Western Graduate&PostdoctoralStudies

Western University

Scholarship@Western

Electronic Thesis and Dissertation Repository

8-17-2015 12:00 AM

On Nonequilibrium Statistical Mechanics

Joshua M. Luczak
The University of Western Ontario

Supervisor
Wayne C. Myrvold
The University of Western Ontario

Graduate Program in Philosophy
A thesis submitted in partial fulfillment of the requirements for the degree in Doctor of Philosophy
© Joshua M. Luczak 2015

Follow this and additional works at: https://ir.lib.uwo.ca/etd

Part of the Philosophy of Science Commons

Recommended Citation

Luczak, Joshua M., "On Nonequilibrium Statistical Mechanics" (2015). *Electronic Thesis and Dissertation Repository*. 3173.
https://ir.lib.uwo.ca/etd/3173

This Dissertation/Thesis is brought to you for free and open access by Scholarship@Western. It has been accepted for inclusion in Electronic Thesis and Dissertation Repository by an authorized administrator of Scholarship@Western. For more information, please contact wlswadmin@uwo.ca.

ON NONEQUILIBRIUM STATISTICAL MECHANICS

(Thesis Format: Monograph)

by

Joshua <u>Luczak</u>

Graduate Program in Philosophy

A thesis submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy

The School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

Abstract

This thesis makes the issue of reconciling the existence of thermodynamically irreversible processes with underlying reversible dynamics clear, so as to help explain what philosophers mean when they say that an aim of nonequilibrium statistical mechanics is to underpin aspects of thermodynamics.

Many of the leading attempts to reconcile the existence of thermodynamically irreversible processes with underlying reversible dynamics proceed by way of discussions that attempt to underpin the following *qualitative* facts: (i) that isolated macroscopic systems that begin away from equilibrium spontaneously approach equilibrium, and (ii) that they remain in equilibrium for incredibly long periods of time. These attempts standardly appeal to phase space considerations and notions of typicality. This thesis considers and evaluates leading typicality accounts, and, in particular, highlights their limitations. Importantly, these accounts do not underpin a large and important set of facts. They do not, for example, underpin facts about the rates in which systems approach equilibrium, or facts about the kinds of states they pass through on their way to equilibrium, or facts about fluctuation phenomena. To remedy these and other shortfalls, this thesis promotes an alternative, and arguably more important, line of research: understanding and accounting for the success of the techniques and equations physicists use to model the behaviour of systems that begin away from equilibrium. Accounting for their success would help underpin not just the qualitative facts the literature has focused on, but also many of the important quantitative facts that typicality accounts cannot.

This thesis also takes steps in this promising direction. It outlines and examines a technique commonly used to model the behaviour of an interesting and important kind of system: a Brownian particle that's been introduced to an isolated homogeneous fluid at equilibrium. As this thesis highlights, the technique returns a wealth of quantitative and qualitative information. This thesis also attempts to account for the success of the model and technique, by identifying and grounding the technique's key assumptions.

Acknowledgements

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Contemporary philosophical discussions of statistical mechanics primarily focus on two issues: securing a foundation for thermodynamics, and interpreting statistical mechanical probabilities. This thesis speaks to the first of these issues.[1] In particular, it speaks to aspects of the foundational debate that make contact with nonequilibrium statistical mechanics.

This thesis has several aims. One of the aims is to help explain what philosophers mean when they say that an aim of nonequilibrium statistical mechanics is to account for aspects of thermodynamics. While there are a number of aspects of thermodynamics whose justification relies on nonequilibrium statistical mechanics, philosophers have tended to focus on the issue of reconciling the existence of thermodynamically irreversible processes with underlying reversible dynamics. Interestingly, and perhaps somewhat surprisingly, this goal is expressed in different ways throughout the litera-

---

[1] For good introductions and discussions on how we should understand the probability distributions introduced into statistical mechanics, see Sklar (1995), Frigg (2008b), Myrvold (2012), and Myrvold (Forthcoming).

ture. Or rather, what's seen as central to achieving this goal differs across authors—whether this is intended or not. To an outsider, these differences may suggest that there is a disagreement or confusion in the literature. It is an aim of this thesis to make the issue of reconciling the existence of thermodynamically irreversible processes with underlying reversible dynamics more clear, and to explain why there is neither a disagreement or confusion in the literature. This is done so as to satisfy the more general aim of helping to explain what philosophers mean when they say that an aim of nonequilibrium statistical mechanics is to underpin aspects of thermodynamics.

Many of the leading attempts to reconcile the existence of thermodynamically irreversible processes with underlying reversible dynamics proceed by way of discussions that attempt to underpin the following *qualitative* facts: (i) that isolated macroscopic systems that begin away from equilibrium spontaneously approach equilibrium, and (ii) that they remain in equilibrium for incredibly long periods of time. These attempts standardly appeal to phase space considerations and notions of typicality. It is an aim of this thesis to consider and evaluate leading typicality accounts, and, in particular, to highlight their limitations. Importantly, these accounts do not underpin a large and important set of facts. They do not, for example, underpin facts about the rates in which systems approach equilibrium, or facts about the kinds of states they pass through on their way to equilibrium, or facts about fluctuation phenomena. To remedy these and other shortfalls, this thesis promotes an alternative, and arguably more important, line of research: understanding and accounting for the success of the techniques and equations physicists use to model the behaviour of systems that begin away from equilibrium. Ac-

counting for their success would help underpin not just the qualitative facts the literature has focused on, but also many of the important quantitative facts that typicality accounts cannot.

This thesis also takes steps in this promising direction. It is an aim of the fourth chapter to outline, examine, and ground the success of a technique commonly used to model the approach to equilibrium and subsequent behaviour of a Brownian particle that's been introduced to an isolated homogeneous fluid at equilibrium. Interestingly, the model is easily generalised, and so is applicable to a wide variety of systems. This chapter attempts to account for the success of the model, by identifying and grounding the technique's key assumptions.

It's standard for philosophical discussions of statistical mechanics to take place within a classical mechanical framework. This thesis follows this convention and eschews quantum mechanical considerations. While the following view will not be defended in this thesis, it is the view of the author that classical statistical mechanics needs no help from quantum mechanics to make sense; classical statistical mechanics can stand on its own.

1.1 Chapter Summaries

1.1.1 Chapter 2

A number of authors claim that an aim of statistical mechanics is to provide a suitable foundation for thermodynamics. With respect to the role nonequilibrium statistical mechanics plays in the foundational project, philosophers have tended to focus on reconciling the existence of thermodynamically irre-

versible processes with underlying reversible dynamics. Interestingly, what's seen as central to this task differs across authors. Or rather, and more accurately, many authors use distinct concepts when they outline what they see as central to reconciling the tension between these macroscopic processes and the underlying dynamics of those systems that give rise to them. Importantly, these differences may indicate, to those unfamiliar with the literature, that there is either a confusion or disagreement in the literature. This chapter highlights these differences. This is done via a discussion of various reversibility concepts and several laws of thermodynamics. It also explains why these differences, in fact, do not amount to there being a confusion or disagreement in the literature, and why well informed authors needn't (and probably don't) worry about eliminating these differences. This chapter encourages thinking of these different statements as short, approximate statements that refer to a collection of relevant ideas central to the reconciliation issue.

This chapter, by way of making the reconciliation issue more clear, helps explain what philosophers mean when they say that an aim of statistical mechanics is to underpin aspects of thermodynamics. And this, in turn, contributes to this thesis's aim of encouraging that we pursue an alternative line of research within the philosophical literature on statistical mechanics: to understand and account for the success of the techniques and equations physicists use to model the behaviour of systems that begin away from equilibrium. By better understanding the goals philosophers have traditionally set themselves, we'll not only be able to better situate, appreciate, and evaluate their attempts to achieve them, but we'll also better appreciate how this proposed line of research fits with contemporary discussions of statistical

mechanics.

1.1.2   Chapter 3

Chapter two discusses the reconciliation of thermodynamically irreversible processes with underlying reversible dynamics with reference to thermodynamics. This chapter discusses the issue from the perspective of statistical mechanics.

One of the aims of this chapter is to outline and review leading attempts to address the reconciliation issue. These attempts are typically part and parcel of works that aim at accounting for why macroscopic systems approach equilibrium and why they remain in equilibrium for incredibly long periods of time. Almost all of the attempts discussed in this chapter appeal to a notion of typicality.

This chapter discusses the most pressing problems with typicality accounts and, where possible, discusses their most promising solutions. It is argued that they have severe limitations, even when we overlook their problems. It is argued that they do not underpin a large and important set of facts that concern the behaviour of systems that begin away from equilibrium. While they may account for why isolated macroscopic systems approach equilibrium and why they remain in equilibrium for incredibly long periods of time, they do not, among other things, underpin facts about the rates in which systems approach equilibrium, or about the kinds of states they pass through on their way to equilibrium, or about fluctuation phenomena. They do not help us form expectations about these aspects of a system's behaviour or help justify the expectations we may already have about them,

having formed them on the basis of experience.

This chapter claims that the limitations of typicality accounts are a symptom of what they are aiming at. By focusing on recovering aspects of thermodynamics, those contributing to the foundational project have merely been in the business of articulating microphysical accounts of irreversible macroscopic behaviour that recover a few qualitative facts. To remedy these shortfalls, this chapter promotes pursuing a different line of research; one in which we attempt to understand and ground the success of the techniques and equations physicists use to model the behaviour of systems that begin away from equilibrium. It is claimed that accounting for their success would help underpin not just the qualitative facts the literature has focused on, but also many of the important quantitative facts that typicality accounts cannot.

This chapter is composed of five sections. The first section outlines and discusses the essence of what many people think of as the first serious attempt to account for the behaviour of systems that begin away from equilibrium. This approach is based on Ludwig Boltzmann's early work on statistical mechanics, when he first offers what has come to be known as Boltzmann's equation and H-theorem. As this section highlights, this work gives rise to one of the very issues that is at the centre of contemporary discussions of statistical mechanics: the reconciliation of thermodynamically irreversible processes with underlying reversible dynamics. The second section outlines and discusses leading ways of accounting for the behaviour of systems that begin away from equilibrium that also addresses the tension between their macrolevel and microlevel descriptions. These approaches are inspired by, and build on, Boltzmann's attempts (circa 1877) to resolve this tension and,

once again, account for the behaviour of systems that begin away from equilibrium. As we'll see, each one incorporates a notion of typicality. The second section also considers worries specific to each kind of typicality account discussed. More general concerns are discussed in the third section. Both sections also discuss, where possible, the most promising solutions to these concerns. The fourth section highlights and discusses the limitations of typicality accounts. It also encourages those working in the field to begin work on understanding and grounding the success of the techniques and equations physicists use to model the behaviour of systems that begin away from equilibrium. The fifth section summarises the material presented in this chapter.

1.1.3 Chapter 4

Chapter three ends by highlighting the limitations of typicality accounts and by suggesting that we pursue a line of research aimed at properly understanding and accounting for the success of the techniques physicists use to model the behaviour of systems that begin away from equilibrium.

This chapter takes steps in this promising direction. It outlines and examines a technique commonly used to model the behaviour of an interesting and important kind of system. More accurately, this chapter outlines and examines a technique commonly used to model the behaviour of a Brownian particle that's been introduced to an isolated homogeneous fluid at equilibrium. As we'll see, the technique returns a wealth of quantitative and qualitative information. This chapter also attempts to account for the success of the model and technique, by identifying and grounding the technique's

key assumptions.

This chapter is composed of four sections. The first section details a way of modelling the Brownian particle's approach to equilibrium and subsequent behaviour. The approach takes its cue from the theory of Brownian motion. It appeals to Langevin dynamics. As this section highlights, the described technique generates a collection of interesting equations that track important aspects of the system's behaviour. These equations provide us with the resources to answer the kinds of questions commonly asked about the behaviour of this kind of system. The second section expands and comments on several aspects of the approach. It also identifies the technique's key assumptions and discusses their motivation. This section also begins a discussion about what would have to be true, or be at least approximately true, at the microscopic level, in order to justify the use of these assumptions, and to account for the success of the equations they help generate. The third section attempts to motivate a crucial microphysical fact, which, it is claimed, justifies the use of the technique's key assumptions, grounds the success of the technique, and underpins the facts the equations it helps generate track. The fourth section summarises the details of this chapter. It also contains some suggestions about the direction of future investigations.

Chapter 2

On Some of the Aims of Nonequilibrium Statistical Mechanics

A number of authors claim that *the aim* of statistical mechanics is to provide a suitable foundation for thermodynamics.[1,2] For example, in a recent review article Roman Frigg writes:

> Thermodynamics (TD) correctly describes a large class of phenomena we observe in macroscopic systems. The aim of statistical mechanics is to account for this behaviour in terms of the dynamical laws governing the microscopic constituents of macroscopic systems and probabilistic assumptions. (Frigg 2008b: p.99)

---

[1]See, for example, Sklar (1995: p.3), Lebowitz (1999: p.346), Uffink (2007: p.923), and Frigg and Werndl (2012b: pp.99-100).

[2]This observation has also recently been made by Wallace (2013).

Others, perhaps more cautiously, claim that it is *a central aim* of statistical mechanics to provide a foundation for thermodynamics. Katinka Ridderbos (2002: p.66), for example, says that

> One of the cardinal aims of the theory of statistical mechanics is to underpin thermodynamic regularities by a theory formulated in terms of the dynamical laws governing the motion of the microscopic constituents of a thermodynamic system.

While Craig Callender (2001: p.540) writes,

> Kinetic theory and statistical mechanics are in part attempts to explain the success of thermodynamics in terms of the basic mechanics.

