The Tyranny of Scales

Robert W. Batterman[*]
Department of Philosophy
University of Pittsburgh

June 20, 2011

Abstract

This paper examines a fundamental problem in applied mathematics. How can one model the behavior of materials that display radically different, dominant behaviors at different length scales. Although we have good models for material behaviors at small and large scales, it is often hard to relate these scale-based models to one another. Macroscale models represent the integrated effects of very subtle factors that are practically invisible at the smallest, atomic, scales. For this reason it has been notoriously difficult to model realistic materials with a simple bottom-up-from-the-atoms strategy. The widespread failure of that strategy forced physicists interested in overall macro-behavior of materials toward completely top-down modeling strategies familiar from traditional continuum mechanics. The problem of the "tyranny of scales" asks whether we can exploit our rather rich knowledge of intermediate micro- (or meso-) scale behaviors in a manner that would allow us to bridge between these two dominant methodologies. Macroscopic scale behaviors often fall into large common classes of behaviors such as the class of isotropic elastic solids, characterized by two phenomenological parameters—so-called elastic coefficients. Can we employ knowledge of lower scale behaviors to understand this universality—to determine the coefficients and to group the systems into classes exhibiting similar behavior?

0

Contents

1   Introduction

In this paper I will focus on a fundamental problem in applied mathematics. This is the problem of modeling across scales. Many systems, say a steel girder, manifest radically different, dominant behaviors at different length scales. At the scale of meters, we are interested in its bending properties, its buckling strength, etc. At the scale of nanometers or smaller, it is composed of many atoms, and features of interest include lattice properties, ionic bonding strengths, etc. To design advanced materials (such as certain kinds of steel) materials scientists must attempt to deal with physical phenomena across 10+ orders of magnitude in spatial scales. According to a recent (2006) NSF research report this "tyranny of scales" renders conventional modeling and simulation methods useless as they are typically tied to particular scales. [11, p. 29] "Confounding matters further, the principal physics governing events often changes with scale, so the models themselves must change in structure as the ramifications of events pass from one scale to another." [11, pp. 29–30] Thus, even though we often have good models for material behaviors at small and large scales, it is often hard to relate these scale-based models to each other. Macroscale models represent the integrated effects of very subtle factors that are practically invisible at the smallest, atomic, scales. For this reason it has been notoriously difficult to model realistic materials with a simple bottom-up-from-the-atoms strategy. The widespread failure of that strategy forced physicists interested in overall macro-behavior of materials toward completely top-down modeling strategies familiar from traditional continuum mechanics.

The problem of the "tyranny of scales" asks whether we can exploit our rather rich knowledge of intermediate micro- (or meso-) scale behaviors in a manner that would allow us to bridge between these two dominant methodologies. Macroscopic scale behaviors often fall into large common classes of behaviors such as the class of isotropic elastic solids, characterized by two phenomenological parameters—so-called elastic coefficients. Can we employ knowledge of lower scale behaviors to understand this universality—to determine the coefficients and to group the systems into classes exhibiting similar behavior? This is related to engineering concerns as well: Can we employ our smaller scale knowledge to better design systems for optimal macroscopic performance characteristics.

The great hope that has motivated a lot of recent research into so-called "homogenization theory," arises from a conviction that the "between-scales"

point of view developed by Kadanoff, Fisher, and Wilson in the renormalization group approach to critical phenomena in fluids and magnets, may very well be the proper methodological strategy to overcome the tyranny of scales. A number of philosophers have recently commented on the renormalization group theory, but they seem to have missed what is truly novel about the methodological perspective that the theory employs.

Philosophical discussions of the applicability of mathematics, in general, have not, in my opinion, paid sufficient attention to contemporary work on this problem of modeling across scales. In many instances, philosophers hold on to some sort of ultimate reductionist picture: Whatever the fundamental theory is at the smallest, basic scale, it will be sufficient *in principle* to tell us about the behavior of the systems at all scales. Continuum modeling on this view, represents an *idealization*—as Feynman has said, "a smoothed-out imitation of a really much more complicated microscopic world." [13, p. 12–12] Furthermore, the suggestion is that such models are *in principle* eliminable.

There is a puzzle however. Continuum model equations such as the Navier-Stokes equations of hydrodynamics or the equations for elastic solids work *despite* the fact that they completely (actually, *almost* completely— this is crucial to the discussion below) ignore small scale or atomistic details of various fluids. The recipe (I call it "Cauchy's recipe") by which we construct continuum models is *safe*: If we follow it we will most always be led to empirically adequate successful equations characterizing the behavior of systems at the macroscopic level. Why? What *explains* the safety of this recipe? Surely this requires an answer. Surely, the answer must have something to do with the physics of the modeled systems at smaller scales. If such an answer cannot be provided, we will be left with a kind of sceptical attitude expressed by Penelope Maddy and discussed below in section 3. The worry is that without such an answer, we should not expect anything like a unified conception of applied mathematics' use of continuum idealizations. If an answer is forthcoming, then we have to face the reductionist picture mentioned above. Will such an answer—an answer that explains the robustness and safety of employing continuum modeling—support the view that continuum models are mere conveniences, only pragmatically justified, given the powerful simplifications gained by replacing large but finite systems with infinite systems? As noted, many believe that a reductionist/eliminitivist picture is the correct one. I maintain that even if we can explain the safety and robustness of continuum modeling, this reductionist picture is mistaken.

It is a mistaken picture not only of how science works but also of how the world works. I suggest that much philosophical confusion about reduction, emergence, atomism, and antirealism follows from the absolute choice between bottom-up and top-down modeling that the tyranny of scales *apparently* forces upon us. As noted, recent work in homogenization theory is beginning to provide much more subtle descriptive and modeling strategies. This new work calls into question the stark dichotomy drawn by the "do it in a completely bottom-up fashion" folks and those who insist that top-down methods are to be preferred.

The next section discusses the proposal that the use of continuum idealizations present no particular justificatory worries at all. Recent philosophical literature has focused on the role of continuum limits in understanding various properties of phase transitions in physical systems such as fluids and magnets. Some authors, particularly Jeremy Butterfield [9] and John Norton [20], have expressed the view that there are no particularly pressing issues here: The use of infinite limits is perfectly straightforwardly justified by appeal to pragmatic considerations. I argue that this view is mistaken and that it misses an important difference in methodology between some uses of infinite limits and those used by renormalization group arguments and homogenization theory.

Section 3 discusses Maddy's concerns. In section 4, I present an interesting historical example involving nineteenth century attempts to derive the proper equations governing the behavior of elastic solids and fluids. A controversy raged throughout that century concerning the merits of starting from bottom-up atomic description of various bodies in trying to arrive at empirically adequate continuum equations. It turns out that the bottom-up advocates lost the debate. Correct equations apparently could only be achieved by eschewing all talk of atomic or molecular structure, advocating instead a top-down approach supplemented, importantly, with experimentally determined data. In section 5, I formulate the tyranny of scales as the problem, just mentioned, of trying to understand the connection between recipes for modeling at atomic scales (Euler's recipe) and Cauchy's recipe appropriate for continuum models. Finally, I present a general discussion of work on homogenization that provides at least the beginning of an answer to the safety question and to the problem of bridging scales between the atomic and the continuum. This research can be seen as allaying Maddy's sceptical worries about a unified applied mathematical methodology regarding the use of continuum idealizations of a certain kind.

2 Steel Beams, Scales, Scientific Method

Let's consider the steel girder in a bit more detail. In many engineering applications steel displays linear elasticity. This is to say that it obeys Hooke's Law—its strain is linearly proportional to its stress. One phenomenological parameter related to its stress/strain (i.e., stiffness) properties is Young's Modulus appearing in the equations of motion for solids as well as in equilibrium and variational equations. At scales of 1 meter to 10^{-3} meters, say, the steel girder appears to be almost completely homogeneous: Zooming in with a small microscope will reveal nothing that looks much different. So for behaviors that take place within this range of scales, the steel girder is well-modeled or represented by the Navier-Cauchy equations:

$$(\lambda + \mu)\nabla(\nabla \cdot \mathbf{u}) + \rho\nabla^2\mathbf{u} + \mathbf{f} = 0. \tag{1}$$

The parameters, λ and μ are the "Lamé" parameters and are related to Young's modulus.

Now jump from this large scale picture of the steel to its smallest *atomic* scale. Here the steel, for typical engineering purposes, is an alloy that contains iron and carbon. At this scale, the steel exhibits highly symmetric crystalline lattice structures. It looks nothing like the homogenous girder at the macroscales that exhibits no crystalline structure. Somehow between the lowest scale of symmetric crystals and the scale of meters or millimeters, the low level symmetric structures must disappear. But that suggests that properties of the steel at its most basic, atomic level cannot, by themselves, determine what is responsible for the properties of the steel at macroscales.

In fact, the story is remarkably complex. It involves appeal to various geometrical properties that appear at *microscales* intermediate between the atomic and the macro[1], as well as a number of other factors such as martensitic transformations.[2] The symmetry breaking is effected by a combination of point defects, line defects, slip dislocations, as well as higher dimensional wall defects that characterize interfacial surfaces. All of these contribute to

---

[1] I call these intermediate scales "microscales" and the structures at these scales "microstructures" following the practice in the literature, but it may be best to think of them as "mesoscopic".

