Turn and Face the Strange... Ch-ch-changes

Philosophical Questions Raised by Phase Transitions

Tarun Menon and Craig Callender

August 2, 2011

Phase transitions are abrupt changes in the macroscopic properties of a system. Examples of the phenomenon are familiar: freezing, condensation, magnetization. Often these transitions are particularly dramatic, as when solid objects composed of the silvery metal gallium vanish into puddles when picked up (the temperature of the hand is just enough to raise gallium's temperature past its melting point). Characterized generally, one finds them inside and outside of physics, in systems as diverse as neutron stars, DNA helices, financial markets, and traffic. In the past half-century, the study of phase transitions and critical phenomena has been a central preoccupation of the statistical physics community. In fact, it is now a truly interdisciplinary area of research. Phase transitions manifest at many different scales and in all sorts of systems, so they are of interest to atomic physicists, materials engineers, astronomers, biologists, sociologists and economists. However, philosophical attention to the foundational issues involved has thus far been limited.

This is unfortunate, because the theory of phase transitions is unusual in many ways and offers a novel perspective that could enrich a number of debates in the philosophy of science. In particular, questions about reduction, emergence, explanation and approximation all arise in a particularly stark manner when considering this phenomenon. Here we'll focus on these questions as they relate to the most studied type of phase transition, namely, transitions between different equilibrium phases in thermodynamics. These are sudden changes between one stable thermodynamic state of matter and another while one smoothly varies a parameter. A paradigmatic example is the change in water from liquid to gas as the temperature is raised or the pressure is reduced.

In the small philosophical commentary on this topic, such changes have provoked many surprising claims. Many have claimed that phase transitions cannot be reduced to statistical mechanics, that they are truly emergent phenomena. The argument for this conclusion hangs on one's understanding of the infinite idealization invoked in the statistical mechanical treatment of phase transitions. In this chapter we'll focus on puzzles associated with this idealization. Is infinite idealization necessary for the explanation of phase transitions? If so, does it show that phase transitions are, in some sense, emergent phenomena? If so, what precisely is that sense? Questions of this sort provide a concrete basis for the exploration of philosophical approaches to reduction and idealization, and they also bear on the ongoing scientific study of these systems.

1 The Physics of Phase Transitions

Phase transitions raise interesting questions about inter-theoretic relationships because they are studied from three distinct theoretical perspectives. Thermodynamics provides a macroscopic, phenomenological characterization of the phenomenon. Statistical mechanics attempts to ground the thermodynamic treatment by explaining how this macroscopic behavior arises out of the interaction of microscopic degrees of freedom. This project has led to the employment of renormalization group theory, a tool first developed in the context of particle physics for studying the behavior of systems under transformations of scale. While renormalization group theory is usually placed under the broad rubric of statistical mechanics, the methods employed are importantly different from the traditional tools of statistical mechanics. Rather than a probability distribution over an ensemble of configurations of a single system, the primary theoretical device of renormalization group theory is the flow generated by the scaling transformation on a space of Hamiltonians representing distinct physical systems. In this section we describe how these three approaches treat the phenomenon of phase transitions, with special attention to the employment of the infinite particle idealization.

1.1 Thermodynamic Treatment

The thermodynamic treatment of phases and phase transitions began in the nineteenth century. Experiments by Andrews, Clausius, Clapeyron, and many others provided data that would lead to developed theories of phase transitions and critical phenomena. Gradually it was recognized that at certain values of temperature and pressure a substance can exist in more than one thermodynamic phase (e.g., solid, liquid), while at other values there can be a change in phase but no co-existence of phases.

For instance, as pressure is reduced or temperature is raised, liquid water transitions to its gaseous phase. At the boundary between these phases, both liquid and gaseous states can coexist; the thermodynamic parameters of the system do not pick out a unique equilibrium phase. In fact, at the triple point of water (temperature 273.16 K and pressure 611.73 Pa), all three phases – solid, liquid and gas – can coexist. The transitions at these phase boundaries are marked by a discontinuity in the density of water. As the pressure is reduced at a fixed temperature, the equilibrium state of water switches abruptly from a high-density liquid phase to a low-density gaseous phase. This is an example of a *first order* phase transition. As the temperature is increased past the critical temperature of 647 K, water enters a new phase. In this regime, there are no longer macroscopically distinct liquid and gas phases, but a homogenous supercritical fluid that exhibits properties associated with both liquids and gasses.

Changing the pressure leads to a continuous change in the density of the fluid; there are no phase boundaries. This supercritical phase allows a transition from liquid to gas that does not involve any discontinuity in thermodynamic observables: raise the temperature of the liquid past the critical temperature, reduce the pressure below the critical pressure (22 MPa for water), then cool the fluid back to below the critical temperature. This path takes the system from liquid to gas without crossing a phase boundary. The transition of a system past its critical point to the supercritical phase is a *continuous* phase transition.

Mathematically, phase transitions are represented by non-analyticities or singularities in a thermodynamic potential. A singularity is a point at which the potential is not infinitely differentiable, so at a phase transition some derivative of the thermodynamic potential changes discontinuously. A classification scheme due to Ehrenfest provides the resources to distinguish between first and second order transitions in this formalism. A first order phase transition involves a discontinuity in the first derivative of a thermodynamic potential. In the liquid-gas first order transition, the volume of the system, a first derivative of the thermodynamic potential known as the Gibbs free energy, changes discontinuously. For a second order phase transition the first derivatives of the potentials are continuous but there is a discontinuity in a second derivative of a thermodynamic potential. At the liquid-gas critical point we see a discontinuity in the compressibility of the fluid, which is a first derivative of volume and hence a second derivative of the Gibbs free energy. Ehrenfest's scheme extends naturally to allow for higher order phase transitions as well. An n-th order transition would be one whose n-th derivative is discontinuous. Contemporary statistical mechanics retains the category of first order phase transitions (sometimes referred to as abrupt transitions), but all other types of non-analyticities in thermodynamic potentials are grouped together as continuous phase transitions.

Continuous phase transitions are often referred to as order-disorder transitions. There is usually some symmetry in the supercritical phase that is broken when we cross below the critical point. This broken symmetry allows for the material to be ordered in various ways, corresponding to different phases. A stark example of the transition between order and disorder is the transition in magnetic materials, such as iron, between paramagnetism and ferromagnetism. At room temperature, a piece of iron is permanently magnetized when exposed to an external magnetic field. In the presence of a field, the minimum energy configuration is the one with the largest possible net magnetic moment reinforcing the field, so the individual dipoles within the iron align to maximize the net moment. This configuration remains stable even when the external field is removed. Materials with this propensity for induced permanent magnetization are called *ferromagnetic*. If the temperature is raised above 1043 K, the ferromagnetic properties of iron vanish. The iron is now *paramagnetic*; it can no longer sustain induced magnetization when the external field is removed. In the stable configuration, there is no correlation between the alignments of neighboring dipoles. In the paramagnetic phase, no direction is picked out as special after the magnetic field is switched off. The material exhibits spatial symmetry. In the ferromagnetic phase, this symmetry is broken. The dipoles line up in a particular spatial direction even after the field is removed. The order represented by this alignment does not survive the transition past criticality.

A simple way to understand this transition between order and disorder is in terms of the minimization of the Helmholtz free energy of the system:

$$F = E - TS \tag{1}$$

Here E is the energy of the system, T is the temperature, and S is the entropy. The stable configuration minimizes free energy. At low temperatures, the energy term dominates, and the low energy configuration with dipoles aligned is favored. At high temperatures, the entropy term dominates, and we get the high entropy configuration with uncorrelated dipole moments. The change in magnetic behavior is explicable as a shift in the balance of power in the battle between the ordering tendency due to minimization of energy and the disordering tendency due to maximization of entropy. As indicated, the paramagnetic-ferromagnetic transition is continuous, not first order. All first derivatives of the free energy are continuous, but second derivatives (such as the magnetic susceptibility $\chi = \frac{\partial^2 F}{\partial H^2}$, where H is the magnetization) are not.

The transition from order to disorder is also represented, following Landau, as the vanishing of an *order parameter*. In the case under consideration, this parameter is the net magnetization M of the system . Below the critical point, you have different phases with distinct values of the order parameter. If we simplify our model of the magnetic material so that the induced magnetization of the dipoles is only along one spatial axis (as in the Ising model), then at each temperature below criticality the order parameter can take two values, related by a change of sign. The magnetization vanishes as we approach the critical point and remains zero in the supercritical phase, corresponding to a disappearance of distinct phases.

The vanishing of the order parameter close to the critical temperature T_c is characterized by a power law:

$$M \propto (-t)^\beta \tag{2}$$

where t is the reduced temperature $(T - T_c)/T_c$. The exponent β characterizes the rate at which the magnetization falls off as the critical temperature is approached. It is an example of a *critical exponent*, one of many that appear in power laws close to the critical point. The experimental and theoretical study of critical exponents has been crucial to recent developments in the theory of phase transitions.

1.2 Statistical Mechanical Treatment

Statistical mechanics is the theory that applies probability theory to the microscopic degrees of freedom of a system in order to explain its macroscopic behavior. The tools of statistical mechanics have been extremely successful in explaining a number of thermodynamic phenomena, but it turned out to be particularly difficult to apply the theory to the study of phase transitions. There were two significant obstacles to the development of a successful statistical mechanical treatment of phase transitions: one experimental and one conceptual.

The experimental obstacle had to do with the failure of mean field theory. This was the dominant approach to the statistical mechanics of phase transitions up to the middle of the twentieth century. The theory is best explicated by considering the Ising model, which represents a system as a lattice of sites, each of which can be in two different states. The states will be referred to as spin up and spin down, in analogy with magnetic systems. However, Ising models have been successfully applied to a number of different systems, including the liquid-gas system near its critical point. The Hamiltonian for the Ising model involves a contribution by an external term, corresponding to the external magnetic field for magnetic systems, and internal coupling terms. The only coupling is between neighboring spins on the lattice. It is energetically favorable for neighboring spins to align with one another and with the external field. This model is supposed to represent the way in which local interactions can produce the kinds of long range correlations that characterize a thermodynamic phase.

In statistical mechanics, all thermodynamic functions are determined by the canonical partition function. The coupling terms in the Hamiltonian make the calculation of the partition function for the Ising model mathematically difficult. To make this calculation tractable, we approximate the contribution of a particular lattice site to the energy of the system by supposing that all its neighbors have a spin equal to the ensemble average. This approximation ignores fluctuations of spins from their mean values. The fluctuations become less relevant as the number of neighbors of a particular lattice site increases, so the mean field approximation works better the higher the dimensionality of the system under consideration. Once the partition function is calculated using this approximation, there is an elegant method due to Landau for determining the critical exponents. Unfortunately, Landau's method gives results that conflict with experiment. For instance, the mean field value for the critical exponent β is 0.5, but observation suggests the actual value is about 0.32. The approximation fails close to the critical point of a magnetic system. In fact, this failure is predicted by Landau theory itself. The theory tells us that as we approach the critical point, the *correlation length* diverges. This is the typical distance over which fluctuations in the microscopic degrees of freedom are correlated. As this length scale increases, fluctuations become more relevant, and the mean field approximation, which ignores fluctuations, weakens. Mean field theory cannot fully describe continuous phase transitions because of this failure near criticality. Another approach is needed for a full statistical mechanical treatment of the phenomenon.

