

Spontaneous Symmetry Breaking in Finite Systems

James D. Fraser

Please cite published version in *Philosophy of Science*
<https://doi.org/10.1086/687263>

Abstract

The orthodox characterisation of spontaneous symmetry breaking (SSB) in statistical mechanics appeals to novel properties of systems with infinite degrees of freedom, namely the existence of multiple equilibrium states. This raises the same puzzles about the status of the thermodynamic limit fuelling recent debates about phase transitions. I argue here that there are prospects of explaining the success of the standard approach to SSB in terms of the properties of large finite systems and consequently, despite initial appearances, the need to account for SSB phenomena does not offer decisive support to claims about the explanatory and representational indispensability of the thermodynamic limit.

1 Introduction

In both classical and quantum theory, systems with infinite degrees of freedom can have properties which are not found in any finite system. There has recently been much debate amongst philosophers of physics about the role these novel properties of infinite systems play in accounting for important physical phenomena.

In statistical mechanics, phase transitions, such as the boiling of water, are associated with non-analyticities in the free energy which only occur in the thermodynamic (or macroscopic) limit. Roughly speaking, this means taking the volume and particle number to infinity, while systems which actually display phase behaviour, such as kettles of water, are evidently finite in extent. Nevertheless, Batterman (2005) claims that the non-analyticities that occur in the thermodynamic limit are needed to successfully rep-

resent and explain phase transitions. Similarly, Ruetsche (2003) argues that unitarily inequivalent representations found in the thermodynamic limit of quantum statistical mechanics play an essential role in accounting for phase structure in quantum systems and need to be taken physically seriously.

This view of the thermodynamic limit has been contested by Butterfield (2011), Norton (2011) and Callender and Menon (2013) however. According to these authors, the thermodynamic limit is predictively and explanatorily successful in phase transition theory because of features infinite systems share with large finite systems—specifically, the non-analytic functions found in the limit are said to provide a good approximation to the values of macroscopic quantities in realistic, finite volume, models. On this view, phase transitions in concrete systems are sharp changes in their macroscopic properties, rather than physical discontinuities.

This paper examines a phenomenon which raises very similar interpretive issues, namely spontaneous symmetry breaking (SSB). As with phase transitions, the definition of SSB in statistical mechanics refers to properties afforded by the thermodynamic limit: in this case the non-uniqueness of the system’s equilibrium state. At first glance, this characterisation of SSB lends new support to claims about the indispensability of the thermodynamic limit. There are seemingly inescapable arguments to the effect that the characteristic features of SSB phenomena cannot occur in finite systems, and consequently the kind of de-idealisation strategy pursued by Butterfield and others in the phase transition debate is foiled. I argue here that this is not in fact the case: there are prospects of making sense of the success of the standard approach to SSB in terms of the behaviour of large finite systems. The implications of this programme for our understanding of the notion of SSB, and philosophical issues with which it has become embroiled, are discussed.

2 Spontaneous Symmetry Breaking in Statistical Mechanics

The notion of SSB is employed in a number of theoretical contexts, not always without controversy. Understanding the nature of SSB in quantum field theory (QFT), and the appeal to spontaneously broken gauge symmetry in standard presentations of the Higgs mechanism in particular, is an ongoing project in the foundations of physics.¹ My focus

¹See Friederich (2013) and references therein for recent critiques of the notion of spontaneously broken gauge symmetry, especially in the standard model. I bracket concerns about gauge symmetry breaking here, restricting my attention to the breaking of non-gauge symmetries in statistical mechanics.

here however is on the more mundane SSB phenomena described by statistical mechanics, the archetype being ferromagnetism. At high temperatures the electron spins inside a ferromagnetic crystal are uncorrelated and the magnetisation is effectively zero, but if the temperature is lowered to some critical value spins suddenly align giving rise to a net magnetisation in a particular direction. Though the electromagnetic interactions which produce these aligned states are isotropic, below the critical temperature the ferromagnet displays stable states which are not rotationally symmetric.

It will be useful for what follows to distinguish two assumptions underlying of the standard definition of SSB in statistical mechanics. First, there is the symmetry breaking part. The emergence of asymmetric behaviour at low temperatures is taken to mean that the system's equilibrium macrostate below the critical temperature is non-invariant under the relevant symmetry transformation (a global rotation in the case of the ferromagnet). Second, there is the spontaneous part. The sense in which this symmetry is broken spontaneously is typically spelled out via a contrast with so-called explicit symmetry breaking.² This describes situations in which asymmetric behaviour is produced by asymmetries in the underlying dynamics. Applying an external magnetic field to a ferromagnet, for instance, gives rise to aligned states, but in this case the rotational symmetry of interactions between spins is explicitly broken by the external field, which takes the form of a term in the system's Hamiltonian that is non-invariant under rotations. SSB is said to occur in the absence of any such non-invariant terms: in other words, the symmetry that is broken at low temperatures is taken to be an exact symmetry of the dynamics.

Taking these two claims together leads, inexorably, to the identification of SSB with the degeneracy of a system's equilibrium states and, just as inexorably, to the need to take the thermodynamic limit. In general, any transformation which leaves a system's Hamiltonian invariant maps its equilibrium states into other equilibrium states. Consequently, an equilibrium state can only be non-invariant under an exact symmetry of the dynamics if it is one of a set of equilibrium states related by the action of the relevant transformation. It turns out however that the equilibrium state of a system with a finite number of degrees of freedom is always unique; degenerate equilibrium states are only found in the thermodynamic limit. Indeed, while it is SSB in quantum theory which is often said to require infinite degrees of freedom, the thermodynamic limit plays an analogous, and equally crucial, role in modelling SSB in quantum and classical statistical mechanics. I explain the basic structure of this account with reference to the classical Ising model here, returning to additional issues raised by quantum theory in §4.2.

²See Castellani (2003) for a discussion of the distinction between spontaneous and explicit symmetry breaking.

In this famous model one takes a cubic (or, more generally, hypercubic) lattice, with N sites and volume V , and associates with each site i a variable $s_i = \pm 1$. The values of s_i can be interpreted as a naive representation of the orientation of spins inside a ferromagnetic crystal: $s_i = 1$ corresponding to a spin aligning ‘up’; $s_i = -1$ corresponding to a spin aligning ‘down’. In keeping with this picture, we can identify the quantity, $m = \sum_i s_i/V$, with the (volume averaged) magnetisation of the system.³ The basic Hamiltonian consists of a single term representing short ranged interactions that make it energetically favourable for spins to align parallel with one another:

$$H = -J \sum_{i,j} s_i s_j, \tag{1}$$

where J is a positive constant parameterising the interaction strength and the sum is taken over ‘nearest neighbour’ spin pairs. The key property of this Hamiltonian for our purposes is that it is invariant under a ‘spin flip’ transformation $s_i \rightarrow -s_i$: it is this reflection symmetry that is spontaneously broken at low temperature.

