Had We But World Enough, and Time... But We Don't!

Justifying the Thermodynamic and Infinite-time Limits in Statistical Mechanics

Patricia Palacios

Munich Center for Mathematical Philosophy

In this paper, I compare the use of the thermodynamic limit in the theory of phase transitions with the infinite-time limit in the explanation of equilibrium statistical mechanics. In the case of phase transitions, I will argue that the thermodynamic limit can be justified pragmatically since the limit behavior (i) also arises before we get to the limit and (ii) for values of N that are physically significant. However, I will contend that the justification of the infinite-time limit is less straightforward. In fact, I will point out that even in cases where one can recover the limit behavior for finite t, i.e. before we get to the limit, one cannot recover this behavior for realistic time scales. I will claim that this leads us to reconsider the role that the rate of convergence plays in the justification of infinite limits and calls for a revision of the so-called Butterfield's principle.

1 Introduction

"Had we but world enough, and time" are the words with which Andrew Marvell begins his passionate poem in which he tells his lover that things would be different if they had infinite space and time. While neither the number of particles in real systems nor the time of measurements are infinite, it is common in statistical mechanics to take the number of particles and time to infinity in order to recover the values of thermodynamic observables. These are called the thermodynamic limit and the infinite-time limit, respectively. When these limits are taken, it is assumed that, at least for the purpose of inferring the values of the thermodynamic observables, the infinite case is rather similar to the finite case (contrary to the situation described by Marvell!). But, is this so?

In the past few years, there has been a fervent controversy around the use of the thermodynamic limit in the statistical mechanical treatment of phase transitions. For some (e.g. Batterman 2005, Batterman 2011, Bangu 2009, Bangu 2011), the assumption of the thermodynamic limit – and so of an infinite system – is not innocuous since it

describes a situation that is qualitatively different from a situation in which there is an arbitrarily large number of particles, generally put in terms of the "singular nature" of the thermodynamic limit. As a consequence, it is claimed that the behavior in the limit is physically real (Batterman 2005) or that phase transitions are not reducible to statistical mechanics (e.g. Batterman 2011, Bangu 2011, Morrison 2012). Others (e.g. Butterfield 2011, Butterfield and Buoatta 2011, Norton 2012) arrive at the opposite conclusion in that they claim that the invocation of the thermodynamic limit can be justified pragmatically. They arrive at this conclusion by arguing that the thermodynamic limit satisfies what Landsman (2013) calls Butterfield's principle, according to which the limit is justified if and only if the same behavior that arises in the limit also arises, at least approximately, "on the way to the limit".

Analogously, the use of the infinite-time limit in the explanation of equilibrium states has also motivated a discussion in the physical and philosophical literature (e.g. Malament and Zabell 1980, Earman and Rédei 1996, Emch and Liu 2013, Sklar 1995). For some (e.g. Malament and Zabell 1980), the fact that the infinite-time limit cannot always be associated with finite time averages is one of the main motivations to abandon the association of phase averages with time averages in the ergodic explanation of equilibrium. Others (e.g. Emch and Liu 2013, Gallavotti 1999) have recognized the problems associated with the traditional ergodic approach but have suggested a pragmatic justification for the infinite-time limit, which relies on the necessity of adapting the time-scale of the analysis to the phenomenon we want to describe.

Despite its importance, the infinite-time limit has been generally left aside from the recent philosophical debate around the use of infinite idealizations in statistical mechanics.[1] In this paper, I will contribute to filling in this gap by comparing the use of the thermodynamic limit in the theory of phase transitions with the infinite-time limit in the explanation of equilibrium statistical mechanics. In the case of phase transitions, I will argue (Section 2) that the thermodynamic limit can be justified pragmatically, since the limit behavior also arises before we get to the limit and for a number of particles N that is physically significant. However, I will contend (Section 3) that the justification of the infinite-time limit is less straightforward. In fact, I will point out that even in cases where one can recover the limit behavior for finite time t, i.e. before we get to the limit, one fails to recover this behavior for realistic time scales. I will argue that this leads us to reconsider the role that the rate of convergence plays in the justification of infinite limits and calls for a revision of the so-called Butterfield's principle. I will end this paper (Section 4) by offering a criterion for the justification of infinite limits based on the notion of controllable approximations.

---

[1]An exception is Norton (2014) who discusses the infinite-time limit in the explanation of reversible processes.

2 The Thermodynamic Limit in the Theory of Phase Transitions

Phase transitions are rapid changes in the phenomenological properties of a system observed, for example, in the transition from liquid water to gas. In recent years, these familiar phenomena have captured the attention of philosophers of science. The reason is that they make particularly salient the importance of infinite idealizations in the recovery of macroscopic behavior from microscopic theories. Indeed, it has been argued that without the thermodynamic limit, which consists in letting the number of particles and the volume go infinity, statistical mechanics cannot recover the values of the thermodynamic observables and, therefore, cannot give a reductive account of phase transitions.