As mentioned, there was also a deeper conceptual obstacle to a statistical mechanics of phase transitions. If one adopts the definition of phase transitions employed by thermodynamics, then *phase transitions in statistical mechanics don't seem possible.* The impossibility claim can be explained very easily. As mentioned above, thermodynamic functions are determined by the partition function. For instance, the Helmholtz free energy is given by:

$$F = -kT \ln Z \tag{3}$$

where k is Boltzmann's constant, T is the temperature of the system, and Z is the canonical partition function:

$$Z = \sum_n \exp\left(-\frac{E_n}{kT}\right) \tag{4}$$

with E_n labeling the different possible mechanical energies of the system. Recall the definition of a phase transition according to thermodynamics:

(Def 1) An equilibrium phase transition is a nonanalyticity in the free energy.

Depending on the context, one might choose a nonanalticity in a different thermodynamic potential; however, that freedom won't affect matters here.

As natural as it is, (Def 1) makes a phase transition seem unattainable in statistical mechanics. The reason is that each of the exponential functions in (4) is analytic, the partition function is just a sum of exponentials, and the free energy essentially is just the logarithm of this sum. Since a sum of analytic functions is itself analytic and the logarithm of an analytic function itself analytic, the Helmholtz free energy, expressed in terms of the logarithm of the partition function, will also be analytic. Hence, there will be no phase transitions as defined by (Def 1). Since the same reasoning can be applied to any thermodynamic function that is an analytic function of the canonical partition function modifications of (Def 1) to other thermodynamic functions won't work either. (For a rigorous proof of the above claims, see Griffiths (1972).)

In the standard lore of the field, this problem was resolved when Onsager in 1944 demonstrated for the first time the existence of a phase transition from nothing but the partition function. He did this rigorously for the two-dimensional Ising model with no external magnetic field. How did Onsager manage the impossible? He worked in the thermodynamic limit of the system. This is a limit where the number of particles in the system N and the volume of the system V go to infinity while the density $\rho = N/V$ is held fixed. Letting N go to infinity is the crucial trick in getting around the "impossibility" claim. The claim depends on the sum of exponentials in (4) being finite. Any finite sum of analytic functions will be analytic. Once this restriction is removed, however, it is possible to find nonanalyticities in the free energy. The apparent lesson is that statistical mechanics can describe phase transitions, but only in infinite particle systems.

It is common to visualize what's going on in terms of the Yang-Lee theorem. The free energy is a logarithm of the partition function, so it will exhibit a singularity where the partition function goes to zero. But the partition function is a polynomial of finite degree with all positive coefficients, so it has no real positive roots. Instead the roots are imaginary and the zeroes of the partition function must be plotted on the complex plane. The Yang-Lee theorem, for a two-dimensional Ising model, says that these zeroes sit on the unit sphere in the complex plane. As the number of particles increases, the zeroes become denser on the unit sphere until at the thermodynamic limit they intersect the positive real axis. Since a real zero of the partition function is only possible in this limit, it is only in this limit that we can have a phase transition (understood as in Def 1).

An alternative definition of phase transitions is sometimes used, one proposed by Lebowitz (Lebowitz, 1999). A phase transition occurs, on this definition, just in case the Gibbs measure (a generalization of the canonical ensemble) is nonunique for the system. This corresponds to a coexistence of distinct phases and therefore a phase transition. Using this alternative definition, however, won't change philosophical matters. The Gibbs measure can only be nonunique in the thermodynamic limit, just as (Def 1) can only be satisfied in the thermodynamic limit. That said, this way of looking at the issue perhaps makes it easier to see the similarities between the foundational issues raised by phase transitions and those raised by spontaneous symmetry breaking.

1.3 Renormalization Group Theory

We mentioned in the previous section that mean field theory fails near the critical point for certain systems because it neglects the importance of fluctuations in this regime. Dealing with this strongly correlated regime required the introduction of a new method of analysis, imported from particle physics. This is the renormalization group method. While mean field theory hews to tools and forms of explanation that are orthodox in statistical mechanics, such as determining aggregate behavior by taking ensemble averages, renormalization group theory introduced a somewhat alien approach with tools more akin to those of dynamical systems theory than statistical mechanics.

To explain the method, we return to our stalwart Ising model. Suppose we coarse-grain a 2-D Ising model by replacing 3×3 blocks of spins with a single spin pointing in the same direction as the majority in the original block. This gives us a new Ising system with a longer distance between lattice sites, and possibly a different coupling strength. You could look at this coarse-graining procedure as a transformation in the Hamiltonian describing the system. Since the Hamiltonian is characterized by the coupling strength, we can also describe the coarse-graining as a transformation in the coupling parameter. Let K be the coupling strength of the original system, and R be the relevant transformation. The new coupling strength is $K' = RK$. This coarse-graining procedure could be iterated, producing a sequence of coupling parameters, each related to the previous one by the transformation R. The transformation defines a flow on parameter space.

How does this help us ascertain the critical behavior of a system? If you look at an Ising system at its critical point, you will see clusters of correlated spins of all sizes. This is a manifestation of the diverging correlation length. Now squint, blurring out the smaller clusters. The new blurry system that you see will have the same general structure as the old one. You will still see clusters of all sizes.

This sort of scale invariance is characteristic of critical behavior. The system has no characteristic length scale. Coarse-graining produces a new system that is statistically identical to the old one. At this point, the Hamiltonian of the system remains the same under indefinite coarse-graining, so it must be a fixed point in parameter space, i.e. a point K_f such that $K_f = RK_f$. The non-trivial (viz. not $K = 0$ or $K = \infty$) fixed points of the flow characterize the Hamiltonian of the system at the critical point, the point at which correlation length diverges and there is no characteristic scale for the system. The critical exponents can be calculated by series expansions near the critical point. Critical exponents predicted by renormalization group methods agree with experiment much more than the predictions of mean field theory.

The same approach can be applied to systems with more complicated Hamiltonians involving a number of different parameters. Some of these parameters will be relevant, which means they get bigger as the system is rescaled. If a system has a non-zero value for some relevant parameter, then it will not settle at a non-trivial fixed point upon rescaling, since rescaling will amplify the relevant parameter and therefore change the couplings in the system. At criticality, then, the relevant parameters must be zero. An example of a relevant parameter for the Ising system is the reduced temperature t. If $t = 0$, the system can flow to a non-trivial fixed point. However, if t is perturbed from zero, the system will flow away from this critical fixed point towards a trivial fixed point. So a continuous transition only takes place when $t = 0$, which is at the critical temperature. Other parameters might turn out to be irrelevant at large scales. They will get smaller and smaller with successive coarse-grainings, effectively disappearing at macroscopic scales. This elimination of microscopic degrees of freedom means that the renormalization group transformation can be irreversible (which would, strictly speaking, make it a semi-group rather than a group), and there can be attractors in parameter space. These are fixed points into which a number of microscopically distinct systems flow. This is the basis of *universality*, the shared critical behavior of quite different sorts of systems. If the systems share a fixed point their critical exponents will be the same, even if their microscopic Hamiltonians are distinct. The differences in the Hamiltonians are in irrelevant degrees of freedom that do not affect the macroscopic critical behavior of the system. Systems that flow to the same non-trivial fixed point are said to belong to the same universality class. The liquid-gas transition in water is in the same universality class as the paramagnetism-ferromagnetism transition. They have the same critical exponents, despite the evident differences between the systems.

The difference between relevant and irrelevant parameters can be conceptualized geometrically. In parameter space, if we restrict ourselves to the hyper-surface on which all relevant parameters are zero, so that the differences between systems on this hyper-surface are purely due to irrelevant parameters, then all points on the hyper-surface will flow to a single fixed point. Perturb the system so that it is even slightly off the hyper-surface, however, and the flow will take it to a different fixed point.

It is significant that the fixed point only appears when the system has no characteristic length scale. This is why the infinite particle limit is crucial to the renormalization group approach. If the number of particles is finite, then there will be a characteristic length scale set by the size of the system. Coarse-graining beyond this length will no longer give us statistically identical systems. The possibility of invariance under indefinite coarse-graining requires an infinite system. The requirement for the thermodynamic limit in renormalization group theory can be perspicuously connected to the motivation for this limit in the standard statistical mechanical story. The correlation length of a system near its critical point can be characterized in terms of some second derivative of a thermodynamic potential. For instance, in a magnetic system the range of correlations between parts of the system is proportional to the susceptibility, a second derivative of the free energy. On the thermodynamic treatment, the susceptibility diverges as we approach the critical point, and according to the statistical mechanical treatment this is impossible unless we are in the thermodynamic limit. This means the correlation length cannot diverge, as is required for renormalization group methods to work, unless the system is infinite.

2 The Emergence of Phase Transitions?

All of the above should sound a little troubling. After all, the systems we're interested in, the systems in which we see phase transitions every day, are not infinite systems. Yet the physics of phase transitions seems to make crucial appeal to the infinitude of the systems modeled. It appears that, according to both statistical mechanics and renormalization group theory, phase transitions cannot occur in finite systems. Additionally, the explanation of the universal behavior of systems near their critical point seems to require the infinite idealization. Considerations of this sort have led many authors to say that phase transitions are genuinely emergent phenomena, suggesting that statistical mechanics cannot provide a full reductive account of phase transitions in finite systems. The eminent statistical mechanic Lebowitz says phase transitions are "paradigms of emergent behavior," (Lebowitz, 1999, p. S346) and the philosopher Liu says they are "truly emergent properties" (Liu, 1999, p. S92).

Needless to say, if this claim is correct, phase transitions present a challenge to philosophers with a reductionist bent. The extent of this challenge depends on how we interpret the claim of emergence. The concept of "emergence" is notoriously slippery, interpreted differently by different authors. We will consider a number of different arguments for phase transitions being emergent, corresponding to varying conceptions of emergence. What these arguments have in common is that they all involve a rejection of what Andrew Melnyk has called "reductionism in the core sense" (Melnyk, 2003, p. 83). This is the intuitive conception of reduction that underlies various more precise philosophical accounts of reduction. A theory T_h reduces to a lower-level theory T_l if all the nomic claims made by T_h can be explained using only the resources of T_l and necessary truths.

This conception is deliberately vague, allowing for various precisifications depending on one's theory of explanation, and how one delineates the explanatory resources available to a particular theory. One possible precisification is Ernest Nagel's account of reduction (Nagel, 1979), which says that T_l reduces T_h if and only if the laws of the latter can be deduced from the laws of the former in conjunction with appropriate bridge laws. In this account the core sense of reduction has been filled out with a logical empiricist theory of explanation according to which the explanatory resources of a theory are the deductive consequences of its law-like statements. It is important to recognize that a reductionist is committed to this account of reduction only in so far as she endorses such a theory of explanation. The proper motivation for Nagel's theory lies in the extent to which it successfully captures the core sense of reduction.

In this chapter we do not endorse any particular account of reduction. Instead we consider three broad ways in which the explanatory connection between a higher level theory and a lower level theory may break down, and examine the extent to which these explanatory breakdowns are manifested in the case of phase transitions. Whether or not we have a genuine explanatory failure in a particular instance will depend on the details of our account of explanation. Often, the reductionist may be able to avoid a counterexample by simply re-conceiving what counts as an adequate explanation.[1] However, certain instances will be regarded as explanatory failures under a wide variety of plausible accounts of explanation, perhaps even under *all* plausible accounts of explanation. The weaker the assumptions about explanation required for the counterexample to work, the stronger the case for emergentism. We can arrange our examples of purported explanatory failure into a hierarchy based on the constraints placed on an account of explanation.