The system’s equilibrium state is represented by a probability distribution over its microstates, $S_V = (s_1, \dots, s_N)$, known as the canonical ensemble:

$$P(S_V) = \frac{e^{-\beta H(S_V)}}{Z}, \tag{2}$$

where $\beta = 1/k_B T$ is the inverse temperature and Z is the partition function. This distribution is clearly unique and, in fact, the canonical ensemble assigns a unique equilibrium state to any system with finite degrees of freedom. Things become more interesting in the thermodynamic limit however. Taking $N, V \rightarrow \infty$, with N/V held constant, generates an infinite lattice with a divergent total energy so (2) is no longer a well defined probability distribution. In the classical context the canonical ensemble is typically generalised to such infinite systems by means of the Gibbs measure formalism.⁴ Roughly speaking, a Gibbs measure is a probability measure over a system’s microstates which behaves locally like the canonical ensemble—in the case of the Ising model, any finite sub-region of the lattice is stipulated to have a probability distribution defined over its associated microstates which takes the form of (2). Stated precisely, this definition

³A note on terminology is perhaps needed here. In statistical mechanics the magnetisation is often identified with the expectation value of the quantity I have called m . Part of the reason for this is that $\langle m \rangle$ and m effectively coincide in the thermodynamic limit, as fluctuations vanish. It will be crucial to my discussion in §4 to keep these quantities distinct however, and I refer to the later quantity as the magnetisation here.

⁴Presentations of the Gibbs measure formalism, and its application to SSB and phase transitions, can be found in Georgii (1988) and Lebowitz (1999).

agrees with the canonical ensemble for systems with finite degrees of freedom, but infinite systems also admit Gibbs measures in this sense.

In fact, there can be a set of distinct Gibbs measures in the thermodynamic limit. An elegant argument (originally due to Peierls) illustrates how this can occur in the case of the Ising model on a two dimensional lattice.⁵ Strictly speaking, the Hamiltonian of a finite Ising model should include an additional term representing the effects of external spins at the boundary of the lattice. If we take all of the spins at this boundary to be ‘up’ it can be shown that, below some critical temperature, the expectation value of s_i at the centre of the lattice is positive no matter how large V is. This means that when we take $V \rightarrow \infty$ the boundary goes to spacial infinity while the expected magnetisation of the system converges to a positive value. Running the same argument with ‘down’ spins at the boundary, we find that there are two distinct Gibbs measures, P_+ and P_- , on the infinite lattice which assign positive and negative values to the magnetisation respectively. Normalised linear combinations of these measures are also Gibbs measures of the system. If we take the thermodynamic limit of a finite lattice with boundary conditions that are invariant under spin flips,⁶ for instance, the Gibbs measure converges to $1/2P_+ + 1/2P_-$. Only P_+ and P_- satisfy the properties required of genuine equilibrium states however: in general, systems with infinite degrees of freedom can have a convex set of Gibbs measures whose extremal elements correspond to macroscopically distinct equilibrium states of the system.⁷

The following picture of SSB phenomena emerges. At high temperatures the Ising model has a single Gibbs measure in the thermodynamic limit corresponding to a macrostate with zero magnetisation that respects the reflection symmetry of the Hamiltonian. Below the critical temperature however the system will enter one of the equilibrium states associated with the P_+ or P_- measure in which spins align in a particular direction, breaking this symmetry. More generally, a symmetry is said to be spontaneously broken at low temperatures if there are a set of distinct equilibrium macrostates in the thermodynamic limit which are mapped into each other under the relevant transformation. A spontaneously broken rotational symmetry for instance, as in the case of a more realistic ferromagnetic system, is associated with an infinite set of equilibrium states related to one another by rotations. Though there are differences in its formal implementation

⁵See Kindermann and Snell (1980), and references therein, for the details. Note that this argument does not go through for a one dimensional classical Ising model and rigorous results are, for the most part, not available in three or more dimensions.

⁶In practice this typically means free or periodic boundary conditions, see Kinderman and Snell (1980, 34-35).

⁷For a discussion of the properties of extremal Gibbs measures which justify equating them with equilibrium states see Georgii (1988, 119).

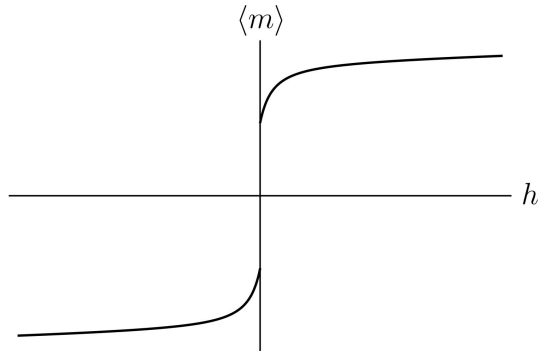


Figure 1: Sketch of the expected magnetisation as a function of the external magnetic field, h , for an Ising model on an infinite two dimensional lattice.

(described in §4.2) this account of SSB survives essentially unchanged in quantum statistical mechanics.

This approach to modelling SSB raises interpretive problems, as we shall soon see, but as a piece of first order science its success is undeniable.⁸ The asymmetric equilibrium states which appear in the thermodynamic limit give an accurate quantitative description of the macroscopic properties of concrete systems at low temperatures. Furthermore, this account clearly gets something right about the qualitative character of SSB phenomena. The availability of multiple equilibrium states at low temperatures, each on an equal footing with respect to the dynamics, captures the most intuitive sense in which ferromagnetic symmetry breaking, and analogous behaviour, is ‘spontaneous’; namely that the system has an equal chance of ‘choosing’ one of a number of possible asymmetric configurations at the critical temperature. The philosophically interesting question is not whether the infinite systems obtained in the thermodynamic limit are genuinely predictive and explanatory, but why.