As one might have expected, the consensus among philosophers and philosophically minded physicists (which will hereafter collectively be referred to as philosophers) is that it is an aim of *nonequilibrium* statistical mechanics to underpin certain aspects of thermodynamics. While there are a number of aspects of thermodynamics whose justification relies on nonequilibrium statistical mechanics, philosophers have tended to focus on the issue of reconciling the existence of thermodynamically irreversible processes with underlying reversible dynamics. Interestingly, what's regarded as central to this task differs across authors. Authors sometimes claim that:

G1. It is a goal of nonequilibrium statistical mechanics to account for thermodynamic behaviour, where we say that an isolated system is behaving thermodynamically if it's in equilibrium or if it's spontaneously approaching equilibrium.[3]

G2. It is a goal of nonequilibrium statistical mechanics to account for irreversible macroscopic behaviour.[4]

G3. It is a goal of nonequilibrium statistical mechanics to show how a time-reversal non-invariant thermodynamics emerges from a time-reversal invariant microphysics.[5]

G4. It is a goal of nonequilibrium statistical mechanics to provide a microphysical justification for the Second Law of thermodynamics.[6]

Importantly, these goals are distinct. They are, however, closely related. For one thing, **G1** and **G2** do not make explicit reference, like **G3** and **G4**, to some microphysics. It's clear, however, from the context in which they appear, that what authors mean when they say that it's a goal of nonequilibrium statistical mechanics to account for thermodynamic behaviour, or irreversible macroscopic behaviour, is that it's a goal to provide a microphysical underpinning of these concepts.

---

[3]See, for example, Lazarovici and Reichert (2014: p.3).

[4]See, for example, Frigg (2011: p.77).

[5]See, for example, Wallace (2013: p.1). Actually, Wallace (2013: p.1) describes something weaker than this. More accurately, he describes the goal as being one in which writers have tried to show how the *apparent* time-irreversibility of thermodynamics can be reconciled with the *apparent* time-reversibility of microphysics. Since this thesis has confined itself to talking about classical mechanics and classical thermodynamics, it seems reasonable to drop the "apparent" qualifier.

[6]See, for example, Frigg (2008a: p.670) and Lazarovici and Reichert (2014: p.4).

One of the aims of this chapter is to help explain what philosophers mean when they say that an aim of nonequilibrium statistical mechanics is to account for certain aspects of thermodynamics. Another is to highlight the ways in which **G1-G4** are distinct, and to highlight the ways in which they are related. A third is to explain why these differences are negligible with respect to the reduction of thermodynamics to statistical mechanics. A fourth is to explain why these distinct but related ways of expressing what's central to reconciling the tension between macrolevel and microlevel descriptions of relevant phenomena do not reveal either a confusion or disagreement in the literature. All of these aims are related. By discussing the ways in which these goals are distinct, related, but ultimately whose differences are negligible with respect to the reduction of thermodynamics to statistical mechanics, this chapter hopes to make the issue of reconciling the existence of thermodynamically irreversible processes with underlying reversible dynamics more clear. If it's successful, then this chapter will have helped articulate what it is those attempting to underpin aspects of thermodynamics using nonequilibrium statistical mechanics are actually trying to achieve. The attempt to satisfy each of these aims also contributes, importantly, to this thesis's larger aim of encouraging the pursuit of a different line of research. By better understanding thermodynamics, its relationship to statistical mechanics, and the facts philosophers have traditionally tried to underpin using statistical mechanics, we'll be able to better situate, and appreciate, this thesis's promotion of understanding and accounting for the success of the techniques physicists use to model the behaviour of systems that begin away from equilibrium.

While it's helpful for this thesis's narrative that this chapter proceeds by way of a discussion about some distinct but roughly equivalent claims about what's central to reconciling the existence of irreversible macroscopic processes with an underlying reversible dynamics, it also serves the purpose of later clarifying the relationship between thermodynamics and statistical mechanics. As a later section explains, the general expectation, among those working on the reduction of thermodynamics to statistical mechanics, is that the laws and concepts of thermodynamics appear in statistical mechanics as complicated, approximate statements that are true under certain conditions. Importantly, it is for this reason that these distinct but related claims do not properly indicate that there is either a confusion or disagreement in the literature.

This chapter consists of three sections. The first section introduces and discusses some concepts, and some thermodynamic laws, that are relevant to ways of interpreting **G1-4**. They are also relevant for understanding the ways in which these goals are distinct but related. Since much of the foundational literature is concerned with recovering *classical thermodynamics*, this discussion focuses on classical thermodynamic concepts and classical thermodynamic laws—though often in their modern guise. This is distinguished from the less orthodox, more formal, axiomatic treatments of thermodynamics that have appeared since the early 1900's, and which are associated with Constantin Carathéodory, and Elliott Lieb and Jakob Yngvason.[7] The results of this discussion are summarised in the second section of this chapter.

---

[7]See Carathéodory (1909), and Lieb and Yngvason (1999). Other axiomatic approaches have been offered by Robin Giles (1964), John Boyling (1972), and Josef-Maria Jauch (1972, 1975), among others.

The third section explains why the differences between **G1-4**, and goals similar to them, are negligible with respect to the reduction of thermodynamics to statistical mechanics, and, relatedly, why these differences do not point to either a confusion or disagreement in the literature. It also explains why authors contributing to the foundational discussion needn't worry, and probably don't worry, about eliminating these differences.

2.1 Classical Thermodynamics

Classical thermodynamics (which is sometimes called orthodox thermodynamics) was developed around 1850. It's usually associated with authors such as Rudolf Clausius, Lord Kelvin (William Thomson), and Max Planck—though some trace its origins back to Sadi Carnot's work on heat engines.[8] Thermodynamics characterises macroscopic systems in terms of macroscopically measurable quantities, e.g. temperature, pressure, volume, etc. It also describes changes in them in terms of heat and work exchanges with an environment. The theory rests on a set of fundamental laws. These laws are intended to be independent of any particular hypothesis concerning the microscopic constitution of macroscopic systems. They have traditionally been understood as generalised statements of experimental facts.

This section outlines and discusses four thermodynamic laws. They are

---

[8]See Mendoza and Carnot (1960). Most authors consider classical thermodynamics to have emerged around 1850, in works that attempted to recast *Carnot's theorem* (see section 2.1.5)—which was originally expressed using terms familiar to the caloric theory of heat—in what we now call classical thermodynamic terms. Other early, important, but lesser known contributors to classical thermodynamics are Émile Clapeyron, William Rankine, and Ferdnand Reech. See Uffink (2001, 2007) for more on the interesting history of thermodynamics.

known as the Minus First Law, the Zeroth Law, the First Law, and the Second Law of thermodynamics. Interestingly, only two of these laws—the First and Second—were stated as laws in original presentations of the theory. While many modern thermodynamics textbooks include the Zeroth Law, explicit presentations of the Minus First Law as a law are, at present, only found in philosophical works. A number of modern textbooks do however include a Third Law of thermodynamics. Unlike the other laws, the Third Law is not relevant to the aims of this chapter, so it will not form part of the discussion.[9] The discussion will focus on highlighting aspects of the theory that bear on **G1-4**. This will be achieved mostly through a discussion of thermodynamic laws. This section will also highlight, along the way, some of the relations between **G1-4**. But before turning to a discussion of these laws, it is important that we first disentangle and discuss some important and closely related concepts.

2.1.1 Time-Reversal Invariance and "Reversibility"

The state of a system is typically represented by a point in some state space Ω. For example, the *thermodynamic state* of a system is typically represented by a point in a state space characterised by a small number of macroscopically measurable parameters.[10] The same system can also be represented in classical phase space, a 6-n-dimensional space in which each point represents the position and momentum $(\mathbf{x_i}, \mathbf{p_i})$ of every particle constituting the

---

[9]See Kardar (2007: Sec.1.10) for an introduction and discussion of the Third Law.

[10]The thermodynamic state of a system need not be characterised by *directly* macroscopically measurable parameters. It can, for example, be characterised using energy and entropy, which have to be measured indirectly via the First and Second laws.

system. Unlike thermodynamic state spaces, points in phase space represent unique microstates. As a consequence, they provide a microphysical, and hence more detailed, picture of the system's state.

A state *history* is a trajectory through state space. That is, a mapping $\sigma : I \subseteq \mathbb{R} \to \Omega$, for some time interval I.

Systems are governed by dynamical laws. They enable us to distinguish a set $\mathcal{D}$ of *dynamically possible* trajectories from *kinematically possible* trajectories.

For any time t_0, we can define a reflection of the time axis around t_0 by

$$t \to t^T = t_0 - (t - (t_0)). \tag{2.1}$$

Standardly, we take $t_0 = 0$, so that $t^T = -t$.

We can talk about a state's *time-reversal*. For a single classical particle, the time-reversal operation

$$(\mathbf{x}, \mathbf{p}) \to (\mathbf{x}, -\mathbf{p}) \tag{2.2}$$

returns time-reversed states. For systems that are constituted by many classical particles, the relevant time-reversal operation reverses each particle's momentum.

We can also talk about the time-reversal invariance of physical laws and theories. Given a time reflection (i.e. (2.1)) and a state reversal operation (such as (2.2)), whose general form we write as

$$\omega \to \omega^T, \tag{2.3}$$

we can define an operation that reverses *state histories*. Define the history-reversal operation

$$\sigma \to \sigma^T \tag{2.4}$$

by

$$\sigma^T(t) = \sigma(t^T)^T. \tag{2.5}$$

So then, if σ includes a sequence of states $\dots \sigma(t_1), \sigma(t_2), \sigma(t_3), \sigma(t_4)$, $\dots$, then the time-reversed history includes a sequence of states $\dots \sigma(t_4^T)^T$, $\sigma(t_3^T)^T, \sigma(t_2^T)^T, \sigma(t_1^T)^T, \dots$.

A physical theory is said to be *time-reversal invariant* (or symmetric under time-reversal) if and only if, whenever a state history σ is dynamically possible, the time-reversed state history σ^T is also dynamically possible. Another way of putting this is to say that the theory is time-reversal invariant if and only if $\mathcal{D}^T \subseteq \mathcal{D}$. A theory is said to be *time-reversal non-invariant* (or asymmetric under time-reversal) if and only if it's not time-reversal invariant. I.e. if there exists at least one state history that is dynamically possible whose time-reversed history is not.[11]

Notice that it's the form of a theory's dynamical laws, given a state reversal operation, that determines whether or not the *theory* is time-reversal invariant. This is why authors sometimes speak of the time-reversal invariance of physical laws rather than the time-reversal invariance of physical theories,

---

[11]Asymmetry should not be confused with the stronger notion of *antisymmetry*. A physical theory is said to be *antisymmetric* under time-reversal if and only if *every* dynamically possible state history has a time-reversed history that is not dynamically possible.

and why they sometimes use the expression "time-reversal invariant theory" interchangeably with "time-reversal invariant dynamical laws".

"Reversible" is a term that has several meanings. Sometimes it's used to refer to time-reversal invariant processes. That is, dynamically possible state histories whose time-reverse state histories are also dynamically possible. This use most commonly appears in philosophy of science and philosophy of physics literature. In these contexts, "irreversible" is used to refer to time-reversal non-invariant processes. That is, dynamically possible state histories whose time-reverse state histories are not, according to the theory, dynamically possible.

In other situations, "reversible" is used to mean, what Jos Uffink (2001: p.316) has called, *recoverable*. Experience suggests that in many cases the transition, by some process, from an initial thermodynamic state s_i to a final thermodynamic state s_f, *cannot be fully undone.* The expression "fully undone" is meant to indicate not just a return of *the system* to its initial thermodynamic state but also *any auxiliary system* with which it interacted, without any additional change associated therewith. Opening a bottle of perfume in a well ventilated room is an example of an irreversible processes in this sense. In this and other such cases, there is no process which starts off from the final state s_f that restores the initial state s_i, completely. As Uffink (2001: pp.316-317) notes, this concept of reversibility differs from the previous one in at least three respects. First, the only thing that matters for reversibility understood as recovery is a return of the initial thermodynamic state. We needn't specify a history-reversal operation that ensures the composite system pass through its reversed sequence of states. In this

respect, this notion of reversibility is weaker than reversibility understood as time-reversal invariance. And irreversibility understood as irrecoverability is a logically stronger notion than irreversibility understood as time-reversal non-invariance. A second difference is its emphasis on *complete* recovery. A complete recovery involves not only the return of the system to its initial state but also the return of any auxiliary system that it interacted with. Reversibility understood as time-reversal invariance is simply concerned with states of the system, and not with the states of systems it interacted with as well. A third difference concerns the concept of possibility that's implicitly invoked. Here, the intended notion of possibility is usually tied to the means available, to beings like us, with our epistemic and physical limitations, to recover the initial thermodynamic state. Irreversible processes, then, are processes that result in some thermodynamic state from which it is not possible, given certain epistemic and physical limitations, to recover the initial thermodynamic state. This differs from reversibility understood as time-reversal invariance in that in that case one is simply making a claim about what is dynamically possible, irrespective of further limitations.

More commonly, however, the word "reversible" is used to mean what more careful authors mean when they say that certain processes are *quasi-static and reversible.* Thermodynamics can be thought to distinguish between two kinds of processes. First, there are ones that take place gently, with no turbulence or friction. Then there are all other processes. Quasi-static processes are processes of the first kind. They are processes that are carried out so slowly that the system is, at every moment, effectively in equilibrium. Reversible processes in this context—that is, when they are distinguished

from quasi-static processes—are understood as they were in the previous sense: as processes that occur to systems that result in thermodynamic states from which it's possible to recover their initial thermodynamic state along with the initial thermodynamic state of any auxiliary system with which it has interacted with, and without there being any other changes associated therewith.

We can already see from the reversibility concepts outlined above that there are a number of ways to interpret **G2**. Importantly, it should be clear that there are ways of interpreting **G2** that are distinct from ways of interpreting **G3**.

Having highlighted some things that could be meant by **G2** and **G3**, we turn now to the task of identifying where these reversibility concepts make contact with the laws of thermodynamics. We'll also look to see where goals such as **G1** and **G4** make contact with the theory, and with reversibility concepts.

2.1.2   The Minus First Law

The Minus First Law first appeared in name in Harvey Brown and Jos Uffink's 2001 article "The Origins of Time-Asymmetry in Thermodynamics: The Minus First Law". Earlier authors, however, both appreciated its content and considered it, like Brown and Uffink, to be more fundamental than the other laws of thermodynamics.[12] Commonly, however, the law is invoked without being flagged as a law.[13] It states:

---

[12]See, for instance, Uhlenbeck and Ford (1963: p.5), Kestin (1979: p.72), and Lebowitz (1994: p.135).