At the bottom of this hierarchy (at least for the purposes of this chapter) is *conceptual novelty*. This is the sort of "irreducibility" involved when there is some natural kind in the higher level theory which cannot be equated to a single natural kind in the lower level theory. It may be the case that the phenomena that constitute the higher level kind can be individually explained by the lower level theory, but the theory does not unite them as a single kind. Conceptual novelty involves a failure of type-type reduction, but need not involve a failure of token-token reduction. In the case of phase transitions, it has been suggested that although one can provide a perfectly adequate explanation of individual transitions using statistical mechanics, the theory does not distinguish these phenomena as a separate kind. For instance, from the perspective of statistical mechanics, the transition from ice to water in a finite system as we cross 273.16 K is not qualitatively different from the transition from cold ice to slightly warmer ice as we cross 260 K, at least if something like the standard story is correct. The only difference is that the thermodynamic potentials change a lot more rapidly in the former situation than in the latter, but they are still analytic, so this is merely a difference of degree, not a difference of kind.

[1] As an example, consider multiple realization, often presented as a failure of reduction. However, it is only a failure if we believe that a lower-level explanation of the higher-level law must be unified, i.e. the explanation must be the same for every instance of the higher-level law. If we are willing to allow for disunified explanation, then we may indeed have a genuine lower-level explanation of the higher-level law, preserving the core sense of reduction.

There are two tacks one can take in response to this observation. The first is that this is a case where statistical mechanics corrects thermodynamics. Just as it showed us that the second law is not in principle exceptionless, it shows us that rigorous separation of phases, the only phenomenon worthy of the name "phase transition", is only possible in infinite systems. This view of the emergence of phase transitions is expressed by Kadanoff when he says that "in some sense phase transitions are not exactly embedded in the finite world but, rather, are products of the human imagination." (Kadanoff, 2009, p. 778) Thermodynamics classifies a set of empirical phenomena as phase transitions, involving a qualitatively distinct type of change in the system. Statistical mechanics reveals that these phenomena have been misclassified. They are not genuinely qualitatively distinct and should not be treated as a separate natural kind. This response does not appear to pose much of a threat to reductionism. It may be true that thermodynamics hasn't been reduced to statistical mechanics in a strict Nagelian sense, but this seems like much too restrictive a conception of reduction. There are many paradigmatic cases of scientific reduction where the reducing theory explains a corrected version of the reduced theory, not the theory in its original form. This correction may often involve dissolving inappropriate distinctions. If this is all there is to the challenge of conceptual novelty, it is not much of a challenge.

However, one might want to resist this eliminativism and reject the notion that thermodynamics has misclassified phenomena. Perhaps the right thing to say is that at the thermodynamic level of description it is indeed appropriate to have a distinct kind corresponding to phase transitions in finite systems. But the appropriateness of this kind is invisible at the statistical mechanical level of description, since statistical mechanics does not have the resources to construct such a class. This is a more substantive challenge to reductionism, akin to cases of multiple realizability. As an analogy, consider that from the perspective of our molecular theory there is no natural kind (or indeed finite disjunction of kinds) corresponding to the category "can opener". It seems implausible that we will be able to delineate the class of can openers using only the resources of our microscopic theory. Yet we do not take this to mean that our microscopic theory corrects our macroscopic theory, demonstrating that can openers do not exist as a separate kind. Can openers do exist. They are an appropriate theoretical kind at a certain level of description. Similarly, the fact that statistical mechanics does not have the resources to delineate the class of finite particle phase transitions need not lead us to conclude that this classification is bogus.

How might the reductionist respond to conceptual novelty of this sort? One response would be to develop a sense of explanation that makes reduction compatible with multiple realization. Even though statistical mechanics does not group phase transitions together the way that thermodynamics does, it is still able to fully explain what goes on in individual instances of phase transition. Perhaps the existence of individual explanations in every case constitutes an adequate explanation of the nomic pattern described by thermodynamics. If this is the case, the core sense of reduction is satisfied. One does not need to look at phase transitions to notice that any claim about the reduction of thermodynamics to statistical mechanics must be based on a conception of reduction that is compatible with multiple realizability. Temperature, that most basic of thermodynamic properties, is not (the claims of numerous philosophers notwithstanding) simply "mean molecular kinetic energy". It is a multiply realizable functional kind. If our notion of reduction precludes the existence of such properties, then the project of reducing thermodynamics cannot even get off the ground.

To us, this seems like the correct response to claims of emergence based on the conceptual novelty of phase transitions. If this is all it takes for emergence, then practically every thermodynamic property is emergent. Perhaps the emergentist is willing to bite this bullet, but we think it is more plausible that that the argument from conceptual novelty to emergence relies on a much too restrictive conception of scientific explanation. It is, however, worth noting another line of response. It may be the case that a class of finite particle phase transitions can be constructed within statistical mechanics that overlaps somewhat (but not completely) with the thermodynamic classification. This would be a case of statistical mechanics correcting thermodynamics, but not by eliminating the phenomenon of phase transitions in finite systems. Instead, statistical mechanics would re-define phase transitions in a manner that preserves our judgments about a number of empirical instances of the phenomenon. If such a re-definition could be engineered, phase transitions would not be conceptually novel to thermodynamics. The prospects for this strategy are discussed in section 3.1.

Let us suppose our conception of reduction is broad enough that mere conceptual novelty does not indicate a failure of reduction. We accept with equanimity that under certain conditions it might be appropriate to model phenomena using a conceptual vocabulary distinct from that of our reducing theory. For instance, at a sufficiently coarse-grained level of description a certain set of thermodynamic transformations is fruitfully modeled as exhibiting singular behavior, and appropriately grouped together into a separate natural kind. However, one might not think that a fully reductive explanation has been given unless one can explain using the resources of the reducing theory *why* this model is so effective under those conditions. Why does modeling a finite particle phase transition as non-analytic work so well at the thermodynamic level of description if finite systems cannot exhibit non-analyticities at the statistical mechanical level of description? If we cannot give such an explanation, we have another potential variety of emergence: *explanatory irreducibility.*

To give an idea of the kind of story we are looking for, consider the infinite idealization involved in explaining the extensivity of certain thermodynamic properties. Many thermodynamic properties are extensive, such as the entropy, internal energy, volume, and free energy. What this means is that if we divide a system into macroscopic parts, the values of those properties behave in an additive way. Loosely put, if we double the size of the system (that is, double internal energy, particle number, volume), then we double that system's extensive properties, e.g., the entropy.[2] Intensive properties, by contrast, don't scale this way; for example, if we double the size of a system, we don't double the pressure. Extensivity and intensivity are features usefully employed by phenomenological thermodynamics. However, when we look at a system microscopically, we quickly see that no finite system is ever strictly extensive or intensive. Correlations exist between the particles in one part of a system and another part. If we want to reproduce the thermodynamic distinction exactly, we are stymied: no matter how large the system, if it's finite, surface effects contribute to the partition function, which will mean that systems' energies and entropies cannot be neatly halved. For instance, if we define the entropy as a function over the joint probability distributions involved (as with the Gibbs entropy), we see that the entropy is extensive only when the two subsystems are probabilistically independent of one another. The only place we can reproduce the sharp distinction is by going to the thermodynamic limit. There we can define a variable f as extensive if f goes to infinity as we approach the thermodynamic limit while f/V is constant in the limit, where V is the volume of the system.[3] *Strictly* speaking, only in infinite systems are entropy, energy, etc. truly extensive.

Does this fact imply that there is a great mystery about extensivity, that extensivity is truly emergent, that thermodynamics doesn't reduce to finite N statistical mechanics? We suggest that on any reasonably uncontentious way of defining these terms, the answer is no. We know exactly what is happening here. Just as the second law of thermodynamics is no longer strict when we go to the microlevel, neither is the concept of extensivity. The notion of extensivity is an idealization, but it is one approximated well by finite particle statistical mechanics. For boxes of length l containing particles interacting via short-range forces, the surface effects scale as l^2 and the volume as l^3. Surface effects become less and less important as the system gets larger. Beings restricted to macroscopic physics would do well to call upon the extensive/intensive distinction, since in most cases the impact of surface effects would be well beyond the precision of measurements made by such beings. Here we see that extensivity in finite systems is conceptually novel to thermodynamics. It does not exist in statistical mechanics. However, leaving the story there is unsatisfactory. We need a further account, from a statistical mechanical perspective, of why this new concept works so well in thermodynamics. And indeed such a story is forthcoming. It relies crucially on the fact that the resolution of our measurements is limited, but this in itself does not, or at least should not, derail the reductionist project. As long as we have a story that explains why beings with such limitations could fruitfully describe sufficiently large systems as extensive – a story in terms of the components of the system and their organization, and how

[2]Strictly speaking, additivity and extensivity are different properties; see Touchette (2002). Since they overlap for many real systems, they are commonly run together; however, it's a mistake to do so in general, for some quantities scale with particle number N (and hence are extensive), yet are not additive.

[3]Some textbooks even go in the other direction, namely, defining the thermodynamic limit as that state wherein entropy and energy are extensive.

relevant quantities scale as the system gets larger – we do not have a genuine challenge to reductionism in the core sense.

The question is whether a similar sort of explanation is available to account for the efficacy of the infinite idealization involved in the statistical mechanical analysis of phase transitions. If there is not, we would have a case for emergence. There would be something about the systems under consideration that could not be accounted for reductively, viz. the fact that their behavior at a phase transition can, under certain conditions, be adequately modeled as the behavior of an infinite system. This feature of finite systems is crucial to understanding their behavior at phase transitions, so if it cannot be explained it would be legitimate to say that phase transitions are emergent. In section 3.2 we examine the possibility of a reductive explanation of the efficacy of the inifinite idealization.

Modelling the behavior of particular systems is not the only function of the infinite idealization in the study of phase transitions. The idealization plays a central role in the renormalization group explanation for universal behavior at the critical point. As we have discussed above, universal behavior is accounted for by the presence of stable fixed points in the space of Hamiltonians, each of which is the terminus of a number of different renormalization flow trajectories. This sort of explanation raises special problems that do not arise when we consider the sort of infinite idealization involved in the assumption of extensivity. There we have a property that, as it turns out, can only be approximated by finite systems. It is only actually instantiated in infinite systems. However, the property itself can be characterized without recourse to the infinite idealization. I could in principle construct an explanation of why a finite thermodynamic system approximates extensive behavior without any appeal to the infinite idealization. The idealization gives us a model of a genuinely extensive system, but it is not essential to an understanding of why it is useful to treat macroscopic finite systems as extensive.

It appears that the situation is different when we consider the renormalization group explanation of universality. There, the infinite idealization plays a different role. Talking about how a particular large finite system approximates the behavior of an infinite system will not be helpful, because universality is not about the behavior of individual systems, finite or infinite. It is a characteristic of classes of systems. The renormalization group method explains why physical systems separate into distinct universality classes, and it explains this in terms of certain structural features of the space of systems, the fixed points of the renormalization flow. It is the existence of these features, and their connection to the phonemenon of universality, that requires the infinite idealization. We might be able to give an account of why a particular large finite system approaches very close a fixed point as it is rescaled, approximating the behavior of an infinite system, but this will not tell us why this behavior *matters*. In order to see the connection between approaching a fixed point and exhibiting universal behavior, we need the infinite idealization. This argument is made in Batterman (2011). We address it in section 4.