Before addressing this issue in earnest it is worth touching on the close connections with the way that phase transitions are represented in statistical mechanics. As has already been mentioned, phase transitions are associated with non-analyticities in macroscopic observables, and ultimately the free energy, which only occur in the thermodynamic limit. The expected magnetisation of the infinite Ising model, for instance, is a dis-

⁸On the other hand, it is worth bearing in mind that the standard account of SSB has limitations which may be important for assessing its interpretive significance. An infinite system cannot always be constructed by taking the volume to infinity in the manner sketched here—many systems, especially those with long ranged interactions, do not have a well defined thermodynamic limit. If such systems display SSB phenomena this will clearly be problematic for the thesis (discussed below) that the thermodynamic limit plays an indispensable role in accounting for this behaviour.

continuous function of the external magnetic field, h , below the critical temperature, corresponding to the first order phase transition observed in real ferromagnets when the direction of an external field is reversed (figure 1). The appearance of these non-analyticities is intimately related to the non-uniqueness of the system's equilibrium state at the phase boundary. In the case of the Ising model, it is the existence of the P_+ and P_- equilibrium states that allows the expected magnetisation to change discontinuously at $h = 0$. Indeed, the non-uniqueness of a system's equilibrium states is sometimes used as an alternative definition of a phase transition (Lebowitz, 1999). It is not surprising, therefore, that SSB and phase transitions pose similar conceptual problems. Nevertheless, the standard account of SSB does raise new issues in the debate about the interpretive status of the thermodynamic limit, as I argue in the next section.

3 Interpreting the Thermodynamic Limit

There is something puzzling about the approach to SSB just described. The infinite volume models obtained in the thermodynamic limit seem to be highly idealised representations of the concrete systems which exhibit SSB phenomena—real ferromagnets clearly have a finite volume and number of constituents. But the explicit appeal to novel properties of infinite systems in the standard definition of SSB seems to problematise an intuitive picture of how idealised models relate to their targets. In calculating the electric field produced by a long charged wire one often treats it as being infinitely long. In so far as this model is predictive and explanatory it is presumably because it captures relevant features of the target system that are preserved when this infinite length idealisation is corrected—we can easily verify in this case that the electric field associated with an infinite charged wire provides a good approximation to the field close to a long finite length wire. The idea that idealised models are successful because the behaviour they aim to describe is robust under 'de-idealisation' is most famously associated with McMullin's notion of Galilean idealisation (McMullin, 1985). The role of the thermodynamic limit in the standard approach to SSB does not seem to fit this framework however. The degenerate equilibrium states which signal the onset of SSB completely disappear if we remove the thermodynamic limit and return N and V to realistic finite values.

As I have stressed, this is essentially the same problem engendered by the appeal to non-analyticities in the free energy in phase transition theory. As the debate about this case shows, there are a number of ways that one might respond.

Some philosophers have concluded that the thermodynamic limit plays an indispensable

role in representing and explaining physical phenomena. There are a number of ways that this claim might be spelled out. On the one hand, Batterman and others take the thermodynamic limit to be an instance of a so-called ineliminable or essential idealisation which is not amenable to the kind of de-idealisation analysis sketched above for the infinite wire model.⁹ In particular, Batterman (2005) claims that, despite its status as an idealisation, the non-analyticities afforded by the thermodynamic limit correspond to genuine physical discontinuities that manifest during a change of phase. The analogous claim in the case of SSB would presumably be that the degenerate equilibrium states found in the thermodynamic limit are real features of macroscopic systems which can only be captured by idealised infinite volume models. Earman (2004) suggests a very different way of understanding how the thermodynamic limit could be representationally indispensable. As he points out, the new mathematical structures found in the limit of infinite volume also occur in a field theoretic context. Assuming that the world is, at base, a quantum field we could take the novel properties that appear in the thermodynamic limit to correspond to features of this more fundamental reality which are missed by models with finite degrees of freedom. On this view the thermodynamic limit turns out not to be an idealisation after all!¹⁰ While these two interpretations clearly differ in important respects they both take novel properties afforded by the thermodynamic limit to have direct physical significance. Call this the indispensabilist reading of the thermodynamic limit.

In the recent debate about phase transition theory however a number of philosophers have argued that a more conservative response to the puzzle is available. As I understand them, Butterfield (2011) and Callender and Menon (2013) want to say that, despite appearances, the thermodynamic limit is an ordinary idealisation, not fundamentally unlike the limit of infinite length in the charged wire case.¹¹ Note that the fact that novel properties obtain in an idealised limit does not, in itself, establish that they are indispensable in any interesting sense. The electric field produced by an infinitely long charged wire has a translation symmetry that is not respected by any finite length

⁹It is this reading of the indispensability of the thermodynamic limit which is sometimes coupled with claims about emergence and the failure of reductive explanation. By contrast, Earman's appeal to field theory is thoroughly reductive in character, which suggests that the connection between the interpretive status of the thermodynamic limit and broader debates about emergence and reduction is more subtle than is sometimes recognised.

¹⁰Thought I do not press this point here, one can question whether either way of understanding the indispensability of the thermodynamic limit really accounts for its success. Appealing to the still inchoate notion of essential idealisation arguably raises a new set of puzzles—see Fraser (2015) and Shesh (2013)—and objections to invoking field theory in the phase transition debate, which apply just as forcefully in the case of SSB, can be found in Butterfield (2011, 1078) and Mainwood (2006, 228-231).

¹¹Norton (2011) gives a slightly different, but closely allied diagnosis of the role of the thermodynamic limit in phase transition theory.

wire, but nobody claims that this property is explanatorily essential, much less physically real. The reason being that the success of the infinite wire model is very plausibly accounted for by other properties which are robust when the infinite length idealisation is removed. Similarly, according to these authors, it is not the novel properties of infinite systems but those they share with large finite systems that are doing the real explanatory and representational work in applications of the thermodynamic limit. Call this the deflationary reading of the thermodynamic limit.

The basic observation motivating this position in the phase transitions debate has been that, while the non-analyticities in the free energy disappear when the thermodynamic limit is removed, this function can still approximate the changes in macroscopic properties of a finite system. Recall that, below the critical temperature, the expected magnetisation of the Ising model becomes a discontinuous function of the external magnetic field in the thermodynamic limit, signalling a first order phase transition (figure 1). General theoretical considerations, and evidence from numerical simulations, suggest that the expected magnetisation of a finite Ising model has an abrupt change in sign at $h = 0$, which becomes increasingly steep as N increases, so that, if the lattice is sufficiently large, it will be very well approximated by the discontinuous magnetisation function of an infinite Ising lattice. We seem to have a simple schema for explaining the success of the orthodox approach to phase transitions in terms of the behaviour of large finite systems: a real phase transition is a sharp but smooth change in a large system's macroscopic properties that is approximated by the non-analytic free energy found in the infinite volume limit to accuracies within acceptable error.