[13]See, for instance, Pauli (1973: p.1) and Sklar (1995: p.20).

> **Minus First Law:** An isolated system in an arbitrary initial state within a finite fixed volume will spontaneously attain a unique state of equilibrium.

Like the other laws of thermodynamics, the Minus First Law intends to capture a phenomenological fact. As Brown and Uffink (2001: p.528) explain, the Minus First Law can be broken into the following three claims:

MFL1. The *existence* of equilibrium states for isolated systems. The defining property of such states is that once they are attained, they remain thereafter constant in time, unless the external conditions are changed. The claim that such states exist is not trivial—it rules out the possibility of spontaneous fluctuation phenomena.

MFL2. The *uniqueness* of the equilibrium state; i.e. for any initial state of an isolated system in a fixed given volume, there is exactly one state of equilibrium.

MFL3. The *spontaneous* approach to equilibrium from nonequilibrium. A nonequilibrium state will typically come about as the result of a removal of internal constraints, such as the rapid displacement of adiabatic walls separating two bodies.[14]

Spelled out like this, it's clear to see that the Minus First Law provides a *characterisation* of thermodynamic equilibrium and the approach to equilibrium. It also reveals that the Minus First Law is asymmetric under time-reversal. The spontaneous motion towards equilibrium is time-asymmetric

---

[14]No indication of the speed of the approach to the new equilibrium state is given of course: thermodynamics provides no equations of motion.

because of what equilibrium states, as characterised by **MFL1**, are: once attained no spontaneous departure from them is possible without intervention from the environment.

The Minus First Law says that isolated systems will either be in a unique state of equilibrium or else be spontaneously approaching equilibrium. Compare this to **G1**. It merely gives this behaviour a name: thermodynamic behaviour. **G1** then, is probably best interpreted as saying that it's a goal of nonequilibrium statistical mechanics to provide a microphysical justification of the Minus First Law. Of course, by providing a microphysical justification of the Minus First Law, one can be thought to underpin irreversible macroscopic behaviour, in line with **G2**—where irreversible macroscopic behaviour is taken to be exhibited by systems that are governed by macroscopic laws that are time-reversal non-invariant—and one can also be thought to underpin a particular interpretation of **G3**; to show how a time-reversal non-invariant thermodynamics emerges from a time-reversal invariant microphysics.

It's worth noting that neither the Minus First Law, the three claims that constitute it, nor **G1** contain adverbs such as "quickly" or "slowly". In fact, none of them makes any reference to the rates in which systems relax to equilibrium. (In fact, *none* of the goals **G1-4** make any reference to the rates in which systems approach equilibrium.) While it's sensible to omit any reference to relaxation rates when articulating the Minus First Law (since relaxation rates vary significantly from system to system, and because generality is a feature of a good law) it's worth bearing in mind, for what's to come in the next chapter, that we have detailed information about the

rates in which many systems approach equilibrium.

2.1.3 The Zeroth Law

The Zeroth Law, like the Minus First Law of thermodynamics, was invoked for a long time without notice. While some early writers both noted and emphasised its importance, it wasn't formally recognised as a law until 1939, when Ralph Fowler and Edward Guggenheim coined the expression "Zeroth Law".[15]

We say that two objects are in thermal contact with each other if it's possible for heat to flow between them. More will be said about heating in the upcoming discussion of the First Law of thermodynamics. For now, it will suffice to say that it's a form of energy transfer. If we bring two objects, A and B into thermal contact, then one of three things can happen as the composite system, subject to the Minus First Law, equilibrates. First, heat could flow from A to B. Second, heat could flow from B to A. Or third, there could be no flow of heat between A and B. The Zeroth Law of thermodynamics says that the third situation is *transitive*.

> **Zeroth Law:** If A is in thermodynamic equilibrium with B when they are in thermal contact, and B is in thermodynamic equilibrium with C when they are in thermal contact, then A is in thermodynamic equilibrium with C.

This law allows us to define an important equivalence relation on thermo-

---

[15]James Clark Maxwell was an early writer who both acknowledged and emphasised the importance of what would later become the Zeroth Law of thermodynamics. See Maxwell (1871: p.32-33), for example.

dynamic equilibrium states.[16] We say that if A and B are thermodynamic equilibrium states, then they are *equitemperature states* if and only if A can be brought into thermal contact with B without the flow of heat.[17] The equitemperature relation is trivially reflexive and symmetric. That it is transitive is a substantive assumption. The equitemperature relation is needed to introduce the concept of thermodynamic temperature and to establish a numerical temperature scale.

The Zeroth Law and, as we'll see, the First Law of thermodynamics are not directly related to **G1-4**—or to any of the reversibility concepts that were introduced earlier. They are, however, needed to properly understand versions of the Second Law that are connected to **G1-4** and to reversibility concepts. It is for this reason that they have been included in this presentation and discussion of classical thermodynamics.

2.1.4 The First Law

Work is a form of energy transfer. One way to increase the internal energy of a system is to do work on it. For example, when we compress springs we increase their internal energy by transferring energy to them. That is, by doing work on them. They retain this energy in the form of *potential energy*. This energy can be converted into *kinetic energy* by allowing *them* to do work on their environment. The energy that can be recovered from

---

[16]A relation is said to be an *equivalence relation*, if it is reflexive, symmetric, and transitive.

[17]The characterisation of equitemperature is not rich enough to provide us with a numerical scale of temperature. In fact, it's not even rich enough to impose a total ordering on the class of equivalent sets. That is, it cannot be used to make comparative claims between sets of nonequivalent equitemperature states.

the compression of an ideal spring that begins in state s_i and that finishes in state s_f is equal to

$$W = -\int_{s_i}^{s_f} \mathbf{F} \cdot d\mathbf{x}. \qquad (2.6)$$

$\mathbf{F}$ is the force opposing the direction of compression.

The same amount of energy could, however, be used to stir a viscous fluid. Suppose, for example, that we attach a compressed ideal spring to a paddle, immerse it in a viscous fluid, and release it. Interestingly, in this case, we readily recognise that we're not able to completely recover as kinetic energy the energy we put into the system as work. Initially, we might think that this energy has been lost. But then we notice that the system has gotten warmer. This suggests, and, in fact, we say, that the energy that went into the system in the form of work was not lost, but was instead converted into heat. What makes this suggestion better than the first is that there is a measurable mechanical equivalent of heat: we can measure the amount of work that is needed to raise a gram of water by 1 degree, for example.

The First Law of thermodynamics can be thought of as a thermodynamic expression of the principle of conservation of energy. It says that if an amount of work W is done on a system, and heat Q passes into it, the internal energy U of the system is changed by an amount

First Law: $\Delta U = Q + W$.

Sometimes the Law is expressed in differential form.

Differential Form: $\mathrm{d}U = đQ + đW$.

Since U is a state function, we use an *exact differential*, dU, to represent change in it. Q and W are not state functions. They are path dependent. That is why they appear as *inexact differentials* in the differential form of the Law.

As we've seen, thermodynamics uses the terms "heat" and "work" in connection with two modes of energy transfer. We can increase the internal energy of a system either by heating it or by doing work on it. It is assumed, at least in classical presentations of the theory, that these two modes of energy transfer can readily be distinguished from one another.

2.1.5   The Second Law

There are several versions of the Second Law of thermodynamics.[18] Classical thermodynamics standardly identifies three. These are: the Kelvin statement, the Clausius statement, and the Entropy statement. This section will begin by presenting each statement. It will then discuss their relation to one another and what is typically meant by **G4**. That is, what authors mean when they say that it's a goal of nonequilibrium statistical mechanics to provide a microphysical justification of *the* Second Law of thermodynamics.

First, there's Kelvin's statement of the Second Law.

> **Kelvin Statement:** No process is possible whose sole result is
> the complete conversion of heat into work.[19] (Kardar 2007: p.9)

---

[18]See Uffink (2001) for an interesting and detailed discussion of the many versions of the Second Law of thermodynamics.

[19]This is a common, modern presentation of Kelvin's statement of the Second Law. As Uffink (2001: p.327-328) notes, this statement of the Law was inspired by what Kelvin considered to be an *axiom* of thermodynamics:

Then there's Clausius's statement.

> **Clausius Statement:** Heat can never pass from a colder body to a warmer body without some other change, connected therewith, occurring at the same time. (Clausius 1856: p.86)

The first statement rules out a perfect engine. The second rules out a perfect refrigerator.

Unlike Kelvin and Clausius's statements of the Second Law, the Entropy statement cannot be stated and at the same time be so easily understood. It will help to have some additional concepts on the table, before presenting this statement of the Law. We begin with Carnot's theorem.

Carnot's Theorem

A *heat engine* works by taking in a certain amount of heat Q_h, from a heat source (e.g. a coal fire), converting a portion of it into work W, and dumping the remaining heat Q_c into a heat sink (e.g. atmosphere).

A heat engine is said to work in a *cycle* if it returns to its initial thermodynamic state. This is similar to, but distinct from, reversibility understood as recovery. In this case, auxiliary systems are allowed to be in different thermodynamic states when the engine returns to its initial thermodynamic state. A heat engine that operates in a cycle undergoes no net change in internal energy.

It is impossible, by means of inanimate material agency, to derive mechanical effect from any portion of matter by cooling it below the temperature of the coldest of the surrounding objects. (Kelvin 1851: p.265)
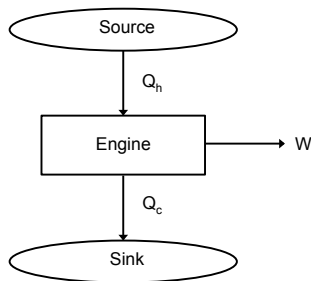
Figure 2.1: A heat engine.

The principle of conservation of energy entails that

$$W = Q_h - Q_c. \tag{2.7}$$

The fraction of heat the engine takes from the heat source and converts into useful work, having dumped the rest, is called the *efficiency* of the engine. It's denoted by η.

$$\eta = \frac{W}{Q_h} = 1 - \frac{Q_h}{Q_c}. \tag{2.8}$$

A *Carnot engine* is any heat engine that operates in a cycle in a quasi-static and reversible manner, and all of its heat exchanges take place either at a source temperature T_h, or at a sink temperature T_c. In these situations, heat sources and heat sinks are treated as being so large that they can supply or absorb quantities of heat without incurring a measurable change in temperature. That is, they are treated as reservoirs. The distinguishing characteristic of the Carnot engine is that all of its heat exchanges occur at only two temperatures.

Carnot's theorem gives us information about the maximum efficiency of heat engines. It has two parts.

> **Carnot's theorem:** 1. Every Carnot engine that operates between a pair of heat reservoirs, with temperatures T_h and T_c, is equally efficient. 2. Moreover, any *other* heat engine operating between these reservoirs is less efficient than a Carnot engine.

When a Carnot engine runs forward, it draws heat from the hot reservoir and converts a portion of it into work.
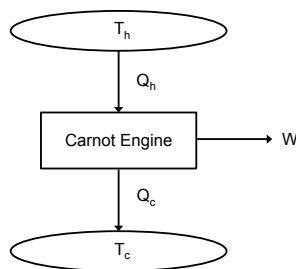
Figure 2.2: A Carnot engine operates between temperatures T_h and T_c, with no other heat exchanges.

Since a Carnot engine is reversible, it can also be run backward and work as a refrigerator. In this case, work is done on the Carnot engine. This is used to move heat from the cold reservoir to the hot reservoir (see Figure 2.3).

Carnot's theorem is proven by deriving a contradiction. Consider a heat engine, operating in a cycle, with efficiency η_{he}, that has been setup to extract an amount of heat Q_h from a hot reservoir, do work $W = \eta_{he}Q_h$ on a Carnot engine operating in reverse, and discard heat $Q_c = (1 - \eta_{he})Q_h$. That is, we
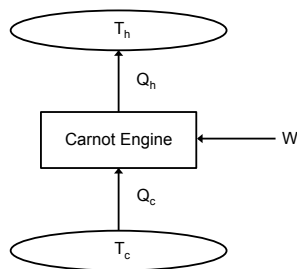
Figure 2.3: A Carnot engine operating in reverse. Heat is transferred from
the cold reservoir to the hot reservoir.

run the engine forwards, in a cycle, and use it to do work on a Carnot engine
that is extracting heat Q'_c from the cold reservoir, and discarding heat Q'_h
into the hot reservoir with efficiency η_{ce}.
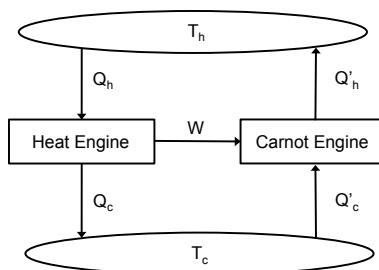
Figure 2.4: A heat engine used to run a Carnot engine in reverse.

The amount of work that is transferred from the heat engine to the Carnot
engine is

$$W = \eta_{he}Q_h = \eta_{ce}Q'_h. \qquad (2.9)$$

This means that the net effect of the process transfers a quantity of heat

$$Q = Q'_h - Q_h = (\frac{\eta_{he}}{\eta_{ce}} - 1)Q_h \qquad (2.10)$$

from the cold reservoir to the hot reservoir. If $\eta_{he} > \eta_{ce}$, then (2.10) is positive. The net result of the process, then, is one in which we have moved heat from the cold reservoir to the hot reservoir. This *contradicts* Clausius's statement of the Second Law.[20] So we conclude that for any heat engine and any Carnot engine, $\eta_{he} \leq \eta_{ce}$. This is the second part of Carnot's theorem. The first part follows as a corollary. Every Carnot engine that operates between a pair of heat reservoirs is equally efficient, since each can be used to run any other one backward.

Thermodynamic Temperature

Carnot's theorem says that every Carnot engine operating between a pair of heat reservoirs has the same (maximum) efficiency. Since their efficiency depends *only* on the temperatures of the reservoirs, they can be used to define a temperature scale. That is, they can be used to define a scale on the class of sets of equitemperature states that was established using the Zeroth Law.