[2] These latter are transformations that take place under cooling when a relatively high symmetry lattice such as one with cubic symmetry loses symmetry to become tetragonal. Some properties of steel girders therefore depend crucially on dynamical changes that take place at scales in between the atomic and the macroscopic. [22, pp. 547–548]

the *homogenization* of the steel we see and manipulate at the macroscale. And of course, in engineering contexts, the macro features (bending properties, for example) are the most important—we do not want our buildings or bridges to collapse.

2.1 Limits, Averages, Continuum Models

A much simpler case than steel involves trying to connect the finite statistical mechanical theory of a fluid at the atomic scale to the thermodynamic continuum theory at macro scales.[3] The relationship between statistical mechanics and thermodynamics has received a lot of attention in the recent philosophical literature. Many of these discussions concern the nature and potential emergence of phase transitions in the so-called thermodynamic limit. [9, 18, 7, 1] The debates revolve around use of the thermodynamic limit and its role in understanding certain features of thermodynamic systems.[4] It will be instructive to consider this case once again in a different context than has been explicitly discussed in the literature. This is the context in which we consider the general problem of upscaling from atomic to laboratory scales, as in the case of the steel girder discussed above. In doing this, I hope it will become clear that many of the recent philosophical discussions miss a crucial feature of the methodology of applying limits like the thermodynamic limit.

Let's begin with Jeremy Butterfield's discussion of the "Straightforward Justification" for the use of infinite limits in physical modeling.

> This Justification consists of two obvious, very general, broadly instrumentalistic, reasons for using a model that adopts the limit $N = \infty$: mathematical convenience, and empirical adequacy (up to a required accuracy). So it also applies to *many* other models that are almost never cited in philosophical discussions of emergence and reduction. In particular, it applies to the many classical continuum models of fluids and solids, that are obtained by taking a limit of a classical atomistic model as the number of atoms

---

[3]Though simpler than the case of understanding how atomic aspects of steel affect its phenomenological properites, this is, itself, a difficult problem for which a Nobel prize was awarded.

[4]This is the limit in which the number of particles N in a system approaches infinity in such a way that the density remains constant—the volume has to go to infinity at the same time as the number of particles.

N tends to infinity (in an appropriate way, e.g. keeping the mass density constant).[5] [9, p. 1080]

He continues by emphasizing two "themes" common to the use of many different infinite models:

> The first theme is abstraction from finitary effects. That is: the mathematical convenience and empirical adequacy of many such models arises, at least in part, by abstracting from such effects. Consider (a) how transient effects die out as time tends to infinity; and (b) how edge/boundary effects are absent in an infinitely large system.

> The second theme is that the mathematics of infinity is often much more convenient than the mathematics of the large finite. The paradigm example is of course the convenience of the calculus: it is usually much easier to manipulate a differentiable real function than some function on a large discrete subset of $\mathbb{R}$ that approximates it. [9, p. 1081]

The advantages of these themes are, according to Butterfield, twofold. First, it may be easier to know or determine the limit's value than the actual value primarily because of the removal of boundary and edge effects. Secondly, in many examples of continuum modeling we have a function defined over the finite collection of atoms or lattice sites that oscillates or fluctuates and so can take on many values. In order to employ the calculus we often need to "have each *value* of the function defined as a limit (namely, of values of another function)." [9, pp 1081–1082] Butterfield seems to have in mind the standard use of *averaging* over a "representative elementary volume" (REV)[6] and then taking limits $N \to \infty$, volume going to zero, so as to identify a continuum value for a property on the macroscale. In fact, he cites continuum models of solids and fluids as paradigm examples:

> For example, consider the mass density varying along a rod, or within a fluid. For an atomistic model of the rod or fluid, that

---

[5]Note that Butterfield explicity claims that the "Straightforward Justification" applies to many cases of classical continuum modeling of solids and fluids, including, I suppose, the case of the steel girder.

[6]I've taken this terminology from [14, p. 1].

postulates N atoms per unit volume, the average mass-density might be written as a function of both position $\mathbf{x}$ within the rod or fluid, and the side-length L of the volume L^3 centred on $\mathbf{x}$, over which the mass density is computed: $f(N, \mathbf{x}, L)$. Now the point is that for fixed N, this function is liable to be intractably sensitive to $\mathbf{x}$ and L. But by taking a continuum limit $N \to \infty$, with $L \to 0$ (and atomic masses going to zero appropriately so that quantities like density do not "blow up"), we can define a continuous, maybe even differentiable, mass-density function $\rho(x)$ as a function of position—and then enjoy all the convenience of the calculus.

So much by way of showing in general terms how the use of an infinite limit $N = \infty$ can be justified—but not mysterious! At this point, the general philosophical argument of this paper is complete! [9, p. 1082]

So for Butterfield most of the discussions concerning the role, and particularly, the justification of the use of the thermodynamic limit in the debates about phase transitions have generated a lot of hot air. The justification, on his view, for employing such limits in our modeling strategies is largely pragmatic—for the sake of convenience. In addition, there is, as he notes, the further concern that the use of such limits be empirically adequate—getting the phenomena right to within appropriate error bounds. Much of his discussion in [9] then concerns showing that the use of such limits can most always be shown to be empirically adequate in this sense. Unfortunately, I think that things are more subtle than the straightforward justification as outlined here admits. In fact, there are very good reasons to think that the use of the thermodynamic limit in the context of the renormalization group (RG) explanation of critical phenomena—one of the cases he highlights—fails to be justified by his own criteria for justification. It is a *different* methodology, one that does not allow for the sort of justificatory story just told. The straightforward story as described above cannot be told for the RG methodology for the simple reason that that story fails to be empirically adequate in those contexts.

In order to see this, I need to say a bit about what the RG argument aims to do. I will then give a simple example of why, one should, almost always, expect the story above involving averaging over a representative volume element (REV) to fail. In fact, the failure of this story is effectively the

motivation behind Wilson's development of the *distinct* RG methodology.

I've discussed the RG in several publications [2, 3, 7] and Butterfield in [9] and Butterfield and Bouatta in [10] present concise descriptions as well. For the purposes here, as noted earlier, I am going to present some of the details with a different emphasis than these other discussions have provided. In particular, I want to stress the role of the RG as part of a methodology for upscaling from a statistical theory to a hydrodynamic/continuum theory. In so doing, I follow a suggestion of David Nelson [19, pp. 3–4] who builds on a paper of Ken Wilson. [26] The suggestion is that entire phases of matter (not just critical phenomena) are to be understood as determined by a "fixed point" reflecting the fact that "universal physical laws [are] insensitive to microscopic details." [19, p. 3] Specifically, the idea is to understand how details of the atomic scale physics gets encoded (typically) into a few phenomenological parameters that appear in the continuum equations governing the macroscopic behavior of the materials. In a sense, these phenomenological parameters (like viscosity for a fluid, and Young's modulus for a solid) characterize the appropriate "fixed point" that defines the class of material exhibiting universal behavior despite potentially great differences in microscale physics.

Let us consider a ferromagnet modeled as a set of classical spins σ_i on a lattice—the Ising model. In this model, neighboring spins tend to align in the same direction (either up or down: $\sigma_i = \pm 1$). Further, we might include the effect of an external magnetic field B. Then the Hamiltonian for the Ising model is given by

$$H[\{\sigma_i\}] = -J \sum_{\langle i,j \rangle} \sigma_i \sigma_j + \mu B \sum_i \sigma_i,$$

with the first sum over nearest neighbor pairs of spins, μ is a magnetic moment. A positive value for the coupling constant J reflects the fact that neighboring spins will tend to be aligned, both up or both down.

For ferromagnets we can define an order parameter—a function of the net magnetization for the system—that exhibits a discontinuity or jump at the so-called critical or Curie temperature, T_c. Above T_c, in zero magnetic field, the spins are not correlated due to thermal fluctuations and so the net magnetization is zero. As the system cools down to the Curie temperature, there is singularity in the magnetization (defined as a function of the free energy). (See figure 1.) The magnetization exhibits power law behavior near
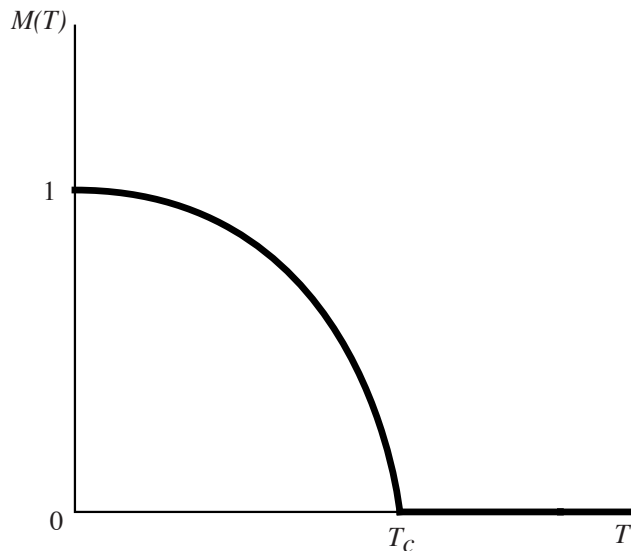
Figure 1: Spontaneous Magnetization at T_c

that singularity characterized by the relation

$$M \propto |t|^\beta,$$

where t is the reduced temperature $t = \frac{T - T_c}{T_c}$. It is a remarkable fact that physically quite distinct systems—magnets modeled by different Hamiltonians, and even fluids (whose order parameter is the difference between vapor and liquid densities in a container)—all exhibit the same power law scaling near their respective critical points: The number β is universal and characterizes the phenomenological behavior of a wide class of systems at and near criticality.[7]