There is a great deal more to be said about this dispute. Perennial philosophical questions about reduction, idealisation and scientific explanation loom large in the debate and the need to account for the universal behaviour observed in disparate systems during continuous phase transitions raises further puzzles about the explanatory role of the thermodynamic limit.¹² The focus of this paper however is on what SSB brings to the table.

The account of SSB detailed in the previous section presents a new challenge for the deflationary reading of the thermodynamic limit and seems to provide stronger grounds for claims about its indispensability. Pointing to the fact that non-analytic functions can approximate analytic ones does not help in this case. The issue now is not why we can legitimately model a ferromagnet's macroscopic properties as changing discontinuously

¹²Batterman has long argued that the role of the thermodynamic limit in renormalization group explanations of the universality of critical exponents is not susceptible to de-idealisation; see, for instance, Batterman (2010). See Callender and Menon (2013) for an attempt to develop a deflationary reading of the thermodynamic limit in this context.

under variations in the external magnetic field but why, in the absence of an external magnetic field, the system can be successfully represented as having multiple available equilibrium states at low temperatures, despite the fact that the equilibrium state of any finite system is unique. Not only has a putative de-idealisation story for the thermodynamic limit not been provided in the case of SSB, there seem to be powerful reasons why no such account can be given even in principle. As we have seen, if a system's equilibrium state is unique it must be invariant under the same transformations as its Hamiltonian. This means that the expected values of macroscopic observables must be consistent with these symmetries; the expected magnetisation of a finite Ising model, for instance, is always zero no matter how large the lattice is, which appears to flatly contradict what we see in real ferromagnets. *Prima facie* then, the description of SSB afforded by the thermodynamic limit cannot approximate behaviour which is already seen in large finite systems as they simply lack the resources needed to accommodate stable symmetry breaking macrostates.

As a result, it appears that the indispensabilist reading of the thermodynamic limit is vindicated: if the sort of de-idealisation account supplied in the phase transition debate is ruled out, it seems the novel properties of infinite systems appealed to in the standard approach to SSB must be taken physically seriously, in one way or another, in order to account for its success.

4 Spontaneous Symmetry Breaking in Finite Systems

This kind of no-go argument against the possibility of providing a de-idealisation story for the thermodynamic limit in the standard account of SSB is misleading however, as I will now argue. SSB behaviour only appears to be impossible in finite systems if the modelling assumptions implicit in the standard definition of SSB are treated as sacrosanct. Relaxing the requirement that the system's equilibrium state at low temperatures is asymmetric (the symmetry breaking part), or the demand that the broken symmetry is exact (the spontaneous part), allows us to see how systems with finite degrees of freedom can exhibit the kind of behaviour observed in real ferromagnets. §4.1 explores the former route. I point out that a system with a unique, symmetric, equilibrium state can still be expected to exhibit asymmetric states over long time periods. This suggests one way of treating the thermodynamic limit as an ordinary idealisation which can be made fairly precise in the case of some simple classical systems. §4.2 introduces the challenges quantum theory raises for understanding how SSB can occur in finite systems—I argue that they are not as insurmountable as is sometimes thought. §4.3 explores the latter route. Landsman (2013) has developed an approach to SSB in finite quantum systems

based on the instability of their equilibrium states under asymmetric perturbations to the dynamics. The upshot of the discussion is that, despite initial pessimism, there is a viable programme for explaining the success of the standard approach to SSB in terms of the properties of finite systems.

4.1 Approach I: Long Lived Asymmetric States

The fact that a finite system's equilibrium state is invariant under the symmetries of its Hamiltonian does not, in fact, rule out SSB type behaviour: an equilibrated finite system can still be expected to exhibit asymmetric states over very long time periods. Once again, the classical Ising model provides a useful reference point here. The canonical ensemble of the Ising model assigns a probability distribution, $P(m)$, to the possible magnetisations of the system. Provided the boundary conditions imposed on a finite lattice do not break the interaction's reflection symmetry, $P(m)$ will also respect this symmetry, which, as has already been noted, implies that the expected magnetisation is zero. This does not mean however that $P(m)$ must take the form of a normal distribution centred at $m = 0$; it can have maxima at positive and negative m . The system would then be expected to spend most of its time in magnetised microstates.

In the case of the Ising model on a two dimensional lattice it can be shown that this is exactly what happens below the critical temperature. As was mentioned in §2, if the boundary conditions are suitably symmetric the Gibbs measure of the system is an equal linear combination of the P_+ and P_- measures. These extremal measures assign probability one to $m = \pm a$ (for some positive a) in the thermodynamic limit, and can be shown to satisfy central limit theorems as $N \rightarrow \infty$.¹³ This means that, below the critical temperature, $P(m)$ has two 'humps' at $m \pm a$, which become increasingly sharply peaked as N increases, finally converging to Dirac measures at $N = \infty$ (figure 2). Though rigorous results are less forthcoming for more realistic models there are general theoretical grounds for believing that analogous behaviour will hold whenever the Gibbs measure is non-unique in the thermodynamic limit. The generic result that statistical fluctuations scale like $1/\sqrt{N}$ (where N is now the number of constituents) implies that the canonical ensemble of a large finite system will have most of its mass in microstates which assign values to macrostate variables close to those associated with the degenerate equilibrium states that appear in the limit.

Consequently, we can expect such systems to spend long time periods in symmetry breaking states. A large finite Ising model below the critical temperature will enter one of the

¹³See Kindermann and Snell (1980, 34-62) and references therein.

