

Discontinuities and Singularities, Data and Phenomena: For Referentialism

The main aim of this paper is to delineate and evaluate two rival positions in the ongoing debate prompted by the employment of infinite idealizations in the treatment of phase transitions. According to the reconstruction of the debate proposed here, the key-disagreement between these two camps revolves around their understanding of a couple of technical notions. The central one is that of a *discontinuity* of a thermodynamical parameter (such as volume or entropy); a closely related notion is that of a *singularity* of a thermodynamical potential (a ‘free energy’ function). Another notion – and what, in fact, ignited the debate – is the role played (if any) by an *infinite (statistical mechanical) system*. The dispute is between what I shall label here ‘Referentialism’, roughly, the claim that discontinuities and singularities do refer to features of thermal systems, and its challengers, the non-Referentialists, who take the opposite view.¹ A more detailed account of these notions and positions will be provided below.

The plan of the paper is as follows. In section 1, I will spell out these three notions in significantly more technical detail than has been done in the philosophical literature so far. I do this for a philosophical purpose, which I take up in sections 2 and 3: to pin down what, exactly, is at stake between Referentialists and their opponents. The two tasks are related, since I suspect that the lack of clarity in the philosophical discussion is due precisely to the absence of a careful parsing of the relevant physics. Then, as I said, in sections 2 and 3 I will draw on these (non-trivial) details and comment on the viability of these rival positions. In essence, I argue that although Referentialism encounters some problems (it turns out not to be a unitary thesis), it survives unscathed by the objections raised against it by its opponents. An important element on which my overall argument relies is the distinction between *data* and *phenomena*, as introduced by Bogen and Woodward (1988).

In slightly more detail, the line of thinking I pursue here is to show that the non-Referentialist criticism is misguided, in two respects. First, in section 2, I articulate a rebuttal of this criticism, based on an analysis of what the Referentialists (following the physicists) mean by a ‘discontinuity’. I maintain that what they say is entirely uncontentious, and has virtually no bearing on any ontological-metaphysical thesis about the constitution of reality (in particular, is entirely consistent with the atomic theory of matter). Second, in section 3, I make a more constructive point (in favor of Referentialism), by appealing to the data-phenomena distinction. I sketch a proposal meant to show how singularities are referential. The key-idea is that they should be understood as referring to, or as ‘picking out’, the very *phenomenon* of a phase change – where (following Bogen and Woodward) a phenomenon is a real, genuine constituent of the world, generating (unavoidably scattered) data-points when observed via instruments. As we will see, the key to disentangling the various issues here is to carefully distinguish between singular *mathematical* points, appearing when we introduce a certain formalism (of free energy), and clusters of *measured* data-points, which come out of the laboratories.

¹ Although irrelevant for the actual content of the discussion here, authors and specific papers can be assigned to each side of the debate (under the usual caveat that they may not agree with this categorization). Thus, Batterman’s (2001; 2005; 2006; 2010; 2016) are illustrative for Referentialism, and similar sympathies can be detected in Morrison’s (2012; 2014) and Bangu (2009; 2015, forthcoming). As for the other camp, I will be discussing mainly Norton (2012), and Menon and Callender (2013), as well as gesturing in passing to Butterfield (2011) and Shech (2013). Ardourel (in press) and P. Palacios can perhaps be added to this group, too.

1.

The following phase-diagram is the typical starting-point in the thermodynamical treatment of the phase transitions.² (The statistical-mechanical treatment follows later on.)

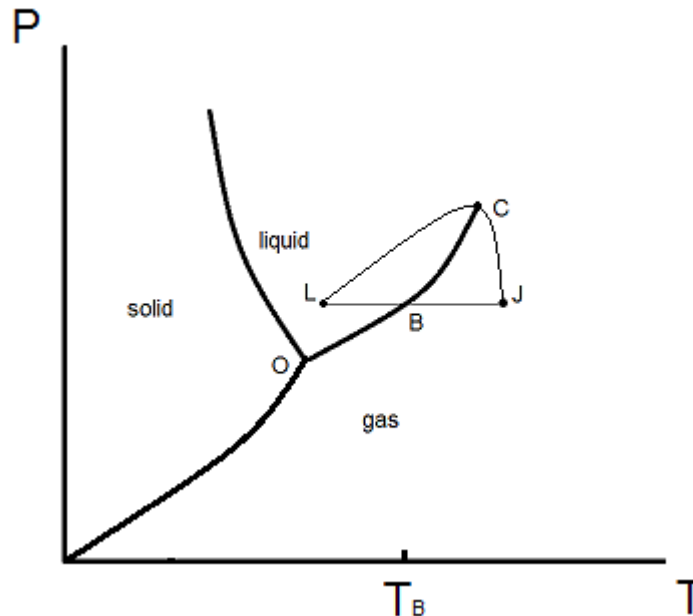


Fig. 1. A phase diagram for water.

Consider a possible system described in Fig. 1 – say, a certain amount of water filling up a glass cylinder with a piston on top.³ The two control parameters are the pressure P in the cylinder, and the temperature T . Point O is the ‘triple’ point (there the water is in all three phases simultaneously). State C represents the ‘critical’ point; above it, we cannot talk of liquid or gaseous states, but of a (supercritical) fluid state. As is clear from Fig. 1, a phase change can be initiated along the line L - B - J as follows. At L , we supply, quasi-statically, heat at constant pressure. The increase in temperature takes us toward point B on OC . Upon reaching B , the temperature of the system remains constant for a while, while the system keeps absorbing heat. Thus, such an L - B - J transition needs ‘latent heat’ to happen. Note that although the state B of the system is represented as a (mathematical) point, this does not mean that the transition (the crossing of OC) is instantaneous, or even very fast; when the system is at B , gaseous bubbles begin to form in the cylinder. At B , when the pressure of these vapors equals the pressure of the water in the cylinder, the system has reached its (so-called) ‘boiling’ point, i.e., the boiling temperature T_B corresponding to a certain value of the pressure. Now the gaseous phase coexists with the liquid phase inside the cylinder.

In order to make the system move out of state B , and complete the transition at constant pressure, we have to keep heating it. After a while, the system’s temperature begins to increase and the system moves out of B passing on to the other side of OC , going toward J . Once the system leaves B , there is no liquid water in the cylinder; it has all turned into vapors and, as we keep heating the cylinder, the temperature of these vapors increases as the system moves toward J . (One also notices that the vapors absorb heat *at a different rate* than the

² Except for emphasizing certain aspects of these phenomena, the presentation and the diagrams follow the textbook treatment (e.g., Zemansky and Dittman (1997)). Section 2 and the diagrams draw on my xxxx (forthcoming).