The RG provides an explanation for this universal behavior; and in particular, it allows one to theoretically determine the value for the exponent β. For the 3-dimensional Ising model, that theoretical value is approximately .31. Experimentally determined values for a wide class of fluids and magnets are found in the range .31–.36. So-called "mean field" calculations predict a value of .5 for β. [26, p. 120] A major success of the RG was its ability to correct mean field theory and yield results in close agreement with experiment. In a mean field theory, the order parameter M is defined to be

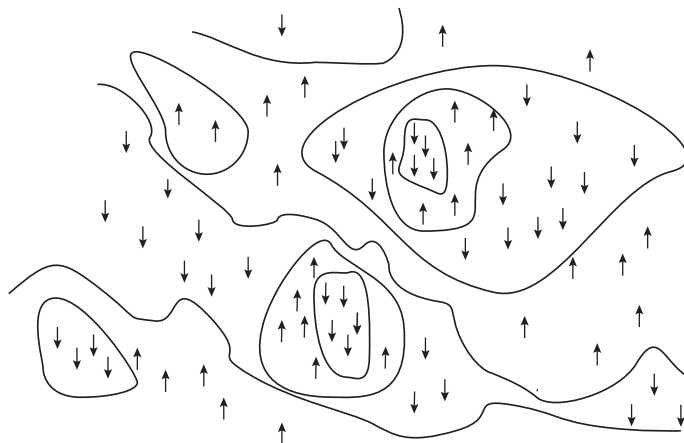---

[7]See [16, 2, 3, 7] for details.

Figure 2: Bubbles within Bubbles within Bubbles . . . [After Kadanoff [15, p. 11–12].]

the magnetic moment felt at a lattice site due to the average over all the spins on the lattice. This averaging ignores any large scale fluctuations that might (and, in fact, are) present in systems near their critical points. The RG corrects this by showing how to incorporate fluctuations at all length scales, from the atomic to the macro, that play a role in determining the macroscopic behavior (specifically the power law dependence—$M \propto |t|^{\beta}$) of the systems near criticality. In fact, near criticality the lattice system will contain "bubbles" (regions of correlated spins—all up or all down) of all sizes from the atomic to the system size. As Kadanoff notes, "[f]rom this picture we conclude that critical phenomena are connected with fluctuations over all length scales between ξ [essentially the system size] and the microscopic distance between particles." [15, p. 12]

So away from criticality, say below the critical temperature, the lattice systems will look pretty much *homogeneous*.[8] For a system with $T \ll T_c$ in figure 1 we would have relatively large correlated regions of spins pointing in the same direction. There might be only a few insignificantly small regions where spins are correlated in the opposite direction. This is what is responsible for there being a positive, nonzero, value for M at that temperature. Now suppose we were interested in describing a large system like this away from

---

[8]Systems above the critical temperature will also appear homogeneous as the spins will be uncorrelated, randomly pointing up and down.
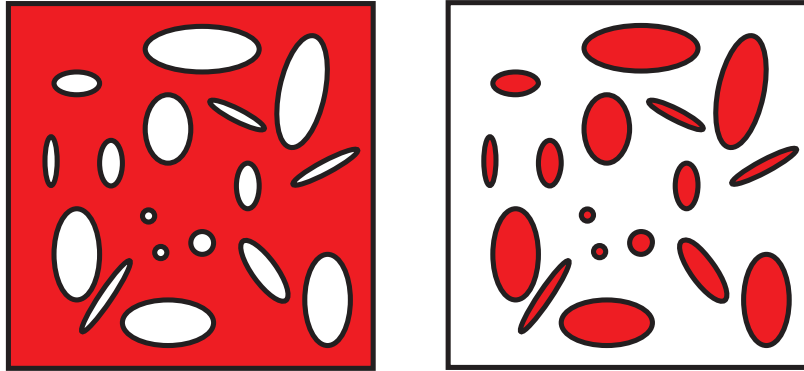
10

Figure 3: 50–50 Volume Mixture

criticality using the continuum limit as understood by Butterfield above. We would choose a representative elementary volume of radius L around a point $\mathbf{x}$ that is small with respect to the system size ξ, but still large enough to contain many spins. Next we would average the quantity $M(N, \mathbf{x}, L)$ over that volume and take the limits $N \to \infty$, $L \to 0$ so as to obtain the proper continuum value and so that we would be able to model the actually finite collection of spins using *convenient* continuum mathematics.

But near the critical temperature (near T_c) the system will look *heterogeneous*—exhibiting a complicated mixture of two distinct phases as in figure 2. Now we face a problem. In fact, it is the problem that effectively undermined the mean field approach to critical phenomena. The averaging method employing a representative elementary volume element misses what is most important. Somehow we will need to know how to weight the different phases as to their import for the macroscopic behavior of the system. In other words, were we to perform the REV averaging, all of the physics of the fluctuations responsible for the coexisting bubbles of up spins and bubbles of down spins would be ignored.

Here is a simple example to see why this methodology will often fail for heterogenous systems. [25, p.11] Consider a composite material consisting of *equal* volumes of two materials one of which is a good electrical conductor and one of which is not. A couple of possible configurations are shown in figure 3. Suppose that the dark, connected phase, is the good conductor. If we were to proceed using the REV recipe, then, because the volume fractions are the same, we would grossly underestimate the bulk conductivity of the ma-

11

terial in the left configuration and grossly underestimate its bulk insulating capacities in the right configuration. REV averaging treats only the volume fraction and completely misses microstructural details that are relevant to the bulk (macroscale) behavior of the material. In this simple example, the microstructural feature that is relevant is the topological connectedness of the one phase vs. the other—that is, the details about the boundaries between the two phases.[9] In more complicated situations, such as the steel girder with which we began, such microstructural features include mesoscale dislocations, defects of various kinds, and martensitic transformations.

The upshot, then, is that the straightforward justification for the use of infinite limits will miss exactly what is important for understanding what is going on for systems at and near criticality. There, they no longer appear homogeneous across a large range of scales. If we are to try to connect (and thereby extract) correct phenomenological macroscopic values for appropriate parameters (e.g., β) we need to consider structures that exist at scales greater than the fundamental/basic/atomic.

This is exactly what the RG does, by including in the calculations, the effects of fluctuations or equivalently, the fact that bubbles within bubbles of different phases appear near criticality. We need methods that tell us how to homogenize heterogeneous materials. In other words, to extract a continuum phenomenology, we need a methodology that enables us to upscale models of materials that are heterogeneous at small scales to those that are homogeneous at macroscales, as is evidenced by the fact that only a very small number of phenomenological parameters are required to characterize their continuum level behaviors. It appears, then, that the straightforward justification of the use of continuum limits needs to be reconsidered or replaced in those contexts where the materials of interest exhibit heterogeneous microstructures. Homogenization theory, and the RG employ a different kind of continuum limit than the straightforward justification that Butterfield's offers.

In section 6 I will say a bit more about the nature and generality of this different methodology. In the next section I'll connect the discussion of the use of continuum limits in modeling to some sceptical concerns raised by Penelope Maddy. Maddy, unlike Butterfield, is worried that the use of

---

[9]One might object that all one needs to do to save the REV methdology would be to properly weight the contribution of the different phases to the overall average. But this is not something that one can do *a priori* or through *ab initio* calculations appealing to details and properties of the individual atoms at the atomic scale.

continuum limits are much too diverse to be captured by a simple set of methodological directives and justifications. While Butterfield's pragmatic justication endorses the eliminability in principle of infinite limits, Maddy apparently wants to deny the latter.

3   Maddy's Concerns

In her interesting paper, "How Applied Mathematics Became Pure," Penelope Maddy presents a strong case for scepticism about whether there can be a unified methodology for applied mathematics. Her primary reasons for this scepticism reflect the reductionist/fundamentalist paradigm mentioned above, as well as the idea that a straightforward correspondence between mathematics and physics is rarely realized in applications. "There are rare cases where [the relation between mathematical structure and physical facts] is something like isomorphism—. . . but most of the time, the correspondence is something much more complex, and all too often, it is something we simply do not yet understand . . . ." [17, p. 33]

As noted, a standard picture of the hierarchy of the sciences tells us that atomistic theories are fundamental and that from them we should be able to account for all physical phenomena. But, Maddy notes, even a cursory glance at contemporary applied mathematics and physics reveals a picture quite unlike this. Solids, fluids, gases, etc. all exhibit macroscopic features that are well-described and accounted for by venerable continuum theories: Continuum mechanics, Fluid mechanics, and Thermodynamics. Yet, at the same time, these systems are composed of atoms, molecules, etc., and obey discrete equations of atomic and molecular dynamics.