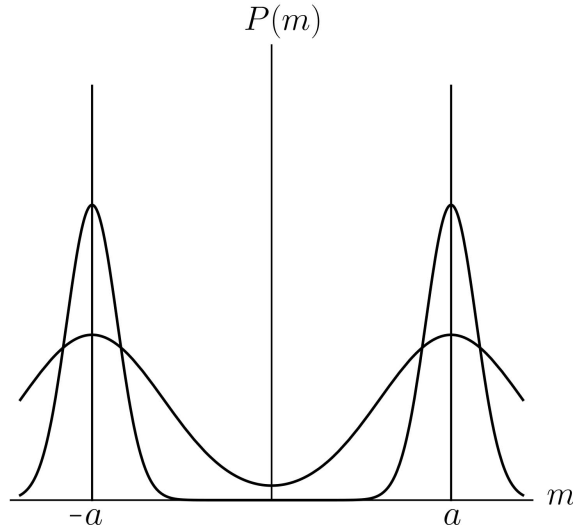


Figure 2: Sketch of $P(m)$ below the critical temperature for an Ising model with small, large and infinite N .

two regions of its state space in which $m \approx \pm a$, with equal probability, and remain there for a very long time. In fact, we can give heuristic arguments—ultimately based on artificial implementations of the system’s dynamics, like the so-called Glauber dynamics—that a finite Ising model will transition between states of positive and negative magnetisation with a time period proportional to $\exp \sqrt{N}$.¹⁴ In the thermodynamic limit the system will remain in an ‘up’ or ‘down’ aligned state forever—they become distinct equilibrium states. But if N is of the order of avagadro’s number the system can still be expected to exhibit a magnetisation in a particular direction for a very long time indeed: much longer than we can feasibly observe real ferromagnets for, and perhaps longer than the age of the universe!

This points to an explanation of the successful appeal to infinite systems in the standard account of SSB in terms of properties they share with large finite systems. On this view, what we actually observe in real ferromagnets at low temperatures are long lived asymmetric states whose macroscopic properties are well approximated by the degenerate equilibrium states found in the thermodynamic limit. The picture of SSB afforded by the thermodynamic limit, in which the system must choose between a set of symmetry breaking equilibrium states, thus provides an appropriate description of the behaviour of large systems over long, but finite, time periods.¹⁵

¹⁴See Lebowitz (1999, 353) and Kinderman and Snell (1980, 55-61).

¹⁵In fact, this way of understanding how the standard account of SSB relates to the behaviour of fi-

4.2 Finite Quantum Systems

As has already been alluded to, SSB in quantum theory has often been thought to provide additional reasons for attaching special interpretive significance to systems with infinite degrees of freedom. The question of whether the approach to SSB in finite systems just sketched in the context of classical systems can be extended to quantum statistical mechanics is, therefore, an urgent one.

In many respects, the way that SSB is modelled in quantum statistical mechanics is analogous to the classical case. The canonical ensemble of a finite quantum system is traditionally represented by a density operator,

$$\hat{\rho} = \frac{e^{-\beta\hat{H}}}{Z}. \quad (3)$$

which, again, is not well defined in the thermodynamic limit. In the quantum context the algebraic formulation of quantum statistical mechanics provides a rigorous framework for generalising the canonical ensemble to infinite systems.¹⁶ The analogue of Gibbs measures are so-called KMS states, which agree with (3) for finite systems, but can also exist in the thermodynamic limit. As in the classical case, the canonical ensemble of a finite system is unique, but a system with infinite degrees of freedom can have a convex set of KMS states whose extremal elements correspond to distinct symmetry breaking equilibrium states.¹⁷

There are peculiarities of quantum theory however which do seem to have important consequences for the possibility of SSB behaviour in finite systems. While classical systems with finite degrees of freedom admit multiple ground states, the corresponding quantum systems typically do not. The classical Ising model has two ground states: one

nite systems is sometimes found in the statistical physics literature. Commenting on the non-uniqueness of the Gibbs measure of the Ising model in the thermodynamic limit, Lebowitz writes:

This means physically that when V is very large the system with “symmetric” [boundary conditions] will, with equal probability, be found in *either* the “+ state” or the “− state”. Of course as long as the system is finite it will “fluctuate” between these two pure phases, but the “relaxation times” for such fluctuations grows exponentially in V , so the either/or description correctly captures the behavior of macroscopic systems. (Lebowitz 353, author’s italics)

A similar discussion is found in Binney et al. (1992, 48-51).

¹⁶See Bratteli and Robinson (2003) and Ruetsche (2011).

¹⁷There has been some debate about how best to characterise SSB in quantum statistical mechanics, see Emch and Lui (2005) and Ruetsche (2011). My discussion here is neutral with respect to this dispute.

with all spins pointing ‘up’, the other with all spins pointing ‘down’. But the quantum Ising model (with a transverse magnetic field)¹⁸ has a single ground state, which is a superposition of the states ψ_+ and ψ_- corresponding to spins aligned ‘up’ and ‘down’ respectively. More generally, a unique quantum ground state can typically be constructed from a superposition of quantum states associated with degenerate minima in the potential if the system has finite degrees of freedom. In the limit of infinite degrees of freedom however, unitarily inequivalent Hilbert space representations of the system’s algebra of observables occur, which can support distinct ground and equilibrium states. In the thermodynamic, ψ_+ and ψ_- become genuine ground states belonging to unitarily inequivalent representations of the infinite quantum Ising model. This has led Ruetsche (2011) to argue that the need to accommodate SSB phenomena in quantum theory compels us to imbue multiple unitarily inequivalent representations with physical content.

The foundational issues raised by the quantum ground state and unitarily inequivalent representations are legion, particularly in the context of QFT. Recall that my ambitions here are limited to understanding the success of the standard account of SSB in statistical mechanics. The key question then is whether the uniqueness of the quantum ground state blocks the extension of the approach to SSB in finite systems described in the previous section to quantum systems.

Intuitively it might seem to: surely at low temperatures the system will remain close to its ground state and no symmetry breaking behaviour will be observed? This is not as obvious as it seems however. Finite quantum systems can support asymmetric states which are stable over very long time periods, despite the uniqueness of their ground states, as can be seen in the case of the quantum Ising model. If the potential of a finite system has multiple minima, the associated asymmetric quantum states can typically be expressed as superpositions of the ground and first excited states: in the quantum Ising model, $\psi_{\pm} = \frac{1}{\sqrt{2}}(\psi_0 \pm \psi_1)$. What happens as $N \rightarrow \infty$ is that the energy gap between these states vanishes. In the case of the quantum Ising model it can be shown that the energy gap between the ground and first excited states, ΔE , decays like $\exp(-N)$, and Koma and Tasaki (1994) prove analogous results for a broader class of lattice models which display degenerate ground states in the thermodynamic limit. This means that, if the system is large, these asymmetric states will have a very slow time evolution. A