In a case of explanatory irreducibility the higher level theory models a particular phenomenon in a conceptually novel manner, and the efficacy of this model cannot be explained by the lower level theory. However, this does not preclude the possibility that the phenomenon can be modeled within the lower level theory in a different way. There may be aspects of the phenomenon (such as, say, its macroscopic similarity to other phenomena) that cannot be captured by the descriptive resources of the theory, but the phenomenon itself can be described by the theory. Consider, for instance, the relationship between neuroscience and folk psychology. It might be argued that the latter is explanatorily irreducible to the former. Perhaps there is no viable neuroscientific account of why the reasons explanations common in folk psychology are successful, but a materialist about the mind could maintain that this is merely because the neuroscientific theory operates at too fine a scale to discern the patterns that ground this sort of explanation. In every token instance covered by the folk psychological explanation, there is nothing relevant going on that isn't captured by neuroscience. It's just that the way neuroscience describes what is going on isn't conducive to the construction or justification of reasons explanations. The patterns which the neuroscientific description fails to see are nonetheless wholly generated by processes describable using neuroscience.

A substance dualist, however, would argue that there is an even deeper failure of reduction going on here. The phenomena and processes described by neuroscience are by themselves inadequate to even generate the kinds of patterns that characterize reasons explanations. This is because the lower level theory does not have the resources to describe a crucial element of the ontological furniture of the situation, the mind or the soul. Here we have more than a mere case of explanatory irreducibility. We may call cases like this, where the lower level theory cannot even fully describe a phenomenon that can be modeled by the higher level theory, examples of *ontological irreducibility*.

This is probably the sense in which the British emergentists conceived of emergence (see McLaughlin (1992) for an illuminating analysis of this school of thought). With reference to phase transitions, this view is perhaps most starkly expressed in Batterman (2005). Batterman argues that the discontinuity in the thermodynamic potential at a phase transition is not an artifact of a particular mathematical representation of the physical phenomenon, but is a feature of the physical phenomenon itself. He says, "My contention is that thermodynamics is correct to characterize phase transitions as real physical discontinuities and it is correct to represent them mathematically as singularities." (ibid., p. 234) If there are genuine discontinuities in physical systems, it seems we could not represent them accurately using only continuous mathematical functions. So, since the statistical mechanics of finite systems does not give us discontinuities, it is incapable of fully describing this physical phenomenon. We can only approach an explanation of the phenomenon by working in the infinite limit. The idealization is a manifestation of the inability of the theory to fully describe the phenomenon of phase transitions in finite systems. We discuss these ideas further in section 3.3.

In the remainder of this paper, we discuss the status of these three notions of emergence – conceptual novelty, explanatory irreducibility and ontological as they apply to both the standard statistical mechanical notion of phase transitions and the treatment of critical phenomena by the renormalization group. These topics are treated separately because, as discussed above, the renormalization group introduces new issues bearing on the topic of emergence and reduction that go beyond issues involving infinite idealization in traditional statistical mechanics.

3 The Infinite Idealization in Statistical Mechanics

In the previous section, we discussed three ways in which the relationship between statistical mechanics and thermodynamics might be non-reductive. There is a hierarchy to these different senses of emergence set by the varying strengths of the assumptions about explanation required in order for them to represent a genuine failure of the core sense of reduction. Conceptual novelty is the weakest notion of emergence, explanatory irreducibility is stronger, and ontological irreducibility is stronger still. In this section, we discuss the case that can be made for phase transitions exemplifying each of these notions of emergence. We conclude that in the domain of ordinary statistical mechanics (excluding the renormalization group), the case for phase transitions being either ontologically or explanatorily irreducible is weak. The case for phase transitions being conceptually novel is stronger, but even here there are questions that can be raised.

3.1 Conceptual Novelty

A natural kind in a higher level theory is conceptually novel if there is no kind in any potential reducing theory that captures the same set of phenomena. Are thermodynamic phase transitions conceptually novel? That is, does the kind 'phase transition' have a natural counterpart kind in statistical mechanics? If we restrict ourselves to finite N systems, it's commonly believed that there is not a kind in statistical mechanics corresponding to phase transitions and that one can only find such a kind in infinite N statistical mechanics. *We believe, to the contrary, that no theory, infinite or finite, statistical mechanical or mechanical, possesses a natural kind that perfectly overlaps with the thermodynamic natural kind.* Yet if one relaxes the demand of perfect overlap, then then there are kinds – even in finite N statistical mechanics – that overlap in interesting and explanatorily powerful ways with thermodynamic phase transitions. Strictly speaking, thermodynamic phase transitions are conceptually novel; more loosely speaking, they are not.

To begin, one might wonder in what sense 'phase transition' is a kind even in thermodynamics. After all, there are ambiguities in the way we define phases. Is glass a super-cooled liquid or a solid? It depends on which criteria one uses and no set seems obviously superior. Be that as it may, the notion of a transition is relatively clear in thermodynamics, and it is defined, as above, as a discontinuity in one of the thermodynamic potentials. Let's stick with this.

Now, is the kind picked out by Def 1 the counterpart of the thermodynamic definition? Despite many claims that it is, Def 1's extension is clearly very different than that given by thermodynamics. To mention the most glaring difference – and on which, more later – there are many systems that do not have well-defined thermodynamic limits. Do they not have phase transitions? One can define words as one likes, but the point is that there are many systems that suffer abrupt macroscopic changes, changes that thermodynamics would count as phase transitions, but which do not have thermodynamic limits. Systems with very long-range interactions are prominent examples. But in fact the conditions on the existence of a thermodynamic limit are numerous and stringent, so in some sense most systems don't have thermodynamic limits. A strong case can be made that Def 1, as a result, provides at best sufficient conditions for a phase transition, and not necessary conditions.

How does finite N statistical mechanics fare? The conventional wisdom is that finite N statistical mechanics lacks the resources to have counterparts of thermodynamics phase transitions. However, we believe that people often assent to this claim too quickly. One of the more interesting developments in statistical mechanics of late has been challenges to ordinary statistical mechanics from the realms of the very large and the very small. These are regimes that test the applicability of normal Boltzmann-Gibbs equilibrium statistical mechanics. The issues arise from the success of statistical mechanical techniques in new areas. In cosmology, statistical mechanics is used not only to explain the inner workings of stars but also to explain the statistical distribution of galaxies, clusters and more. In these cases, the force of interest is of course the gravitational force, one that is not screened at short distances like the Coulomb force. Systems like this do not have a well-defined thermodynamic limit, often aren't approximately extensive, suffer negative heat capacities, and more (see Callender (2011) for discussion). There has also been an extension of statistical mechanical techniques to the realm of the small. Sodium clusters obey a solid-like to liquid-like "phase transition", Bose-Einstein condensation occurs, and much more. These atomic clusters have been surprisingly amenable to statistical mechanical treatment, yet they too don't satisfy the conditions for the application of the thermodynamic limit. Physically, one way to think about what is happening here is that in small systems a much higher proportion of the particles reside on the surface, so surface effects play a substantial role in the physics. As a result, these systems also raise issues about extensivity, negative specific heats, and much more.[4]

These systems are relevant to our concerns here for a very simple reason: they appear to have phase transitions, yet lack a well-defined thermodynamic limit, so Def 1 seems inadequate. Orthogonal to our philosophical worries about

[4] For the thermodynamic limit to exist, two conditions on the potential in the Hamiltonian must be satisfied, one on large distances, one on small distances. These extensions can be viewed as challenges in either length scale. In another sense, however, one can view both types of systems as unified together as "small" systems. If we define a system as "small" if its spatial extension is less than the range of its dominant interaction, then even galactic clusters are small.

reduction, there are also purely physical motivations for better understanding thermodynamic phase transitions from the perspective of finite statistical mechanics. Naturally, some physicists appear motivated by both issues, the conceptual and the physical:

> Conceptually, the necessity of the thermodynamic limit is an objectionable feature: first, the number of degrees of freedom in real systems, although possibly large, is finite, and, second, for systems with long-range interactions, the thermodynamic limit may even be not well defined. These observations indicate that the theoretical description of phase transitions, although very successful in certain aspects, may not be completely satisfactory. (Kastner 2008, 168)

As a result of this motivation, there are already several proposals for finite-particle accounts of phase transitions. These are sometimes called *smooth phase transitions.* The research is ongoing, but what exists already provides evidence of the existence of thermodynamic phase transitions in finite systems. There are many different schemes, but we'll concentrate on the two most well-known.

3.1.1 Back Bending

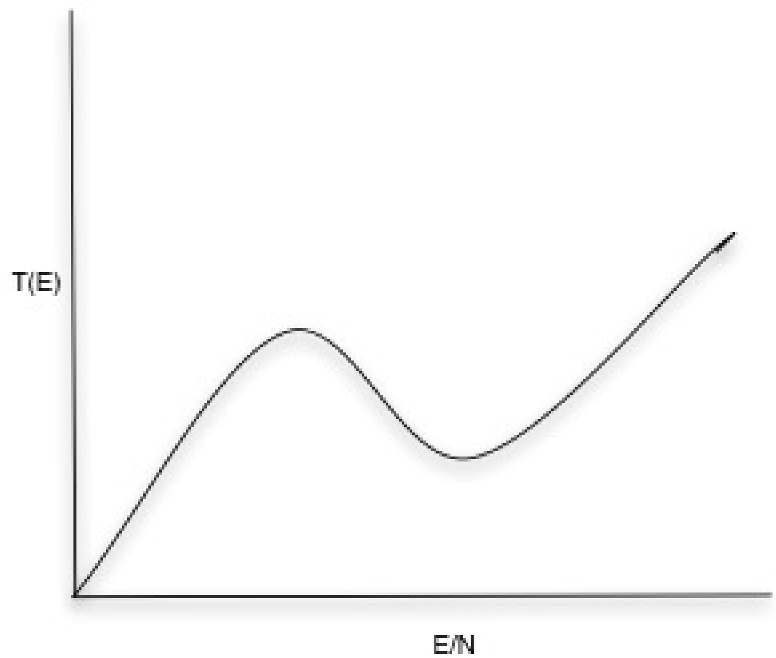
Inspired in part by van der Waals theory and its S-shaped bends, this theory has been developed by Wales and Berry (1994), Gross and Votyakov (2000) and Chomaz, Gulminelli, and Duflot (2001). Unlike in the traditional theory of phase transitions, here the authors work with the microcanonical ensemble, not the canonical ensemble. The general idea is that the signatures of phase transitions of different orders are read off from the curvature of the microcanonical entropy, $S = k_b \ln \Omega(E)$, where $\Omega(E)$ is the microcanonical partition function. In particular, if written in terms of the associated caloric curve, $T(E) = 1/\partial_E \ln[\Omega(E)]$, we can understand a first-order transition as a "back-bending" curve, where for a given value of $T(E)$ one can have more than one set of values for E/N (see figure 1). For our illustrative purposes we'll use this as our definition:

(Def 2) A first-order phase transition occurs when there is "back bending" in the microcanonical caloric curve.

(Def 2) is equivalent to the entropy being convex or the heat capacity being negative for certain values. As expected, back-bending can be seen in finite-N systems. So with Def 2 we have an alternative criterion of phase transitions that nicely characterizes phase transitions even in systems that do not have thermodynamic limits. We hasten to add that the theory is not exhuasted by a simple definition. Rather, the hope – which has to some extent been realized –is that it and its generalizations can predict and explain both continuous phase transitions and also phase transitions in systems lacking a thermodynamic limit.