³ The graph is a projection of the complete three-dimensional phase diagram, also including a volume axis V , onto the P - T plane.

liquid). The transition took place at constant pressure; however, the piston was pushed up, so the volume has increased. Thus, the density of the substance in the cylinder – now, vapors – has dropped massively.

Let us focus on the volume now. It undergoes a change: it had a certain value at the beginning, right when the system entered state B. Then, after the piston moved up and the system (only vapors now) left B, the volume reads a higher value. The difference between the two values of the volume is the reason why the physicists say that at B, the transition point in the diagram, the volume undergoes a *discontinuity*. An L-B-J transition like this one is then (fittingly) called ‘discontinuous’ – and also ‘first order’ (we will see below why). There are other parameters of the system, such as its entropy S , that behave in a similar fashion. They are ‘discontinuous’ too, having one value at the beginning of the transition and another value at the end (while the system is still in state B).

Before we move on, let us note in passing that another kind of transition from one equilibrium phase to another is possible. Suppose we go along path LJ again, only this time toward the critical point C (by first gradually increasing both the pressure and the temperature, then decreasing them both to reach J again.) The important novelty is that at, and near, C, the liquid and gaseous phases are virtually indistinguishable – and not coexistent, as in the case of the previous transition happening at B. At (and above) C, the water in the cylinder manifests ‘critical opalescence’ (it becomes milky and cloudy), and the ‘signature’ of the phase change is now given by other parameters becoming discontinuous, such as the heat capacity (some parameters diverge, i.e., become infinite). A system transitioning from one equilibrium phase to another in this way is said to undergo a ‘continuous’ phase change. (There’s no latent heat needed to complete the transition, so the halting noticed in the L-B-J process does not occur.) These continuous phase changes and, more generally, the phenomena taking place ‘at criticality’, will be left out of this discussion.

In order to characterize the L-B-J types of transitions within thermodynamics, we need to introduce another important concept, that of ‘free energy’, a (so-called) thermodynamic ‘potential’. (There are in fact two free energies, A and G , ‘Helmholtz’ and ‘Gibbs’, defined, respectively, as $A = U - TS$ and $G = H - TS$, where U is the internal energy of the system, H is its enthalpy $= U + PV$, T the absolute temperature, and S the entropy.)

Now, let us consider the two equilibrium phases independently. A free energy function, for liquid and gaseous phases is represented in Fig. 2 as the segments D(K)E and F(K)N, respectively. Considered separately, each is well defined and has continuous derivatives at every point, including K. (The shapes of G and A are roughly the same when plotted against the temperature, so ‘free energy’ refers to both.)

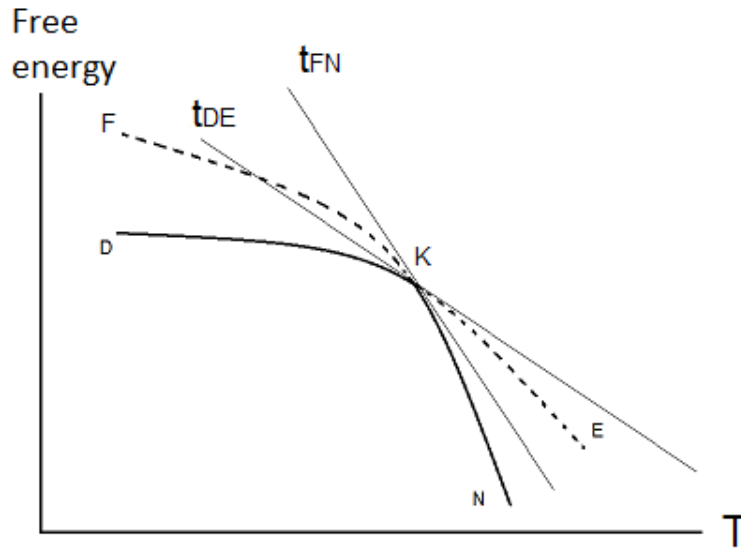


Fig. 2. Free energy plotted against temperature

Now consider the manipulation of a system shown in Fig. 2. Only the thick, un-dotted lines have physical significance. We start at D, in the liquid phase, and increase the temperature while keeping the pressure constant. The passage from the liquid phase to the vapor phase takes place at K. When entering K, the liquid phase has a certain volume. Here, the free energy function, characterizing the liquid phase (the segment DK) does not continue smoothly into the dotted segment KE. Instead, it switches to the other free energy function, the segment KN, which characterizes the vapor phase. When exiting K, the volume (now of vapors) occupied is larger than the volume occupied when the substance entered K. Once again, although K is represented as a mathematical point, this switch is not instantaneous, or even extremely fast, as the system spends time at K; this state corresponds to state B in Fig. 1.

From a purely mathematical perspective, we see that as the system passes through K, the tangent (first derivative) of DE at K (t_{DE} in Fig. 2) changes into t_{FN} , the tangent of FN at K. We are now in the realm of mathematical abstraction and this change is indeed ‘sudden’. Thus, if we take the whole path DN to represent the free energy function for the system undergoing an evolution from state D to state N, to say that it is not smooth at point K is to say that this function has *two* (first) derivatives associated with it. Point K is thus a ‘kink’, a sharp turn – mathematically, it is a point where the free energy function of the system is in fact *not* differentiable, or *non-analytic*. Equivalently, point K is a ‘singularity’, since there the curvature of the free energy is infinite.

Moreover, mathematically speaking, since $A = U - TS$ and $G = H - TS$, then $S = -(\partial A/\partial T)_V$ and $V = (\partial G/\partial P)_T$. These discontinuities of entropy and volume now correspond to the *mathematically* discontinuous change of the first derivatives at K (in Fig. 2). Fig. 3 below summarizes these observations:

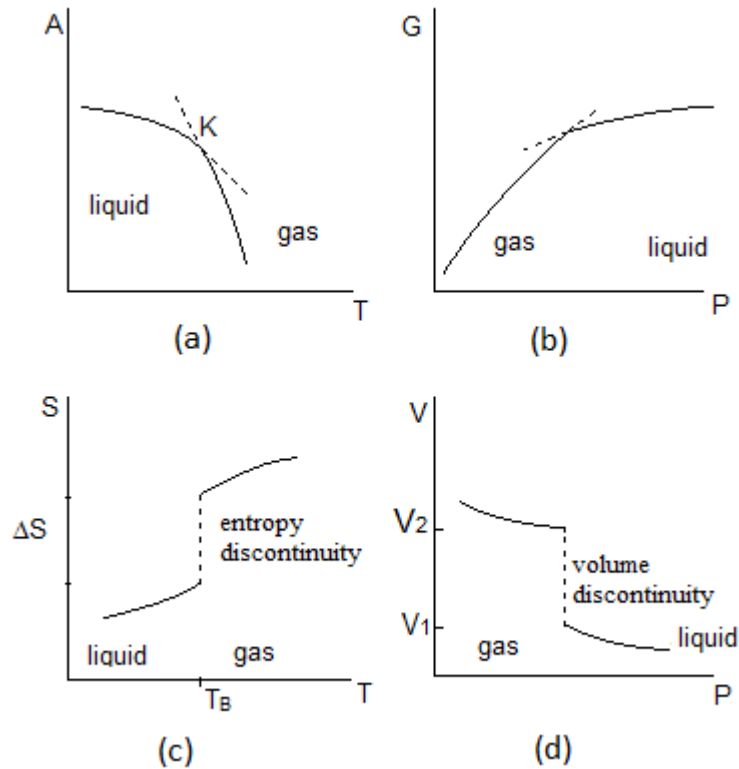


Fig. 3

Several clarifications about Fig. 3 are needed. The *singularities* of the free energies (shown in (a) and (b)) correspond to *discontinuities* of the entropy and volume (shown in (c) and (d)). The entropy discontinuity (in (c)) gives the amount of latent heat $T_B \Delta S$. The dotted segments in (a) and (b) correspond to the dotted segments in Fig 2, where the system is in meta-stable states.

The introduction of the notion of free energy allows the following characterization of a L-B-J phase transition between two equilibrium phases. Such a transition occurs when certain parameters, such as volume (or entropy), exhibit discontinuities – or, since these parameters are also the first-order derivatives of the free energy, when such derivatives are discontinuous as explained in Fig. 2; equivalently, such a transition occurs when free energy has a non-analytic point.⁴

At this moment, the thermodynamical treatment of the discontinuous (first-order) phase changes is (almost) complete. Yet, before we look at the statistical mechanical treatment of such a phase change, let me clarify that the description I provided so far mixed, for the purpose of streamlining the presentation, an idealized thermodynamic description, on one hand, and the description of some experimental procedures, on the other. Thus, there are features that are ignored in this mixed description. For one thing, heating (at finite rate) does not exactly produce an equilibrium; for another, temperature has limited precision in a finite-size system, volume fluctuations will occur, and the temperature of the phase transition cannot be specified with infinite precision in the experiment – in particular, point B is not really a

⁴ The current classification originates in, but refines, P. Ehrenfest's: if the first-order derivative of the free energy is discontinuous, the transition is first-order (discontinuous). A transition is continuous when a second (or higher)-order derivative (and the parameter it represents) is discontinuous, or when certain parameters diverge to infinity.

point from the experimental view. (All this will become clear later on, I hope, but placing the warning here is helpful.)

Let us now move on to the statistical mechanical treatment of a phase change. This is important because the very *raison d'être* for statistical mechanics was to provide an account of the macroscopic (thermodynamical) behavior in microscopic terms – under the assumption that thermodynamical systems (the water in the cylinder) is actually made of an enormous, but finite, number of molecules. Needless to say, the idea was (and is) an epochal success, and one – in fact the most important – concept ensuring the connection between the two levels is that of the (canonical) *partition function* Z , defined as $Z = \sum_i e^{-E_i/k_B T}$. It⁵ is written in terms of Boltzmann's constant k_B and thus it is a micro-level notion. It assumes a (quantum) microscopic constitution of systems, being a sum over all the possible microstates of energy E_i . It also provides expressions for free energies (A and G), and thus links with all other macroscopic quantities, since they can be obtained by differentiation from the free energy (as we saw). One way to get the free energy from the partition function is as $A = -k_B T \ln Z$.

Now, given this relation, we encounter an interesting puzzle. Mathematically, we see that a non-analyticity (singularity) of A (or G) is a non-analyticity of Z , but Z cannot harbor such singular points, since it is a finite sum of exponentials, and each of these is itself an analytic function, which means that Z itself is analytic. If we get back to Fig. 2, this is to say that one and the same Z cannot describe both phases. It thus looks like statistical mechanics is unable (in the strongest, mathematical, sense) to acknowledge the existence of phase transitions as they are described above (within thermodynamics) – a major problem for thermal physics.

However, this is not what one finds in statistical mechanics textbooks (fortunately). The way around this problem relies on employing the following conceptual maneuver. First, consider an idealized version of the statistical mechanical system under scrutiny, a system composed of an infinite number of particles N (which occupies an infinite volume V , and has a finite density N/V .) For this *infinite system* – that is, obtained by taking the $N \rightarrow \infty$ limit, or the 'thermodynamic limit' – it has been proved that the existence of singularities *is* possible.⁶ As Ruelle explains:

The physical systems to which the thermodynamic formalism applies are idealized to be actually infinite, i.e. to fill \mathbf{R}^{ν} (where $\nu = 3$ in the usual world). This idealization is necessary because only infinite systems exhibit sharp phase transitions. Much of the thermodynamic formalism is concerned with the study of *states* of infinite systems. (2004, 2)

⁵ A number of technicalities had to be left out of this presentation. Z is typically evaluated for a *homogeneous* system at fixed volume, but the discussion here involved a *heterogeneous* system, i.e., a system where liquid and gas coexist (at point B), and pressure instead of volume is the control parameter. Thus, a calculation of F from Z gives an unphysical function $p(V, T)$ with an unstable region (see figure 8.6.2 in Reif (1965, 308)). In fact, the evaluation of the full partition function in these conditions does not yield the spontaneous symmetry-breaking characteristic for most phase transitions (as the analysis of the Ising model reveals for a magnetic system). So, to work with the appropriate partition function, one needs to confine the sum to only part of the state space (as well as make a few other further assumptions).

⁶ More conditions have to be satisfied for the existence of the thermodynamic limit; see Liu (1999, S96); Goldenfeld (1992, 25-28) discusses examples where the limit does not exist. Sklar (1993, 78–81) lists several reasons why the limit is generally useful (e.g., it makes the effect of fluctuations vanish). The idea of such a limit was first proposed by H. Kramers, in the late 1930s. L. Onsager (1944) was able to show that an infinite 2-dimensional Ising model undergoes a continuous phase transition (near criticality, the specific heat diverges logarithmically). Yang and Lee (1952) proved that the limit yields discontinuous transitions too (it generalizes to other types of transitions).

Let us take stock. Phase transitions of the L-B-J type (in Fig. 1) are characterized by the discontinuous behavior of certain parameters (such as the volume, as seen in Fig. 3(d)) – and it is this behavior that in fact suggests their label (‘discontinuous’ transitions). When we employ a more general mathematical representation in terms of free energy functions, these discontinuities are captured as singularities of these functions (as seen in Fig. 2 and Fig 3 (a), (b)). The behavior of the first derivatives of these free energies at such singular points is the source of the other label for these phenomena (‘first-order’ transitions). Once this macro-level thermodynamical mathematical description is imported within statistical mechanics, it turns out that the existence of such singularities can be exhibited only by an infinite system (i.e., in the ‘thermodynamic limit’). But, according to the atomic theory of matter, statistical systems are finite.