We are faced with the following situation: Either the continuum models at large length scales are eliminable *in principle* or they are not. If they are eliminable, then any correspondence between those continuum models and systems is unnecessary. After all, in that case all reference to continuum scale properties will just be paraphrases of lower scale properties. If they are ineliminable there is a clear sense in which correspondence between system and model fails—continuum models speak of continua, not of discrete particles.

To argue that continuum models are eliminable requires demonstrating how such an elimination is to be carried out for even a remotely realistic case. Maddy is sceptical that this sort of story can be told. It seems to me that Maddy's worry is reasonable, given the "tyranny of scales" mentioned

in the introduction.[10] If the "principal physics" can change with scales, it is difficult to see how one will be able to carry out such an eliminativist story.

Maddy thinks that if one cannot tell such a story, then there is also little reason to believe that the various different applications of continuum models can be unified into a coherent methodology. After discussing a number of examples in which continuum models are employed she concludes:

> It appears unlikely that any general uniform account of how mathematics applies to the world could cover the wide variety of cases [she considers]. To take just a few of the examples [she has discussed in her article], the point particle model of an ideal gas works effectively for dilute gases because the occupied volume is negligibly small compared to the total volume; the "continuum hypothesis" works effectively in fluid dynamics because there is a suitable "plateau" between volumes too small to have stable temperature and volumes too large to have uniform temperature; . . . .[17, p. 34]

I believe that Maddy's scepticism about a unified methodology for applied mathematics hinges upon an assumption that only a tale about correspondence between model and system can yield a unified methodology.[11] Since that doesn't seem to be in the offing, we appear to be out of luck: Continuum models are neither eliminable nor are they capable of being understood in a unified way. As she says, "[g]iven the diversity of the considerations raised to delimit and defend these various mathematizations [the cases mentioned above], anything other than a patient case-by-case approach would appear singularly unpromising." [17, p. 35]

I believe there is a way out, and that an interesting story can be told—one that in effect will tell us how to bridge the scale separation between discrete molecular modeling and continuum models in a principled way. But, and this is the main point of the paper, being able to tell such a story in no way vindicates the elimitivist or reductionist attitudes of those who take continuum models and theories to be merely pragmatic conveniences.

---

[10]Note, again, that Butterfield does not seem to be moved by this sort of concern. He believes that continuum modeling is merely pragmatically justified as a shortcut for the much more complicated, but in principle story to be told from the fundamental atomic/small scale point of view.

[11]See [5] for a discussion of this kind of correspondence or mapping relation.

In the next section I consider an interesting historical example that should give pause to those who think continuum models are ultimately unnecessary. This is the story of deriving appropriate continuum equations for the behavior of elastic solids and gave rise to a controversy that lasted for most of the nineteenth century.

4 Bridging Across Scales: An Historical Controversy

Why are the Navier-Stokes equations named after Navier and Stokes? The answer is not as simple as "they both, independently, arrived at the same equation." In fact, there are differences between the equation Navier first came up with and that derived by Stokes. The differences relate to the assumptions that each employed in his derivation, but more importantly, these different assumptions actually lead to different equations. Furthermore, the difference between the equations was symptomatic of a controversy that lasted for most of the nineteenth century. [12, p. 86]

While the Navier-Stokes equation describes the behavior of a viscous fluid, the controversy has its roots in the derivation of equations for the behavior of an elastic solid. I intend to focus on the latter equations and only at the end make some remarks about the fluid equations.

The controversy concerned the number of material constants that were required to describe the behavior of elastic solids. According to Navier's equation, a single constant marked a material as isotropic elastic. According to Stokes and Green, two constants were required. For anisotropic elastic materials (where symmetries cannot be employed) the debate concerned whether the number of necessary constants was 15 or 21. This dispute between, respectively, "rari-constancy" theorists and "multi-constancy" theorists, depended upon whether one's approach to the elastic solid equations started from a hypothesis to the effect that solids are composed of interacting molecules, or from the hypothesis that solids are continuous.

Navier's derivation began from the hypothesis that the deformed state of an elastic body was to be understood in terms of *forces acting between individual particles or molecules that make up the body.* Under this assumption, he derived equations containing only one material constant ϵ.

Navier's equations for an elastic solid are as follows [12, p. 80]:

$$\epsilon \triangle u + 2\frac{\partial \Theta}{\partial x}\rho X \;=\; \rho\frac{\partial^2 u}{\partial t^2}, \tag{2}$$

$$\epsilon \triangle v + 2\frac{\partial \Theta}{\partial y}\rho Y \;=\; \rho\frac{\partial^2 v}{\partial t^2}, \tag{3}$$

$$\epsilon \triangle w + 2\frac{\partial \Theta}{\partial z}\rho Z \;=\; \rho\frac{\partial^2 w}{\partial t^2}. \tag{4}$$

Here ϵ, Navier's material constant, reflects the molecular forces that turn on when external forces are applied to the body. x, y, z are the coordinates representing the location of a *material point* in the body.[12] u, v, w are the displacement components in the directions x, y, z; X, Y, Z represent the external accelerations (forces) in the directions x, y, z; $\triangle = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$ is the Laplace operator; $\Theta = \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}$; and ρ is the material density.

Cauchy also derived an equation for isotropic elastic materials by starting from a molecular hypothesis similar to Navier's. However, his equation contains the correct number of material constants (two). It is instructive to write down Cauchy's equations and to discuss how, essentially, a *mistaken, inconsistent* derivational move on his part yielded a more accurate set of equations than Navier.

Cauchy's equations for an elastic solid are as follows [12, p. 81] (Compare with equation (1).):

$$(R+A)\triangle u + 2\frac{\partial \Theta}{\partial x}\rho X \;=\; \rho\frac{\partial^2 u}{\partial t^2}, \tag{5}$$

$$(R+A)\triangle v + 2\frac{\partial \Theta}{\partial y}\rho X \;=\; \rho\frac{\partial^2 v}{\partial t^2}, \tag{6}$$

$$(R+A)\triangle w + 2\frac{\partial \Theta}{\partial z}\rho X \;=\; \rho\frac{\partial^2 w}{\partial t^2}, . \tag{7}$$

R, A are the two material constants. Cauchy noted, explicitly, that when $A = 0$ his equations agree with Navier's when $R = \epsilon$. [12, p. 81] How

---

[12]Note that in continuum mechanics, generally, a material point or "material particle" is *not* an atom or molecule of the system, rather it is an imaginary region that is large enough to contain many atomic subscales (whether or not they really exist) and small enough relative to the scale of field variables characterizing the impressed forces. Of course, as noted, Navier's derivation *did* make reference to atoms.

did Cauchy arrive at a different equation than Navier, despite starting, essentially, from the same molecular assumptions about forces? He did so by assuming that, despite the fact he is operating under the molecular hypothesis, he can, in his derivation replace certain summations by integrations. In effect, he actually employs a *continuum condition* contradictory to his fundamental starting assumption.[13]

George Green, in 1839, published a study that arrived at the correct equations—essentially (5)–(7)—by completely eschewing the molecular hypothesis. He treated the entire body as composed of "two indefinitely extended media, the surface of junction when in equilibrium being a plane of infinite extent."[14] He also assumed that the material was not crystalline and, hence, isotropic. Then using a principle of the conservation of energy/work he derived, using variational principles of Lagrangian mechanics, his multi-constant equation.

Finally, following the discussion of Todhunter and Pearson [24], we note that Stokes' work supported the multi-constancy theory in that he was able to generalize his equations for the behavior of *viscous fluids* to the case of elastic solids by making no distinction between a viscous fluid and a solid undergoing permanent—plastic—deformation. "He in fact draws no line between a plastic solid and a viscous fluid. The formulae for the equilibrium of an isotropic plastic solid would thus be bi-constant. [24, p. 500.] This unification of continuum equations lends further support to the multi-constancy theory.

Of course one might wonder whether the debate between the rari-constancy and multi-constancy theorists could be decided empirically. Todhunter and Pearson examine a number of experimental "demonstrations" that supposedly support the multi-constancy point of view. But interestingly, they raise good reasons for not taking them to be decisive. They are responding to a rather flippant dismissal of rari-constancy theories by Thomson and Tait:

> The only condition that can be theoretically imposed upon these coefficients (the 21 of Green) is that they must not permit w (the work) to become negative for any values, positive or negative, of the strain components .... Under Properties of Matter we shall see that an untenable theory (Boscovich's) falsely worked out by

---

[13]See [24, p. 224, pp. 325–327.] for details. Note also, how this limiting assumption yields *different and correct* results in comparison to the finite atomistic hypotheses.

[14]Cited in [24, p. 495.]

mathematicians, has led to relations among the coefficients of elasticity which experiment has proved to be false.[15]

Todhunter and Pearson point out two things. First, the discrete/atomic hypothesis of Navier and Cauchy is not actually the point particle theory of Boscovich, and they find no mathematical mistakes in their derivations.[16] Nevertheless, there are affinities with Boscovich's atomistic theory. Second, and more importantly, the experimental evidence to which Thomson and Tait appeal, is suspect for interesting reasons.