If η_{hc} is the efficiency of a Carnot engine operating between reservoirs H and C, define the *thermodynamic temperature T* by

$$\frac{T_c}{T_h} =_{df} 1 - \eta_{hc}. \qquad (2.11)$$

This defines the thermodynamic temperature of any reservoir up to an

---

[20]Where Clausius's statement of the Second Law is taken as an axiom.

arbitrary scale factor.

The Carnot Cycle

Functioning Carnot engines perform *Carnot cycles*. These cycles are usually divided into four stages. This subsection describes one of these cycles for an engine working between two reservoirs whose working substance is an ideal gas.

The first stage involves *expansion at constant temperature*. This is otherwise know as an *isothermal expansion*. The engine absorbs heat, Q_h, from the hot reservoir at temperature T_h, and does work on the environment. The second stage involves *adiabatic expansion*. That is, the gas continues to expand but now does so without exchanging heat with the environment.[21] It does work on the environment as it cools to temperature T_c. The third stage involves *compression at constant temperature*. This is otherwise known as an *isothermal compression*. At this stage, an external agent does work on the gas. The engine expels heat, Q_c, into the cold reservoir at temperature T_c. The fourth and final stage of the cycle involves *adiabatic compression*. This stage also involves an agent doing work on the gas, raising its temperature to T_h. The compression occurs without heat being exchanged with the environment.

---

[21]A system is *adiabatically isolated* if and only if it can't exchange heat with the environment. An *adiabatic process* (e.g. an adiabatic expansion) is one in which the system exchanges no heat with the environment.
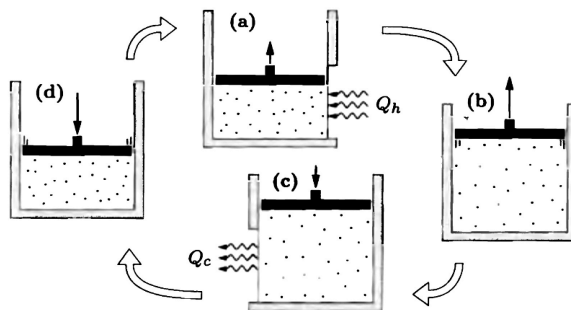
Figure 2.5: The four stages of a Carnot Cycle. (A) isothermal expansion.
(B) adibatic expansion. (C) isothermal compression. (D) adiabatic
compression. (Schroeder 2000: p.126)

Thermodynamic Entropy

The following equality holds for a Carnot cycle:

$$\oint \frac{\dah Q}{T} = \frac{Q_h}{T_h} - \frac{Q_c}{T_c} = 0. \tag{2.12}$$

If we assume that any quasi-static and reversible cycle can be represented
by a collection of Carnot processes, that is, by a path through the system's
thermodynamic state space that alternates between isothermal and adiabatic
segments, then for *any* thermodynamic system

$$\oint_{qsr} \frac{\dah Q}{T} = 0. \tag{2.13}$$

The subscript attached to the path integral signifies that the cycle is
performed in a quasi-static and reversible manner. If (2.13) did not hold, then

we could construct a quasi-static and reversible heat engine that operated in a cycle with an efficiency that was different from the efficiency of a Carnot engine, in violation of Carnot's theorem.

It follows from (2.13) that there exists a state function S such that, for any quasi-static and reversible process, and any two thermodynamic states a and b,

$$\int_a^b \frac{đQ}{T} = S_b - S_a. \tag{2.14}$$

This state function is called the *thermodynamic entropy* of the system. It can also be written in differential form.

$$dS = \left(\frac{đQ}{T}\right)_{qsr} \tag{2.15}$$

For any heat engine operating in a cycle that is less efficient than a Carnot engine, it must be the case that

$$\oint \frac{đQ}{T} < 0. \tag{2.16}$$

Equations (2.13) and (2.16), when written together, state that for any cycle,

$$\oint \frac{đQ}{T} \leq 0. \tag{2.17}$$

This is known as Clausius's theorem.[22] It's often expressed in differential form.

---

[22]This is also sometimes referred to as the "Inequality of Clausius".

$$\bar{d}Q \leq TdS. \tag{2.18}$$

The Entropy statement of the Second Law follows from this result.

> **Entropy Statement:** The thermodynamic entropy of an adiabatically isolated system cannot decrease. That is, $dS \geq 0$.

Relations Between Statements of the Second Law

Three statements of the Second Law have been presented: the Kelvin statement, the Clausius statement, and the Entropy statement. One may wonder then what could be meant by **G4**. That is, one may wonder what is meant by those who say that it is a goal of nonequilibrium statistical mechanics to provide a microphysical justification of *the* Second Law of thermodynamics.

Such a claim is likely to have been informed by two thoughts. First, most authors are aware of the *equivalence* of these statements. Their equivalence will be highlighted shortly. Second, it's reasonable to think that many authors have the Entropy statement in mind when they state this goal. The reason for this has to do with its content. It, unlike the other statements, formally identifies a thermodynamic quantity that tracks irreversible macroscopic behaviour: thermodynamic entropy. Since (as we'll see) philosophers typically have claims such as **G1-3** in mind whenever they claim something like **G4**, it seems reasonable to interpret them as saying that it's a goal of nonequilibrium statistical mechanics to provide a microphysical justification of why the thermodynamic entropy of adiabatically isolated systems cannot decrease.

Interestingly, one might see the satisfaction of this goal as a way of also satisfying **G2** and **G3**—where these claims are suitably interpreted. That is, where we associate increases in thermodynamic entropy with irreversible macroscopic behaviour (**G2**), and where we take the Entropy statement of the Second Law as being a (or perhaps *the*) time-reversal non-invariant aspect of thermodynamics that is referred to in **G3**.

Returning to the equivalence of statements of the Second Law, we find that the Kelvin and Clausius statements are equivalent if we insist that all temperatures have the same sign.[23] Their equivalence is typically displayed by highlighting that a violation of one of these statements leads to a violation of the other, and vice versa.

Consider an engine $\overline{C}$ that violates the Clausius statement of the Second Law by transferring Q heat from a colder reservoir to a hotter one. Now consider a heat engine that operates between these reservoirs that transfers heat Q_h from the hotter reservoir and dumps Q_c into the colder one. The *combined system* takes $Q_h - Q$ from the hot reservoir, produces work equal to $Q_h - Q_c$, and dumps $Q_c - Q$ into the colder reservoir. If we adjust the output of the heat engine such that $Q_c = Q$, the net result is a perfectly efficient engine $\overline{K}$, in violation of Kelvin's statement. See Figure 2.6.

Now consider an engine $\overline{K}$ that absorbs Q heat from the hot reservoir

---

[23]As Uffink (2001: p.329) notes, Tatyana Ehrenfest-Afanassjewa (1925, 2002) noticed that the two formulations only become equivalent when we add an extra axiom to thermodynamics, namely that all temperatures have the same sign. If we allow systems with negative absolute temperature, which is not forbidden by the standard laws of thermodynamics, then one can distinguish between these two formulations. As Uffink (2001: p.329) continues, this observation became less academic when Norman Ramsey (1956) gave concrete examples of physical systems with negative absolute temperatures (e.g. nuclear spin systems.
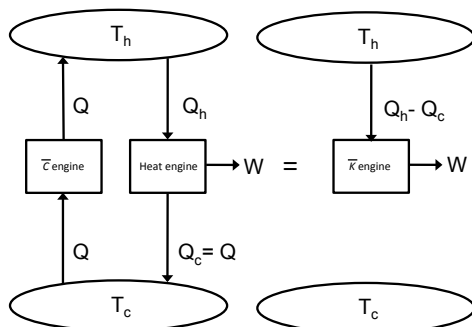
Figure 2.6: An engine violating the Clausius statement $\overline{C}$ can be connected to a heat engine, resulting in a combined device $\overline{K}$ that violates the Kelvin statement.

and converts it entirely into work. This engine violates Kelvin's statement. The work output by this engine can be used to run a heat engine backwards (i.e. it can work as a refrigerator), with the net outcome of transferring heat from a colder body to a hotter body, in violation of Clausius's statement.
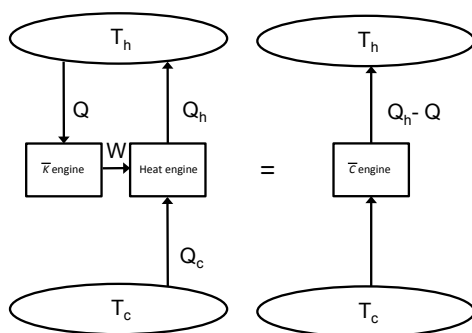
Figure 2.7: An engine violating the Kelvin statement $\overline{K}$ can be connected to a refrigerator, resulting in a violation of the Clausius statement.

Since Kelvin and Clausius's statements are equivalent, it remains to show

that they are equivalent to the Entropy statement. But since Kelvin and Clausius's statements are equivalent, we only need to show that one of them is equivalent to the Entropy statement. And since earlier subsections effectively detailed how the Entropy statement followed from the Clausius statement, all that remains is for us to show that the Clausius statement follows from the Entropy statement.

Consider a violation of the Entropy statement. That is, consider a heat engine operating in a cycle between two reservoirs with different temperatures whose thermodynamic entropy decreases, i.e. $dS < 0$. Then, for such a process,

$$\mathrm{d}Q > T dS. \tag{2.19}$$

So

$$\oint \frac{\mathrm{d}Q}{T} > 0. \tag{2.20}$$

This means that we have heat passing from a colder body to a warmer body without some other change, connected therewith, occurring at the same time. That is, we have a violation of Clausius's statement.

2.2   Summary of Results

Through a discussion of various reversibility concepts and the laws of thermodynamics, this chapter highlighted ways in which the goals **G1-4** are distinct but related. This was done so as to make the issue of reconciling the exis-

tence of thermodynamically irreversible processes with underlying reversible dynamics more clear, and to highlight what philosophers are attempting to underpin using nonequilibrium statistical mechanics. It was also done so that we might, after the coming section and next chapter, better appreciate the relationship between thermodynamics and statistical mechanics. This discussion also contributed to the task of explaining what philosophers mean when they say than an aim of nonequilibrium statistical mechanics is to account for certain aspects of thermodynamics.

To these ends, an earlier section suggested that **G1** is probably best interpreted as saying that a goal of nonequilibrium statistical mechanics is to provide a microphysical account of the Minus First Law. Earlier sections also discussed other ways of interpreting this goal and highlighted its connections to **G2-4**. It was noted that if one were to provide a microphysical account of the Minus First Law, one could also be thought to fulfil the goal of accounting for irreversible macroscopic behaviour, in line with **G2**—where irreversible macroscopic behaviour is taken to be exhibited by systems that are governed by time-reversal non-invariant macroscopic laws—and to satisfy the goal expressed by a particular interpretation of **G3**, where the Minus First Law is thought to be at least one aspect of the theory that is responsible for its time-reversal non-invariance.

Earlier sections highlighted that **G2** could be interpreted in various ways, depending on how we interpret "irreversible macroscopic behaviour". It could be understood as macroscopic behaviour that is governed by some time-reversal non-invariant law. **G2** could also be understood as saying that a goal of nonequilibrium statistical mechanics is to account for why certain

processes render particular initial thermodynamic states irrecoverable. **G2** could also be interpreted as a request for a microphysical account of the Minus First Law, but it could also be interpreted as a request for a microphysical account of the Second Law of thermodynamics.

How we interpret **G3** depends on what aspects of the theory are thought to be captured by the expression "time-reversal non-invariant thermodynamics". Again, this could be seen as saying that a goal of nonequilibrium statistical mechanics is to provide a microphysical account of the Minus First Law, but it could also be interpreted as saying that a goal of nonequilibrium statistical mechanics is to provide a microphysical account of the Second Law of thermodynamics. It could also be understood as saying that it is a goal of nonequilibrium statistical mechanics to provide a microphysical account of both laws.

Earlier sections suggested that **G4** is probably best interpreted as saying that it is a goal of nonequilibrium statistical mechanics to provide a microphysical account of the Entropy statement of the Second Law of thermodynamics. This is despite its equivalence to Kelvin and Clausius's statements of the law. The reason is that, unlike Kelvin and Clausius's statements, the Entropy statement formally identifies a thermodynamic quantity that tracks irreversible macroscopic behaviour: thermodynamic entropy. It was also indicated that one might see the satisfaction of this goal as a way of satisfying **G2** and **G3**—where these goals are suitably interpreted. That is, respectively, where we associate increases in thermodynamic entropy with irreversible macroscopic behaviour, and where we take the Entropy statement of the Second Law as being a (or perhaps *the*) time-reversal non-invariant

aspect of thermodynamics that is requested by **G3**.

2.3 Not Strict Interpretations, Suitable Analogues

Having highlighted that **G1-4** are distinct, and that each goal could be made more precise—often in a number of ways—one may wonder whether this points to either a confusion in the literature or a disagreement about what is central to reconciling the existence of thermodynamically irreversible processes with underlying reversible dynamics. One might also think that the appropriate way for the discussion to proceed is for contributors to first carefully specify what it is they're attempting to use nonequilibrium statistical mechanics to accomplish.

This chapter will now end by elaborating a little bit on the relationship between thermodynamics and statistical mechanics, and by explaining why there is neither confusion nor disagreement in the literature. It will also explain why well informed authors needn't worry about, and probably don't worry about, eliminating the differences between expressions of goals such as **G1-4**.