2.

As introduced at the outset, the debate I will engage with here concerns what I called Referentialism – the idea, as Batterman puts it, that

...phase transitions [are] *real* physical discontinuities and [thermodynamics] is correct to represent them mathematically as singularities. (2005, 234)

I emphasized the word ‘real’ in this quote, since it marks the contentious element, to which the non-Referentialists reacted vigorously. One such well-known reaction is Norton’s:

If the atomic theory of matter is true, then ordinary thermal systems of finitely many components cannot display discontinuous changes in their thermodynamic properties. The changes they manifest are merely so rapid as to be observationally indistinguishable from discontinuous behavior. Indeed, if we could establish that the phase transitions of real substances exhibit these discontinuities, we would have refuted the atomic theory of matter, which holds that ordinary thermal systems are composed of finitely many atoms, molecules, or components. (2012, 225)

Real phase transitions cannot exhibit the discontinuities on pain of contradicting the atomic theory of matter, and, were the discontinuities established factually, the atomic theory would fall. (2012, 228)

So, the concern here, and not only Norton’s, is prompted by the worry that Referentialism clashes with the atomic theory of matter.⁷ The reasoning in the quote above, otherwise valid (a *modus tolens*), goes as following. If such changes in the thermodynamics properties (e.g., the volume) are truly discontinuous (i.e., if the system somewhat ‘jumps’ from having one value for its volume to another value, if not instantly presumably very rapidly – see below), then the only way available for statistical mechanics to account for them is by assuming that the physical systems are composed of an infinite number of molecules. Since we cannot accept that (otherwise we would conflict with the atomic theory of matter), the antecedent has to be rejected (“ordinary thermal systems of finitely many components cannot display discontinuous changes in their thermodynamic properties”). This amounts to a rejection of Referentialism: ‘discontinuous change in a thermodynamic property’ (e.g., volume) has no referent in reality.

An astute reader of the detailed presentation in section 2 has perhaps already spotted the problems with the argument in the quote above: not only is it incomplete, but it also departs from the way the Referentialists (following the physicists) use the notion of a ‘discontinuity’.

⁷ See also Shech (2013, 1175). Butterfield (2011, 1077) has a similar reaction.

I will begin with some comments on this second point, although it is rather difficult to speculate on what non-Referentialists take a ‘discontinuity’ to be. From the quote above, it seems to correspond to an extremely “rapid” change, so fast that it is “observationally indistinguishable from discontinuous behavior”. To see that this is just *not* what a discontinuity is taken to be in this context in the thermodynamics textbooks, we have to recall the discussion of Fig. 3(d) above. We begin in the gas phase and increase pressure. The process takes place at a constant temperature. The volume decreases until it reaches the point marked as V_2 . At this moment, the substance in the cylinder is characterized by a certain value of the temperature, a certain pressure and volume (V_2). Qualitatively, right before the system has volume V_2 , there are only vapors in the cylinder. The system keeps absorbing the heat we supply (such that the temperature remains constant), and what we observe next is that as the system leaves state V_2 , the vapors *gradually* turn into liquid; droplets form first, observable on the glass walls of the cylinder. This condensation process (at the micro-level also called ‘nucleation’) ends in the state where there is only liquid in the cylinder – this is the state of the system as it has volume V_1 . The temperature is kept constant during this time, and we soon notice that no matter how much we increase the pressure, this value V_1 of the volume will not quite decrease any further – the compressibility of liquids is almost zero. During this process, physicists say that the system exhibits a discontinuity ‘in the volume’, in the precise sense that for a certain value of pressure and temperature, the substance in the cylinder starts by occupying a larger volume V_2 and ends up by occupying a smaller one V_1 . This is then the precise sense in which one speaks of a discontinuity in this thermodynamic parameter (volume). As such, it is as real as anything can be in physics, and should cause no concern.

I said that it is difficult to speculate about what may be responsible for the confusion here, but one possibility is that although we can unproblematically speak of a discontinuity in the volume, the non-Referentialists do not realize that the change itself, from one state (of volume V_2) to another (of volume V_1) does happen *continuously* – hence, once again, one should not be worried that atomism is in danger. In fact, the textbook description of this isothermic condensation process given by Reif (1965, 310) is explicit on this (my emphases):

If under these circumstances a fraction ξ of the mole of substance is in the gaseous phase, then the total molar volume V_{tot} is given by

$$V_{\text{tot}} = \xi V_2 + (1 - \xi)V_1$$

In the course of the phase transformation the total molar volume changes *continuously* from the value V_2 for the gas to the value V_1 for the liquid as the fraction ξ changes *continuously* from 1 to 0. In this process there will, of course, be a change of entropy and an associated latent heat.⁸

Let us now turn to Menon and Callender’s (2013, 215-6) discussion of the discontinuity issue. I find their presentation still somewhat ambiguous when it comes to what they take a discontinuity to be, and more generally on the distinction between discontinuities and singularities.

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.

⁸ I replaced Reif’s notation V_B , V_A with V_2 and V_1 , respectively.

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 5.3).

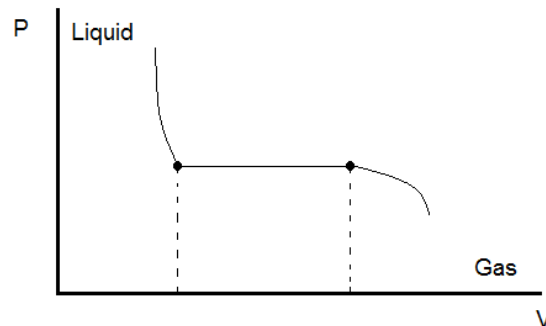


Figure 5.3 P - V diagram for a liquid-gas system at a phase transition.

The two black dots are coexistence 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 coexistence 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 coexistence point it does not instantaneously switch to a liquid phase. It continues in its gaseous phase, but this supersaturated vapor is meta-stable. (...)