Def 2 is rather striking when one realizes that it is equivalent to a region of negative heat capacities appearing. The reader familiar with the van Hove

Figure 1: Backbending of the caloric curve.



theorem may be alarmed, for that theorem forbids back-bending in the thermodynamic limit. Since our concerns are about the finite case, this in itself isn't troubling. But if one hopes that this definition goes over to the infinite N definition in the thermodynamic limit, where ensemble equivalence holds for many systems, this might be a problem: the canonical ensemble can never have negative heat capacity, whereas the microcanonical one can, and yet they're equivalent for "normal" short-range systems in the thermodynamic limit. Doesn't "ensemble equivalence" in the infinite limit squeeze out these negative heat capacities? No, for one must remember that ensemble equivalence holds, where it does, only when systems aren't undergoing phase transitions. This is a point originally made by Gibbs (1902). And indeed, ensemble inequivalence can be used as a marker of phase transitions. What is happening is that the microcanonical ensemble has structure that the canonical ensemble cannot see; the regions of back-bending (or convex entropy, or negative heat capacity) are missed by the canonical ensemble. Yet since the canonical ensemble is equivalent to the microcanonical – if at all – only when no phase transition obtains, there is no opportunity for conflict with "equivalence" results.

This remark provides a clue to the relation between Def 1 and Def 2, and a way of thinking about the first as a sub-species of the second. When there is back-bending there is ensemble inequivalence. From the perspective of the canonical ensemble for an infinite system, this is where a nonanalyticity appears in the thermodynamic limit. It can "see" the phase transition in that case; but when finite it is blind to this structure. Def 2 can then be seen as more general, since it triggers the nonanalyticity seen in infinite systems and captured by Def 1, but also applies to finite systems.

Many more interesting facts have recently been unearthed about the relationships among back-bending, non-concave entropies, negative heat capacity, ensemble inequivalence, phase transitions, and non-extensivity. We refer the reader to Touchette and Ellis (2005) for discussion and references. For rigorous connections between Def 1 and Def 2, see Touchette (2006).

3.1.2 Distribution of Zeros

This approach grows directly out of the Yang-Lee picture. The Yang-Lee theorem is about the distribution of zeroes of the grand canonical ensemble's partition function in the complex plane. A critical point is encountered when this distribution "pinches" the real axis, and this can only occur when the number of zeroes is infinite. Fisher and later Grossmann then provided an elaborate classification of phase transitions in terms of the distribution of zeros of the canonical partition function in the complex temperature plane. Interested in Bose-Einstein condensation, nuclear fragmentation and other "phase transitions" in small systems, a group of physicists at the University of Oldenburg sought to extend this approach to the finite case (see Borrmann, Mülken, and Harting (2000)). For our purposes, we can define their phase transitions as

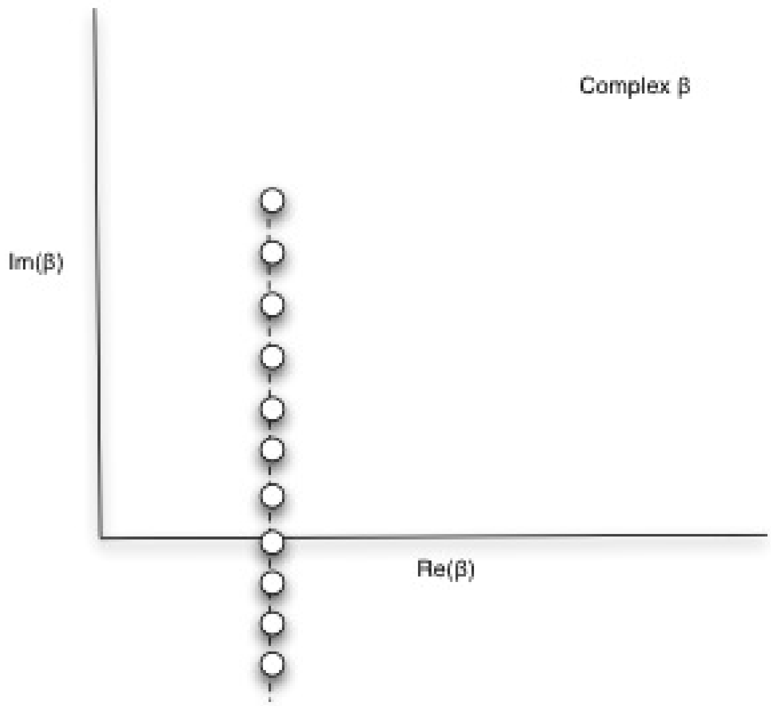
(Def 3) A phase transition occurs when the zeroes of the canonical partition function align perpendicularly to the real temperature axis and the density scales with the number of particles.

The distribution of zeros of a partition function contains a lot of information. The idea behind this approach is to extract three parameters (α,γ,τ_1) from the partition function that tell us about this distribution: τ_1 is a function of the number of zeros in the complex temperature plane, and it is positive for finite systems; γ is the crossing angle between the real axis and the line of zeros; and α is determined from the approximate density of zeros on that line. What happens as we approach a phase transition is that the distribution of zeros in the complex temperature plane "line up" and gradually gets denser and straighter as N increases.[5]

We stress that, as with the previous group, the physicists involved do not offer a stray definition but rather a a comprehensive theory of phase transitions in small systems. In particular, the Oldenburg group can use this theory to not only predict whether there is a phase transition, but also to identify the correct order of the transition. Their classification excels when treating Bose-Einstein condensation, as it reproduces the space dimension and particle number dependence of the transition order.

[5] A small movie of this occurring for small magnetic clusters is available at http://smallsystems.isn-oldenburg.de/movie.gif.

Figure 2: Distribution of zeroes in the complex inverse temperature ($\beta = 1/kT$) plane.



Like the approach using Def 2, the present approach works for both finite and infinite systems. For finite systems, $\tau_1$is always positive and we look for cases where $\alpha = \gamma$: these correspond to first-order transitions in finite systems. More complicated relations between α and γ correspond to higher-order transitions. For infinite systems, phase transitions of the first-order occur when $\alpha = \gamma = \tau_1 = 0$ and for higher-order when $\alpha > 0$. So the scheme includes the Def 1 case as a sub-species. One can then view Def 3 – or more accurately, the whole classification scheme associated with (α,γ,τ_1) – as a wider, more general definition of phase transitions, one including small systems, with Def 1 as a special case when the thermodynamic limit is legitimate.

What is the relationship between Def 2 and Def 3? It turns out that they are almost equivalent. Indeed, if one ignores a class of systems that may turn out to be unphysical, they are demonstrably equivalent; see Touchette (2006)[6]. The rich schemes of which these definitions form a part may not be equivalent, but on the question of what counts as a phase transition they will largely agree.

[6]This paper shows that yet another definition, one based on a bimodality of the energy distribution, is almost equivalent to Def 3. However, the bimodality definition is equivalent to Def 2, so the demonstration links Def 2 and Def 3.

As a result of the work on finite-N definitions – and while duly recognizing that it is very much ongoing – it seems to us that statistical mechanics is hardly at a loss to describe phase transitions in finite systems. The situation instead seems to us to be more subtle. No definition in statistical mechanics, infinite or finite, exactly reproduces the extension picked out by thermodynamics with the kind 'phase transition.' What one judges the best definition then hangs on what extension one wants to preserve. If focusing on thermodynamic systems possessing thermodynamic limits, then Def 1 is fine. Then the kind 'phase transition' is conceptually emergent relative to finite-N statistical mechanics. But if impressed by long-range systems, small systems, non-extensive systems, and "solidlike-to-liquidlike" mesoscopic transitions, then one of the finite-N definitions is necessary. Relative to these definitions, the kind 'phase transition' is not conceptually novel. If one wants a comprehensive definition, for finite and infinite, then the schemes described provide the best bet. Probably none of the definitions provide necessary and sufficient conditions for a phase transition that overlaps perfectly with thermodynamic phase transitions. That, however, is okay, for thermodynamics itself doesn't neatly characterize all the ways in which macrostates can change in an "abrupt" way.

In any case, we don't believe that conceptual novelty by itself poses a major threat to reductionism. After all even a (too) strict Nagelian notion of reduction can accommodate conceptual novelty (as long as the novel higher level kind is expressible as a finite disjunction of lower level kinds). Conceptual novelty is only a problem when you don't have explanatory reducibility of the conceptually novel kind, a question to which we now turn.

3.2 Explanatory Irreducibility

Explanatory irreducibility occurs, we said, when the explanation of a higher-level phenomenon requires a conceptual novelty, yet the reducing theory does not have the resources to explain why the conceptual novelty is warranted.[7] Where

[7] There are some potential connections between "explanatory irreducibility" and notions in the literature on idealization. In particular, depending upon how one understands Galilean idealization, it's possible that a conceptual novelty is explanatorily irreducible just in case it is not a "harmless" Galilean idealization. Coined by McMullin, a Galilean idealization in a scientific model is a deliberate distortion of the target system that simplifies, unifies or generally makes more useful or applicable the model. Crucially, a Galilean idealization is also one that allows for controlled "de-idealization." In other words, it allows for adding realism to the model (at the expense of simplicity or usefulness, to be sure) so that one can see that the distortions are justified by convenience and are not ad hoc. Idealizations like this are sometimes dubbed "controllable" idealizations and are widely viewed as harmless. What to make of such non-Galilean idealizations is an ongoing project in philosophy of science. One prominent idea – see, e.g., Cartwright (1983) or Strevens (2009) – is that the model may faithfully represent the significant causal relationships involved in the real system. The departure from reality need not then accompany a corresponding lack of faith in the deliverances of the model. It is possible that we could understand the standard explanation of phase transitions as a distortion that nonetheless successfully represents the causal relationships of the system. Perhaps the thermodynamic limit is legitimatized by the fact that surface effects aren't a difference-maker (in the sense of Strevens) in the systems of interest. We'll leave this line of thought to others to develop.

phase transitions are especially interesting, philosophically, lies in the fact that, at first glance, they seem to be a real-life and prominent instance of explanatory irreducibility. To arrive at this claim, let's suppose that the finite-N definitions surveyed above are theoretically inadequate. Assume that Def 1 is employed in the best explanation of the phenomena. Then we've already seen that no finite-N statistical mechanics can suffer phase transitions so understood. If the "reducing theory" is finite-N statistical mechanics, then we potentially have a case of explanatory irreducibility. But should the reducing theory be restricted to finite-N theory?

One quick way out of difficulty would be to include the thermodynamic limit as part of the reducing theory. However, this would be a cheat. The thermodynamic limit is, we believe, the production of another phenomenological theory, not a piece of the reducing theory. The ontology of the classical reducing theory is supposed to be finite N classical mechanics. Such a theory has surface effects, fluctuations, and more, but the thermodynamic limit squashes these out. More importantly, the ontology of the system in the thermodynamic limit is not the classical mechanics of billiard balls and the like. A quick and interesting way to see this point is to note that the thermodynamic limit is mathematically equivalent to the continuum limit (Compagner, 1989). The continuum limit is one wherein the size and number of particles is decreased without bound in a finite-sized volume. When thermodynamic emerges from this limit, it is emerging from a theory describing continuous matter, not atomistic matter. New light is shed on all that is regained in the thermodynamic limit if we see it as regained in the continuum limit. For here we do not see properties emerging from an atomic microworld behaving thermodynamically, but rather properties emerging from a continuum, a realm well "above" the atomic. For this reason, with respect to the reduction of thermodynamics to statistical mechanics, we do not see proofs that thermodynamic properties emerge in the thermodynamic limit as cases whereby thermodynamic properties are reduced to mechanical properties.

