally are validated by a model's formal features as a mathematical object, they can relate different models in quite unexpected ways. This allows for the transfer of warrant from one model to another, even in cases where both represent different classes of physical systems. The resulting phenomenon of cross-model justification, so I want to suggest, is not merely another form of 'moulding' a mathematical model to concrete empirical situations;

rather, cross-model justification fulfills a normative function by giving cohesion and stability to actual practices of modelling and simulation.

References

[1] Daniela M. Bailer-Jones: "When scientific models represent", *International Studies in the Philosophy of Science*, 17 No. 1, 2003, pp. 59-74.

[2] Rodney J. Baxter: *Exactly Solved Models in Statistical Mechanics*, New York: Academic Press 1982.

[3] Marcel Boumans: "Built-in justification", in Margaret Morrison and Mary S. Morgan (eds.): *Models as Mediators. Perspectives on Natural and Social Science*, Cambridge: Cambridge UP 1999, pp. 68-96.

[4] Nancy Cartwright: *The Dappled World. A Study of the Boundaries of Science*, Cambridge: Cambridge UP 1999.

[5] David Chalmers: "A Computational Foundation for the Study of Cognition" (1993), online document at http://consc.net/papers/computation.html, accessed 6 June 2006.

[6] Florian Gebhard: *The Mott Metal-Insulator Transition: Models and Methods. (Springer Tracts in Modern Physics, Vol. 137)*, Berlin: Springer 1997.

[7] Axel Gelfert: "Mathematical Rigor in Phyiscs: Putting Exact Results in Their Place", *Philosophy of Science*, 72, 2005, pp. 735-748.

[8] Ronald N. Giere: "Using Models to Represent Reality", in Lorenzo Magnani, Nancy J. Nersessian and Paul Thagard (eds.): *Model-Based Reasoning in Scientific Discovery*, New York: Plenum Publishers 1999, pp. 41-57.

[9] Robert B. Griffiths: "Rigorous results and theorems", in Cyril Domb and Melville S. Green (eds.): *Phase Transitions and Critical Phenomena*, New York: Academic Press, 1972, pp. 8-109.

[10] Michael Heidelberger: "Applying Models in Fluid Dynamics", *International Studies in the Philosophy of Science*, 20, 2006, pp. 49-67.

[11] R. I. G. Hughes: "The Ising model, computer simulation, and universal physics", in Margaret Morrison and Mary S. Morgan (eds.): *Models as Mediators. Perspectives on Natural and Social Science*, Cambridge: Cambridge UP 1999, pp.97-145.

[12] Paul Humphreys: *Extending Ourselves*, Oxford: Oxford UP 2004.

[13] Ernst Ising: "Beitrag zur Theorie des Ferromagnetismus", *Zeitschrift für Physik*, 31, 1925, pp. 253-258.

[14] Margaret C. Morrison: "Modelling Nature: Between Physics and the Physical World", *Philosophia Naturalis*, 35, 1998, pp. 65-85.

[15] Margaret Morrison: "Models as autonomous agents", in Margaret Morrison and Mary S. Morgan (eds.): *Models as Mediators. Perspectives on Natural and Social Science*, Cambridge: Cambridge UP 1999, pp. 38-65.

[16] Margaret Morrison and Mary S. Morgan (eds.): *Models as Mediators. Perspectives on Natural and Social Science*, Cambridge: Cambridge UP 1999.

[17] Margaret Morrison and Mary S. Morgan: "Models as mediating instruments", in Margaret Morrison and Mary S. Morgan (eds.): *Models as Mediators. Perspectives on Natural and Social Science*, Cambridge: Cambridge UP 1999, pp. 10-37.

[18] Mark Steiner: "The Application of Mathematics to Natural Science", *The Journal of Philosophy*, 86 No. 9, 1989, pp. 449-480.

[19] Mauricio Suárez: "Theories, Models, and Representations", in Lorenzo Magnani, Nancy J. Nersessian and Paul Thagard (eds.): *Model-Based Reasoning in Scientific Discovery*, New York: Plenum Publishers 1999, pp. 75-83.

[20] Michael A. Tusch, Yolande H. Szczech, and David E. Logan: "Magnetism in the Hubbard model: An effective spin Hamiltonian approach", *Physical Review B*, 53 No. 9, 1996, 5505-5517.

[21] Eric Winsberg: "Sanctioning Models: The Epistemology of Simulation", *Science in Context* 12, 1999, pp. 275-292.

[22] Eric Winsberg: "Simulated Experiments: Methodology for a Virtual World", *Philosophy of Science* 70, 2003, pp. 105-125.