In fact, Thomson and Tait appeal to properties of wires and plates to show the "utter worthlessness" of the rari-constancy theory. [24, p. 499]. Todhunter and Pearson, while no friends of rari-constancy theory, point out that wires and thin plates are far from the three dimensional isotropic bodies to which the rari-constancy theorists apply their results. "These bodies possess owing to their method of manufacture not only a cylindrical or planar system of elasticity, but extremely often an initial state of stress, both conditions which lead rari-constant theorists to bi-constant formulae." [24, p. 499]

These historical debates represent just the tip of the iceberg of the complexity surrounding both theoretical and experimental work on the behavior of the supposedly simpler, isotropic, cases of elastic solids. Nevertheless, the multi-constancy theory wins the day for appropriate classes of structures. And, derivations that start from atomic assumptions fail to arrive at the correct theory. It seems that here may very well be a case where a continuum point of view is actually required: Bottom up derivation from atomistic hypotheses about the nature of elastic solid bodies fails to yield correct equations governing the macroscopic behavior of those bodies. There are good reasons, already well understood by Green and Stokes, for eschewing such reductionist strategies.

de Boer reflects on the reasons for why this controversy lasted so long and was so heated:

> Why was so much time spent on molecular theory considerations, in particular, by the most outstanding mechanics specialists and mathematicians of the epoch? One of the reasons must have been

---

[15]Cited in [24, p. 498].

[16]Of course, Cauchy did illegitimately replace sums with integrals, but that mistake is not present in Navier's derivation.

18

the temptation of gaining the constitutive relation for isotropic and anisotropic elastic continua directly from pure mathematical studies and simple mechanical principles; It was only later realized that Hooke's *generalized law* is an assumption, and that the foundation of the linear relation had to be supported by experiments. [12, pp. 86–87]

The upshot of this discussion is reflected in de Boer's emphasis that the constitutive equations or special force laws (Hooke's law) are dependent, for their very form, on experimental results. So a simple dismissal of continuum theories as "in principle" eliminable and merely pragmatically justified, is mistaken. Of course, the phenomenological parameters, like Young's modulus (related to Navier's ϵ), must encode details about the actual atomistic structure of elastic solids. But it is naive, indeed, to think that one can, in any straightforward way *derive or deduce* from atomic facts what are the phenomenological parameters required for continuum model of a given material. It is probably even more naive to think that one will be able to *derive or deduce* from those atomic facts what are the actual values for those parameters for a given material.

This historical discussion, and the intense nineteenth century debate between the rari- and multi- constancy theorists apparently supports the view that there is some kind of fundamental incompatibility between small scale and continuum modeling practices. That is, it lends support to the *stark choice* one must apparently make between bottom-up and top-down modeling suggested by the tyranny of scales.

A modern, more nuanced, and better informed view challenges this consequence of the tyranny of scales and will be discussed in section 6. However, such a view will not, in my opinion, bring much comfort to those who believe the use of continuum models or idealizations are only pragmatically justified and in principle eliminable. A modern statement supporting this point of view can be found in [22]:

> . . . [M]any material properties depend upon more than just the identity of the particular atomic constituents that make up the material. . . . [M]icrostructural features such as point defects, dislocations, and grain boundaries can each alter the measured macroscopic 'properties' of a material. [22, pp. 5–8]

It is important to reiterate that, contrary to typical philosophical usage,

"microstructural features" here is not synonymous with "atomic features"! Defects, dislocations, etc. exist at higher scales.

The historical controversy outlined in this section serves to support Maddy's conclusions that continuum models appear to be necessary for applied mathematics and yet, somehow, remain detached from reality. In the next section I will further develop the stark dichotomy between bottom-up modeling and top-down modeling as a general philosophical problem arising between different recipes for applying mathematics to systems that exist across a wide range of scales.

5 Euler's and Cauchy's Recipes

5.1 Euler

Applied mathematical modeling begins with an attempt to write down an equation governing the system exhibiting the phenomenon of interest. In many situations, this aim is accomplished by starting with a general dynamical principle such as Newton's second law: $\mathbf{F} = \mathbf{ma}$. Unfortunately, this general principle tells us absolutely nothing about the material or body being investigated and, by itself, provides no model of the behavior of the system. Further data are required and these are supplied by so-called "special force laws" or "constitutive equations."

A recipe, due to Leonhard Euler, for finding an appropriate model for a system of particles proceeds as follows [27]:

1. Given the class of material (point particles, say), determine the kinds of special forces that act between them. Massive particles obey the constitutive gravitational force: $\mathbf{F}_G = G\frac{m_i m_j}{r_{ij}^2}$. Charged particles additionally will obey the Coulomb force law: $\mathbf{F}_E = k_e \frac{q_i q_j}{r_{ij}^2}$.

2. Choose Cartesian coordinates along which one decomposes the special forces.

3. Sum the forces acting on each particle along the appropriate axis.

4. Set the sum for each particle i equal to $m_i \frac{d^2 x}{dt^2}$ to yield the total force on the particle.

This yields a differential equation that we then employ (= try to solve) to further understand the behavior of our point particle system. Only rarely (for very few particles or for special symmetries) will this equation succumb to analytical evaluation. In many instances, further simplification employing mathematical strategies of variable reduction, averaging, etc. enable us to gain information about the behavior of interest.

5.2   Cauchy

As we saw in section **??**, Cauchy played a critical role in the derivation of equations for elastic solids. He derived what is now called the "Cauchy momentum equation," which is the continuum analogy of Newton's second law $\mathbf{F} = \mathbf{ma}$:

$$\rho \left(\frac{\partial \mathbf{v}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{v} \right) = \nabla \cdot \sigma + \mathbf{f}.^{17}$$

This is the general dynamical principle one employs to find continuum model equations. It is the starting point for what I will call "Cauchy's" recipe. As with Euler's recipe, one now needs special force laws or constitutive equations. For a solid body we ask whether its stress is linearly related to its strain: Does it obey Hooke's law for an elastic solid? For a fluid, we ask about its viscosity, its compressibility, and the nature of surface pressures acting on it. These constitutive equations then enable us to derive equations like those of Navier, Green, and Stokes that tell us how different materials will behave under various kinds of manipulations.

5.3   Controversy

A question of pressing concern, particularly given Maddy's worries, is why the Cauchy recipe should work at all. We've seen in the historical example, that it does, and in fact, we've seen that were we simply to employ an Euler type (point particle) recipe, we would not arrive at the correct results. In asking why the Cauchy recipe works on the macroscale, we are asking about the relationship between the dynamical models that track the behavior of individual atoms and molecules at the scale of nanometers and equations like those of Navier, Stokes, Cauchy, and Green that are applicable at the

---

[17]ρ is the density, σ is the stress tensor, $\mathbf{v}$ is the velocity vector field, $\mathbf{f}$ are body forces acting on the material—usually just gravity.

scale of millimeters. Put slightly differently, we would like an account of why it is *safe* to use the Cauchy momentum equation in the sense that it yields correct equations with the appropriate (few) parameters for broadly different classes of systems—from elastic solids to viscous fluids?

From the point of view of Cauchy's recipe, one derives the equations for elastic solids, or the Navier-Stokes equations, independently of *any* views about the molecular or atomic makeup of the medium. In the nineteenth century there were doubts about whether matter was atomistic.

To ask for an account of why it is *safe* to use the Cauchy recipe for constructing macroscale models is to ask for an account of the robustness of that methodology. The key physical fact is that the bulk behaviors of solids and fluids are *almost* completely insensitive to the actual nature of the physics at the smallest scale. The "almost" here is crucial. The atomic details that we don't know (and, hence, don't explicitly refer to) when we employ Cauchy's recipe are *encoded* in the small number of phenomenological parameters that appear in the resulting equations—Young's modulus, the viscosity, etc. So the answer to the safety question, will involve showing how to determine the "fixed points" characterizing broad classes of macroscopic materials— fixed points that are characterized by those phenomenological parameters. Recall the statement by Nelson cited above in section 2.1. In the context of critical phenomena and the determination of the critical exponent β, this upscaling or connection between the Euler recipe and the Cauchy recipe, is accomplished by the renormalization group. In that context, the idea of a critical point, and related singularities plays an important role. But Nelson's suggestion is that upscaling of this sort should be possible even for classes of systems *without* critical points. We would like to understand why, for instance, Young's modulus is the appropriate phenomenological parameter for classifying solids as linear elastic, despite rather severe differences in the atomic structure of members of that class. Finding answers to questions of this latter type, is the purview of so-called "homogenization" theory, of which one can profitably think the RG to be a special case.

In the next section, I will spend a bit more time on the RG explanation of the universality of critical behavior, filling in some gaps in the discussion in section 2.1. And, I will try to say something about general methodology of upscaling through the use of homogenization limits.