If this is right, then we have a potential case of explanatory irreducibility. The best explanation of the phenomenon of phase transitions contains an idealization whose efficacy cannot be explained from the perspective of finite-N theory. So are phase transitions actually explanatorily irreducible? The answer hangs on whether de-idealization can be achieved within finite-N statistical mechanics. We believe that it can be. We have already hinted at one possibility. If one could show that one or more of the finite-N definitions approximate in a controlled way Def 1, then we could view Def 1 as "really" talking about one of the other definitions. Indeed, this seems very much a live possibility with either Def 2 or Def 3 above. However, suppose we believe that this is not possible. Is there any other way of de-idealising the standard treatment of phase transitions? We believe that there is, and both Butterfield (2011) and Kadanoff (2009) point toward the right diagnosis.

Before getting to that, however, notice that the actual practice of the science more or less guarantees that *some* finite-N approximation must be available. In recent years there has been an efflorescence of computational models of statistical mechanical phenomena (see Krauth (2006)). Since we cannot simulate an infinite system, these models give an inkling of how we might approximate the divergences associated with critical behavior in a finite system. Consider, for instance, the Monte Carlo implementation of the Ising model (see, for instance, Wolff (1989)). The Monte Carlo method involves picking some probabilistic algorithm for propagating fluctuations in the lattice configuration of an Ising system as time evolves. Each run of the simulation is a random walk through the space of configurations, and we study the statistical properties of ensembles of these walks.

It might be argued that the system size in these simulations is effectively infinite, since the lattice is usually implemented with periodic boundary conditions. However, this periodicity should be interpreted merely as a computational tool, not as a simulation of infinite system size. The algorithm is supposed to study the manner in which fluctuations propagate through the lattice, but the model will only work if the correlation length is less than the periodicity of the system. If fluctuations propagate over scales larger than the periodicity, we will have a conflict between the propagation of fluctuations and the constraints set by the periodicity of boundary conditions. So the periodic boundary conditions should be interpreted as setting an effective system size. The model is only useful as long as the correlation length remains below this characteristic length scale. Unfortunately, the periodic boundary conditions also mean that the model is not accurate at the critical point, only close to it. As the correlation length approaches system size in a real system, surface effects become relevant, and the simulation neglects these effects.

Nonetheless, the Monte Carlo method does allow us to see how Ising systems approach critical behavior near the critical point. For instance, models exhibit the increase of correlation length as the critical point is approached, and the associated slow-down of equilibriation (due to the increased length over which fluctuations propagate). As we construct larger and larger systems, the model is precise closer and closer to the critical point, and we can see the correlation length get larger. We can also model the non-equilibrium phenomenon of avalanches, where the order parameter of the system changes in a series of sharp jumps as the external parameter in the Hamiltonian is varied. As an example, the magnetization of a magnetic material exhibits avalanches as the external field is tuned. The avalanches are due to the way in which fluctuations of clusters of spins trigger further fluctuations. At the critical point, we get avalanches of all sizes. Again, the approach to this behavior can be studied by examining the how the distribution of avalanches changes as the system approaches the critical point. These are just some examples of how finite models can be constructed to examine the behavior of a system arbitrarily close to the critical point. These models fail sufficiently close to criticality because they do not adequately deal with boundary effects. However, they do give an indication of how the behavior of large finite systems can be seen as smoothly approximating the behavior of infinite systems.

We now turn to a more explicit attempt to understand the idealization. Butterfield (2011, §3.3 and §7) thinks the treatment of phase transitions doesn't occasion any great mystery. We agree, and reproduce his mathematical analogy (with slight modifications) to illustrate the point. Consider a sequence of real functions g_N, where N ranges over the natural numbers. For each value of N, the function $g_N(x)$ is continuous. It is equal to -1 when x is less than or equal to $-1/N$, increases linearly with slope N when x is between $-1/N$ and $1/N$, and then stays at 1 when x is greater than or equal to $1/N$. The slope of the segment connecting the two constant segments of the function gets steeper and steeper as N increases.

While every member of this sequence of functions is continuous, the limit of the sequence $g_\infty(x)$ is discontinuous at $x = 0$. Now consider another another sequence of real functions of x, f_N. These are two-valued functions, defined as follows:

$$f_N(x) = \begin{cases} 0, & g_N(x) \text{ is continuous at } x \\ 1, & g_N(x) \text{ is discontinuous at } x \end{cases}$$

Given these definitions, $f_N(x)$ is the constant zero function for all N. If we just look at the sequence of functions, we would expect the limit of the sequence f_N as $N \to \infty$ to also be constant. However, if we construct $f_\infty(x)$ from $g_\infty(x)$ using the above definition, we will not get a constant function. The function will be discontinuous; it will take on the value 1 at $x = 0$. If one focuses only on f_N without paying attention to how it is generated from g_N, the behavior in the limit will seem mysterious and inexplicable given the behavior at finite N.

Imagine that we represent a physical property in a model in terms of $f_N(x)$ taking on the value 1, where N is a measure of the size of the physical system. This property can only be exemplified in the infinite-N limit, of course. And if we restricted ourselves to considering f_N when trying to explain the property, we would be at a loss. No matter how big N gets, as long as it is finite there is no notion of being nearer or further away from the property obtaining. We might conclude that the property is emergent in the infinite limit, since we can't "de-idealize" as we did in the case of extensivity and show how a finite system approximates this property. However, this is only because we are not paying attention to the $g_N(x)$. Realizing the relationship between f_N and g_N allows us to account for the behavior of f_N in the infinite limit from a finite system perspective, since there is a clear sense in which the functions g_N approach discontinuity as N approaches infinity.

We might put the point as follows. Suppose we have a theory of some physical property that utilizes the predicates g, N, and x. Suppose further that we are particularly interested in the rapid increase in $g_N(x)$ around $x = 0$ when N is large. Rather than analyze of $g_N(x)$ for particular finite values of N, it might make sense from a computational perspective to work with the infinite idealization $g_\infty(x)$, where the relevant behavior is stark and localized at $x = 0$. We may introduce a new "kind" represented by the predicate f that picks out the phenomenon of interest in the infinite limit. This kind is conceptually novel to the $\{g, N, x\}$ framework. Indeed, one can imagine a whole theory written in terms of f, without reference to g. Using such a theory it could be difficult to see how f is approximated by some function of finite-N. Because f is two-valued, the property it represents will appear to just pop into existence in the infinite limit without being approximated in any way by large finite systems. Restricted to f (and hence $g_\infty(x)$), one wouldn't have the resources present to explain how f emerges from the shape of g when N is finite.

This is precisely what happens in phase transitions. As Butterfield shows, the example of f and g translates nicely into the treatment of phase transitions. The magnetization in an Ising model behaves like $g_N(x)$, where N is the number of particles and x is the applied field. For finite systems, the transition of the system between the two phases of magnetization occurs continuously as the applied field goes from negative to positive. In the infinite case, the transition is discontinuous. The sequence of functions f_N isolate one aspect of the behavior of the functions g_N – whether or not they are continuous. If we just focus on this property, it might seem like there is entirely novel behavior in the infinite particle case. The shape of $f_\infty(x)$ around $x = 0$ is not in any sense approximated or approached by f_N as N gets large. If it is the case that large finite systems can be successfully modeled as infinite systems, this might seem to be a sign of explanatory irreducibility. The success of the infinite particle idealization cannot be explained because the infinite particle function is not the limit of the finite particle function sequence f_N. The illusion of explanatory irreducibility is dispelled when we realize that any explanation involving f_∞ can be rephrased in terms of g_∞, and the latter function does not display inexplicably novel behavior. It is in fact the limit of the finite particle functions g_N. As N increases, g_N approaches g_∞ in a well-defined sense. At a sufficiently large but finite system size N_0, the resolution of our measuring instruments will not be fine-grained enough to distinguish between $g_{N_0}(x)$ and $g_\infty(x)$. We have an explanation, much like the one we have for extensivity, of the efficacy of the infinite idealization.

Recognizing that the predicate f only picks out part of the information conveyed by the predicate g dissolves the mystery. The new predicate is useful when we are working with the idealization, but it makes de-idealization a more involved process. To see the connection between a phase transition defined via Def 1 and real finite systems, one must first "undo" the conceptual innovation and write the theory as a limit of nascent functions. At that point one can then see that the idealization is an innocent simplification and extrapolation of what happens to certain physical curves when N grows large.[8]

3.3 Ontological Irreducibility

Ontological irreducibility involves a very strong failure of reduction, and if any phenomenon deserves to be called emergent, it is one whose description is ontologically irreducible to any theory of its parts. Batterman argues that phase transitions are emergent in this sense (Batterman, 2005). It is not just that we

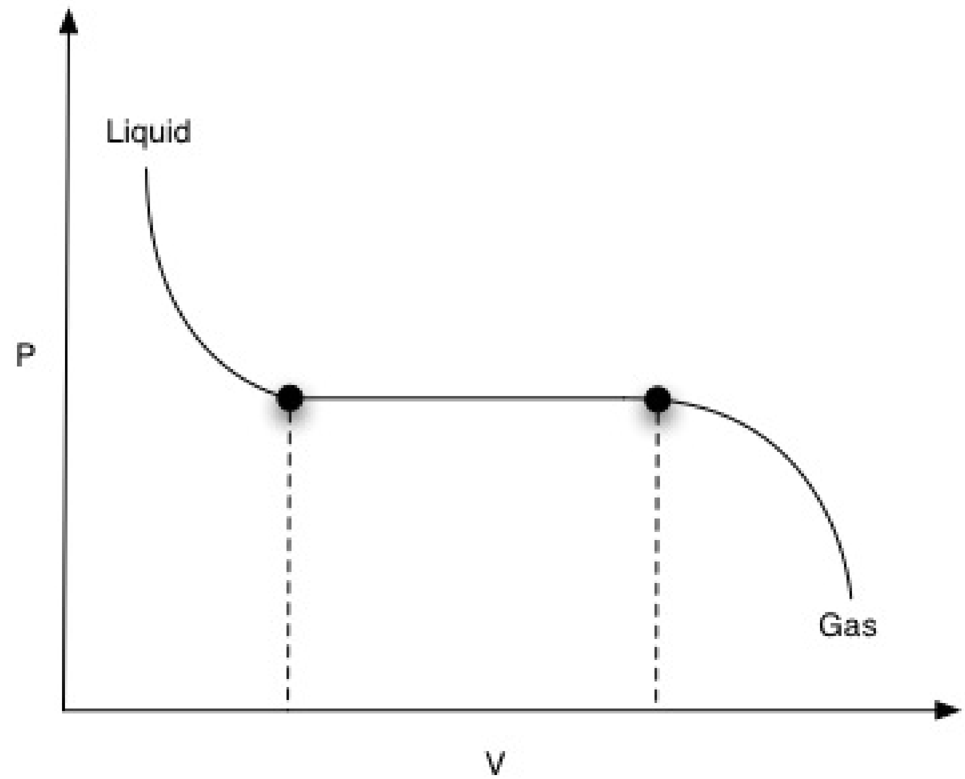
[8]Thanks to Jim Weatherall for kick-starting our thinking of phase transitions as delta functions that can be approximated by analytic functions, and to Jeremy Butterfield for kindly letting us use an advance copy of his 2010.

do not know of an adequate statistical mechanical account of them, we cannot construct such an account. Phase transitions, according to this view, are cases of genuine physical discontinuities. The discontinuity is there in nature itself. The thermodynamic representation of these phenomena as mathematical singularities is quite natural on this view. It is hard to see how else to best represent them. However, canonical statistical mechanics does not allow for mathematical singularities in thermodynamic functions of finite systems, so it does not have the resources to adequately represent these physical discontinuities. If the density of a finite quantity of water does as a matter fact change discontinuously at a phase transition, then it seems that statistical mechanics is incapable of describing this phenomenon, so the thermodynamics of phase transitions is genuinely ontologically irreducible.