In this section, I will explain the apparent need for the thermodynamic limit in the statistical mechanical treatment of phase transitions and I will argue – in the same vein as Butterfield (2011)– that, despite some claims about the "singular nature" of the thermodynamic limit, this idealization can be justified pragmatically.[2]

2.1 The Problem of Phase Transitions

In thermodynamics, phases correspond to regions of the parameter space where the values of the parameters uniquely specify equilibrium states. Phase boundaries, in contrast, correspond to values of parameters at which two different equilibrium states can coexist. The coexistence of different equilibrium states at phase boundaries expresses itself as discontinuities of thermodynamic quantities, which are related to the first derivatives of the free energy with respect to a parameter such as pressure or temperature. If the system intersects a phase boundary when going from one phase to another, i.e., encounters a discontinuity in a macroscopic observable, the system is said to undergo a first-order phase transition. If the system moves from one phase to another without intersecting a phase boundary, the system is said to undergo a continuous phase transition, in which case there are no discontinuities involved in the macroscopic observables, but there are divergences in the second derivatives of the free energy.

In the statistical mechanical treatment of phase transitions, which is generally constructed on the basis of Gibbs' canonical ensembles, one can describe phase transitions in terms of discontinuities or divergencies of the derivatives of the free energy by invoking the thermodynamic limit. However, it appears that one cannot do so without the infinite limit. In fact, in the canonical ensemble, the free energy is defined as the logarithm of the partition function Z:

$$F = -k_B T \ln Z, \tag{1}$$

---

[2]Since the goal here is to relate the problem of the thermodynamic limit in the theory of phase transitions with the infinite-time limit in the explanation of equilibrium, I will be deliberately brief in my exposition of the problem of phase transitions. A more detailed treatment of these topics can be found, for instance, in Kadanoff (2009), Butterfield (2011), Batterman (2001), and Butterfield and Buoatta (2011).

where k_B is the Boltzmannian constant. The partition function is the sum over all states accessible to the system:

$$Z = \sum_i e^{-\beta H_i}, \tag{2}$$

where $\beta = \frac{1}{k_b T}$ and H_i is the energy associated to a particular microstate i. Since the Hamiltonian is usually a non-singular function of the degrees of freedom, it follows that the partition function is a sum of analytic functions. As a consequence, neither the free energy nor its derivatives can have the singularities that characterize phase transitions in thermodynamics. Taking the thermodynamic limit, which consists of letting the number of particles and the volume of the system go to infinity, i.e., $N \to \infty$, $V \to \infty$, in such a way that N/V remains constant, allows one to recover those singularities and provide a rigorous definition for the phenomena that turns out to be empirically adequate.

Since we assume that real systems have a finite number of degrees of freedom, the question that arises is how can one justify the use of the thermodynamic limit in the statistical mechanical treatment of phase transitions, notwithstanding the fact that we know that it relies on an infinite idealization. In other words, what is the justification that we have for applying a theory that uses the thermodynamic limit to finite systems. One might think that what explains the success of the theory is that it provides us with a mathematical model that approximates the behavior of finite systems. The question is then under which conditions are we allowed to arrive at that conclusion.

Butterfield (2011, p. 1077) says that we could arrive at that conclusion if we could show that the value of the quantity f evaluated on the limit system $v(f_\infty)$, when $N = \infty$, is equal to the limit of the sequence of values obtained as N tends to infinity, $\lim_{N \to \infty} v(f_N)$ (this is the case of non singular limits). Moreover, he argues that in cases when the limit $\lim_{N \to \infty} v(f_N)$ is not well defined, one could arrive at the same conclusion if the values of the quantities evaluated on a system with "actual" N_0 approximate the values of the quantities evaluated on the limit system: $v(f_\infty) \approx v(f_{N_0})$.

According to Butterfield, this would allow us to give a "straightforward justification" of the limit in terms of convenience and correctness. The limit will be mathematically convenient, because it will make calculations easier (it is generally easier to deal with infinite sums than with large but finite ones) and it will be empirically correct from the point of view of finite statistical mechanics, because this mathematical model will give us an approximation of the values of the quantities that we obtain for large but finite N.

Unfortunately, as Butterfield himself recognizes, there are at least three difficulties that seem to prevent one from giving a straightforward justification to the thermodynamic limit in the treatment of phase transitions.

(i) The first difficulty, pointed out most notably by Batterman in a series of papers (2002, 2005, 2011), concerns the so-called "singular nature" of the thermodynamic limit. According to Batterman, a limit is singular "if the behavior in the limit is of a *fundamentally different character* than the nearby solutions one obtains as $\epsilon \to 0$" (2005, p. 2), where $\epsilon \to 0$ is taken as the "limiting behavior", which means

that $v(f_\infty) \not\approx v(f_{N_0})$. Batterman argues that the thermodynamic limit is *singular* in the previous sense because even if we take N to be arbitrarily large, as long as it is finite, the derivatives of the free energy will never display a singularity. It is important to note that he arrives at that conclusion by assuming that the singular behavior of a quantity is qualitatively different from its analytic behavior. As will be seen in the next section, this assumption is far from trivial.

(ii) The second difficulty regards the apparently essential role of the thermodynamic limit in the renormalization group approach. In order to give an account of the quantitative behavior of continuous phase transitions, it was necessary to incorporate renormalization group (RG) techniques. These techniques rely on the existence of non-trivial fixed points, which are points in a space of Hamiltonians at which different renormalization trajectories arrive after repeated iterations of a renormalization group transformation (details elsewhere, e.g, in Goldenfeld 1992, Wilson and Kogut 1974). It has been claimed (Batterman 2011, Morrison 2012) that the thermodynamic limit is "ineliminable" in this approach, because no matter how large we take N to be, as long as it is finite, the RG trajectory will not converge towards a non-trivial fixed point. This is supposed to follow from the fact that finite systems cannot display a divergence in the correlation length and therefore cannot present a loss in the characteristic length scale, which is necessary to define non-trivial fixed points in the space of Hamiltonians.