¹⁸In order to obtain a non-trivial quantum version of the Ising model a transverse magnetic field term has to be added, so the basic Hamiltonian is:

$$\hat{H} = -J \sum_{i,j} \hat{\sigma}_i^z \hat{\sigma}_j^z - h_x \sum_i \hat{\sigma}_i^x,$$

where $\{\hat{\sigma}^x, \hat{\sigma}^y, \hat{\sigma}^z\}$ are pauli matrices. Henceforth I refer to this model simply as the ‘quantum Ising model’.

quantum Ising model which is prepared in the ψ_+ state will remain in a magnetised state for an arbitrarily long time, for arbitrarily large N . In fact, the existence of these slowly precessing asymmetric states has been put forward in the physics literature as a way of understanding how quantum systems can exhibit SSB, most notably in Anderson’s famous discussion of the electric dipole moment of ammonia and sugar (Anderson 1972).¹⁹

This suggests that the macroscopic properties of a large quantum system can be well approximated by the symmetry breaking equilibrium states found in the thermodynamic limit over long time periods, despite the uniqueness of the system’s ground state. How this claim can be developed in detail is not entirely clear however—there remains the question of how the system enters an asymmetric state in the first place, and why we do not see the true quantum ground state.²⁰ It may be that the canonical ensemble of a finite quantum system which displays degenerate equilibrium states in the thermodynamic limit has most of its mass in symmetry breaking pure states at low temperatures, and the de-idealisation story sketched above for the classical Ising model can be straightforwardly extended to quantum systems.²¹ But it has been suggested by some authors that one must appeal to additional resources, specifically the system’s coupling to the environment, to account for the occurrence of SSB behaviour in finite quantum systems. The next section examines an alternative approach to SSB in finite systems in this vein which has been explicitly developed in the context of quantum statistical mechanics.

4.3 Approach II: Instability Under Asymmetric Perturbations

Landsman (2013) has put forward a different approach to SSB in finite quantum systems. It turns out that the equilibrium states of a large quantum system can be very unstable under small asymmetric perturbations to the dynamics. If we relax the assumption that the symmetry which is broken at low temperatures is exact, this provides another way of understanding how SSB phenomena can manifest in finite systems.

¹⁹See also, Anderson (1997, 175-182) for a similar discussion of finite quantum ferromagnetic systems.

²⁰Some authors claim that there is a close connection here with the measurement problem, see Landsman (2013) and Emch and Lui (2005). This is something of a red herring in my view however, and I keep the question of how SSB in quantum systems should be understood separate from broader interpretive issues about quantum theory.

²¹To my knowledge it remains a conjecture that the kind of bifurcation seen in the canonical ensemble of the classical Ising model at low temperatures also occurs in finite quantum systems.

It can be shown that adding a small perturbation to the Hamiltonian of a large quantum Ising model that favours ‘up’ aligning states causes the system’s ground state to abruptly shift towards ψ_+ , and converge to it as $N \rightarrow \infty$, with analogous statements holding for the system’s equilibrium state below the critical temperature. One way to see why this occurs is to consider the effective Hamiltonian governing the ground and first excited states. In the ψ_{\pm} basis,

$$H = \begin{pmatrix} 0 & -\Delta E \\ -\Delta E & 0 \end{pmatrix}, \quad (4)$$

where ΔE is, again, the energy difference between the ground and first excited state, which, as we have seen, decays exponentially in N . Consequently, any term added to the upper diagonal element, no matter how small, will dominate the effective Hamiltonian for sufficiently large N , shifting the ground state close to ψ_+ . Of course, the same argument applies to perturbations favouring ‘down’ aligning spins. More generally, this type of instability under asymmetric perturbations can be expected to hold if the energy gap between ground and first excited states vanishes sufficiently quickly as $N \rightarrow \infty$.²²

This provides another way of explaining the success of the standard account of SSB by way of the behaviour of large finite systems. In many cases, modelling the symmetry which is broken at low temperatures as an exact invariance of the target system’s dynamics is an idealisation. In concrete systems, small, otherwise empirically negligible, asymmetric perturbations are very likely to be physically instantiated. In the case of ferromagnetic crystals such perturbations might originate in the system’s coupling to its environment, background magnetic fields, or defects in the lattice structure. At low temperatures these tiny asymmetries have a dramatic effect on the behaviour of the system, shifting it into an equilibrium state whose macroscopic properties are very well approximated by one of the degenerate equilibrium states found in the thermodynamic limit. Given our ignorance of the details of the perturbations affecting a real ferromagnet there is an obvious epistemic sense in which a ferromagnet has an equal chance of aligning in any particular direction, and it is appropriate to represent the system as having a number of available equilibrium states below the critical temperature.

While Landsman’s discussion is grounded in quantum theory, it is plausible that this kind of approach to SSB in finite systems can also be applied to classical systems. If the expected magnetisation of a large classical Ising model below the critical temperature is well approximated by the discontinuous function found in the thermodynamic limit,

²²Landsman (2013) and Koma and Tasaki (1994) show that this is the case for a number of other models, but the claim that this behaviour holds whenever there are degenerate equilibrium states in the thermodynamic limit is conjectural. This issue is discussed further in §5.

as Butterfield and others argue, it will be very sensitive to changes in the external magnetic field. As a result, a tiny background magnetic field will shift the system's equilibrium state into a magnetised state which is well approximated by one of the symmetry breaking Gibbs measures found in the thermodynamic limit.²³ This is, of course, a much less general kind of perturbation to the dynamics than those considered by Landsman, but it does suggest that the equilibrium states of finite classical systems can display the same kind of instability properties as their quantum counterparts. SSB behaviour in finite classical systems may also be produced by small asymmetries arising from environmental effects then.

A puzzle might seem to arise here. We have now seen two, apparently quite different, ways of producing the phenomenological features of SSB phenomena in finite systems: the first based on long lived asymmetric states and the second on asymmetric perturbations to the dynamics. If both accounts apply to the same model, how are we to understand what actually happens in a concrete system which displays SSB? There is no real difficulty here however. Following Lui (2003) we can distinguish between the features of a model which make SSB possible and the physical mechanism which brings about asymmetric behaviour in a particular system that model is used to represent. A finite volume model can display both long lived asymmetric states and the kind of instability properties cited by Landsman, yet the way that SSB arises in different physical instantiations of the model may differ from case to case. Furthermore, there is also no reason to treat these two approaches to SSB in finite systems as in competition. Each can be viewed as getting around the apparent impossibility of SSB in finite systems by rejecting the symmetry breaking and spontaneous part of the standard definition of SSB respectively, but it is, of course, also possible to reject both. It is plausible that in many concrete system's environmental perturbations and the stability properties of asymmetric states both play a role in producing SSB type phenomena.²⁴

5 Philosophical Implications

The picture of how SSB phenomena manifest in concrete systems which emerges from the foregoing discussion is a complex one, and many of the details remain open to further investigation. Nevertheless, the findings of the previous section have important im-

²³Binney et al (1992, 48-51) gives a characterisation of SSB in finite classical systems along these lines.