Perhaps the nucleation of droplets from zero radius could be seen as an example of a physical discontinuity. (2013, 216)

What they write here leaves open what they deny: that a genuine discontinuity in the volume exists (Norton's claim), or that the change takes place discontinuously. If the former, I am afraid that they are, as we saw, mistaken; if the latter, they are right, the change is continuous indeed. Although I am inclined to read the passage charitably, I should note that Referentialists have never said otherwise. Nor did they object to the idea that a process like nucleation takes place, at the microscopic level. In fact, it is hard to see why nucleation is brought in discussion here at all – and that it is makes one suspect that the above reading excessively charitable. If talking about nucleation is meant to provide – on the Referentialist's behalf, as it were – an example of a genuine discontinuity (as if the one in the volume was not so), then this is beside the point: as we saw, this is simply *not* the sense in which the Referentialists use the notion. As we saw, these phase transitions are not called 'discontinuous' for reasons having to do with nucleation. Sure, if one wishes to characterize the nucleation micro-process as 'discontinuous', one is, I suppose, free to do so, but this has virtually nothing to do with how the term is used by the Referentialist in this context in thermodynamics. These remarks also apply to the possible suggestion that when Norton talks about very "rapid" changes, he has in mind nucleation – since, indeed, nucleation *is* a very rapid process, of the order of 10^{-3} seconds; see Carpenter and Bahadur (2015) – but, again, this is just irrelevant here.

This point links the discussion to another (correct) remark in the Menon and Callender account above, namely that the transition between the two states is *not* “instantaneous” (or “so rapid as to be observationally indistinguishable from discontinuous behavior”, as Norton described it.) After all, the *relevant* process essentially involves latent heat (something which neither Norton, nor Menon and Callender mention in their papers), and this is the reason why the issue of the rapidity of the process cannot even arise.

Another aspect where both Norton’s and Menon and Callender’s accounts pose even more serious difficulties is the notion of a singularity. I mentioned as a shortcoming of Norton’s argument above that it is incomplete. This is so since it passes from the discontinuities to the introduction of infinite systems without mentioning the intermediate notion, the singularity – the mathematical representation of the physical discontinuity in the formalism of the free energies.⁹ As we recall, it is its presence in the partition function that is the essential element requiring the introduction of the infinite system. By not keeping track of the distinction, the objection he raises against Referentialism misses the point: while the discontinuities are trivially real (if we use the term carefully, and in line with its use in thermodynamics), and hence unobjectionable, the singularities do have a more problematic referential status, as we will see below.

I will return to this in a moment, not before examining what Menon and Callender have to say about singularities. They note that

This representation [Fig. 5.3] certainly involves a mathematical singularity: as the pressure is reduced, the volume of the system changes discontinuously.

This is, again, rather unclear: what does ‘involves’ mean here? Where, and what, is the singularity in Fig. 5.3? Strictly speaking, there is no singularity there – and, just to be clear (if this is what they meant), none of the black dots is such a singular point. These points do mark (as is otherwise correctly said in their quote) states where the system is in both gas and liquid phases (since they belong to the horizontal coexistence line), but they have virtually nothing to do with genuine singularities as precisely defined in section 2, i.e., points like K in Fig. 2 (and Fig. 3(a)), or K* in Fig. 3(b). However, if one insists to find a singularity in Fig. 5.3, I suppose one can perhaps say, by stretching this notion, that it is (somehow) to be identified not with a point, but with the whole segment in between the two black dots (including them). It is then this whole line that is ‘projected’ into a point like K (or K*). These are properly speaking singular points, appearing when we introduce the formalism of the free energy.

Their next remark is also puzzling, since it mentions ‘an artifact of representation’ in relation to the isotherm in Fig. 5.3. (“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.”) Insofar as the two black dots are not singularities, what exactly is the ‘artifact’ here? Here is a way to clarify what they may have had in mind. We saw that when it comes to singularities, displaying a diagram such as the one in Fig. 5.3 is rather misleading. Yet, there is a sense in which one can talk about how a singularity may be regarded as an ‘artifact of representation’. To see why and how, we need to focus on diagrams of a different

⁹ It is because of this ‘gap’, I suspect, that one may (legitimately) begin to wonder how the very idea of an inconsistency between the existence of discontinuities and the finiteness of the statistical mechanical system may arise – if the systems are finite, which they are, their volumes (for instance) can’t have discontinuities. We will see below that this inconsistency is only apparent.

kind, such as the one drawn in Fig. 4 below – which, moreover, should be examined next to the one in Fig. 2. In terms of what is represented in Fig. 2, it is crucial to realize that such a diagram, in which we see continuous curves, and their intersection (at K), is *not* a reflection of what is actually *measured* experimentally. For an illustration of what is measured, i.e., a direct plotting of the actual data-points, see Fig. 4 below.

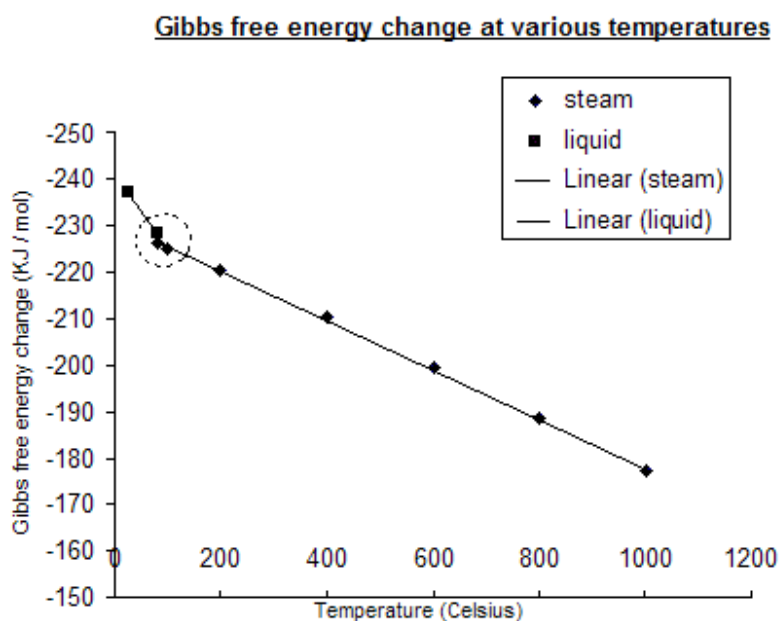


Fig. 4. An example of actual Gibbs free energy variation depending on temperature. I added the dotted circle to the original diagram, to be found here¹⁰: https://www.doitpoms.ac.uk/tlplib/fuel-cells/pem_efficiency.php (accessed 09/20/2017)

The relation between Fig. 4 and Fig. 2 is as follows. Just as in Fig. 2, in Fig. 4 the system begins in the liquid phase, and the free energy is plotted as connecting the two square dots. As the temperature increases, the system transitions to the vapor (steam) phase, and the free energy for this phase is drawn as a continuous line connecting diamond points. The transition takes place *at* the cluster of data-points (inside the dotted circle in Fig. 4), containing steam states (one square point) and liquid states (two diamond points). A perhaps more precise way to put it is that a cluster of measurement points is taken at the transition, and from this cluster of data, together with the data points further away from the transition, the point of the transition is inferred. Crucially, no matter how good the resolution of a measurement is, i.e., no matter how many, and how precise additional measurement points we take in the laboratory (here only three are shown inside the region), the transition will still appear in the form of a cluster (and this is so indeed because of the unavoidability of ‘fluctuations’ that Menon and Callender mention, while they are noticeably absent in Fig. 5.3¹¹). It is a cluster like this (in Fig. 4) that is represented in idealized form in Fig. 2 as a singular point (K). Now

¹⁰ As is explained on the website, the data in the graph is collected for hydrogen fuel cells. The aim of the measurement was to see what factors influence their efficiency. The conclusion represented in the graph is that the energy from a fuel cell reaction decreases proportionally with the temperature at which the reaction takes place. In this measurement (expressed in molar quantities), the change in free energy (shown on the ordinate axis) is $\Delta g_f = \Delta h_f - T\Delta_s$, where Δh_f is the difference between the enthalpy h_f of the products and that of the reactants (the same holds for entropy Δ_s).