(iii) The third difficulty is the problem of generality. Even if we could show that in some cases the values of the quantities that successfully describe phase transitions in the limit "$N = \infty$" approximate the values of the quantities evaluated for large but finite N, there remains the question of whether this is so in all cases in which the thermodynamic limit is used to describe the phenomena of phase transitions. Landsman (2013) argues, for instance, that for the case of quantum systems displaying spontaneous symmetry breaking and the classical limit $\hbar \to 0$ of quantum mechanics, the situation is different and much more challenging than in classical phase transitions.

2.2 Butterfield's Principle and Butterfield's Solution to the Problem of Phase Transitions

The difficulties mentioned in the previous section have motivated controversial claims. For instance, it has been argued that the need for the thermodynamic limit in the theory of phase transitions and, especially, in the theory of continuous phase transitions implies the failure of the reduction of thermodynamics to statistical mechanics (Batterman 2011, Morrison 2012, Bangu 2009). Moreover, it has been argued that as a consequence of the "singular" nature of the thermodynamic limit, one should conclude that the singularities that describe phase transitions in the limit are *physically real* (Batterman 2005).

Independently of whether these conclusions actually follow from the problems pointed out above, the fact is that, in light of those difficulties, the use of the thermodynamic

limit in the theory of phase transitions appears as conceptually puzzling and requires an explanation.

So the question is: can we restore a straightforward justification for the thermodynamic limit despite the objections mentioned above? Butterfield (2011) actually argued that we can. In order to support his view, he presents a series of examples to show that the "qualitative" difference between the behavior of the relevant quantities in the limit and close to the limit is only apparent, since it is the consequence of focusing on the wrong quantities. Let me summarize his argument. Consider a sequence of functions:

$$g_N(x) = \begin{cases} -1 & \text{if } x \leq -1/N \\ Nx & \text{if } -1/N \leq x \leq 1/N \\ 1 & \text{if } x \geq 1/N \end{cases}$$

As N goes to infinity, the sequence converges pointwise to the discontinuous function:

$$g_\infty(x) = \begin{cases} -1 & \text{if } x < 0 \\ 0 & \text{if } x = 0 \\ 1 & \text{if } x > 0 \end{cases}$$

If one introduces another function f, such that

$$f = \begin{cases} 1 & \text{if } g \text{ is discontinuous} \\ 0 & \text{if } g \text{ is continuous} \end{cases}$$

then one will conclude that the value of f_∞ at the limit $N = \infty$ is *fundamentally different* from the value when N arbitrarily large but finite: $f_\infty \not\approx f_N$. Consequently we will conclude that the thermodynamic limit is "singular" in Batterman's sense. However, if one looks at the behavior of the function g one will see that the limit value of the function is approached smoothly and therefore that the limit system is not "singular" in the previous sense. Thus, if one looks only at the quantity f, one will not be able to see what is revealed when one looks at the behavior of the quantity g, namely that the limit is actually an approximate description of the behavior before we get to the limit. According to Butterfield, this is exactly what happens with classical phase transitions, and, for typical examples of phase transitions, he seems right. Consider the paramagnetic–ferromagnetic transition in magnetic materials. This transition is characterized by the divergence of a second derivative of the free energy - the magnetic susceptibility χ - at the critical point. If we introduce a quantity that represents the divergence of the magnetic susceptibility and attribute a value 1 if the magnetic susceptibility diverges and 0 if it does not (analogous to the function f in Butterfield's example), then we might conclude that the limit quantities have values that are considerably different from the values of the quantities for arbitrarily large but finite N. However, if we focus on the behavior of a different property, namely the magnetic susceptibility χ itself, we will arrive at a different conclusion. In fact, the magnetic susceptibility χ, which is the quantity that is physically meaningful, is defined as the derivative of the magnetization with respect to

an external magnetic field $\chi = \partial M / \partial H$. As N grows, the change in the magnetization becomes steeper and steeper, and the quantity smoothly approaches a divergence in the limit (analogous to the function g). This means that in statistical mechanics one can, in principle, find finite systems that have values of the magnetic susceptibility χ that approximate the thermodynamic behavior.

I take it as a moral of Butterfield's argument that the "singular" nature of a limit is not in conflict with the straightforward justification of the limit explained in section 2.1. [3] In fact, this is an argument that suggests that in finite systems, the relevant quantities that serve to describe phase transitions, such as the magnetic susceptibility, will display values that approximate the values obtained in the limit: $v(\chi_\infty) \approx v(\chi_{N_0})$, for a large but finite N_0. This satisfies the criterion for straightforward justification exposed in the previous section.

In the literature this criterion for restoring a straightforward justification has been taken as equivalent to what has been sometimes called the "Butterfield principle". As Landsman (2013) puts it:

> Butterfield's Principle is the claim that in this and similar situations, where it has been argued (by other authors) that certain properties emerge strictly in some idealisation (and hence have no counterpart in any part of the lower-level theory), "there is a weaker, yet still vivid, novel and robust behaviour [...] that occurs before we get to the limit, i.e., for finite N. And it is this weaker behaviour which is physically real." (p. 383)

Here "novel and robust" represent the behavior that is novel and robust with respect to the behavior of systems with finite N: in the case of phase transitions that is the discontinuities and singularities in the derivatives of the free energy. And the word "weak " is meant to emphasize that the behavior that arises before one gets to the limit only approximates the behavior that is observed in the limit.