Why think phase transitions are physically discontinuous? Batterman appeals to the qualitative distinction between the phases of fluids and magnets. Yet describing the distinction between the phases as "qualitative" is potentially misleading. It is true that the different phases of certain systems appear macroscopically distinct to us. A liquid certainly seems very different from a gas. However, from a thermodynamic perspective the difference is quantitative. Phases are distinguished based on the magnitudes of certain thermodynamic parameters. The mere existence of distinct states of the system exhibiting these different magnitudes does not suggest that there is any discontinuity in the transition between the systems. This is a point about the mathematical representation, but the lesson extends to the physical phenomenon. While it is true that the phases of a system are macroscopically distinct, this is not sufficient to establish that the physical transition from one of these phases to the other as gross constraints are altered involves a physical discontinuity.

In order to see whether there really is a discontinuity that is appropriately modeled as a singularity we need to understand the dynamics of the change of phase. So we take a closer look at what happens at a first order phase transition. Consider the standard representation of an isotherm on the liquid-gas P-V diagram at a phase transition (figure 3). The two black dots are co-existence points. At these points the pressure on the system is the same, but the system separates into two distinct phases: low-volume liquid and high-volume gas. The two co-existence points are connected by a horizontal tie-line or Maxwell plateau. On this plateau, the system exists as a two-phase mixture. It is here that the dynamics of interest takes place. However, the representation above is too coarse-grained to provide a full description of the behavior of the system at transition. This representation certainly involves a mathematical singularity: as the pressure is reduced, the volume of the system changes discontinuously. But a closer look at how the transition takes place demonstrates that this is just an artifact of the representation, and not an accurate picture of what is going on at the transition. The P-V diagram ignores fluctuations, but fluctuations are crucial to the transition between phases. The process by which this takes place is nucleation. When we increase the pressure of a gas above the co-existence point it does not instantaneously switch to a liquid phase. It continues in its gaseous phase, but this super-saturated vapor

Figure 3: $P-V$ diagram for a liquid-gas system at a phase transition.



is meta-stable. Thermal fluctuations cause droplets of liquid to nucleate within the gaseous phase. In this regime, the liquid phase is energetically favored, and this encourages the expansion of the droplet. However, surface effects at the gas-liquid interface impede the expansion. When the droplet is small surface effects predominate, preventing the liquid phase from spreading, but if there is a fluctuation large enough to push the droplet over a critical radius the free energy advantage dominates and the liquid phase can spread through the entire system. A full account of the gas-liquid transition will involve a description of the process of nucleation, a non-equilibrium phenomenon that is not represented on the equilibrium P-V diagram in figure 3.

Perhaps the nucleation of droplets from zero radius could be seen as an example of a physical discontinuity. However, an analysis of this process is not beyond the reach of finite particle statistical mechanics. We can study the nucleation of a new phase using the Ising model. As the external field crosses zero, simulations of the model show that initially local clusters of spins flip. Some of these clusters are too small, so they shrink back to zero, but once there is a large enough cluster – a critical droplet – the flipping spreads across the entire system and the new phase takes over. All of this is observable in a simple finite particle Ising system, so the phenomenon of nucleation can be described by statistical mechanics without having to invoke the thermodynamic limit. If it is the case that physical discontinuities cannot be accurately described by statistical mechanics, then we have good reason for believing there are no such discontinuities in the process of phase transition.

Even if we grant that phase transitions involve a physical discontinuity and can only be accurately represented by a mathematical singularity, the ontological irreducibility of the phenomenon does not follow. Very recently it has been shown that the microcanonical entropy, unlike the canonical free energy, can be non-analytic for finite systems. And indeed, a research program has sprung up based on this discovery that tries to link singularities of the microcanonical entropy to thermodynamic phase transitions (Franzosi, Pettini, and Spinelli (2000), Kastner (2008)). That program demonstrates that non-analyticities in the entropy are associated with a change in the topology of configuration space. Consider the subset of configuration space M_v that contains all points for which the potential energy per particle is lower than v. As v is varied, this subset changes, and at some critical values of v the topological properties of the subset change. This topology change is marked by a change in the Euler characteristic. For finite systems, there is a non-analyticity in the entropy wherever there is a topology change. For infinite systems there is a continuum of points at which the topology changes , so a straightforward identification of phase transitions with topology change is inappropriate.[9] Nevertheless, it is widely believed that there is some connection between these finite nonanalyticities and thermodynamic phase transitions.

This is a fledgling research program and there are still a number of open questions. It is unclear what topological criteria will be necessary and sufficient to define phase transitions, if any such criteria can be found. What is important for our purposes is that it is clear that the microcanonical ensemble does exhibit singularities even in the finite particle case, and that there is a plausible research program attempting to understand phase transitions in terms of these singularities. As such, it is certainly premature to declare that phase transitions are ontologically irreducible even if they involve genuine physical discontinuities. Statistical mechanics might well have the resources to adequately represent these discontinuities without having to advert to the thermodynamic limit.

[9]The problem with identifying these singularities with phase transitions in thermodynamics is that as N grows the order of the phase transition also increases, roughly as $N/2$. These transitions are far weaker than the ones encountered in thermodynamics, and in any case, unobservable in real noisy data unless N is really small.

4 The Infinite Idealization in the Renormalization Group

We have argued that there is good reason to think the use of the infinite limit in the statistical mechanical description of phase transitions does not show that the phenomenon is either ontologically or explanatorily irreducible. Here we examine whether similar claims can be made about the way the infinite idealization is used in renormalization group theory. While this theory is usually included under the broad rubric of statistical mechanics, there are significant differences between renormalization group methods and the methods characteristic of statistical mechanics. Statistical mechanics allows us to calculate the statistical properties of a system by analyzing an ensemble of similar systems. Renormalization group methods enter when correlations within a system extend over scales long enough to make straightforward ensemble methods impractical (see Kadanoff (2010a) for more on this distinction). The properties of the system are calculated not from a single ensemble but from the way in which the ensemble changes upon re-scaling. In statistical mechanics, the infinite idealization is important for the effect it has on a single ensemble (allowing non-analyticities, for instance). In renormalization group theory, the infinite idealization is important because it allows unlimited re-scaling as we move from ensemble to ensemble. The apparent difference in the use of the idealization suggests the possibility of significant philosophical distinctions. It will not do to blithely extend our conclusions about statistical mechanics to cover renormalization group theory.

We distinguish two different types of explanation that utilize the renormalization group framework. The first is an explanation of the critical behavior of particular systems, and the second is the universal behavior of classes of systems. The first type of explanation does not raise any fundamentally new issues that we did not already consider in our discussion of the explanatory reducibility of phase transitions in statistical mechanics. The second type of explanation does raise significant new issues, since we move from the examination of phenomena in particular systems to phenomena characterizing classes of systems. Batterman (2011) argues that the renormalization group explanation of universality is a case of explanatory irreducibility. While we might be able to tell a complex microphysical story that explains why a particular finite system exhibits certain critical behavior (the first type of explanation), we cannot account for the fact that many microscopically distinct systems exhibit identical critical behavior (the second type of explanation) without using the infinite idealization.

We begin with a brief discussion the first type of explanation: the renormalization group applied to the critical behavior of individual systems. We know from theory and experiment that there are large-scale correlations near the critical point, and that mean field theory does not work in these conditions. We need a method that can handle systems with long correlation lengths, and this is exactly the purpose that the renormalization group method serves. We idealize the correlation length of the system as infinite so that it flows to a fixed point under rescaling, and then calculate its critical exponent by examining the behavior of the trajectory near the fixed point.

This raises the question of why a system with a large correlation length can be successfully represented as a system with an infinite correlation length. If we have no explanation of the success of this idealization, we have a case of explanatory irreducibility. However, when we are focusing on the behavior of a particular system, any irreducibility in the renormalization group theory is inherited from orthodox statistical mechanics. The justification of the infinite correlation length idealization will coincide with the justification for the infinite system size idealization. Why does the renormalization group method need the infinite limit? Because it relies on the divergence of the correlation length at the critical point, which is impossible in a finite system. Why does the correlation length diverge? Because it is related to the susceptibility, which is a second derivative of the free energy and diverges. Why does the susceptibility diverge? Because there is a non-analyticity in the free energy. Explaining why (or whether) this non-analyticity exists takes us back to the statistical mechanical definition of phase transitions. If statistical mechanics can explain phase transitions reductively then the renormalization group does not pose an additional philosophical problem *when we focus on its application to particular systems.* It is true that the system must be idealized in order to employ renormalization group theory, but that idealization can be justified outside renormalization group theory.

The more interesting case is the second type of explanation, the explanation of universality. Without the renormalization group method, we might examine the behavior of individual finite system and discover that a number of such systems, though microscopically distinct, exhibit strikingly similar macroscopic behavior near criticality. However, this would not tell us why we should *expect* this macroscopic similarity, and so it is not really a satisfactory explanation of universality. The renormalization group method gives us a genuine explanation: when the correlation length diverges, there is no characteristic length scale. If the relevant parameters for the system vanish, as they do at criticality, the system will flow to a fixed point under repeated re-scaling. Fixed points can function as attractors, leading to similar critical behavior for a number of different systems.

If the system size is finite, the system will not flow to a fixed point. We might be able to show that a number of distinct large finite systems flow to points in system space that are very close to each other, but once again all that we have done is *revealed* the universality of critical (or near-critical) behavior. We haven't *explained* it. There is a generic reason to expect distinct infinite systems to flow to stable fixed points, but without mentioning fixed points there does not seem to be a generic reason to expect distinct finite systems to flow to points that are near each other. So it seems that fixed points play an indispensable role in the explanation of universal behavior. We can't "de-idealize" and remove reference to fixed points in the explanation, the way we can for non-analyticities in particular systems. Think back to Butterfield's example described in section 3.2. In that example, the apparent explanatory irreducibility of the behavior of f_∞ was resisted by re-phrasing our explanations in terms of g_∞, a function whose behavior in the limit isn't novel. In the case of the renormalization group, it seems that this move is unavailable to us. Fixed points are a novel feature that only appear in the infinite limit. There does not seem to be a clear sense in which the renormalization flow of finite systems can approximate a fixed point. A point is either a fixed point for the flow or it isn't; it can't be "almost" a fixed point. And unlike Butterfield's example, there does not seem to be a way of re-phrasing the explanation of universality in terms that *are* approximated by large finite systems.