6   A Modern Resolution

To begin, consider a problem for a corporation that owns a lot of casinos. The CEO of the corporation needs to report to the board of trustees (or whomever) on the expected profits for the corporation. How is she to do it? Assuming (contrary to fact) that casino gaming is fair, she would present to the board a Gaussian or normal probability distribution showing the probabilities of various profits and losses, with standard deviations that would allow for statistical predictions as to expected profits and losses. The Gaussian distribution is a function characterized by two parameters—the mean μ and the variance σ^2. Where will the CEO get the values for the mean and variance? Most likely by empirically investigating the actual means and variances displayed over the past year by the various casinos in the corporation. How does she know that these two parameters—properties of collections of casinos offering different and varied kinds of games (roulette, poker, blackjack, slots, etc.)—are the correct ones with which to make the presentation to the board? In effect, by using the same reasoning applied in the RG to determine that the order parameter M scales as $|t|^\beta$ near criticality.[18] There we exploit the fact that near criticality the system is composed of collections of bubbles of up and down spins that all, collectively, contribute to the behavior of the system at the macroscale. We have, here a similar picture. See figure 4.

Furthermore, a similar picture is possible regarding the upscaling of our modeling of the behavior of the steel girder with which we started. Compare the two cases, figure 5, noting that here too only a very small number of phenomenological parameters are needed to model the continuum/macroscale behavior. (E is Young's modulus and I is the area moment of inertia of a cross-section of the girder.)

The general problem of justifying the use of Cauchy's recipe to determine the macroscopic equation models involves connecting a statistical/discrete theory in terms of atoms or lattice sites to a hydrodynamic or continuum

---

[18]For details see [6, 23]. These references stress similarities between RG methods and a certain understanding of the central limit theorem. The difference between averaging and homogenization is related to the difference between the law of large numbers and the central limit theorem: Averaging or first order perturbation theory "...can often be thought of as a form (or consequence) of the law of large numbers." Homogenization or second order perturbation theory ... "can often be thought of as a form (or consequence) of the central limit theorem." [21, pp. 6–7]

Gaussian distribution

casino scale lengths

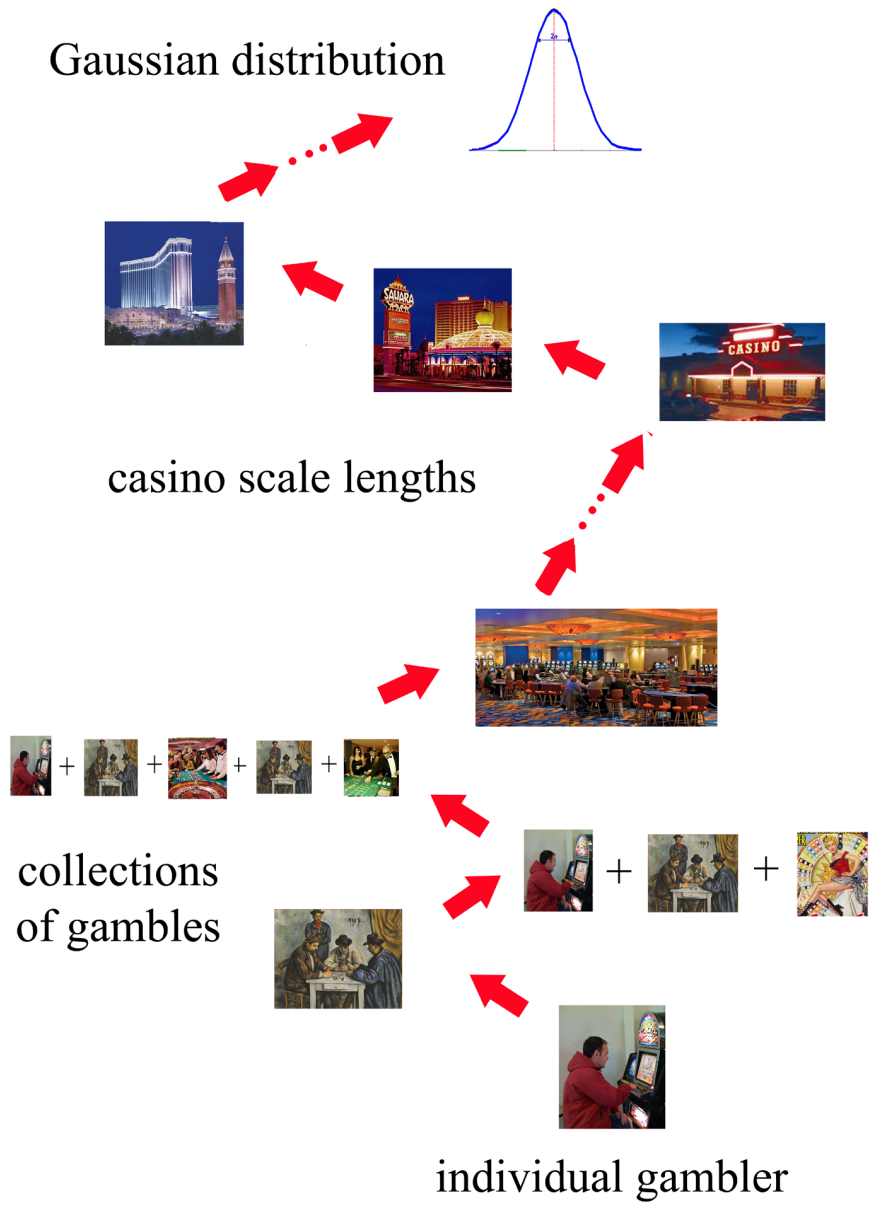collections
of gambles

individual gambler
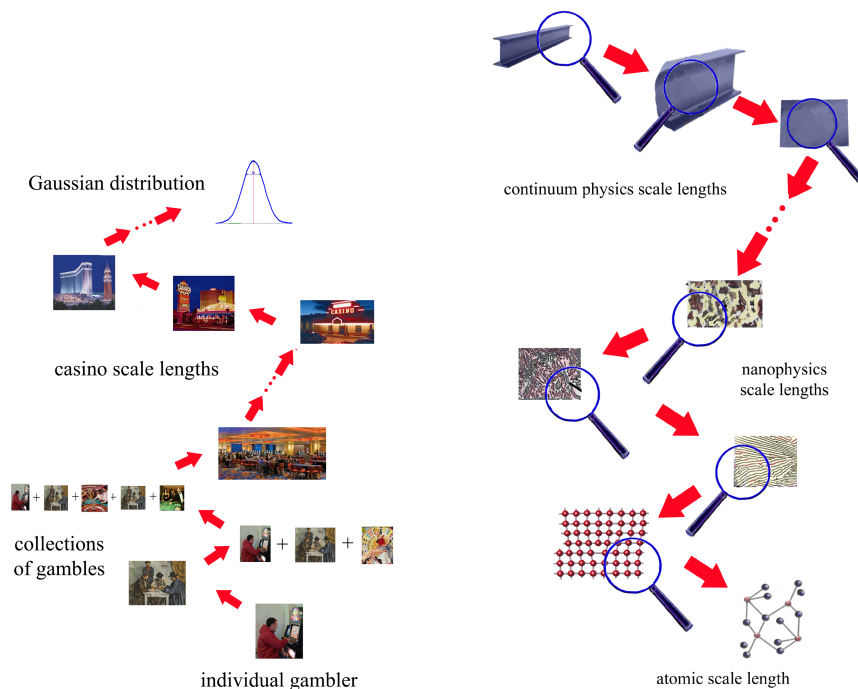
Figure 4: Gambles within Gambles within Gambles . . .

Figure 5: Gaussian and Steel—Few (Macro) Parameters: $[\mu, \sigma^2]$; $[E, I]$

theory. Much effort has been spent on this problem by applied mathematicians and materials scientists. And, as I mentioned above, the RG argument that effectively determines the continuum behavior of systems near criticality is a relatively simple example of this general homogenization program.

In hydrodynamics, for example Navier-Stokes theory, there appear density functions, $\rho(\mathbf{x})$, that are defined over a continuous variable $\mathbf{x}$. These functions exhibit no atomic structure at all. On the other hand, for a statistical theory, such as the Ising model of a ferromagnet, we have seen that one defines an order parameter (a magnetic density function) $M(\mathbf{x})$ that is the average magnetization in a volume surrounding $\mathbf{x}$ that contains many lattice sites or atoms. The radius of the volume, L, is intermediate between the lattice constant (or atomic spacing) and the correlation length ξ: $(a \ll L \ll \xi)$. As noted in section 2.1 this makes the order parameter depend upon the length L. [26, p. 123]

A crucial difference between the hydrodynamic (thermodynamic) theory and the statistical theory is that the free energy in the former is determined

25

using the *single* magnetization function $M(\mathbf{x})$. In statistical mechanics, on the other hand, the free energy is "a weighted average over all possible forms of the magnetization $M(\mathbf{x})$." [26, p. 123] This latter set of functions is parameterized by the volume radius L. On the statistical theory due originally to Landau, the free energy defined as a function of $M(\mathbf{x})$, takes the following form:

$$F = \int \left([\nabla M(\mathbf{x})]^2 + RM^2(\mathbf{x}) + UM^4(\mathbf{x}) - B(M(\mathbf{x})) \, d^3(\mathbf{x}), \quad (8)\right)$$

where R and U are (temperature dependent) constants and B is a (possibly absent) external magnetic field. [26, p. 122] This theory predicts the wrong value, $1/2$, for β–the critical exponent. The problem, as diagnosed by Wilson, is that while the Landau theory can accommodate fluctuations for lengths $\lambda < L$ in its definition of M as an average, it cannot accommodate fluctuations of lengths L or greater.