Norton (2012) also seems to take this as a criterion when he suggests that most of the controversy around phase transitions is dissolved after one recognizes that this theory does not require idealizations (i.e. systems that provide inexact descriptions of the target system) but only approximations (i.e. inexact description of the target system) of the behavior of systems with *very large* number of particles.

It is important to note, however, that if one wants to transform Butterfield's criterion into a *principle*, one needs to show not only that this criterion is necessary for giving

---

[3]One needs to recognize that this only solves the first of the problems pointed out above and does not allow us to conclude that the same argument applies to other cases of phase transitions (problem (iii)), or to explain the role of the thermodynamic limit in renormalization group techniques (problem (ii)). These other problems have been studied extensively, for example, by Batterman 2011, Morrison 2012, Norton 2012 and Butterfield himself (Butterfield 2011, Butterfield and Buoatta 2011. Since I do not have space to discuss these other issues here, I will restrict my analysis to the cases in which one can actually show that the values of the quantities that describe phase transitions in the thermodynamic limit are approximately the same as the values of the quantities before we get to the limit. The paradigmatic examples are the paramagnetic-ferromagnetic transition described above and the liquid-vapor transition at the critical point in which the compressibility behaves analogously to the magnetic susceptibility.

a straightforward justification of infinite idealizations (which seems hard to deny), but also sufficient. In this respect, it is surprising that little attention has been given to the rate of convergence in the justification of infinite limits.

More to the point, if we assume that the limit will be justified if we can prove that the idealized mathematical model is just an approximation of the behavior of realistic systems, it does not suffice to show that the behavior of phase transitions can be recovered for large but finite N, but it must also be shown that it is recovered rapidly enough, which means for values of N that are physically significant, i.e. for $N_0 \approx 10^{23}$. This means that for an ϵ sufficiently small

$$||g_{N_0} - g_\infty|| \leq \epsilon, \tag{3}$$

where $N_0 \approx 10^{23}$. The size of ϵ will be determined by pragmatic considerations, such as measurement precision, estimation of finite-size effects, etc.[4] In the examples discussed here, such an ϵ exists. For instance, one can show that the value of the magnetic susceptibility χ_{N_0} for $N_0 = 10^{23}$ is negligibly close to the limit value $\lim_{N \to \infty} \chi_N$. Therefore, one can be confident that the idealized model for phase transitions is a good approximation of realistic systems. Actually, Butterfield (2011) seems to realize this, when he points out that one needs to show that the values of the relevant quantities in the limit should approximate the values of the quantities for "actual" or realistic values of N (p. 15). However, this condition is lost in what has been later considered as the "Butterfield Principle", which only refers to the values of the quantities for large but finite N and does not say anything about realistic values of N. Sure enough, in the examples of phase transitions Butterfield refers to, the values of the quantities for realistic N are so close to the values obtained in the neighborhood of the limit that distinguishing between such values does not seem to be crucial. However, this is not necessarily the case in other examples of infinite limits. Indeed, we will see next that in the infinite-time limit the values of the relevant quantities for very large but finite time can vary significantly from the values obtained for realistic t.

3 The Infinite-time Limit in the Ergodic Explanation of Equilibrium

The infinite-time limit, which consists in letting time go to infinity $t \to \infty$, has played an important role in statistical mechanics and, like the thermodynamic limit, has also been matter of controversy in the philosophical literature (e.g. Malament and Zabell 1980, Earman and Rédei 1996, Emch and Liu 2013, Sklar 1995).

In this section, I will first discuss the role of the infinite-time limit in the explanation of equilibrium in Gibbsian statistical mechanics and I will then expose the difficulties for giving a straightforward justification of the limit. Contrary to the case of the thermodynamic limit in the theory of phase transitions, I will argue that these difficulties are not related to whether or not one can recover the limit values of the relevant quantities

---

[4]I thank an anonymous referee for pointing this out

for finite t, i.e. before we get to the limit, but rather to whether or not one can recover those values for realistic t. This will reveal the important role of the rate of convergence in the justification of infinite limits.

3.1 The Problem of the Infinite-Time Limit

In order to understand the use of the infinite-time limit in the Gibbs' framework, one needs to become familiar with the Gibbs formalism. The most important concept here is the notion of *ensemble*, defined as an infinite collection of systems governed by the same Hamiltonian but distributed differently over the phase space Γ. An ensemble can also be understood as a probability distribution ρ over Γ, which reflects the probability of finding the state of a system in a certain region of Γ. The uniform probability distribution on an hypersurface Γ_E of this space Γ is referred to as the *microcanonical ensemble*, where the energy and the number of particles are constant. There is a phase function $f : \Gamma_E \to \mathbb{R}$ associated with each relevant physical quantity. The expectation values of those functions will correspond to *phase averages*, defined as follows:

$$\langle f \rangle = \int_{\Gamma_E} f \rho \, d\Gamma_E \tag{4}$$

Phase averages play an important role in this approach because they correspond to the values of the macroscopic quantities measured in experiments. In fact, if we measure the macroscopic quantities of a gas in equilibrium which is enclosed in some container, we will observe that these values coincide with the values predicted by Gibbs' phase averages, even if we do not have any information about the microscopic configuration of the gas.