Those attempting to provide a foundation for thermodynamics recognise that macroscopic systems can be described in various ways, by various physical theories. A dilute gas and its behaviour, for example, can be described in thermodynamic terms. It can also be described at a microscopic level, and its behaviour, we think, can be modelled by describing the behaviour of the entities that constitute it. At the microscopic level, a mechanical theory like

statistical mechanics is used to describe the system's behaviour. Since we think both descriptions are equally valid, the two theories' predictions had better be consistent when applied to the same phenomena. Since statistical mechanics is thought to be the more fundamental theory, the concern is often transformed into whether we can recover thermodynamics from statistical mechanics. Philosophers often talk about reduction in these contexts. The worry is otherwise stated in terms of whether one theory *reduces* to another. This is a way of looking at what those engaged in the foundational discussion are discussing. While reduction is a philosophically hairy notion, the main idea shared by many theories of reduction, and implicitly (but sometimes explicitly) endorsed by those contributing to foundational discussions of thermodynamics, is that one theory reduces to another if we can use the reducing theory to construct an *analogue* of the laws and concepts of the theory to be reduced. What philosophically well-minded contributors do not expect is a logical deduction from statistical mechanics (the reducing theory) to thermodynamics (the theory to be reduced). They also do not expect the laws of the reduced theory to be laws of the reducing theory, nor do they expect the concepts used by the former to always be applicable at the level of the latter. The general expectation is instead that the laws and concepts of thermodynamics appear in statistical mechanics as complicated, approximate statements that are true under certain conditions.[24] To expect anything more would be, in the words of Craig Callender (2001), to take thermodynamics too seriously.

---

[24]Many authors hold this kind of view. See, for example, Frigg and Werndl (2012a: 919-920), Frigg and Werndl (2012b: p.100), Lavis (2005: p.255), Callender (2001), and Callender (1999).

With these considerations in mind, it's reasonable to read goals such as **G1-4** not as ones that are confused, or as ones that are taking a firm stand on what exactly is needed to reconcile the tension between thermodynamics and statistical mechanics. Instead, it's reasonable to read them as short approximate statements that refer to a collection of relevant ideas that are central to reconciling the existence of thermodynamically irreversible processes with underlying reversible dynamics. A collection of relevant ideas, that is, such as those discussed in previous sections, that come from various ways of unpacking goals that resemble **G1-4**. What contributors to the foundational discussion presumably want from nonequilibrium statistical mechanics are, among other things, statistical mechanical stories that, at a macroscopic level, do justice to the aspects of thermodynamics touched on by claims such as **G1-4**.

These considerations also motivate why well informed contributors needn't bother eliminating the differences between these claims. If they are understood as referring to a collection of important and relevant ideas, and if an appropriate reduction of thermodynamic concepts to statistical mechanical ones involves locating suitable analogues, then it's unnecessary to homogenise them.

Chapter 3

An Evaluation of Foundational Accounts

The previous chapter discussed one of the aims of statistical mechanics: to reconcile the existence of thermodynamically irreversible processes with underlying reversible dynamics. It focused on one side of the issue, where it makes contact with thermodynamics. This chapter takes up another side, where it makes contact with statistical mechanics.

One of the aims of this chapter is to outline and review some of the leading attempts to reconcile this conflict. These attempts are typically part and parcel of works that aim at accounting for why isolated systems that begin away from equilibrium spontaneously approach equilibrium and why they remain in equilibrium for incredibly long periods of time. In addressing this goal, these works hope to provide a foundation for aspects of thermodynamics and satisfy goals such as **G1-4**. This chapter considers the most pressing problems facing these accounts and, where possible, presents and discusses

their most promising solutions. It highlights that they have some severe limitations, even when we overlook their problems. It reveals that they do not underpin a large and important set of facts that concern the behaviour of systems that begin away from equilibrium. While they may satisfy goals such as **G1-4** and account for why isolated systems that begin away from equilibrium spontaneously approach equilibrium, and why they remain in equilibrium for incredibly long periods of time, they do not underpin facts about the rates in which systems approach equilibrium, or facts about the kinds of states they pass through on their way to equilibrium, or facts about fluctuation phenomena. They also do not help us form expectations about these things or help justify the expectations we may already have about them, having formed them on the basis of experience.

This chapter claims that the limitations of these accounts are a symptom of what they are aiming at. By focusing on recovering certain aspects of thermodynamics, those contributing to the foundational project have merely been in the business of articulating microphysical accounts of irreversible macroscopic behaviour that recover a few qualitative facts. To remedy this situation, this chapter encourages pursuing a different line of research: understanding and accounting for the success of the techniques and equations physicists use to model the behaviour of systems that begin away from equilibrium. If we're able to properly understand these techniques and equations, and are able to account for their success, then we will not only be able to underpin the kinds of qualitative facts the literature has focused on, but we'll also be able to underpin the important quantitative facts that leading typicality accounts cannot.

The following chapter takes steps in this promising direction. It outlines, discusses, and attempts to ground the success of a way of modelling the approach to equilibrium and subsequent behaviour of a Brownian particle that's been introduced into an isolated homogeneous fluid at equilibrium. The next chapter also highlights how the model contributes to the satisfaction of goals such as **G1-4**, and to the underpinning of important quantitative facts.

This chapter is composed of five sections. The first section outlines and discusses the essence of what many people think of as the first serious attempt to account for the behaviour of systems that begin away from equilibrium. This approach is based on some of Ludwig Boltzmann's early work on statistical mechanics, when he first offers what has come to be known as Boltzmann's equation and H-theorem. As we'll see, this work gives rise to one of the very issues that is at the centre of contemporary discussions of statistical mechanics: the reconciliation of the existence of thermodynamically irreversible processes with underlying reversible dynamics. The second section outlines and discusses the leading ways—found in the contemporary foundational literature—to account for the behaviour of systems that begin away from equilibrium. These approaches are inspired by, and build on, Boltzmann's later attempts (circa 1877) to resolve this issue and, once again, account for the behaviour of systems that begin away from equilibrium.[1] Each one incorporates, in some way or other, a notion of typicality.

---

[1]This presentation and discussion of Boltzmann's work is couched in the language of modern statistical mechanics. It also makes use of its technical resources. While this presentation does not do justice to the messy and tangled history of statistical mechanics, or to Boltzmann's original works, it is in line with how these ideas are now standardly presented. It also happens to be the cleanest way to present the ideas most relevant to

The second section also discusses worries specific to each typicality account. More general concerns are discussed in the third section. Both sections also discuss, where possible, the most promising solutions to these concerns. The fourth section highlights and discusses the limitations of typicality accounts. It also promotes pursuing lines of research that aim at more than merely recovering a few qualitative facts and satisfying goals such as **G1-4**. The suggestion is that we make sense of, and try to ground the success of, the techniques and equations commonly used to model the behaviour of systems that begin away from equilibrium. Understanding and accounting for their success will provide us with a microphysical underpinning of the aspects of the behaviour of systems that begin away from equilibrium not accounted for by typicality accounts. It will also provide us with the means to satisfy goals such as **G1-4**. The chapter ends, in the fifth section, with a brief summary.

3.1   Boltzmann's Early Work

Ludwig Boltzmann famously attempted to explain why isolated systems that begin in some nonequilibrium state spontaneously approach equilibrium and why they remain in equilibrium for incredibly long periods of time. In 1872, Boltzmann considered how the distribution of velocities of the molecules of a dilute gas could be expected to change under collisions and argued that there

---

the aims of this chapter and thesis. Readers interested in the gritty details are encouraged to consult Klein (1973), Brush (1986), Sklar (1995), von Plato (1994), Cercignani (1998), Ehrenfest and Ehrenfest (2002), Uffink (2007), Brown, Myrvold, and Uffink (2009), and Uffink (2014) for discussions of Boltzmann's work on statistical mechanics and its tangled history. For Boltzmann's original works, see Boltzmann (1909). And see de Regt (1996), Blackmore (1999), and Visser (1999) for discussions of the philosophical and methodological shifts in Boltzmann's thinking.

was a unique distribution—now called the Maxwell-Boltzmann distribution—that was stable under collisions.[2]

> **Maxwell-Boltzmann Distribution:** the fraction of molecules having velocity lying within the velocity-space element $\mathrm{d}^3 v$ about **v** is proportional to

$$e^{-mv^2/2kT}\ \mathrm{d}^3 v, \qquad (3.1)$$

> where k is the Boltzmann constant and T is temperature.

Boltzmann further argued that a gas that initially had a different distribution would move toward the Maxwell-Boltzmann distribution. To argue that this was the unique (equilibrium) distribution, which would be approached starting from some other distribution, Boltzmann defined a quantity, which we now call H, showed that it reached a minimum value for the Maxwell-Boltzmann distribution, and argued that it would *monotonically* decrease to its minimum.[3] This result is now known as Boltzmann's H-theorem. It's a straightforward consequence of Boltzmann's transport equation. Importantly, the result is asymmetric under time-reversal.

The next subsection fleshes this picture out by presenting and discussing aspects of the theorem. The subsection after that highlights its logic and limitations using a simple toy model that shares salient features with dilute gases.[4]

---

[2]See Boltzmann (1872).

[3]The quantity we call H was originally denoted E in Boltzmann's early work. See Boltzmann (1872).

[4]The interested reader should consult Uffink (2007) and, especially, Brown et al. (2009) for the details of Boltzmann's H-theorem, and for an in depth discussion of it. For a good

3.1.1 Boltzmann's H-theorem

Boltzmann's H-theorem was based on a model of a gas consisting of n qualitatively identical, hard, spherical molecules, in a container with perfectly elastic walls. He supposed the gas was sufficiently dilute so that only binary collisions needed to be taken into account in the dynamics.

The distribution (or density) function $f(\mathbf{x}, \mathbf{v}, t)$ is defined such that $f(\mathbf{x}, \mathbf{v}, t)\mathrm{d}^3x\mathrm{d}^3v$ is the number of molecules within a volume element d^3x about $\mathbf{x}$ and with velocity lying within the velocity-space element d^3v about $\mathbf{v}$. Boltzmann's transport equation determines how $f(\mathbf{x}, \mathbf{v}, t)$ evolves in time.

Brown et al. (2009: p.175) explain that Boltzmann's first derivation of the transport equation made use of three noteworthy assumptions. Two of these were harmless, the third became controversial. First, Boltzmann assumed that for the initial state of the gas, the momentum distribution is isotropic. That is, $f_0(\mathbf{x}, \mathbf{v}, t) = f_0(\mathbf{x}, v, t)$. Second, Boltzmann assumed that no external forces act on the gas and that the distribution is independent of position at all times. That is, $f(\mathbf{x}, \mathbf{v}, t) = f(\mathbf{v}, t)$. It follows from these considerations that the expression for $\partial f/\partial t$ depends only on collisions, so that the transport equation can be expressed as a balance equation, in which losses are subtracted from gains during collisions. The third and most famous assumption concerned collision numbers. We now refer to this assumption, following Paul and Tatiana Ehrenfest, as the *Stoßzahlansatz*.[5] Here's how Brown et al. (2009: p.175) describe it:

_____

heuristic derivation of Boltzmann's H-theorem see Dorfman (1999).

[5]See Uffink (2007) and Brown et al. (2009) for more on the *Stoßzahlansatz*.

Suppose we have two groups of molecules heading towards each other. Choose one of the groups and call it the *target* group. Consider the collection S of all the spatial volumes (cylinders) swept out by the target molecules in time δt such that if any molecule of the second group is found in S there is bound to be a collision. Note that S depends on the relative velocities of the molecules in the two groups; it is not defined solely in relation to the target molecules! Then suppose that the number of collisions that actually occur is the total volume associated with S multiplied by the density of molecules of the second, *colliding* group. This holds *only if the spatial density of the colliding molecules in S is the same as in any other part of space.*

Since the target region S is defined in relation to the relative velocities of the molecules from the two groups, and since, from the second assumption, the distribution function is position independent, the assumption depends on the relationship between the momenta of the target and colliding molecules before collision. In fact, the *Stoßzahlansatz* is tantamount to the claim that the density $F(v_1, v_2, t)$ of pairs of molecules which are about to collide within the interval $[t, t + \Delta t]$ with velocities v_1 and v_2 is given by the product of the single molecule densities. That is,

$$F(v_1, v_2, t) = f(v_1, t)f(v_2, t). \tag{3.2}$$

Thus the molecules entering into but not out of collision are uncorrelated in velocity.

In the same work, Boltzmann also introduced the H-functional

$$H[f_t] = \int f(\mathbf{x}, \mathbf{v}, t) \ln f(\mathbf{x}, \mathbf{v}, t) \mathrm{d}^3 x \mathrm{d}^3 v \qquad (3.3)$$

which is large for narrow distributions and small for wide ones. Boltzmann treated the density function as continuous and differentiable, since the number of molecules in a gas is very large.

If we assume that the *Stoßzahlansatz* holds at time t_0, then it follows from Boltzmann's transport equation that

$$\frac{dH}{dT} \leq 0. \qquad (3.4)$$

Equality holds when the gas reaches equilibrium, coinciding with the Maxwell-Boltzmann distribution. This is Boltzmann's H-theorem. One of its assumptions is the validity of the *Stoßzahlansatz at all times*. This ensures the *monotonic* behaviour of H over time. It also ensures that it's a temporally asymmetric result. The theorem can be thought to provide a microphysical underpinning of the Minus First Law. It can also be thought to provide an underpinning of the Entropy Statement of the Second Law of thermodynamics, if we associate thermodynamic entropy with $-H$. It can also be thought to satisfy **G2**.

3.1.2 The Kac Ring

This subsection highlights the theorem's logic and limitations by discussing a simple, explicitly solvable toy model called the Kac ring. The Kac ring first appeared in a series of lectures given by Mark Kac at the University of

Colorado in 1959.[6] It's similar, in some important respects, to a dilute gas.

In the model, N *sites* are arranged around a circle, forming a one-dimensional *periodic lattice*. Sites are joined to their neighbours by an *edge*, and $0 < n < N$ of the edges carry a marker. Each site is occupied by either a black ball or a white ball.[7] The balls and markers are similar to the molecules that comprise a gas.
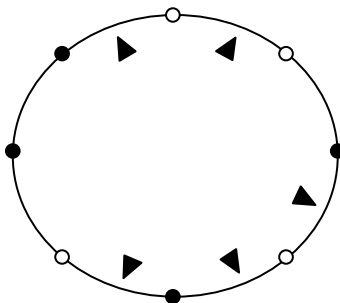
Figure 3.1: A Kac ring with $N = 8$ lattice sites and $n = 5$ markers.