So there is a strong *prima facie* case that universality is explanatorily irreducible. However, we do not believe that the case stands up to scrutiny. To see how it fails, we begin by showing that we *can* explain why finite systems exhibit universal behavior near criticality. However, this explanation does require the full resources of the renormalization group method, including fixed points. So it is not an explanation of the sort that we were contemplating above, one that does away with reference to fixed points. We will argue that this should not actually trouble the reductionist, but first we present the explanation.

Consider an Ising system extending over a finite length. When the system is rescaled, the separation between the nodes on the lattice increases. Since we are keeping the system size fixed, this means the number of nodes will decrease. So unlike the infinite system case, for a finite system the number of nodes is a parameter that is affected by rescaling. If the number of nodes is N, we can now think of $1/N$ as a relevant parameter (as defined in section 1.3). When we restrict ourselves to the infinite case, we are considering a particular hyper-surface of this new parameter space where $1/N$ is set to 0. However, since $1/N$ is a relevant parameter, perturbing the system off this hyper-surface (i.e. switching from the infinite to a finite system) will take the system away from the critical fixed point. This should be cause for concern. It seems there is no hope for an explanatory reduction. If even a slight perturbation off the $1/N = 0$ hyper-surface changes the critical behavior, how can we think of finite systems as approximating the behavior of infinite systems? As Kadanoff says, "if the block transformation ever reaches out and sees no more couplings in the usual approximation schemes... that will signal the system that a weak coupling situation has been encountered and will cascade back to produce a weak coupling phase [a trivial fixed point with $K = 0$]." (Kadanoff, 2010b, p. 47)

However, all is not lost. The difference between the behavior of finite and infinite systems depends on the correlation length. When the correlation length is very small relative to the system size, the finite system behaves much like the infinite system. The values of thermodynamic observables will not differ substantially from their values for an infinite system. The behavior of the finite system will only exhibit a qualitative distinction when the correlation length becomes comparable to the system size. This phenomenon is known as finite size cross-over (see Cardy (1996, ch. 4) for a full mathematical treatment). It is a manifestation of the fact that the behavior of the system is sensitive to the large-scale geometry of the system only when the correlation length is large enough to be comparable to the system size. The cross-over is controlled by the reduced temperature. As long as this parameter is above a certain value (given by an inverse power of the system size), the correlation length will be small enough that no distinction between finite and infinite systems will be measurable. It is only below the cross-over temperature that finite-size effects become significant and the system flows away from the critical point. For a large system, the cross-over temperature will be very small, and its difference from the critical temperature $t = 0$ may be within experimental error. So for a sufficiently large system, it is plausible that the infinite size approximation will work all the way to criticality. Renormalization group theory itself predicts this. A similar point is made in Butterfield (2011).

Cross-over theory also provides tools for estimating the changes to critical behavior that come from changing the geometry of the system by limiting its size. Adding system size as a parameter gives us a new scaling function for the susceptibility, a description of how the susceptibility changes with changes in relevant parameters. As described above, this scaling function gives a behavior for the susceptibility similar to the infinite limit as long as the ratio of correlation length to system size is low. It also allows us to predict the behavior of the susceptibility when this ratio becomes close to one. The susceptibility of a finite system will not diverge; it will have a smooth peak. The height of the peak of susceptibility scales as a positive power of size of the system. So for a large system, the susceptibility will be large but not infinite. In addition, the location of the peak shifts, and this shift scales as an inverse power of the size. This means that for a large system the difference between the critical temperature (the temperature at the critical fixed point of the infinite system) and the temperature at which it attains maximum susceptibility is very small. So for a macroscopic system, cross-over theory explains why it is a good approximation to treat the susceptibility as diverging at the critical point.

The point of this discussion is that we can tell an explanatory story about the circumstances under which particular large finite systems can be treated like infinite systems. If the cross-over temperature is sufficiently small, then limitations of our measurement procedures might make it difficult or even impossible to distinguish that the system does not flow to the critical point. However, this explanatory story does make reference to fixed points in system space. So the worry is that it is not a fully reductive account. We may have explained why individual finite systems can be successfully idealized as flowing to the critical fixed point, but have we accounted for the existence of the critical fixed point? We are taking for granted in our explanation the topological structure of system space, a topological structure that is to a large extent determined by the behavior of *infinite* systems.[10]

This is true, but does it lead to explanatory irreducibility? Is it illicit to include the topological structure of system space among the explanatory resources of our lower level theory? It would be if this structure involved an idealization whose efficacy could not be accounted for within the lower level theory. Isn't an

[10]Our thanks to Robert Batterman for pushing us on this point.

irreducible infinite idealization involved in the postulation of a renormalization flow with fixed points? It is not. As we have seen, the renormalization flow can be defined for all systems, finite and infinite alike, since $1/N$ can be introduced as a relevant parameter. Fixed points will appear on the hyper-surface where $1/N = 0$. There is no infinite idealization involved here. Of course, we are talking about infinite systems and how they behave under the renormalization flow, but this should not be problematic from a reductive point of view. The problem would arise if we *model* finite systems as infinite systems without explanation. But at this stage, when we are setting up the space and determining its topological characteristics, we are not modeling particular systems. In so far as finite systems are represented in our description of the space, they are represented as finite systems, and infinite systems are represented as infinite systems.

So the topological structure of the space can be described without problematic infinite idealization. When we try to explain the universality of critical behavior in finite systems, we do have to employ the infinite idealization, but as we have seen, this idealization is not irreducible if we can use the topological structure of system space in our reductive explanation. We can de-idealize for particular systems, and see why they can be treated as if they flow to the critical point. Understanding the behavior of infinite systems is crucial to explaining the behavior of finite systems, since we only get the fixed points by examining the behavior of infinite systems, but this in itself does not imply emergence. We agree with Batterman (2011) that mathematical singularities in the renormalization group method are information sources, not information sinks. We disagree with his contention that the renormalization group explanation requires the infinite idealization, and is thus emergent. It requires consideration of the behavior of infinite systems, but it does not require us to idealize any finite system as an infinite system. Any actual infinite idealizations in a renormalization group explanation can be de-idealized using finite-size cross-over theory. Locating fixed points does not require an infinite *idealization*, it just requires that our microscopic theory can talk about infinite systems, and indeed it can.

5 Conclusion

Phase transitions are an important instance of putatively emergent behavior. Unlike many things claimed emergent by philosophers (e.g., tables and chairs), the alleged emergence of phase transitions stems from both philosophical and scientific arguments. Here we have focused on the case for emergence built from physics. We have found that when one clarifies concepts and digs into the details, with respect to standard textbook statistical mechanics, phase transitions are best thought of as conceptually novel, but not ontologically or explanatorily irreducible. And if one goes past textbook statistical mechanics, then an argument can be made that they're not even conceptually novel. In the case of renormalization group theory, consideration of infinite systems and their singular behavior provides a central theoretical tool, but this is compatible with an explanatory reduction. Phase transitions may be "emergent" in some sense of this protean term, but not in a sense that is incompatible with the reductionist project broadly construed.*

References

Batterman, R. W. (2005). Critical phenomena and breaking drops: Infinite idealizations in physics. *Studies in History and Philosophy of Modern Physics*, *36*, 225-244.

Batterman, R. W. (2011). Emergence, Singularities, and Symmetry Breaking. *Foundations of Physics*, *41*(6), 1031-1050.

Borrmann, P., Mülken, O., & Harting, J. (2000). Classification of Phase Transitions in Small Systems. *Physical Review Letters*, *84*, 3511-3514.

Butterfield, J. (2011). Less is Different: Emergence and Reduction Reconciled. *Foundations of Physics*, *41*(6), 1065-1135.

Callender, C. (2011). Hot and Heavy Matters in the Foundations of Statistical Mechanics.
(Forthcoming in *Foundations of Physics*)

Cardy, J. (1996). *Scaling and Renormalization in Statistical Physics.* Cambridge: Cambridge University Press.

Cartwright, N. (1983). *How the Laws of Physics Lie.* Oxford: Clarendon Press.

Chomaz, P., Gulminelli, F., & Duflot, V. (2001). Topology of event distributions as a generalized definition of phase transitions in finite systems. *Physical Review E*, *64*, 046114.

Compagner, A. (1989). Thermodynamics as the continuum limit of statistical mechanics. *American Journal of Physics*, *57*, 106-117.

Franzosi, R., Pettini, M., & Spinelli, L. (2000). Topology and phase transitions: Paradigmatic evidence. *Physical Review Letters*, *84*, 2774-2777.

Gibbs, J. W. (1902). *Elementary Principles in Statistical Mechanics.* New York: Charles Scribner's Sons.

Griffiths, R. (1972). *Phase Transitions and Critical Phenomena* (Vol. 1; C. Domb & M. Green, Eds.). London: Academic Press.

Gross, D. H. E., & Votyakov, E. V. (2000). Phase transitions in "small" systems. *The European Physical Journal B - Condensed Matter and Complex Systems*, *15*, 115-126.

Kadanoff, L. (2009). More is the Same; Mean Field Theory and Phase Transitions. *Journal of Statistical Physics*, *137*, 777-797.

Kadanoff, L. (2010a). Relating Theories via Renormalization.
(Available at http://jfi.uchicago.edu/~leop/AboutPapers/RenormalizationV4.0.pdf)

Kadanoff, L. (2010b). Theories of Matter: Infinities and Renormalization.
(arXiv:1002.2985)

Kastner, M. (2008). Phase transitions and configuration space topology. *Reviews of Modern Physics*, *80*, 167-187.

*Our thanks to Daniel Arovas, Robert Batterman, Jeremy Butterfield, Sarang Gopalakrishnan, John Norton and Larry Sklar for helpful comments on an earlier draft of this chapter.

Krauth, W. (2006). *Statistical Mechanics: Algorithms and Computations*. Oxford: Oxford University Press.

Lebowitz, J. L. (1999). Statistical mechanics: A selective review of two central issues. *Reviews of Modern Physics*, *71*, S346-S347.

Liu, C. (1999). Explaining the Emergence of Cooperative Phenomena. *Philosophy of Science*, *66*, S92-S106.

McLaughlin, B. P. (1992). The Rise and Fall of British Emergentism. In A. Beckermann, H. Flohr, & J. Kim (Eds.), *Emergence or Reduction?* (p. 49-93). Berlin: Walter de Gruyter.

Melnyk, A. (2003). *A Physicalist Manifesto: Thoroughly Modern Materialism*. Cambridge: Cambridge University Press.

Nagel, E. (1979). *The Structure of Science: Problems in the Logic of Scientific Explanation*. Indianapolis: Hackett.

Strevens, M. (2009). *Depth: An Account of Scientific Explanation*. Harvard: Harvard University Press.

Touchette, H. (2002). When is a quantity additive, and when is it extensive? *Physica A: Statistical Mechanics and its Applications*, *305*, 84-88.

Touchette, H. (2006). Comment on "First-order phase transition: Equivalence between bimodalities and the Yang-Lee theorem". *Physica A: Statistical Mechanics and its Applications*, *359*, 375-379.

Touchette, H., & Ellis, R. S. (2005). Nonequivalent ensembles and metastability. In C. Beck, G. Benedek, A. Rapisarda, & C. Tsallis (Eds.), *Complexity, Metastability and Nonextensivity* (p. 81-88). World Scientific.

Wales, D. J., & Berry, R. S. (1994). Coexistence in Finite Systems. *Physical Review Letters*, *73*, 2875-2878.

Wolff, U. (1989). Collective Monte Carlo Updating for Spin Systems. *Physical Review Letters*, *62*, 361-364.