The question that has puzzled physicists and philosophers of science is why phase averages coincide with values measured in real physical systems. The answer is not clear. First of all, this formalism is built upon the notion of *ensemble*, which is a fictional entity that does not make direct reference to the behavior of a single system. Second, phase averages do not tell us anything about the dynamics, i.e. they do not give us information about how the system – at the microscopic level – behaves in time. Third, this formalism does not explain why the experimental values always correspond to the average values and are not spread around the mean.[5]

The most intuitive explanation for the success of phase averages consists of associating them with *time averages* $\langle f \rangle_t$. Time averages have a clearer physical meaning because they make reference to the fraction of time that the system spends in the regions of the phase space associated to the mean values of the macroscopic observables. In other words, if we assume that measurements take some time, then we might think that we succeed in measuring phase averages because they correspond to the average values that actually occur during the time of measurement. And here is when the infinite-time limit comes into scene. In order to associate phase averages with time averages, one generally needs to introduce the infinite-time limit. For example, the Birkhoff theorem tells us

---

[5]See Frigg 2008, Uffink 2007 and van Lith 2001 for a more detailed description of the problems associated with the Gibbs formalism.

that if we define the invariant mean of time $\langle f \rangle_t$ of time dependent functions $f(t)$ (which is nothing but the same phase function f) as

$$\langle f \rangle_t = \lim_{T \to \infty} \frac{1}{T} \int_0^T f(t)\, dt, \tag{5}$$

it follows that for almost all points (except on a set with measure zero):

(i) $\langle f \rangle_t$ exists for every integrable function f in Γ_E.

(ii) If the system is ergodic, then $\langle f \rangle = \langle f \rangle_t$.

Note that in this approach, in order to derive the equivalence between phase averages and time averages $\langle f \rangle = \langle f \rangle_t$, one needs to assume that the system is ergodic, which means that as time evolves the dynamic trajectories pass through every set of nonzero measure in Γ_E.[6] The assumption of ergodicity has been itself a matter of controversy in the foundations of statistical mechanics, but for the sake of brevity I will leave this discussion aside and focus instead on the appeal to the infinite-time limit for the justification of equilibrium.

The introduction of the infinite-time limit in the definition of time averages is far from trivial, especially if one thinks that the original motivation for relating phase averages with time averages is the belief that the latter have a clearer physical meaning. In fact, we know that measurements do not take an infinite amount of time: so, what is that justifies the use of the infinite-time limit in this context? One might try to give a straightforward justification for the limit along the lines of Butterfield's principle by saying that even if the measurement times are short with respect to human macroscopic scales, they are *very long* with respect to the microscopic time scales, i.e. time of collision between particles, and therefore they are well approximated by infinite time averages. [7] If so, one might think that one has good reason to consider the infinite-time average as a mathematical model that approximate the values obtained in finite time measurements and will have good reason to give a pragmatic justification for it. For example, that it allows us to wash out fluctuations we deem irrelevant, that it is mathematically convenient and that it allows us not having to decide in advance how long the time of measurement should be. Unfortunately, there are difficulties that prevent us from arriving at this conclusion as quickly as we would like.

(i) The first is that even if the limit defined in (4) exists, it does not mean necessarily that it describes a system in equilibrium. Uffink (2007, p. 92) expresses this difficulty pointing out that generally:

$$\lim_{T \to \infty} \frac{1}{T} \int_0^T f(t)\, dt \neq \lim_{t \to \infty} f(t), \tag{6}$$

---

[6]Strictly speaking, this theorem was formulated in terms of metric transitivity instead of ergodicity. Metric transitivity is a property of dynamical systems that captures the same idea as ergodicity but in measure theoretic sense. For more details see Uffink 2007[sec.6], van Lith 2001[ch. 7]

[7]One can find arguments in this direction, for example, in Gallavotti 1999, Emch and Liu 2013.

where the right-hand side describes a constant value of a physical quantity $f(t)$ and the left-hand side represents an average value of the same quantity. Note that for periodical motions the left-hand side exists whereas the right-hand side does not.

(ii) Second, there is a problem related to the apparent indispensability of the infinite limit in the derivation of the equivalence between phase and time averages (this problem is similar but not equivalent to the problem of "singular" limits discussed above in the context of the thermodynamic limit). We saw that Birkhoff's theorem states that one can derive the equivalence between phase averages and time averages after taking the infinite time limit, but this theorem does not tell us anything about how these two averages are related for large but finite times. Frigg expresses this point as follows: "... the infinity is crucial. If we replace infinite time averages by finite ones (no matter how long the relevant period is taken to be), then the ergodic theorem does not hold any more and the explanation is false." (2008, p. 147)

(iii) Finally, there is the difficulty that even if one can show that the limit in (4) converges, this does not imply that it converges rapidly enough to be empirically meaningful. Measurement times generally take a very short time with respect to human macroscopic time scales. Thus, in order to show that the infinite time average is a good approximation for finite time averages, one needs to prove that the infinite time average is approached within realistic measurement time scales.