The system evolves on a discrete set of ticks $t \in \mathbb{Z}$ from state t to state $t + 1$ in the following way: each ball moves in a clockwise direction to its

---

[6]The purpose of these lectures was to furnish an introduction to probability theory and its applications to an audience that had little knowledge of these subjects. In a lecture on classical statistical mechanics, Kac (1959: p.99) used the ring model to introduce his audience to the statistical mechanical treatment of irreversible phenomena. The Kac ring is often used to introduce this idea. It's also used to introduce some related ideas. See, for example, Bricmont (1995: Appendix 1), Bricmont (2001), Dorfman (1999: Sec.2.3), Kac (1959: Ch.3 Sec.14-15), Gottwald and Oliver (2009), Schulman (1997: Sec.2.1), and Thompson (1972: Sec.1.9).

[7]Similar descriptions of the model can be found in Bricmont (1995: Appendix 1), Bricmont (2001: p.10), Dorfman (1999: Sec.2.3), Kac (1959: p.99), and Gottwald and Oliver (2009: Sec.3).

nearest neighbour. When a ball passes a marker its colour changes. This is analogous to changes in the velocities of molecules of a gas as they collide with one another.

The Kac ring has a number of interesting features. Analogous to Boltzmann's dilute gas, its microdynamics are symmetric under time-reversal. Any reversed sequence of states is compatible with the dynamics of the system. The system is strictly periodic, and so displays recurrence. The ring, after a series of ticks, returns to its initial state. After N ticks, each ball has reached its initial site and changed colour n times. If n is even, the initial state recurs. If n is odd, it takes at most $2N$ ticks for the initial state to recur. Recurrence is a property also shared by contained gases.

Let $B(t)$ denote the total number of black balls and $b(t)$ the number of black balls that pass a marker on the next tick. Similarly, let $W(t)$ denote the number of white balls and $w(t)$ the number of white balls that pass a marker on the next tick. It follows that

$$B(t+1) = B(t) + w(t) - b(t) \qquad (3.5)$$

and

$$W(t+1) = W(t) + b(t) - w(t). \qquad (3.6)$$

We can study the difference between the number of black and white balls at various times.

$$\Delta(t) = W(t) - B(t) \qquad (3.7)$$

and

$$\Delta(t+1) = W(t+1) - B(t+1) = \Delta(t) + 2b(t) - 2w(t). \qquad (3.8)$$

The system is said to be in equilibrium when $W(t) \approx B(t)$. W, B, and Δ are *macroscopic* quantities. They describe global features of the system's state; many different microstates give rise to the same macroscopic quantities. In contrast, w and b give local information about individual sites. They cannot be computed without knowing the location of each marker and the colour of the ball at each site. Importantly, the *evolution* of W, B, and Δ cannot be determined using only macroscopic state information.

This limitation, however, can be overcome if we make the following non-dynamical assumption: suppose that the fraction of white or black balls that change colour *at each tick* is equal to the probability μ that an edge has a marker on it, where μ is equal to the number of markers, n, divided by the number of edges, N. That is,

$$\mu = \frac{n}{N} = \frac{w(t)}{W(t)} = \frac{b(t)}{B(t)}. \qquad (3.9)$$

This assumption is analogous to the assumption that the *Stoßzahlansatz* holds *at all times*. We too have posited a continued independence. Here, the colour of each ball is taken, *at each tick*, to be probabilistically independent of whether there is a marker in front of it. It's important to note that we have also introduced a time-reversal non-invariant element into the model. There are sequences of states of the system that are compatible with the assumption whose temporal reverse is not. For example, consider a ring that

has a white ball at each site and whose markers at some initial time have been randomly distributed. Now let the system evolve for one tick. All and only the balls that have changed colour have passed a marker. The assumption, which holds for this sequence of states, does not hold for its time reverse. In the later case, ball colours are not independent of marker locations. Black balls are found at all and only those sites that have a marker in front of them.

Importantly, this collision assumption enables us to express (3.8) as

$$\Delta(t+1) = \Delta(t) + 2\mu B(t) - 2\mu W(t) = (1 - 2\mu)\Delta(t). \qquad (3.10)$$

This yields

$$\Delta(t) = (1 - 2\mu)^t \Delta(0), \qquad (3.11)$$

which is analogous to Boltzmann's equation. Since $0 < \mu < 1$, (3.11) tells us to expect $|\Delta(t)| \to 0$ as $t \to \infty$. This is the analogue of Boltzmann's H-theorem. This result, like Boltzmann's H-theorem, is asymmetric under time-reversal. Moreover, this result, like Boltzmann's, is inconsistent with the system's dynamics. More will be said about this in a moment.

Returning to the H-theorem, we may ask how Boltzmann arrived at his result having only assumed an underlying dynamics that is symmetric under time-reversal. The answer is that he didn't, and two famous objections have shown that he couldn't. These are known as the reversibility objection and the recurrence objection. The former is usually credited to Josef Loschmidt and the latter to Ernst Zermelo.[8]

---

[8]See Uffink (2007) and Brown et al. (2009) for more on these objections.

The reversibility objection notes that for any set of trajectories of the molecules of a gas, the time-reversed trajectories are also possible. This is a straightforward consequence of assuming that the microdynamics is symmetric under time-reversal. It shows that not *all* microstates of the gas at *any* time lead to a *monotonic* decrease of H. The recurrence objection applies to classical systems with bounded phase spaces. That is, to systems, such as Boltzmann's dilute gas, with total fixed energy. If we consider a small open neighbourhood of the initial state, and ask, will the state of the system, after it leaves that neighbourhood, ever return to it? Then the answer, which makes use of Henri Poincarè's recurrence theorem, is *yes*, it will, for almost all initial phase space-points, i.e. for all except a set of Lebesgue measure zero.

The same objections apply to the Kac ring. Its underlying dynamics are symmetric under time-reversal, so not *all* microstates of the system compatible with its macroscopic properties at *any* time lead to a *monotonic* decrease of $|\Delta(t)|$. This is an instance of the reversibility objection. Moreover, because the system is strictly periodic, no initial microstate will yield a *monotonic* decrease of $|\Delta(t)|$. This is an instance of the recurrence objection.

To derive Boltzmann's original, asymmetric, result, one needs more than what is given by simply applying Newton's laws of motion to molecular collisions. For Boltzmann, it was the assumption that the *Stoßzahlansatz* holds *at all times*. This is analogous to assuming that (3.9) holds *at each tick* when deriving (3.11), a monotonic and temporally asymmetric result.

In the wake of concerns about Boltzmann's H-theorem, many began to wonder how irreversible thermodynamic behaviour could emerge from an

underlying reversible dynamics. That is, one of the very issues that is at the centre of philosophical discussions of statistical mechanics emerged from worries about Boltzmann's original attempt to account for the behaviour of systems that begin away from equilibrium. These concerns have also influenced many to regard the Second Law of thermodynamics as an approximate statement that holds with high probability. These concerns have also influenced many to endorse a weakened version of it, and to seek justification for a revised statistical version of Boltzmann's H-theorem, along the lines that, for macroscopic systems, H will *probably* decrease to a minimum, and stay there for long periods of time. Most physicists now endorse something like the following statement.

> **Revised Second Law:** Although fluctuations will occasionally result in heat passing spontaneously from a colder body to a warmer body, these fluctuations are inherently unpredictable and it is impossible for there to be a process that *consistently* and *reliably* harnesses these fluctuations to do work.

3.2 Typicality Accounts

In the years that have followed Boltzmann's derivation of the H-theorem, a number of authors have attempted to reconcile the existence of thermodynamically irreversible behaviour with an underlying reversible dynamics. They have offered accounts that aim at satisfying goals such as **G1-4**. More directly, however, these accounts aim at explaining why isolated systems spontaneously approach equilibrium and why they remain in equilibrium

for incredibly long periods of time. An interesting and important class of these accounts have been inspired by, and build on, Boltzmann's later attempts (circa 1877) to justify the H-theorem. Each of these accounts takes the lessons learnt from the reversibility and recurrence objections seriously. They attempt to show that irreversible macroscopic behaviour is, in some sense, typical, and they often try to recover a revised version of the Second Law of thermodynamics. Intuitively speaking, something is typical if it happens in the "vast majority" of cases. Typical lottery tickets, for example, lose. Typicality accounts usually focus on accounting for the behaviour of Boltzmann-like gases. More often than not, though, they are thought to be applicable to a wide variety of systems.

Standardly, authors claim that Boltzmann (1877) inspired the modern typicality approach. Joel Lebowitz (1993a, b) and Shelly Goldstein (2001) are often credited as having ushered in modern forms of the view.[9] Typicality views have become increasingly popular, since the appearance of their work.

This section outlines several of the most promising and influential typicality accounts.[10,11] It also presents and discusses a number of their most

---

[9]See also, Lebowitz (1999) and Goldstein and Lebowitz (2004). Jean Bricmont (1995) and Roger Penrose (1989) are sometimes, but much less often, also credited with establishing modern forms of the view.

[10]To be fair, some of the accounts presented in this section do not trade under the name "typicality account". At least not by those endorsing them. Both the ergodic view and the epsilon ergodic view fall under this category. Roman Frigg and Charlotte Werndl, for example, have appealed to epsilon ergodicity to explain why certain systems spontaneously approach equilibrium and why remain in equilibrium for long periods of time. (See Frigg and Werndl 2011, and Frigg and Werndl 2012b.) Interestingly, they both explicitly distance themselves from typicality when they articulate their view. As this section highlights, however, both (what will be called) the ergodic view and the epsilon ergodic view can straightforwardly be read through a typicality lens. Since these views are subject to the same limitations standard typicality accounts face, it is both fitting and convenient to present them as a kind of typicality account.

[11]A less promising and less influential typicality account that is not discussed is con-

worrying concerns. It also discusses their most promising solutions. The next section continues this discussion by focusing on more general, but no less worrying, concerns with typicality accounts. It also discusses their most promising solutions. A later section highlights the limitations of typicality accounts. It notes, in particular, that even if we overlook the most pressing concerns with typicality accounts, that these accounts will not have achieved very much. Their shortfalls will be used to motivate a different line of research: to encourage philosophers to consider, understand, and account for the success of the techniques and equations physicists use to model the behaviour of systems that begin away from equilibrium.

3.2.1 Some Background

Consider, once again, Boltzmann's dilute gas. It's an isolated system, consisting of n classical particles, each with three degrees of freedom, that are confined to a container with perfectly elastic walls. The system has a fixed total energy, E. The microstate of the system is specified by a point x, in $6n$-dimensional phase space, Γ. Γ is endowed with Lebesgue measure μ_L. The system is governed by Hamilton's equations of motion, which define a measure preserving flow ϕ_t on Γ. This means that for every measurable set A, and every time t, $\mu(\phi_t^{-1}(A)) = \mu(A)$. The system's initial microstate at

---

nected to the view that systems are what Frigg (2009: p.84) calls "globally entropy increasing". For more on this view see Frigg (2009: pp.84-87), and the references therein. Other less promising and less influential typicality accounts that are not discussed in this thesis, or rather, accounts that could straightforwardly be offered within the typicality framework but that will not be discussed in this thesis, include the view that relevant systems are *mixing* and the view that relevant systems have an *ergodic decomposition*. For work related to these views, see Frigg and Werndl (2012b: Sec.4.3). See also Sklar (1995: pp.53-59), Bricmont (2001: p.16), and Lavis (2008: Sec.2).

time t_0, $x(t_0)$, evolves into $x(t) = \phi_t x(t_0)$ at time t. For isolated Hamiltonian systems, such as Boltzmann's gas, energy is conserved. So the motion of the system is confined to its 6n-1-dimensional energy hypersurface Γ_E. The measure μ_L can be restricted to Γ_E, inducing an invariant measure μ on Γ_E.

Thermodynamic systems, as noted in the previous chapter, are characterised by a set of macrovariables: volume, temperature, pressure, etc.[12] Macrostates of the system correspond with the macrovariables obtaining certain values, or with some range of values, if we take into consideration our means of distinguishing between them.[13] Each macrostate M_i, $i = 1, 2, \ldots, m$ (where m is finite) corresponds to a set of macro-regions Γ_{M_i} that consist of all $x \in \Gamma$ that take the macrovariable values characteristic of M_i. Together, the Γ_{M_i} form a partition of Γ_E. That is, they do not overlap but jointly cover Γ_E. Symbolically, $\Gamma_{M_i} \cap \Gamma_{M_j} = \emptyset$ for all $i \neq j$ and $i, j = 1, \ldots, m$, and $\Gamma_{M_i} \cup \ldots \cup \Gamma_{M_m} = \Gamma_E$.[14]

M_p is the initial macrostate of the system and M_{eq} is the equilibrium macrostate. M_{eq} is characterised as the macrostate whose associated macrovariables remain (approximately) constant in time.[15] For a Boltzmann-like gas, M_{eq} is standardly identified as the macrostate whose temperature, pressure,

---

[12]There are typically a variety of sets one can use to characterise a thermodynamic system. Physicists often choose sets based on the problems they're dealing with.

[13]It's standardly assumed that each macrostate compatible with the constraints placed on the system can be characterised using macrovariables that take on well defined values, or that are at least compatible with some well defined range of values.

[14]As usual, '∩', '∪', and '∅' denote set theoretic union, intersection, and the empty set, respectively.

[15]See Werndl and Frigg (2014) for an interesting discussion of the consequences of this characterisation of equilibrium when viewed through the lens of Boltzmannian statistical mechanics; they prove a new theorem establishing that equilibrium thus defined corresponds to the largest macroregion of a system's phase space energy hypersurface. Also, notice the similarity of this characterisation of equilibrium with the thermodynamic characterisation outlined by Brown and Uffink (2001).

and volume remain (approximately) constant through time. Corresponding to these macrostates are, respectively, the macro-regions Γ_{M_p} and $\Gamma_{M_{eq}}$.

The Boltzmann entropy of a macrostate M_i is standardly defined as $S_B(M_i) =_{df} k \log[\mu(\Gamma_{M_i})]$, where k is again Boltzmann's constant. Since a system is in exactly one macrostate at a time, we define the Boltzmann entropy of a *system* at a time t, $S_B(t)$, as the entropy of the system's macrostate at t. That is, $S_B(t) =_{df} S_B(M_{x(t)})$ where $x(t)$ is the system's microstate at t and $M_{x(t)}$ is its corresponding macrostate.