> A sure sign of trouble in the Landau theory would be the dependence of the constants R and U on L. That is, suppose one sets up a procedure for calculating R and U which involves statistically averaging over fluctuations with wavelengths $\lambda < L$. If one finds R and U depending on L, this is proof that long-wavelength fluctuations are important and Landau's theory must be modified. [26, p. 123]

The RG account enables one to exploit this L-dependence and eventually derive differential equations (RG) for R and U as functions of L that allow for the calculation of the exponent β in agreement with experiment. The key is to calculate and compare the free energy for different averaging sizes L and $L + \delta L$. One can proceed as follows[19]: Divide $M(x)$ in the volume element into two parts:

$$M(\mathbf{x}) = M_H(\mathbf{x}) + mM_{fl}(\mathbf{x}). \quad (9)$$

M_H is a hydrodynamic part with wavelengths of order ξ and M_{fl} is a fluctuating part with wavelength between L and $L + \delta L$. The former will be effectively constant over the volume.

---

[19]Details in [26, pp. 125–127]

By performing a single integral over m—the scale factor in (9)—we get an interative expression for the free energy for the averaging size $L + \delta L$, $F_{L+\delta L}$, in terms of the free energy for the averaging size L:

$$e^{-F_{L+\delta L}} = \int_{-\infty}^{\infty} e^{-F_L} dm. \qquad (10)$$

In effect, one finds a step by step way to include all the fluctuations—all the physics—that play a role near criticality. One moves from a statistical theory defined over finite N and dependent on L to a hydrodynamic theory of the continuum behavior at criticality. "Including all of the physics," means that the geometric structure of the bubbles within bubbles picture gets preserved and exploited as one upscales from the finite discrete atomistic account to the continuum model at the scale of ξ—the size of the system. That is exactly the structure that is wiped out by the standard REV averaging, and it is for that reason that Landau's mean field theory failed.

6.1   Homogenization

Continuum modeling is concerned with the effective properties of materials that, in many instances, are microstructrually heterogeneous. These microstructures, as noted, are not always to be identified with atomic or lowest scale "fundamental" properties of materials. Simple REV averaging techniques often assume something like that, but in general the effective, phenomenological properties of materials are not simple mixtures of volume fractions of different composite phases or materials. Many times the microstructural features are geometric or topological including (in addition to volume fractions) "surface areas of interfaces, orientations, sizes, shapes, spatial distributions of the phase domains; connectivity of the phases; etc." [25, p. 12] In trying to bridge the scales between the atomic domain and that of the macroscale, one needs to connect rapidly varying local functions of the different phases to differential equations characterizing the system at much larger scales. Homogenization theory accomplishes this by taking limits in which the local length (small length scale) of the heterogeneities approaches zero in a way that preserves (and incorporates) the topological and geometric features of the microstructures.

Most simply, and abstractly, homogenization theory considers systems at two scales: ξ, a macroscopic scale characterizing the system size, and a microscopic scale, a, associated with the microscale heterogenieties. There
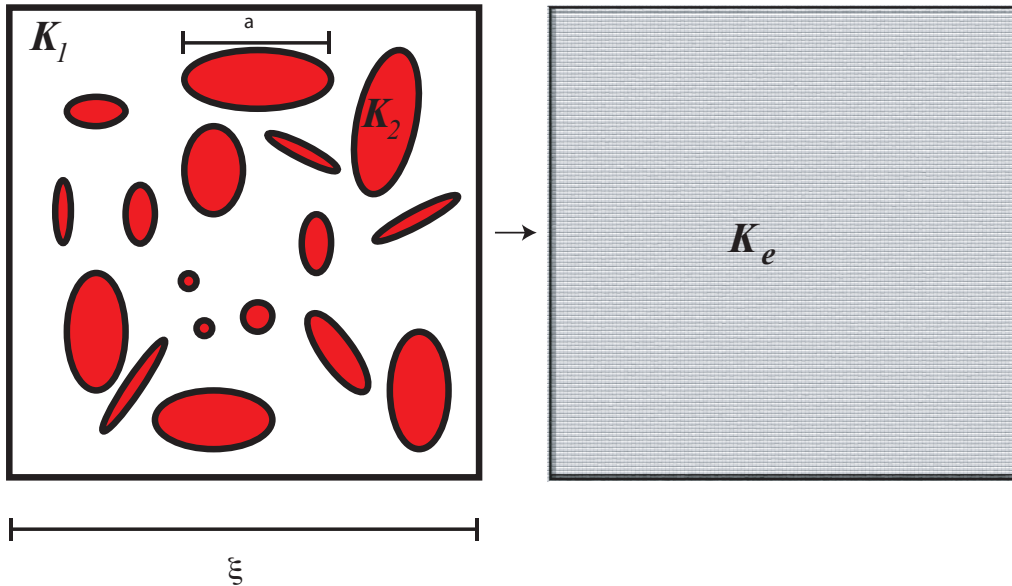
Figure 6: Homogenization Limit [After Torquato [25, pp. 2, 305–306].]

may also be applied external fields that operate at yet a third scale Λ. If the microscale, a, is comparable with either ξ or Λ, then the modeler is stuck trying to solve equations at that smallest scale. However, as is often the case, if $a \ll \Lambda \ll \xi$, then one can introduce a parameter

$$\epsilon = \frac{a}{\xi}$$

that is associated with the fluctuations at the microscale of the heterogeneities—the local properties. [25, pp. 305–306] In effect, then one looks at a family of functions u_ϵ and searches for a limit $u = \lim_{\epsilon \to 0} u_\epsilon$ that tells us what the effective properties of the material will be at the macroscale.

   Figure 6 illustrates this. The left box shows the two scales a and ξ with two phases of the material K_1 and K_2. The homogenization limit enables one to treat the heterogenous system at scale a as a homogenous system at scale ξ with an effective material property represented by K_e. For an elastic solid like the steel girder, K_e would be the effective stiffness tensor and is related experimentally to Young's modulus. For a conductor, K_e would be the effective conductivity tensor that is related experimentally to

28

the parameter σ—the specific conductance—appearing in Ohm's law:

$$\mathbf{J} = \sigma\mathbf{E},$$

where $\mathbf{J}$ is the current density at a give location $\mathbf{x}$ in the material and $\mathbf{E}$ is the electric field at $\mathbf{x}$.

Let me end this brief discussion of homogenization by highlighting what I take to be a very important concept for the general problem of upscaling. This is the concept of an order parameter and related functions. The notion of an order parameter was introduced in our discussion of continuous phase transitions in thermodynamics, and the statistical mechanical explanations of certain of their features, In effect, the order parameter is a microstructure (mesoscopic scale) dependent function introduced to codify the phenomenologically observed transition between different states of matter. As we have seen, the magnetization M represented in figure 1 is introduced to reflect the fact that at the Curie temperature the systems goes from and unordered phase, above T_c to an ordered phase, below T_c. In this context, the divergences and nonanalyticities at the critical point play an essential role in determining the fixed point that characterizes the class of systems exhibiting the same scaling behavior: $M \propto |t|^\beta$. But, again following Nelson's suggestion, entire classes of systems such as the class of linear elastic solids are also characterized by " fixed points" represented by a relatively few phenomenological parameters like Young's modulus.

It is useful to introduce an order-like parameter in this more general context of upscaling where criticality is not really an issue. For example, consider the left image in figure 6. In upscaling to get to the right image, one can begin by defining indicator functions for the different phases as a function of spatial location. [25, pp. 24-25] For instance, if the shaded phase occupies a volume V_s in the space and V_u is its complement, an indicator or characteristic function of that phase is given by

$$\chi^s(\mathbf{x}) = \begin{cases} 1, & \text{if } \mathbf{x} \in V_s \\ 0, & \text{otherwise} \end{cases}$$

One can also introduce indicator functions for the interfaces or boundaries between the two phases.[20] Much information can then be determined by investigating n-point probability functions expressing the probabilities that

---

[20]These will be generalized distribution functions.

n locations $\mathbf{x}_1, \ldots, \mathbf{x}_n$ are to be found in regions occupied by the shaded phase.[21]

$$S_n^s(\mathbf{x}_1, \ldots, \mathbf{x}_n) = Pr\left\{\chi^s(\mathbf{x}_1) = 1, \ldots, \chi^s(\mathbf{x}_n = 1)\right\}.$$

In this way many features *other than* simple volume fraction that exist at microscales can be represented and employed in determining the homogenization limit for complex heterogeneous systems. The introduction of such field variables, correlation functions, etc., allow us to characterize the heterogeneous structures above the atomic scales. In some cases, such as the bubbles within bubbles structure of the different phases at a continuum phase transition, much of this additional apparatus will not be necessary. (Though, of course, it is essential to take into consideration that structure in that particular case.) But for many more involved upscaling problems such as steel, the additional mathematical apparatus will be critical in determining the appropriate effective phenomenological theory at the continuum level. As we have seen these microstructures are critical for an understanding of how the phenomenological parameters at the continuum scale emerge.