²⁴Environmental perturbations may be responsible for shifting the system into an asymmetric state, while the stability properties of this state explain why we continue to see a symmetry breaking macrostate over a long time period.

plications for the interpretive issues raised by SSB. In particular, the suggestion that the success of the standard account of SSB compels us to reify novel properties of infinite systems has been undermined.

Recall that SSB seemed to provide a stronger case for claims about the representational indispensability of the thermodynamic limit than phase transitions. Not only does the orthodox characterisation of SSB appeal to novel properties of infinite systems, there seemed to be a no-go argument against the possibility of SSB behaviour in models with finite degrees of freedom which ruled out the kind of de-idealisation strategy pursued by Butterfield and others in the phase transition debate. I take the foregoing discussion to show that this is not the case. While SSB defined as the existence of degenerate equilibrium states cannot occur in finite systems this does not show that the kind of behaviour observed in real ferromagnets is impossible. The modelling assumptions underlying this definition, which I dubbed the symmetry breaking and spontaneous part, ought not to be treated dogmatically. If a broader view is taken, and these assumptions are weakened in various ways, we can see how SSB behaviour can occur in finite systems.

More than this though, we have seen that there are prospects of accounting for the predictive and explanatory success of infinite volume systems in the standard approach to SSB in terms of features they share with large finite systems. There is clearly work still to be done here. The results discussed in this paper have, for the most part, been limited to the Ising model. It is ultimately a conjecture that the key statements about large finite models can be generalised to the full range of systems to which the standard account of SSB is applied. Furthermore, for quantum systems in particular, there remain open questions about what form the resulting account can be expected to take. Much the same can be said of the account supplied in the phase transition debate however—the claim that the free energy function found in the thermodynamic limit approximates that of a large finite system is based on heuristic arguments, and has not been shown to hold in general. In both cases, once a framework for de-idealising the thermodynamic limit is on the table the claim that novel properties of infinite systems must be treated as physically real in order to make sense of the relevant physics loses its bite. If the standard account of SSB is to be used to support claims about the physical significance of novel properties of infinite models reasons need to be given for thinking that the programme put forward in §4 cannot account for the predictive and explanatory success of the thermodynamic limit in this context.²⁵

²⁵There are two main ways that the indispensabilist might proceed here. First, they might argue that the approach developed in §4 does not recover the right explanandum—the fact remains that finite systems cannot accommodate behaviour that is both ‘symmetry breaking’ and ‘spontaneous’ in the strict sense introduced in §2. If this kind of objection is to go through however, reasons need to be given for thinking that the concrete phenomena which stand in need of explanation, like ferromagnetism, actually

This is not, of course, to say that the indispensabilist reading of the thermodynamic limit has been refuted here. While the approach discussed in this paper offers defenders of the deflationary reading a way of responding to the challenge posed by SSB, and thereby weakens one *prima facie* motivation for the indispensabilist position, this is not, in itself, an argument against it. Furthermore, as I have emphasised, the dispute over the interpretive status of the novel properties of infinite systems is a multifaceted issue which cannot be settled by examining a single application of the thermodynamic limit.

With this caveat in mind, I close by considering the consequences of embracing a deflationary reading of the thermodynamic limit for our understanding of SSB and related issues. I have argued that there are two, mutually compatible, ways to accommodate symmetry breaking behaviour in finite systems, corresponding to the rejection of the symmetry breaking and spontaneous part of the standard definition of SSB. Both routes seem to have revisionary implications. On the first approach the relevant symmetry is not really broken in the system's full time evolution. While a finite Ising system can display magnetised states over long time periods, as $t \rightarrow \infty$ it will spend an equal amount of time in 'up' and 'down' aligned states, consistent with the fact that the expected magnetisation is zero. The second approach, on the other hand, muddies the conventional distinction between spontaneous and explicit symmetry breaking. Explicit symmetry breaking, you will recall, refers to asymmetric behaviour produced by asymmetries in the dynamics, which is exactly what the perturbations appealed to in Landsman's approach are.²⁶ Furthermore, in both cases the question of what counts as SSB behaviour does not seem to have a precise answer. The notion of a long lived asymmetric state is obviously a vague one, and while there is clearly a qualitative difference between the production of magnetised states by tiny environmental effects and via the application of an external magnetic field the distinction is not sharp. Admitting that both approaches can play a role in producing symmetry breaking behaviour further compounds this situation.

This result is not as radical as it might seem however. While it may have implications

satisfy these requirements, beyond the fact that the standard approach defines SSB in this way, which simply begs the question. Second, they might try to find successful applications of the standard account of SSB in which the strategies put forward in §4 fail. The spontaneous breaking of gauge symmetries might seem to present a possible avenue of attack along these lines. It is not clear that we can make sense of gauge symmetries being broken by environmental perturbations, as in Landsman's approach, for instance. As noted in footnote 1 however, the notion of spontaneously broken gauge symmetry is itself controversial. Whether pressure can be put on the deflationary reading by delving into this issue remains open to further investigation.

²⁶Though, as I suggested in §4.3, we can plausibly still recover one sense in which the behaviour of a ferromagnet is 'spontaneous' on Landsman's approach, namely that spins have an equal chance of aligning in any particular direction.

for the ontic status of SSB phenomena, the deflationary reading of the thermodynamic limit, as I understand it, need not have revisionary consequences for scientific practice, and certainly does not amount to an injunction that physics abandon the standard definition of SSB or cease their appeals to infinite systems.²⁷ There are instructive parallels here with Callender and Menon’s (2013) discussion of extensive quantities in statistical mechanics. Roughly speaking, a quantity is said to be extensive if it behaves additively when the system is divided into sub-systems. Strictly speaking, no quantity satisfies this definition when the system is finite; boundary effects spoil the additivity of quantities like the entropy and the distinction between extensive and intensive quantities only really applies in the thermodynamic limit. But, though the notion of extensivity does not carve nature at its joints, there is clearly a sense in which it is appropriate and useful to treat entropy and other observables as extensive in situations in which boundary effects are negligible. Similarly, though SSB phenomena might not form a genuine natural kind by the deflationist’s lights, §4 provides a framework for explaining why the standard definition of SSB can be fruitfully applied to a range of physical phenomena.²⁸ Recall that the interpretive puzzle, as I originally posed it, was why the standard approach to SSB is successful. Accordingly, if successful, the de-idealisation programme discussed in this paper will justify the current practice of statistical mechanics, not undermine it.