Typicality views attempt to account for why systems beginning in M_p end up in M_{eq}, and why they remain in M_{eq} for incredibly long periods of time. They also attempt to show that $S_B(t)$ generally increases to its maximum value as the system evolves from M_p to M_{eq} and that it remains (approximately) at this value for incredibly long periods of time. The later is done so as to claim that $S_B(t)$ is the microphysical analogue of thermodynamic entropy.

3.2.2 The Dominance View

One kind of typicality account, which we'll call the dominance view, attempts to account for the approach to equilibrium and an associated entropy increase by arguing for them on the basis of a technical result that applies to Boltzmann-like gases known as the "dominance of the equilibrium macrostate".

The first ingredient in the account is an interpretation of the measure over microstates on Γ_E—standardly taken to be Lebesgue measure—as a typicality measure. Typicality measures represent the *relative size of sets of states.*

They are usually understood to be a kind of normalised measure. Typical states show a certain property if the measure of the set that corresponds to this property is one or close to one. What it means to be a typicality measure is one of the pressing general concerns that will be discussed in a later section. The second, closely related ingredient is the idea that Lebesgue measure is a good (perhaps, the appropriate) measure of what's typical. That is, that it's a *good* typicality measure. Justifying why Lebesgue measure is a good choice of typicality measure is another pressing general concern that will be discussed in a later section.

The next ingredient is the recognition that under certain circumstances, and for Boltzmann-like gases, M_{eq} is the largest of all M_i.[16] In fact, for large n, M_{eq} is enormously larger than any other region.[17] For example, the ratio $\mu(M_{eq})/\mu(M_i)$ is of the magnitude of 10^n, where M_i is a nonequilibrium macrostate of an ordinarily prepared system.[18]

Since, for large n, Γ_E is almost entirely taken up by equilibrium microstates, the dominance of the equilibrium macrostate is often taken to imply that equilibrium *microstates* are typical with respect to Γ_E and Lebesgue measure μ.[19] Some authors claim that its dominance explains why systems exhibit the following *qualitative* facts: that isolated macroscopic systems that begin away from equilibrium spontaneously approach equilibrium and that

---

[16]See Frigg (2011: pp.89-90) and Uffink (2007: pp.974-983) for a discussion of the circumstances in which this result holds. See Werndl and Frigg (2015) for an argument that attempts to generalise this result.

[17]See Ehrenfest and Ehrenfest (2002: p.30).

[18]See Frigg (2011: p.80), Goldstein (2001: p.43), and Penrose (1989: p.403).

[19]See, for example, Frigg (2011: p.81), Bricmont (2001: p.146), Goldstein (2001: p.43), and Zanghì (2005: p.191 & p.196).
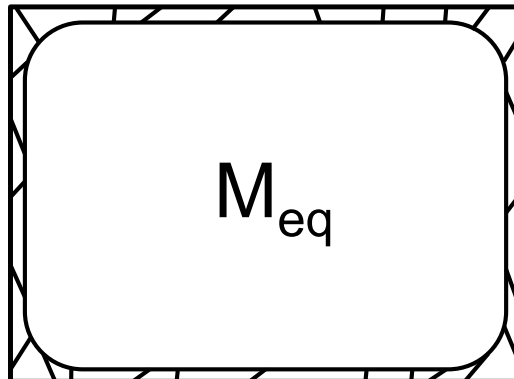
Figure 3.2: The dominance of the equilibrium macrostate on the energy hypersurface.

they stay there for incredibly long periods of time.[20] Some authors also claim that its dominance explains why thermodynamic entropy generally increases. Naturally, this account is seen by some as a way of fulfilling goals such as **G1-4** without introducing a time-reversal non-invariant element.

Frigg (2011: p.81) claims that such a view may be read into the following passage:[21]

> . . . reaching the equilibrium distribution in the course of the tem-

---

[20]See Frigg (2011: Sec.4.3) for a list of authors that endorse this kind of view.

[21]Frigg (2011: p.81) also claims that the dominance view may be read into the following passage by Goldstein, but this quote is taken from an article that is standardly thought to endorse what the next subsection labels "the unspecified dynamical view".

> Suppose a system, e.g. a gas in a box, is in a state of low entropy at some time. Why should its entropy tend to be larger at a later time? The reason is basically that states of large entropy correspond to regions in phase space of enormously greater volume than those of lower entropy. (Goldstein 2001: p.49)

poral evolution of a system is inevitable due to the fact that the overwhelming majority of microstates in the phase space have this distribution; a fact often not understood by the critics of Boltzmann... (Zanghì 2005: p.196; Frigg's translation)

The view can also be found in Price (1996: pp.39-40).

Thus it seems to me that the problem of explaining why entropy increases has been vastly overrated. The statistical considerations suggest that a future in which entropy reaches its maximum is not in need of explanation; and yet that future, taken together with the low-entropy past, accounts for the general gradient.

The most troubling concern for the dominance view is its lack of a connection to the dynamics.[22] Nothing that has been said so far *guarantees*, or even makes it *likely*, that the microstates of systems prepared into some nonequilibrium macrostate evolve into states that constitute the equilibrium macrostate. It also doesn't follow from the view presented above that systems whose microstates do evolve into the equilibrium macrostate remain in it for even short periods of time. Or even that states that *begin* in equilibrium stay there! The idea that systems remain in states away from equilibrium is perfectly consistent with the idea that equilibrium microstates are typical. So too is the idea that they fluctuate wildly in and out of equilibrium. The question then of what kinds of dynamical considerations need to be added to the view, along with any restrictions that need to be placed on the kinds of initial microstates we allow—on the basis of available preparation procedures—so

---

[22]Other concerns are discussed in Frigg (2009) and Frigg (2011). See also Uffink (2014).

that it can underpin the sought after qualitative facts, and satisfy goals such as **G1-4**, leads us onto discussing other typicality accounts.

3.2.3   The Unspecified Dynamical View

Aware of concerns with the dominance view—including, most troublingly, the concern mentioned above, that measure theoretic considerations cannot, on their own, tell us *anything* about the behaviour of systems that begin away from equilibrium—many supporters of typicality incorporate into their view some claim about a system's dynamics and initial state.[23] Standardly, they claim that given any reasonable account of the dynamics, typical initial microstates are taken by the dynamics into the equilibrium macrostate and that they remain there for incredibly long periods of time. Let's call this the unspecified dynamical view. The classic and often cited expression of this view is found in Goldstein (2001: pp.43-44):

> $[\Gamma_E]$ consists almost entirely of phase points in the equilibrium macrostate $[M_{eq}]$, with ridiculously few exceptions whose totality has volume of order $10^{-10^{20}}$ relative to that of $[\Gamma_E]$. For a non-equilibrium phase point $[x]$ of energy E, the Hamiltonian dynamics governing the motion $[x(t)]$ would have to be ridiculously special to avoid reasonably quickly carrying $[x(t)]$ into $[M_{eq}]$ and keeping it there for an extremely long time—unless, of course, $[x]$ itself were ridiculously special.[24]

---

[23]For a list and discussion of other concerns, see Frigg (2009) and Frigg (2011). See also Uffink (2014).

[24]This quote has been modified to be consistent with the formalism that has already been introduced. Its content has not been altered.

The unspecified dynamical view is, at present, the most promising and influential kind of typicality view. One of its ingredients is, like before, an interpretation of the measure over microstates on Γ_E as a typicality measure. Another is that it's a good choice of typicality measure. But here, unlike the dominance view, which treats the relevant typicality property as being a microstate that corresponds to the equilibrium macrostate, the relevant property is (or is something close to) being a microstate that has been prepared in an ordinary way, that evolves under the dynamics into a microstate that corresponds to the equilibrium macrostate (if it wasn't already in that macrostate), and that remains in the equilibrium macrostate for long periods of time.

When supporters of this kind of view discuss the dynamics of relevant systems, it's common for them to write very general and often noncommittal things. The description offered above by Goldstein is a good example. Here's another example, offered by Frigg (2008b: p.114):[25]

> If we now assume that the system's state drifts around more or less "randomly" on $[\Gamma_E]$ then, because $[M_{eq}]$ is vastly larger than any other macro region, sooner or later the system will reach equilibrium and stay there for at least a very long time.

As Frigg (2008b: p.114) rightly notes, the qualification "more or less randomly" is essential. If its motion were just right, then it could avoid ever wandering into the equilibrium macrostate. Or else, it could, among other things, wander in and out of this state in a very erratic fashion, thereby giving rise to a kind of anti-thermodynamic behaviour.

[25]See Lazarovici and Reichert (2014: p.7), also.

While some are dissatisfied by these descriptions, and think that a more detailed account of the dynamics is necessary, supporters of this view often see the generality of their position as both reasonable and virtuous.[26] Their point is often that there are many different dynamical properties a system could possess that each have the consequence that thermodynamic behaviour is typical, given the dominance of the equilibrium macrostate. And that the dynamics would need to be very precise indeed to avoid this consequence, since we're unable to prepare systems into states that would otherwise yield anti-thermodynamic behaviour.

Those unsatisfied with the unspecified dynamical view often try to supplement it by claiming that relevant systems possess some dynamical property found on the ergodic hierarchy.[27] Properties standardly appealed to include: ergodicity and epsilon-ergodicity.[28]

---

[26]Uffink (2007), Frigg (2009), Frigg (2011), Frigg and Werndl (2011), and Uffink (2014) express dissatisfaction at the manner in which generality is obtained by this view. Frigg and Werndl (2012b: p.101), for example, say:

> Just saying the relevant Hamiltonians possess the dynamical property of TD-likeness has no explanatory power—it would be a pseudo-explanation of the *vis dormitiva variety. The challenge is to identify in a non-question begging* way a dynamical property (or, indeed, properties) that those Hamiltonians whose flow is TD-like have.

Lazarovici and Reichert (2014: p.31-34), for example, see the generality of the position as both reasonable and virtuous.

[27]See Frigg, Berkovitz, and Kronz (2014) and Berkovitz, Frigg, and Kronz (2006) for more on the ergodic hierarchy.

[28]Some authors appeal to some form of mixing, but very few, if any, claim that relevant systems are K-systems or Bernoulli.

3.2.4   The Ergodic View

While no one contributing to the contemporary discussion appears to endorse what we'll call the ergodic view, it both frequently appears in contemporary discussions and is regarded by many to capture a position carved out by some of the ideas Boltzmann expressed in his later writings on statistical mechanics. Understanding this view makes it easier to understand a closely related view that *is* taken seriously, so it's worth discussing here.

Boltzmann conjectured that

> The great irregularity of the thermal motion and the multitude of forces that act on a body make it probable that its atoms, due to the motion that we call heat, traverse all positions and velocities which are compatible with the principle of [conservation of] energy. (Quoted in Uffink (2007: p.40))

This has come to be known as the *ergodic hypothesis.* As stated, it cannot be correct, as the trajectory is a one-dimensional continuous curve and so cannot fill a space of more than one dimension. But it can be true that almost all trajectories eventually enter every open neighbourhood of every point on the energy surface.[29]

A *dynamical system* $\langle \Gamma, \mathcal{S}, \mu, \phi_t \rangle$ with a measure preserving evolution (i.e. for every measurable set A, and every time t, $\mu(\phi_t^{-1}(A)) = \mu(A)$) is said to

---

[29]Boltzmann argued, on the basis of the ergodic hypothesis, that the long-run fraction of time that a system spends in a given subset of the energy surface is given by the measure that Josiah Gibbs was to call microcanonical. See Uffink (2007: Sec.4.1 & Sec.6.1-6.2) for a good discussion of ergodic theory and Boltzmann's ergodic hypothesis. For a classic discussion of Boltzmann's hypothesis see Ehrenfest and Ehrenfest (2002).

be *ergodic* if and only if, for any set A, such that $\mu(A) > 0$, the set of initial points that never enters A has zero measure.[30]

An important part of the ergodic view is the following two-part result established by George David Birkhoff (1931a,b).[31]

For any measure-preserving dynamical system,

1. For any $A \in \mathcal{S}$, and where

$$\chi_A(x) = \begin{cases} 1, & \text{if } x \in A \\ 0, & \text{if } x \notin A, \end{cases} \tag{3.12}$$

the limit

$$\langle A, x_0 \rangle_{time} = \lim_{T \to \infty} \frac{1}{T} \int_0^T \chi_A(T_t(x_0)) dt \tag{3.13}$$

exists for almost all points x_0. That is, except for a set of measure zero.

2. If the dynamical system is ergodic, then

$$\langle A, x_0 \rangle_{time} = \mu(A) \tag{3.14}$$

for all $A \in \mathcal{S}$ and almost all $x_0 \in \Gamma$.

The first part of the theorem says that it makes sense to talk about the long-run fraction of time the system spends in A. Or rather, it makes sense to talk about it for almost all initial microstates. The second part says that the long-run fraction of time the system spends in A is equal to the phase space volume of A—where we ignore those sets whose limit does not exist.

Again, noting the dominance of the equilibrium macrostate, it follows

---

[30]As usual, $\mathcal{S}$ is a σ-algebra on Γ.

[31]It should be noted that this result is appropriated by supporters of typicality. There is no indication that Birkhoff was or would be a supporter of typicality.

that almost all initial conditions lie on solutions that spend most of the time in equilibrium and only show relatively short fluctuations away from it.

If we add to this result the assumption (or can show that) relevant systems are ergodic, and if we again interpret the measure on Γ_E as a typicality measure, and hold that it's a good measure of what's typical, then it follows, given ordinary preparation procedures, that typical initial states approach equilibrium, and that they remain in equilibrium for incredibly long periods of time. Again, just like the views that have come before, this view could be seen as a way of fulfilling goals such as **G1-4** without introducing a temporally asymmetric element.