The main lesson to take from this all-too-brief discussion is that physics at these micro/meso-scopic scales need to be considered. Bottom-up modeling of systems that exist across a large range of scales is not sufficient to yield observed properties of those systems at higher scales. Neither is complete top-down modeling. After all, we know that the parameters appearing in continuum models must depend upon details at lower scale levels. The interplay between the two strategies—a kind of mutual adjustment in which lower scale physics informs upper scale models and upper scale physics corrects lower scale models—is complex, fascinating, and unavoidable.[22]

---

[21]See [25] for a detailed development of this approach.

[22]Much of my work in recent years, though often framed in terms of relations between theories, can and should be understood as supporting this more complex point of view. Relations between theories that are asymptotically related to one another do not fit with a simple all or nothing dichotomy between "fundamental" theories that get the ontology correct, and "theories emeritus" that are false, but still useful. The latter term is Belot's. See [2, 8, 4].

7 Conclusion

The problem of the tyranny of scales has been formulated as one of seeing if it is possible to exploit microstructural scale information (intermediate between atomic scales and macroscopic scales) to bridge between two dominant and apparently incompatible modeling strategies. These are the traditional bottom-up strategies associated with a broadly reductionist account of science and pure top-down strategies that held sway in the nineteenth century and motivated the likes of Mach, Duhem, Maxwell, *et al.* Despite great progress in understanding the physics of atomic and subatomic particles, the persistence of continuum modeling has led to heated debates in philosophy about emergence, reduction, realism, etc. We have canvassed several different attitudes to the apparent ineliminability of continuum level modeling in physics. On the one hand, there is the view of Butterfield and others, that the use of continuum limits represents nothing more than a preference for the mathematical convenience of the infinite. On the other hand, Maddy holds that such continuum models, employing infinite idealizations, are apparently necessary. But she despairs about finding a unified understanding of why and how that is possible.

I have tried to do two things in this paper. First, I've tried to show that neither of these attitudes are ultimately acceptable. Regarding Maddy's scepticism, there is hope that with homogenization theory a unified methodology can be formulated for the diverse uses of continuum limits in various modeling instances. One does not need to think of those cases as in someway floating free from reality. Regarding Butterfield *et al.*'s belief that continuum models are simply mathematical conveniences posing no real philosophical worries, I've tried to show that this fails to respect some rather deep differences between kinds of continuum modeling. In particular, the strategies employed in the renormalization group and in homogenization theory differ significantly from those employed in standard representative elementary volume (REV) averaging scenarios. The significance of Wilson's renormalization group advance was exactly to point out *why* such REV methods fail and *how* that failure can be overcome. The answer, as we have seen, is to pay attention to "between" scale structures as in the case of the bubbles within bubbles picture of what happens at phase transitions. Incorporating such structures—features that cannot be understood as averages over atomic level structures—is exactly the strategy behind upscaling attempts that connect Euler type modeling recipes to Cauchy type continuum recipes.

Homogenization lets us give an answer to why the use of Cauchy's recipe is safe and robust. It provides a satisfactory justification for the use of such continuum models, but not one that is "straightforward" or pragmatically motivated.

Secondly, I've tried to focus attention on a rather large subfield of applied mathematics that should be of interest to philosophers working on specific issues of modeling, simulation, numerical methods, and idealizations. In addition, understanding the nature of materials in terms of homogenization strategies can inform certain questions about the nature of physical properties and issues about realism. For instance, we have seen that many materials at macroscales are characterized by a few phenomenological parameters such as the elastic constants. Understanding the nature of materials requires understanding why these constants and not others are appropriate, as well as understanding from where the constants arise. One important lesson is that many of these material defining parameters are not *simply* dependent upon the nature of the atoms that compose the material. There is a crucial link between *structure* at intermediate scales and observed *properties* at the macroscale.

It may do to end with an nice statement (partially cited earlier) from Rob Phillips' excellent book *Crystals, Defects, and Microstructures* expressing this point of view.

> Despite the power of the idea of a material parameter, it must be greeted with caution. For many features of materials, certain 'properties' are not *intrinsic.* For example, both the yield strength and fracture toughness of a material depend upon its internal constitution. That is, the measured material response can depend upon microstructural features such as the grain size, the porosity, etc. Depending upon the extent to which the material has been subjected to prior working and annealing, these properties can vary considerably. Even a seemingly elementary property such as the density can depend significantly upon that material's life history. The significance of the types of observations given above is the realization that many material properties depend upon more that just the identity of the particular atomic constituents that make up that material. ... [M]icrostructural features such as point defects, dislocations, and grain boundaries can each alter the measured macroscopic 'properties' of a mate-

rial. [22, pp. 5–8]

References

[1] Sorin Bangu. Understanding thermodynamic singularities: Phase transitions, data, and phenomena. *Philosophy of Science*, 76(4):488–505, 2009.

[2] Robert W. Batterman. *The Devil in the Details: Asymptotic Reasoning in Explanation, Reduction, and Emergence*. Oxford Studies in Philosophy of Science. Oxford University Press, 2002.

[3] Robert W. Batterman. Critical phenomena and breaking drops: Infinite idealizations in physics. *Studies in History and Philosophy of Modern Physics*, 36:225–244, 2005.

[4] Robert W. Batterman. Response to Belot's "Whose devil? Which details?". *Philosophy of Science*, 72(1):154–163, 2005.

[5] Robert W. Batterman. On the explanatory role of mathematics in empirical science. *The British Journal for the Philosophy of Science*, doi = 10.1093/bjps/axp018:1–25, 2009.

[6] Robert W. Batterman. Reduction and renormalization. In Gerhard Ernst and Andreas Hüttemann, editors, *Time, Chance, and Reduction: Philosopohical Aspects of Statistical Mechanics*. Cambridge University Press, 2010.

[7] Robert W. Batterman. Emergence, singularities, and symmetry breaking. *Foundations of Physics*, 41(6):1031–1050, 2011.

[8] Gordon Belot. Whose devil? which details? *Philosophy of Science*, 71(1):128–153, 2005.

[9] Jeremy Butterfield. Less is different: Emergence and reduction reconciled. *Foundations of Physics*, 41(6):1065–1135, 2011.

[10] Jeremy Butterfield and Nazim Bouatta. Emergence and reduction combined in phase transitions. http://philsci-archive.pitt.edu/id/eprint/8554, April 2011.

[11] J. T. Oden (Chair). Simulation based engineering science—an NSF Blue Ribbon Report. www.nsf.gov/pubs/reports/sbes_final_report.pdf, 2006.

[12] Reint de Boer. *Theory of Porous Media: Highlights in Historical Development and Current State.* Springer, Berlin, 2000.

[13] R. Feynman, R. Leighton, and M. Sands. *The Feynman Lectures on Physics*, volume 2. Addison-Wesley, Reading, Massachusetts, 1964.

[14] Ulrich Hornung, editor. *Homogenization and Porous Media*, volume 6 of *Interdisciplinary Applied Mathematics.* Springer, New York, 1997.

[15] Leo P. Kadanoff. Scaling, universality, and operator algebras. In C. Domb and M. S. Green, editors, *Phase Transtions and Critical Phenomena*, volume 5A. Academic Press, 1976.

[16] Leo P. Kadanoff. *Statistical Physics: Statics, Dynamics, and Renormalization.* World Scientific, Singapore, 2000.

[17] Penelope Maddy. How applied mathematics became pure. *The Review of Symbolic Logic*, 1(1):16–41, 2008.

[18] Tarun Menon and Craig Callender. Going through a phase: Philosophical questions raised by phase transitions. 2011.

[19] David R. Nelson. *Defects and Geometry in Condensed Matter Physics.* Cambridge University Press, Cambridge, 2002.

[20] John Norton. Approximation and idealiztion: Why the difference matters. http://philsci-archive.pitt.edu/8622/, 2011.

[21] Grigorios A. Pavliotis and Andrew M. Stuart. *Multiscale Methods: Averaging and Homogenization.* Texts in Applied Mathematics. Springer, New York, 2008.

[22] Rob Phillips. *Crystals, Defects, and Microstructures: Modeling Across Scales.* Cambridge University Press, Cambridge, 2001.

[23] Ya. G. Sinai. *Probability Theory: An Introductory Course.* Springer-Verlag, Berlin, 1992. Trans.: D. Haughton.

[24] Isaac Todhunter and Karl Pearson (Ed.). *A History of the Theory of Elasticity and of the Strength of Materials from Galilei to Lord Kelvin*, volume 1: Galilei to Saint-Venant 1639–1850. Dover, 1960.

[25] Salvatore Torquato. *Random Heterogeneous Materials: Microstructure and Macroscopic Properties.* Springer, New York, 2002.

[26] Kenneth G. Wilson. Critical phenomena in 3.99 dimensions. *Physica*, 73:119–128, 1974.

[27] Mark Wilson. Mechanics, classical. In *Routledge Encyclopedia of Philosophy.* Routledge, 1998.