¹¹ In his (2001, 550) Callender notes, having in mind not Fig. 5.3., but most likely something like Liu’s Figure 4 in (Liu 1999, S104), that “because of the fluctuations we don’t actually *measure* perfect singularities” (my emphasis).

we can see in what sense we can talk of ‘artifacts’ of representation: the singular point (i.e., K , or K^*) doesn’t represent a physical transition *point*, but a *cluster* of data-points.¹² Thus, the singularity is brought into existence, so to speak, or ‘manufactured’, by the way the transition process is represented in the formalism of free energy: as the (idealized) representation of a cluster, or the intersection of two curves (again, idealized as continuous).

The previous discussion settles, I hope, the issue of *discontinuities* – there should be no question as to the viability of Referentialism. However, we also saw that we face yet another, immediately related issue, that of *singularities*.

3.

In order to address this issue, it is useful to distinguish further between two versions of Referentialism: a ‘strong’ and a ‘weak’ form. While both versions regard singularities as referential, they are not referential in the same way. It is this difference that I’ll try to articulate below. This investigation will prompt some related remarks on the introduction of the infinite systems, and the non-Referentialists take on it.

The strong form of Referentialism is perhaps best illustrated by an earlier paper by Batterman, who claims that

real systems exhibit physical discontinuities . . . genuine physical discontinuities—real singularities in the physical system. (2005, 235-6; italics in original)

It is the latter part of the quote that I will focus on now: the ideas that singularities are ‘real’, and that they are ‘in the physical system’. Given the remarks above (in sect. 1 and 2), there should be no concern about the first idea, that the singularities are real, in so far as they correspond to *clusters of states*, or regions in the graphs collecting measurement data. A singularity is thus real – unlike Santa Claus, and more like the center of mass of a (compact) physical object: a mathematical, dimensionless point, serving a well-defined theoretical purpose. This leads to the second idea, namely that singularities are, somehow, ‘in’ the physical system. This does not sound right, since we saw that the direct measurements do not reveal to us any such *points*; and, if it is the measurements that tell us what is ‘in the system’, then a singularity is not ‘there’. In fact, Batterman is aware of this, and acknowledges that

I [did] speak [in my (2005)] rather sloppily of genuine physical singularities. It is best to think instead in terms of some kind of genuine *qualitative* change in the system at a given scale (2010, 22)

This is now unobjectionable: such a change does take place ‘in the system’, and is surely real (this is the reason we are fascinated by the topic in the first place!) It is, however, concessions like this that open the door for a weaker form of Referentialism, whose central tenets I intimated (without using this label) in my (2009). I will quickly rehash the proposal here.¹³ In a nutshell, the key-thought is that a singularity refers to the *phenomenon* of a phase change, where this term is used in the sense of Bogen and Woodward’s (1988) now well-known distinction between *data* and *phenomena*.¹⁴

¹² As Liu (1999, S103) correctly observed: “the transition is neither ‘smooth’ nor ‘singular’”.

¹³ Part of what follows draws on Bangu (2009, sect. 5) as well as some points made in Bangu (2015).

¹⁴ Interestingly, and perhaps not accidentally, one central example used in the 1988 paper to introduce the distinction was that of a phase transition: the melting point of lead. Although quite old, the distinction still draws

The starting point is Liu's remarks that, according to Fig. 4, the singularities in Fig. 2 (and Fig. 3(a), (b)) are "artifacts" (1999, S104), "fictions of [thermodynamics]," which "do not exist in reality" (2001, S336). This means, presumably, that they are not observable (not measurable). But now, following Bogen and Woodward, we can say that Fig. 4 displays a set of *data* (points) gathered from measurements. This is, however, only the first step in the scientific investigation. The role of data is, next, to offer a basis for the *inference* to what they call "phenomena" (1988, 309; 311; 313, etc.)¹⁵ This inference is needed since the collected data points unavoidably include extraneous, distorting factors ('noise' – errors due to the impurities in the sample of substance used, or due to the fact that the measurement is always made away from perfect equilibrium, etc.) The key-thought here is that although measurement (experiment) is the most direct form of access to a physical system, it only allows this access in an approximate way; it does not guarantee access to the stable, enduring, characteristic properties of the substance of interest. The results of the inference are the *phenomena*; unlike the clusters of *data* points, they are not "idiosyncratic to specific experimental contexts" (1988, 317). Phenomena are robust, invariant over various experimental setups and, unlike the raw data, reproducible (Bogen and Woodward 1988, 317; 326; Woodward 2011). It is such a *phenomenon* of phase change that *is* displayed in Fig. 2 (and Fig. 3a, b) – while markedly *absent* in Fig. 4 (where, again, we only have data).

There may be no need to insist on this, since the data/phenomena distinction is now well-known, but before we move on it is important to keep in mind that Bogen and Woodward's use of the term 'phenomenon' does not have the typical meaning one encounters in the philosophy of science literature. On one hand, data are collected from measurements (observations), but "phenomena for the most part cannot be observed and cannot be reported by observational claims" (1988, 343; 306). This use is different from that of most authors writing about phenomena – as precisely what is observed (sometimes with the naked eye), or what is to be 'saved' by postulating theoretical structures (planetary positions, eclipses, etc.) In this framework, phenomena are what generates the (always approximate, to-be-improved-upon) sets of data; they are physically real patterns of behavior (Bogen and Woodward 1988, 321). Yet, on the other hand, phenomena are not at the extreme end of the continuum of difficulty of access by instruments – according to the way I interpret the distinction here, it makes no sense to talk about a phenomenon as an approximation, while it is in the nature of data to be approximate, and, upon improving our instruments, to become more and more precise.¹⁶ So, when it is claimed that phenomena are "detected" (1988, 306), the qualification that they are "detected *through the use of data*" (1988, 306; my emphasis) is essential. A more accurate thermometer cannot (finally!) show us a phenomenon, but it would of course provide better data. They, in turn, would provide evidence that a genuine phenomenon exists.