In the remainder of this paper, I will focus mainly on problem (iii), because it is this problem that reveals the most important difference between the thermodynamic limit discussed in Section 2 and the infinite-time limit.

3.2 The Dog-Flea Model and a Straightforward Justification for the Infinite-Time Limit

In order to understand under which conditions one could give a straightforward justification for the infinite-time limit it is useful to consider a toy model. The toy model that can best help us to grasp these conditions is the Dog-Flea model, invented by Tatjana Ehrenfest-Afanassjewa and Paul Ehrenfest (Ehrenfest and Ehrenfest-Afanassjewa 1907). A version of the model is as follows. Consider two dogs, Poomba and Woori, that share a population of N fleas. Assume further that N is even and that the fleas are labeled by an index from 1 to N. The macroscopic observables of the model are n and m, representing the number of fleas in Poomba and Woori, respectively. A microscopic description of the system corresponds to the specification of the positions of all fleas in each dog. The time evolution of the system is described like this: At every second, a number from 1 to N is taken randomly from a bag and announced. When hearing its name, the corresponding flea jumps immediately from the dog it pestered to the other. Let $p(x_t = n)$ denote the probability that there are n fleas on Poomba at time t, then the model predicts that in the long run ($t \to \infty$), and independently of the

11

initial distribution $p(x_{t=0} = n)|n = 1, 2, ..., N$, the process leads to a time-invariant probability distribution that is symmetric around the value $p = N/2$ and it is very peaked at that value, all the more so when N is large. It is important to emphasize that the model admits only one time-invariant probability distribution, which is the same as the distribution in classical probability theory that in a sequence of N trials of a fair coin, exactly p heads come up. In this way, the model illustrates quite nicely that under certain statistical assumptions, it is possible to obtain the properties that characterize equilibrium. And, analogously to the case described above, the equilibrium distribution is defined in the limit $t \to \infty$.

Following the strategy used in the previous section, we might think that the asymptotic distribution will approximate the behavior of a finite time measurement if the measurement time (macroscopic time scale) is *very long* with respect to the time that it takes for a flea to jump from one dog to the other (microscopic time scale), which is here one second. If this is the case, we might also say that the infinite-time limit is justified pragmatically, since it is mathematically convenient and it enables us to wash out fluctuations.

An advantage of the Dog-Flea model is that it allows us to perform computer simulations to test our hypothesis. Emch and Liu (2013, sec. 3.4) present the results of these simulations for two different time scales:

(a) The first run consists of 10^2 iterations.

(b) The second run consists of 10^4 iterations.

In both cases, the number of fleas is $N = 100$. Remarkably, even if the two macroscopic time scales are long with respect to the microscopic scale (a single iteration), the results for (a) are significantly different from the results obtained for (b). Whereas (a) exhibits values of n, m that are constantly changing, (b) exhibits equilibrium behavior (with chaotic fluctuations) that is in good agreement with the equilibrium distribution obtained in the infinite time limit.

Based on these results, we should conclude that the time invariant distribution (for $t \to \infty$) gives us a good approximation for the values of the macroscopic observables in (b) but not in (a). Accordingly, we can say that we are *justified* in using the limit distribution for describing the situation for (b), but not for (a). Note, that this justification is not related with whether or not the system approaches the equilibrium values in a finite time, but rather with whether or not the system approaches those values in a time that is short with respect to the time of measurement. This obliges us to consider the *convergence rate*, which represents the rapidity at which the limit is reached. In the first case (a), the convergence is not rapid enough. Indeed, the system will eventually approach equilibrium, in a long but finite time, but since this time is much longer than the measurement time, the average values of the observables will not coincide with the values predicted by the time invariant distribution. Therefore, the asymptotic average value will not provide a good approximation of the values measured during that time.

3.3 The Importance of the Rate of Convergence

For the present discussion, the important lesson of the Dog-Flea model is that talking about "long time" is useless unless we specify the relevant time scales of the problem under investigation. In this sense, if we want to justify the infinite-time limit in the explanation of equilibrium, it does not suffice to argue that the time of measurement is "very long" with respect to microscopic time scales, but rather we need to specify the rate of convergence and guarantee that the asymptotic value will be reached within the time scales that we are interested in. As one might suspect, specifying the convergence rate is not a trivial task. To give a more precise idea, let $\langle f(T) \rangle$ represent the average value calculated at time T, that is:

$$\langle f(T) \rangle = \frac{1}{T} \int_0^T f(t)dt. \tag{7}$$

Then in order to determine the convergence rate, one needs to find a finite $\epsilon(T)$ such that:

$$||\langle f(T) \rangle - \langle f \rangle_t || \leq \epsilon(T), \tag{8}$$

where $\langle f \rangle_t$ is the time invariant mean defined in (4). Even in simple models, to obtain definite values of $\epsilon(T)$ is often difficult in both theory and practice, and to demonstrate that this value is very small, i.e, $\epsilon(T) \approx 0$, for realistic measurement times is even harder. More importantly, it is perfectly conceivable to have a situation in which the values of the functions are constantly changing so that the time needed to attain the time average is of the order of the recurrence time, i.e. the time necessary to visit the entire surface Γ_E. One can estimate that the recurrence time for a small sample of diluted hydrogen gas is unimaginably longer than the age of the universe, and this time is even longer if we consider more complicated systems. In situations like this, there might well exist a finite $\epsilon(T)$ that satisfies eq.(7). However, the time for which ϵ is sufficiently small will be much longer than realistic measurement times, which means that for realistic time scales T', say $2/10 \, sec.$, $\langle f(T') \rangle \not\approx \langle f \rangle_t$.[8]

The previous argument just tells us that, even if we could demonstrate that the asymptotic average will be reached within finite but very large times (or in other words "on the way to the limit" as in Butterfield's principle), this does not imply that the asymptotic average will be reached for realistic t and, therefore, it does not imply that we can interpret the limit as giving us a good approximation of the systems that we are interested in. This has an important philosophical consequence because it tells us that the so-called "Butterfield's principle" is not sufficient to justify the limit in this case.