On the other hand, revising our understanding of the nature of SSB phenomena in the manner countenanced by the deflationist may have an important knock on effect on broader philosophical issues. I briefly touch on one here, namely the interpretive problems posed by unitarily inequivalent representations in quantum theory. Ruetsche (2011) uses the way SSB is modelled in statistical mechanics to motivate what she calls a coalescence approach to the interpretation of quantum theory in the limit of infinite degrees of freedom. She argues that in order to explain SSB phenomena in quantum statistical mechanics we must move beyond a single Hilbert space and attribute physical content to multiple unitarily inequivalent Hilbert space representations found in the thermodynamic limit. The deflationist, of course, will deny this. If the programme sketched in §4 is taken seriously then all of the relevant physical information about what goes on

²⁷It is important to distinguish the kind of de-idealisation programme discussed in this paper from attempts to offer a new definition of a phase transition in terms of the resources of finite systems, surveyed in Callender and Menon (2013). I take the deflationary reading of the thermodynamic limit to be a claim about why infinite systems, and definitions which appeal to them, are successful in statistical mechanics. Whether an alternative definition of SSB in finite systems can be provided, and what methodological and metaphysical implications this might have, is, I think, a further question.

²⁸There is clearly a great more to be said here. One might wonder, for instance, whether it is possible to understand the role that the notion of SSB has played in fruitful analogies between condensed matter physics and QFT within the de-idealisation programme sketched in §4. What impact, if any, the discussion of this paper has on broader methodological issues of this kind is a topic for further investigation.

in a real ferromagnet at low temperatures can be captured by a finite quantum model, and therefore a single Hilbert space. What this means for the larger debate about the philosophical significance of unitarily inequivalent representations is a matter for further investigation. Unitarily inequivalent representations, and the notion of SSB, recur in the context of QFT and raise a gamut of issues which have not been touched on here. But the discussion of this paper does suggest that there is scope to push back against the claim, championed by Ruetsche, that multiple unitarily inequivalent representations must be taken ontologically seriously because of the indispensable role they play in scientific explanations.

References

- Anderson, Philip. W. (1972) “More is different” *Science*, 177 (4047), 393-396.
- Anderson, Philip. W. (1997). *Concepts in solids*, World Scientific.
- Batterman, Robert (2005) “Critical phenomena and breaking drops: Infinite idealizations in physics” *Studies in History and Philosophy of Modern Physics*, 36, 225-244.
- Batterman, Robert (2010), “On the explanatory role of mathematics in empirical science” *British Journal for the Philosophy of Science*, 61(1), 1-25.
- Binney, J. J., Dowrick, N. J., Fisher, A. J. and Newman, M. E. J. (1992) *The Theory of Critical Phenomena: An Introduction to the Renormalization Group*, Oxford University Press
- Bratteli, Ola and Robinson, Derek (2003) *Operator Algebras and Quantum Statistical Mechanics, Vol II*, Springer.
- Butterfield, Jeremy (2011) “Less is Different: Emergence and Reduction Reconciled” *Foundations of Physics*, 41 (6), 1065-1135.
- Castellani, Elena (2003). “On the meaning of symmetry breaking” in Brading, K. and Castellani, E. (ed) *Symmetries in physics: Philosophical reflections*, Oxford University Press, Oxford
- Callender, Craig and Menon, Tarun (2011) “Turn and Face the Strange... Ch-ch-changes: Philosophical Questions Raised by Phase Transitions” in Batterman, R. (ed) *The Oxford Handbook of Philosophy of Physics*, Oxford University Press, Oxford.
- Earman, John (2004). “Curie’s Principle and spontaneous symmetry breaking” *International Studies in the Philosophy of Science*, 18(2-3), 173-198.

- Emch, Gerard G. and Lui, Chuang (2005) “Explaining Quantum Spontaneous Symmetry Breaking” *Studies in the History and Philosophy of Modern Physics*, 36, 137-163.
- Fraser, James D. (2015) “Essential Idealisations and the Nature of Scientific Explanation” Unpublished manuscript
- Friederich, Simon (2013). “Gauge symmetry breaking in gauge theories—in search of clarification” *European journal for philosophy of science*, 3(2), 157-182.
- Georgii, Hans-Otto (1988) *Gibbs Measures and Phase Transitions*, W. De Gruyter, Berlin
- Landsman, N. P. (2013). “Spontaneous symmetry breaking in quantum systems: Emergence or reduction?” *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 44(4), 379-394.
- Lebowitz, Joel L. (1999) “Statistical Mechanics: A Selective Reivew of Two Central Issues” *Reviews of Modern Physics*,. 71, 346-357
- Lui, Chuang (2003) “Spontaneous Symmetry Breaking and Chance in a Classical World” *Philosophy of Science* 70: 590-608
- Koma, Tohru and Tasaki, Hal (1994) “Symmetry breaking and finite-size effects in quantum many-body systems.” *Journal of statistical physics* 76.3-4 745-803.
- Kindermann, Ross and Snell, Laurie (1980) *Markov Random Fields and their Application*, American Mathematics Society.
- Mainwood, Paul. (2005) “More is Different? Emergent Properties in Physics” <http://philsci-archive.pitt.edu/8339/>
- McMullin, Ernan. (1985) “Galilean idealization” *Studies in History and Philosophy of Science*, 16(3), 247-273.
- Norton, John. D. (2012) “Approximation and Idealization: Why the Difference Matters”. *Philosophy of Science*, 79(2), 207-232.
- Shech, Elay (2013) “What Is the Paradox of Phase Transitions?” *Philosophy of Science*, 80(5), 1170-1181.
- Ruetsche, Laura (2003) “A Matter of Degree: Putting Unitary Inequivalence to Work” *Philosophy of Science* 70 (5):1329-1342.
- Ruetsche, Laura (2011) *Interpreting Quantum Theories*, Oxford University Press, Oxford