In my (2009), I noted that this contrast (between the robustness of phenomena and the transience of the data) implies a difference in what I called 'epistemic relevance'. Data surely have epistemic relevance, but only secondarily; they are what is 'shaped into' the phenomena.¹⁷ It is, however, the phenomena, that possess primary epistemic relevance, due to

philosophers' of science attention, eliciting both acceptance and criticism. See Woodward (2011) for, as far as I know, his latest restatement and defense of it.

¹⁵ The specific inference procedures vary with the area of scientific research, but a common feature is that they rely on statistical techniques (1988, 311).

¹⁶ This alludes to Norton's insistence on approximations, to be discussed in a moment.

¹⁷ This way of speaking is meant to emphasize that (weak) Referentialism squares remarkably well with Kadanoff's remark that "in some sense, phase transitions are not exactly embedded in the finite world but, rather, are products of the human imagination." (2009, 778). I submit that the sense in question here is captured

their robustness and replicability. It is the phenomena that are accounted for, i.e., explained, predicted, derived, etc., by science (Bogen and Woodward 1988, 305–306; Woodward 2000, S163).

So, as I suggested above already, this weaker form of Referentialism appeals to the data/phenomena framework to illuminate the referential status of singularities. In light of the data/phenomena distinction, what we have in Fig. 4 is *data*. Yet, it is *phenomena* that constitute the focus of scientific interest – and we find them in diagrams like the one in Fig. 2. It is then natural to say that the role of a singularity is to refer to a phase change *phenomenon* – again, in the specific sense of Bogen and Woodward (1988), where they take a thoroughly *realist* stance on phenomena, position which, perhaps needless to say once again, I do share.¹⁸ Thus, singularities are referential at the *phenomenal*/unobservable level, while not at the *data*/observable level. This characterization implies that a singularity does not refer to a directly ascertainable, measurable feature of the physical system (as I suspect that the strong form of Referentialism would have it), but to an indirect feature, introduced via a specific mathematical representational framework (here, in terms of free energy). Thus, while a singularity lacks *observational* significance indeed – note the concession made to the non-Referentialists!¹⁹ – it still retains full *physical* significance.

Since, according to Bogen and Woodward, the scientific enterprise *in general* consists in “[moving] from claims about data to claims about phenomena” (1988, 314), and since (in this particular context) singularities are the (mathematical) tool to do this, i.e., to refer to phenomena, the mere record of scattered data points in Fig. 4 is simply the wrong place to look (and draw conclusions): one will not find there what physicists are actually interested to account for, the *phenomenon* of a phase change.

So, let us take stock for now: the weaker form of Referentialism I am trying to articulate here distinguishes between phase transitions considered at the observational, or data level, on one hand, and phase transitions understood as phenomena on the other – and claims that singularities are referential in this latter sense.

precisely by understanding a phase transition as a phenomenon – whose shaping from data requires, as we saw, ‘human imagination’, although not in an unrestrained way, but, as we also saw, in the form of mathematical modelling. This is perhaps an unnecessary clarification, since this ‘unrestrained way’ was surely not how Kadanoff intended the remark. (I should add the caveat that this is my reading of this remark, and that as far as I know neither Batterman, nor Kadanoff ever endorsed its interpretation along these lines.)

¹⁸ The case for realism about phenomena is made in section VIII of their 1988 paper.

¹⁹ In connection to this, let me note that, as a referee pointed out (and I am almost quoting the suggestion here), it is possible that in the above quote Norton did not mean “rapid” in a temporal sense. So, the proposal then is to interpret him more broadly as saying that the rate of change of a thermodynamic quantity, with respect to temperature, pressure, etc. is so large as to be effectively infinite. If so, there would still be wiggle room for his view, since after all, the very concept of a gas’ volume is well-defined only up to a certain degree of precision associated for example with the atomic granularity of the container walls. Then the more general lesson here would be that observability is problematic: physical concepts are only well-defined up to a finite margin of precision, and that would then seem to support the view that the strict mathematical discontinuities of the thermodynamic representation of the system cannot completely reflect the behavior of the real physical quantities, which are defined only up to a limited margin of precision. I do not discount the possibility that Norton meant something else by what he wrote, but I believe this rendering is a bit too charitable; it is pretty clear from the context that the claim should be understood in a temporal sense. And, after all, the principal sense of the word is, according to the Oxford dictionary, the temporal one: ‘happening in a short time’. See <https://en.oxforddictionaries.com/definition/rapid>. Hence the remarks above don’t change the situation too much in my view.

Against this background, let me comment on how this position and distinction relate somewhat naturally to another distinction, drawn by Norton (2012), between what he calls – in connection to the infinite systems implicated in this discussion – “approximations” and “idealizations”. On my reading, the point of his paper is to make clear the difference between “two sorts of analytic activity” (2012, 228) that physicists and philosophers perform when studying these issues. On one hand, we can think in terms of inexactly described properties of real, finite systems. On the other, we may be tempted to “promote” these approximations (approximate descriptions) to “idealizations” (2012, 211). This is a conceptual move consisting first in introducing a limit physical system (i.e., an infinite system, in this case), and then in attempting to describe its properties. To apply this distinction here, we can say that the property of undergoing a phase transition is only approximated by a real, finite system, but will be realized exactly by the limit system.

Now, this second way of thinking (in terms of the analysis of the properties of an infinite system) may be embraced by some physicists (and some philosophers)²⁰; yet, as Norton shows (with the help of several cleverly designed examples), “the infinite systems often have properties very different from those of the finite system” (2012, 215). There is thus no guarantee “that the limit property and limit system will agree” (2012, 215), and thus “there is a real difficulty facing the use of the strong thermodynamic limit as an idealization. One cannot assume that the limit of well-behaved finite systems will be a well-behaved infinite system.” (2012, 216). Hence, “since an infinite system can carry unexpected and even contradictory properties, the latter practice [the investigation of the properties of infinite systems] carries considerably more risk.” (2012, 228)

As the examples that Norton marshals show, this risk cannot be denied indeed – and if this general cautionary note is the central lesson to draw here, I am in agreement. However, one also gets the impression that for Norton this second ‘analytic activity’ lacks a solid epistemic grounding, and thus, in the end, it should be avoided. On the analysis I outlined above, it is perhaps clear that I do not believe that this is so. Once we see how singularities play an indispensable (referential) role, the introduction of the infinite system creates no conceptual complications. For a (weak) Referentialist, the introduction of such a system is nothing more than a consequence of recognizing the epistemic role of the singularity: a mathematical device needed to understand how a mathematical property of a certain function can be demonstrated. It is a mathematical way to bring a mathematical point into existence (so to speak). Thus, the introduction of such a system incurs no ontological commitment – if this was a serious concern, since, as far as I can tell, no author, physicist or philosopher, has claimed that these infinite systems have physical reality. Batterman, for one, has always been clear that “fluids *are* composed of a finite number of molecules” (2006, 903; his emphasis), that “water in real tea kettles consists of a finite number of molecules” (2010, 7), etc. Thus, what he says about infinities in the context of first-order phase transitions – “I want to champion the manifestly outlandish proposal that despite the fact that real systems are finite, our *understanding* of them and their behavior requires, in a very strong sense, the idealization of infinite systems and the thermodynamic limit.” (2005, 231; my emphasis) – is in perfect harmony with the weak Referentialist approach proposed above. Thus, from the perspective sketched here, it is