Boltzmann himself was aware of the problem of the rate of convergence in the justification of the infinite-time limit, and in order to reconcile this limit with the rapid approach to equilibrium, assumed that the "the macroscopic observables, had

---

[8]For a quantitative estimation, see Gallavotti 1998

an essentially constant value on the surface of given energy with the exception of an extremely small fraction ϵ of cells" (1874 [quoted in Gallavotti 1999, p. 16]). Unfortunately, this assumption is not uncontroversial, and to some extent it does not really solve the problem. In fact, even if we accept the premise postulating, for example, that the functions satisfy symmetry conditions, we still need an argument to associate phase averages with time averages. In other words, we still need an argument that allows us to conclude that the system does not spend so much time in the small fraction of cells that differ from the mean phase values. Ironically, this seems to beg the question, in that it brings us back to the original problem for which the infinite time limit entered the picture, namely the problem of deriving the equivalence between phase and time averages.

Different alternatives have been offered in the literature to deal with this and the other problems associated with infinite time averages. Maybe the most radical was the proposal by Malament and Zabell (1980), where they argue that one can explain the empirical adequacy of phase averages without appealing to time averages at all. Their argument is based on two assumptions: i) the system exhibits small dispersion with respect to the phase average (analogously to Boltzmann's assumption), and ii) the microcanonical measure represents the probability of finding a system in a particular region of the phase space. According to them, these two assumptions taken together lead to the conclusion that the probability that phase functions are always close to their phase averages is very large, without making any reference to infinite time averages. Even if this view looks appealing, two main criticisms have been raised in the literature. The first is that in order to justify assumption (ii), they invoke a version of ergodicity, which is an hypothesis that has been questioned in the foundations of statistical mechanics (e.g. Earman and Redei 1998, Frigg 2008, van Lith 2001). The second, which is more important for us, is that they justify assumption (i) based on Khinchin-Lanford dispersion theorems, which tell us that for functions that satisfy strong symmetry conditions, the dispersions from the mean will go to 0 in the thermodynamic limit. The appeal to the thermodynamic limit would not be problematic, if we could demonstrate that – like the case of phase transitions – there is a straightforward justification for it. Unfortunately, the use of the thermodynamic limit in this context appears to be less straightforward than in the case of phase transitions, because Butterfield's principle is not enough to justify the limit. In fact, for realistic $N \approx 10^{23}$, one can estimate, based on Khinchin's theorem, that the probability that there is a relative deviation from the mean of more than a tiny ϵ is very small, but not sufficiently small to discard that these states will occur in nature. This means that one cannot regard (at least not without risks) the asymptotic results obtained in this and other similar theorems as providing us with a good approximation of the behavior of realistic systems. This problem is also referred in the literature as the measure-epsilon problem (See Uffink 2007, van Lith 2001 and Frigg 2008).

An alternative approach can be found in Earman and Redei (1996). They do not invoke ergodicity for the explanation of the success of phase averages, but they are

quite sympathetic towards the explanatory role of "ergodic-like behavior". According to them, ergodic-like behavior only requires weak mixing behavior with respect to a set of finite observables. It is important to note that the definition of mixing offered by them still requires the appeal to the infinite-time limit. Interestingly for what we are discussing here, they explicitly include *rapid convergence* as an additional condition for the explanation of equilibrium. To justify this assumption they suggest (although not necessarily endorse) two possible routes: a) The first is to make reference to matter-of-fact initial states. b) The second is to assume that systems are subjected to perturbations from outside that act as a kind of 'stirring' mechanism which rapidly drive the observed values of the macroscopic quantities.[9] Even if one should not discard that some progress can be done in each of these lines of research, one should recognize that they are methodologically complicated since they oblige us to the consider specific features of the systems of interest.

The explanation of the empirical success of phase averages is still an open problem in the foundations of statistical mechanics. Although there is some skepticism in the philosophical literature towards the idea of explaining this success via infinite time averages, the infinite-time limit continues playing an important role in physics. It is far beyond the scope of this paper to offer a final assessment for the appeal to the infinite-time limit in the explanation of equilibrium. However, it suffices for our purposes to have shown that much of the problems for providing a justification for such a limit come from the conceptual and methodological difficulties to specify the rapidity at which the limit is approached. I argued that this has an important consequence for the current philosophical literature on infinite limits, because it teaches us an important lesson about the role of the convergence rate in the justification of infinite limits.

4 Conclusion: Infinite Limits as Controllable Approximations

Although there is no consensus regarding the status of infinite limits in physics, it seems reasonable to interpret these idealizations as mathematical models that approximate the behavior of finite systems. The question that one needs to ask, however, is under which conditions are we allowed to arrive at that conclusion. In the debate on phase transitions, it is often assumed that we are allowed to interpret the infinite limit as providing an approximation of finite systems as long as the behavior that arises in the limit also arises, at least approximately, "on the way to the limit", which is what we called here the " Butterfield principle". However, in this paper I argued that in the case of the infinite-time limit this condition is not sufficient to justify the limit. This is because in this case the values of the relevant quantities "before we get to the limit", that is for finite but very large t, can take values significantly different from the values obtained for realistic time scales t.

---

[9] A review of this attempts can be found in Lanford 1973

The above result leads us to a revision of Butterfield's principle that would apply more generally than the original formulation. A proposal is as follows:

> We can justify infinite limits, when $x \to \infty$, as being mathematical models that approximate the behavior of real finite systems, iff (i) the behavior that arises in the limit also occurs, at least approximately, before we get to the limit, i.e., for finite x., and (ii) it also arises for *realistic values* of x.

A concept that captures the main idea of the previous statement is the notion of *controllable* approximations. Emch and Liu (2002, p. 526) define controllable approximations as the ones in which the deviations of the model with respect to realistic systems can be quantitatively estimated. When no such estimation can be given, the approximation is said to be uncontrollable. Uffink (2007, p. 109) makes this notion more precise, suggesting that in the case of controllable approximations involving infinite limits one has control over how large the value of the parameter must be to assure that the infinite limit is a reasonable substitute for a finite system. Since we are interested in the behavior of realistic systems, I claim that this "control" should also involve a specification of the rate of convergence. This will allow us to warrant that the limit is reached for realistic values of the parameters and therefore that it is a good approximation of the target systems. In the cases of phase transitions analyzed in Section 2, the thermodynamic limit appears to be controllable in this sense. However, for what has been argued in Section 3, we do not seem to be in the position of deriving the same conclusion for the case of the infinite-time limit.

Acknowledgments

References

Bangu, S. (2009). Understanding thermodynamic singularities: Phase transitions, data, and phenomena. *Philosophy of Science*.

Bangu, S. (2011). On the role of bridge laws in intertheoretic relations. *Philosophy of Science*, *78(5)*, 1108-1119.

Batterman, R. W. (2001). *The devil in the details: Asymptotic reasoning in explanation, reduction, and emergence.* Oxford University Press.

Batterman, R. W. (2005). Critical phenomena and breaking drops: Infinite idealizations in physics. *Studies in History and Philosophy of Science Part B: Studies in History and Modern Physics*, *36(2)*, 225-244.

Batterman, R. W. (2011). Emergence, singularities and symmetry breaking. *Foundations of Physics*, *41(6)*, 1031-1050.

Butterfield, J. (2011). Less is different: emergence and reduction reconciled. *Foundations of Physics*, *41(6)*(1065-1135).

Butterfield, J., & Buoatta, N. (2011). Emergence and reduction combined in phase transitions. *Frontiers of Fundamental Physics*, *FFP11*.

Earman, J., & Rédei, M. (1996). Why erogic theory does not explain the success of equilibrium statistical mechanics. *The British Journal for the Philosophy of Science*.

Ehrenfest, P., & Ehrenfest-Afanassjewa, T. (1907). Über zwei bekannte einwände gegen das boltzmannsche h-theorem. *Hirzel*.

Emch, G., & Liu, C. (2013). *The logic of thermostatical physics*. Springer Science & Business Media.

Frigg, R. (2008). A field guide to recent work on the foundations of statistical mechanics. In *The ashgate companion to contemporary philosophy of physics*. Ashgate.

Gallavotti, G. (1999). *Statistical mechanics: A short treatise*. Springer-Verlag Berlin Heidelberg.

Goldenfeld, N. (1992). *Lectures on phase transitions and the renormalization group*. Westview Press.

Kadanoff, L. P. (2009). More is the same: Phase transitions and mean field theories. *Journal of Statistical Physics*, *137(5-6)*.

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 Modern Physics*, *44(4)*(379-394).

Lanford, O. (1973). Entropy and equilibrium states in classical statistical mechanics. In A. Lenard (Ed.), *Statistical mechanics and mathematical problems*. Springer-Verlag.

Malament, D., & Zabell, L. (1980). Why gibbs phase averages work: the role of ergodic theory. *Philosophy of Science*.

Morrison, M. (2012). Emergent physics and micro-ontology. *Philosophy of Science*.

Norton, J. D. (2012). Approximation and idealizations: Why the difference matters. *Philosophy of Science*, *79(2)*, 207-232.

Norton, J. D. (2014). Infinite idealizations. *European Philosophy of Science-Philosophy of Science in Europe*.

Sklar, L. (1995). *Philosophical issues in the foundations of statistical mechanics*. Cambridge University Press.

Uffink, J. (2007). Compendium of the foundations of classical statistical physics. In J. Butterfield & J. Earman (Eds.), *Philosophy of physics*. Amsterdam: North Holland.

van Lith, J. (2001). *Stir in sillness: A study in the foundations of equilibrium statis-*

tical mechanics. (Unpublished doctoral dissertation). University of Utrecht.

Wilson, K., & Kogut, J. (1974). The renormalization group and the e expansion. *Physics Reports*, *12(2)*, 75-199.