²⁰ Among physicists, Norton (2012, 219) singles out Langford (1975) and Ruelle (2004; the passage I also cited above). He also points out that in Le Bellac, Mortessagne, and Batrouni (2004) “any condition on an actual infinity of components and the corresponding behavior of that infinite system” is “conspicuously absent”. As an aside comment, I am not confident that this absence is the result of these authors’ deliberate choice; a perusal of their monograph did not convince me that they were concerned with conceptual clarity at the level Norton suggests they were. In fact, it is Ruelle’s treatise that seemed to me to deliberately aspire to this.

perhaps only the (over)enthusiastic tone that I find somewhat unwarranted. After all, everything here is, as one may say, business as usual. Scientists *in general*, as a matter of overarching methodological principle, want to understand (explain, derive, predict, etc.) phenomena – here: represented as singularities – so there is really nothing ‘outlandish’ about what is going on in this case: they collect the data, shape them into a phenomenon, then deploy the theoretical tools to account for it – which naturally leads to a mathematical treatment involving infinities.²¹

References

- Ardourel, V. (in press). The infinite limit as an eliminable approximation for phase transitions. *Studies in History and Philosophy of Modern Physics*.
- Bangu, S. (forthcoming) Phase Transitions. In *The Routledge Companion to the Philosophy of Physics*. E. Knox and A. Wilson (eds.) Routledge.
- Bangu, S. (2009). Understanding Thermodynamic Singularities. Phase transitions, Data and Phenomena *Philosophy of Science* 76.4: 488-505
- Bangu, S. (2015). Why does Water Boil? Fictions in Scientific Explanation In *Recent Developments in the Philosophy of Science. EPSA 13 Helsinki*. U. Mäki *et al.* (eds.) Pp. 319-330. Springer
- Batterman R. (2001). *The Devil in the Details*. Oxford University Press.
- Batterman, R. (2005). Critical phenomena and breaking drops: Infinite idealizations in physics. *Studies in History and Philosophy of Modern Physics*. 36: 225–244.
- Batterman, R. (2006). Hydrodynamic vs. molecular dynamics: intertheory relations in condensed matters physics. *Philosophy of Science*. 73: 888–904.
- Batterman, R. (2010). On the explanatory role of mathematics in empirical science. *British Journal for Philosophy of Science*. 61: 1–25
- Batterman, R. (ed.) (2013) *Oxford Handbook for the Philosophy of Physics*. Oxford University Press.
- Batterman, R. (2016) Philosophical Implications of Kadanoff's work on the Renormalization Group. *Journal of Statistical Physics*. DOI: 10.1007/s10955-016-1659-9
- Bogen, J., and J. Woodward (1988). Saving the Phenomena. *Philosophical Review* 97: 303–352.
- Butterfield, J. (2011). Less is different: Emergence and reduction reconciled. *Foundations of Physics*, 41, 1065–1135.

²¹ I deliberately refrained from engaging with what has recently been dubbed the ‘paradox’ of phase transitions (Shech 2013), originating in Callender (2001). See Bangu (2009; 2015) for a detailed take on it, including the comments on the several attempts of finite statistical mechanics to deal with (first-order) phase transitions. For very recent work on this, see Ardourel (in press).

- Callender, C. (2001). Taking thermodynamics too seriously. *Studies in History and Philosophy of Modern Physics*, 32, 539–553.
- Callender, C., & Menon, T. (2013). Turn and face the strange. Ch-ch-changes: Philosophical questions raised by phase transitions. In Batterman (2013).
- Carpenter, K., Bahadur V. (2015). Saltwater icephobicity: Influence of surface chemistry on saltwater icing *Scientific Reports* 5. Article number: 17563.
- Goldenfeld, N. (1992) *Lectures on Phase Transitions and the Renormalization Group*. Perseus Books.
- Kadanoff, L. (2009). More is the same; mean field theory and phase transitions. *Journal of Statistical Physics* 137: 777–797.
- Lanford, O. (1975). Time Evolution of Large Classical Systems. In J. Moser (ed.) *Dynamical Systems, Theory and Applications: Lecture Notes in Theoretical Physics*, vol. 38, pp. 1–111: Springer.
- Le Bellac, M., Mortessagne, F. and Batrouni, G. (2004). *Equilibrium and Non-equilibrium Statistical Thermodynamics*. Cambridge University Press.
- Liu, C. (1999). Explaining the emergence of cooperative phenomena. *Philosophy of Science*, 66, S92–S106.
- Liu, C. (2001). Infinite systems in SM explanations: Thermodynamic limit, renormalization (semi-) groups, and irreversibility. *Philosophy of Science*, 68, S325–S344.
- Morrison, M. (2012). Emergent physics and micro-ontology. *Philosophy of Science*, 79, 141–166.
- Morrison, M. (2014). *Reconstructing Reality: Mathematics, Models and Simulation*. Oxford Univ. Press
- Norton, J. (2012). Approximation and idealization: Why the difference matters. *Philosophy of Science*, 79(2), 207–232.
- Onsager, L. (1944). *Phys. Rev.* 65, 11749.
- Reif, F. (1965). *Fundamentals of Statistical and Thermal Physics*. McGraw-Hill.
- Ruelle, D. (2004). *Thermodynamic Formalism*. 2nd ed. Cambridge University Press.
- Shech, E. (2013). What is the paradox of phase transitions? *Philosophy of Science*, 80(5), 1170–1181.
- Sklar, L. (1993). *Physics and Chance*. Cambridge Univ. Press
- Woodward, J. (2000) Data, Phenomena, and Reliability. *Philosophy of Science* 67 (Proceedings): S163–S179.

Woodward, J. (2011). Data and phenomena: a restatement and defense. *Synthese*, 182:165–179.

Yang, C. N., and T. D. Lee (1952), “Statistical Theory of Equations of State and Phase Transitions. I. Theory of Condensation” *Physical Review* 87: 404–409.

Zemansky, M., Dittman, R. (1997). *Heat and thermodynamics* (7th ed.). McGraw-Hill.