Standardly, this view is thought to be troubled by two problems.[32] First, it turns out to be extremely difficult to prove that anything that closely resembles a realistic system is ergodic. As Frigg (2008b: p.124) notes, not even a system of n elastic hard balls moving in a cubic box with hard reflecting walls has been proven to be ergodic for arbitrary n. In fact, it has only been proven in cases where $n \leq 4$. What's more, there are dynamical systems that exhibit thermodynamic behaviour that are provably not ergodic. Jean Bricmont (2001), for example, highlights that the Kac ring model and a system of n uncoupled anharmonic oscillators of identical mass exhibit thermodynamic behaviour and that both are provably not ergodic. Other well behaved, non-ergodic systems include: solids, and systems comprised of non-interacting point particles.[33] Examples such as these have led many to

---

[32]For discussions of these and other related problems, see Friedman (1976), Sklar (1995: Ch.5), Earman and Rédei (1996), van Lith (2001), Emch and Liu (2002: Ch.7-9), Uffink (2007: Sec.6.1), and Frigg (2008b: Sec.3.2.4.3).

[33]See Uffink (1996: p.381) and Uffink (2007: p.1017). In a solid, say an ice cube, the molecules are tightly locked to their lattice site, and the phase point can access only a

think that ergodicity does not provide a satisfactory explanation of why isolated systems spontaneously approach equilibrium and why they stay there for long periods of time.[34] The second worry standardly thought to trouble the ergodic view is known as the "measure zero" problem. While, properly speaking, the problem is usually levelled at a view closely related to what we're calling the ergodic view, in which the measure on Γ_E is simply regarded as a probability measure—whose probabilities require interpretation—rather than as a typicality measure, it, in this context, effectively amounts to a request to justify why we can neglect atypical microstates. Since Birkhoff's results apply to all but a set of states with measure zero, one would like to have reason to neglect them. That is, one would like to have reason to treat them as atypical states. The mere fact that the measure labels them as such does not provide the justification that's needed; sets that have measure zero according to some measures have finite measure according to others. So, the concern goes, what singles out this measure (or an equivalence class of measures) as the one to use to determine which sets are negligible? Since this is really just an instance of a general worry that will be discussed later, further discussion of it will be left till then.

minute region of the energy hypersurface.

[34]See, for example, Earman and Rédei (1996) and van Lith (2001). In fact, some—often those who endorse something close to the unspecified dynamical view—take an even stronger position on the ergodic enterprise. David Albert (2000: p.70), for example, thinks that

> . . . the prodigious effort that has over the years been poured into rigorous proofs of ergodicity is nothing more nor less—from the standpoint of the foundations of statistical mechanics—than a waste of time.

3.2.5   The Epsilon Ergodic View

There is a view, or rather, an interpretation of a view, endorsed by some of the discussions leading contributors, that's closely related to the ergodic view. It will be called the epsilon ergodic view. See Frigg and Werndl (2011, 2012a, & 2012b).[35] Unlike earlier views, this view restricts itself to accounting for the behaviour of realistic gases. Its advocates do not claim to be offering an account that speaks to the behaviour of macroscopic systems, generally. The view is primarily aimed at overcoming the first of the worries standardly levelled at the ergodic view. Its supporters hope to show that an important class of realistic systems are epsilon ergodic, and that this dynamical property, in conjunction with many of the ideas that have been introduced previously, accounts for why they spontaneously approach equilibrium, and why they spend large amounts of time in equilibrium.

To introduce epsilon ergodicity, it's helpful to begin, as Frigg and Werndl (2012b: p.104) do, by first defining ε-ergodicity—a distinct but related notion. A dynamical system $\langle \Gamma_E, \mu_E, \phi_t \rangle$ is ε-ergodic, where $\varepsilon \in \mathbb{R}$ and $0 \leq \varepsilon < 1$, if and only if there is a set $Z \subset \Gamma_E$, with $\mu(Z) = \varepsilon$, and with $\phi_t(\hat{\Gamma}_E) \subseteq \hat{\Gamma}_E$ for all $t \in \mathbb{R}$, where $\hat{\Gamma}_E =_{df} \Gamma_E \setminus Z$, such that the system $\langle \hat{\Gamma}_E, \mu_{\hat{\Gamma}_E}, \phi_t^{\hat{\Gamma}_E} \rangle$ is ergodic, where $\mu_{\hat{\Gamma}_E}(\cdot) =_{df} \mu_E(\cdot)/\mu_E(\hat{\Gamma}_E)$ for any measurable set in $\hat{\Gamma}_E$ and where $\phi_t^{\hat{\Gamma}_E}$ is ϕ_t restricted to $\hat{\Gamma}_E$. Naturally, a 0-ergodic system is simply an ergodic system. We say that a dynamical system is *epsilon ergodic* if and only if there exists a very small ε (i.e. $\varepsilon << 1$) for which the system is

---

[35]While Frigg and Werndl explicitly distance themselves from typicality in their 2011 paper and again in their 2012b chapter, and regard themselves as offering a competing account, their work can be straightforwardly read through a typicality lens. What's more, given their later work (see Frigg and Werndl 2012a, and Werndl 2013), it's reasonable to think that they would permit such a reading.

ε-ergodic.[36]

Since an epsilon ergodic system is ergodic on $\Gamma_E \setminus Z$, the set of initial states that lie in $\Gamma_E \setminus Z$ spend most of the time in equilibrium and only show relatively short fluctuations away from it. If ε is very small in comparison to $\mu_E(\Gamma_{M_P})$, then almost all initial microstates spend most of the time in equilibrium and only show relatively short fluctuations away from it. Here, "almost all" has been weakened to include all initial states except ones that form a set of measure ε. If this measure is understood as a typicality measure, then it follows that typical initial states exhibit thermodynamic behaviour. Once again, this has been achieved without introducing a temporally asymmetric element. So the result can be seen as a way of satisfying goals such as **G1-4**.

Supporters of the epsilon ergodic view claim that the kinds of systems usually called upon to show that ergodicity does not provide a sufficient explanation for why isolated macroscopic systems spontaneously approach equilibrium and why they remain in equilibrium for incredibly long periods of time cuts no ice in this context. As Frigg and Werndl (2012b: p.107) argue,

> Clearly, solids are not gases and hence can be set aside. Similarly, uncoupled harmonic oscillators and the Kac ring model are irrelevant because they seem to have nothing to do with gases. The properties of ideal gases are very different from the properties of real gases because there are no collisions in ideal gases and collisions are essential to the behaviour of gases. So while ideal gases

---

[36]Peter Vranas (1998) first introduced the concept of epsilon ergodicity to the foundations of statistical mechanics, but he put it to a different use.

may be expedient in certain context, no conclusion about the dynamics of real gases should be drawn from them. Hence, the well-rehearsed examples do not establish that there is a gas-like system which behaves TD-like while failing to be ergodic.[37]

What's more, they often cite a small collection of rigorous results along with some numerical studies to support the idea that real gases are, in fact, epsilon ergodic.[38]

3.3 Some General Concerns About Typicality Accounts

This section discusses three worries commonly thought to trouble the typicality accounts presented above. It also discusses, where possible, their most promising solutions. Two of these worries were flagged earlier. The first worry concerns interpreting Lebesgue measure as a typicality measure. The second worry grants that it's a typicality measure and asks, why should we think it's a good measure of what's typical? The third worry concerns extending the results gained from the study of Boltzmann-like gases to other

---

[37]This conclusion is a bit odd. The conclusion Frigg and Werndl have in fact argued for is more narrow: that the well-rehearsed examples do not establish that there is a *real gas system* which behaves TD-like while failing to be ergodic. Their conclusion seems to rely on a pretty dim view of idealised models and toy models. The Kac ring, for example, is thought to share salient features with real gases. In fact, it's often called on, as it was in Section 3.2.1, to highlight certain things about real gases *because* it has these features. And yet, Frigg and Werndl's remarks imply that it's not gas-like!

[38]See, for example, Frigg and Werndl (2011: Sec.7), Frigg and Werndl (2012b: Sec.7.6), and Frigg and Werndl (2012a: Sec.6).

macroscopic systems.[39]

3.3.1 Typicality Measures

All of the typicality accounts presented above trade on the assumption that Lebesgue measure can be interpreted as a typicality measure. One might wonder: what justifies interpreting this measure as a typicality measure? A promising answer to this question has recently been offered by Charlotte Werndl (2013).[40] Werndl offers a set of conditions that measures ought to satisfy so as to count as typicality measures and argues that these conditions are satisfied by Lebesgue measure. This set builds on conditions that are often implicitly, but sometimes explicitly, endorsed in the literature.

As mentioned earlier, a typicality measure represents the relative size of sets of states. Any function T that describes the size of sets of states should satisfy the standard axioms of a measure. That is, $T(\emptyset) = 0$, $T(A) \geq 0$ for any measurable set A, and $T(\bigcup_{i \in \mathbb{N}} A_i) = \sum_{i \in \mathbb{N}} T(A_i)$ whenever $A_i \cap A_j = \emptyset$ for all $i, j, i \neq j$. Since typicality measures are thought to represent the *relative* size of sets of states, they are usually normalised, i.e., $T(\Gamma_E) = 1$. It was also mentioned earlier that typical states show a certain property if the measure of the set that corresponds to this property is one or close to one. Formally, typical states of Γ_E have property P relative to a typicality measure T if and only if $T(\Gamma_E \setminus D) \leq \beta$, where $D \subseteq \Gamma_E$ consists of all of the states in Γ_E that have property P and β is a very small real number,

---

[39]For a different discussion of these concerns, and other general concerns, see Frigg (2011: Sec.4.4). See Lazarovici and Reichert (2014: Sec.5) for a reply to the concerns raised by Frigg, and others.

[40]Pitowsky (2012) also attempts to justify typicality measures. See Werndl (2013: Sec.6) for a critique of this attempt.

possibly 0. D is called the typical set and $\Gamma_E \setminus D$ is called the atypical set. We'll look at the conditions Werndl suggests for the case in which a typical state possesses some property that it shares with a collection of states that together form a set of measure one.[41]

The first condition Werndl (2013: p.474) notes is that the measure has to be invariant under the dynamics. What's typical at some time has to be typical at both earlier and later times. Call this condition **C1**. This condition is standardly endorsed by supporters of typicality.

> **C1.** Typicality measures are invariant under the dynamics.

Since we're unable to prepare systems into particular microstates, it's common to place a probability distribution (or class of distributions) over the set of microstates compatible with our preparation procedures. Such a distribution gives the probability that a system has been prepared in some microstate. Werndl (2013: p.474) assumes that these probability distributions, p, are *translation-continuous*. That is,

$$\lim_{\|\tau\| \to 0} p(\mathrm{Tr}(A, \tau)) = p(A) \text{ for all open sets } A, \tag{3.15}$$

where $\|\tau\|$ is the standard Euclidean norm in $\mathbb{R}^n$ and $\mathrm{Tr}(A, \tau)$ is the set A translated by τ. That is,

$$\mathrm{Tr}(A, \tau) = \Gamma \cap \{x + \tau | x \in A\}. \tag{3.16}$$

---

[41]See Werndl (2013: Sec.8) for a set of conditions that are relevant for a more liberal kind of typicality measure; where typical states possess some property that they share with a collection of states that together form a set of measure close to one.

We'll return to this assumption later. Next, Werndl (2013: p.475) assumes that there is a class P of initial probability distributions of interest, which she characterises as,

> **Distributions of Interest:** The initial probability distributions of interest are a class P of translation-continuous probability distributions where for every open set A there is a $p \in P$ which $p(A) > 0$.

The second part of this characterisation intends to capture the possibility that, for any arbitrary open region of phase space, we cannot exclude the possibility that there might be some way of preparing the system such that there is a positive probability that it ends up in this region. While this assumption does not impose a condition that a measure has to satisfy in order to count, by Werndl's lights, as a typicality measure, it does place a restriction on the kinds of distributions with which these conditions have to be satisfied.

The next two conditions are motivated by the idea that typicality measures should be related to initial probability distributions of interest. They are:

> **C2.** If $p(A) = 0$ for all probability distributions of interest for some measurable A, then $T(A)=0$.

And,

> **C3.** If $T(A)=1$ for some measurable A, then $p(A) = 1$ for all probability distributions of interest.

C2 is the requirement that if a set of states has probability zero for all initial probability distributions of interest, the set is atypical. **C3** is the requirement that whenever a set of states is typical, it has probability one for all initial probability distributions of interest.

Werndl (2013: p.475) claims that a measure is a typicality measure whenever it satisfies conditions **C1-3**, for probability distributions characterised as ones of interest. Werndl then argues that Lebesgue measure satisfies these conditions, and so concludes that it can be interpreted as a typicality measure.[42]

While it may appear that the above considerations only provide us with the resources to argue for the permissibility of interpreting Lebesgue measure as a typicality measure, they, in fact, also provide us with the resources to argue that it provides a *good* account of what's typical. This is good news because one might have worried *first*, that these measures were not appropriately connected to experience, and so do not capture what's typical of actual systems, and *second*, that there are measures that can be placed on Γ_E that conflict with what's typical according to Lebesgue measure.

The first of these worries is remedied by the connection of measures to probability distributions of interest. They are related to our means of preparation of systems, and the rationale for assuming that they are translation-continuous has to do with feasible preparation procedures—as we'll see in a moment. The second concern is avoided because of a result due to Malament and Zabell (1980). Malament and Zabell have shown that a probability measure p is translation-continuous if and only if p is absolutely continuous with

---

[42]See Werndl (2013: Appendix A) for the details of this argument.

respect to the Lebesgue measure $(p \ll \mu)$, where μ_2 is said to be absolutely continuous with respect to μ_1 if and only if

$$\text{if } \mu_1(A) = 0 \text{ for a measurable set } A, \text{ then } \mu_2(A) = 0 \qquad (3.17)$$

for any two measures μ_1 and μ_2 defined on $\langle \Gamma, \mathcal{S} \rangle$.

So then, the success of Werndl's proposal effectively boils down to the plausibility of her characterisation of probability distributions of interest, and, relatedly, the assumption that initial probability distributions of interest are translation-continuous. Regarding the later condition, Werndl cites the motivation Malament and Zabell (1980) offer in support of this assumption. They write:

> Given two measurable sets on the constant energy surface, if one
> is but a small displacement of the other, then it seems plausible
> to believe that the probability of finding the exact microstate of
> the system in one set should be close to that of finding it in the
> other. (Malament and Zabell 1980: p. 346)

Unfortunately, this rationale does not uniquely support translation-continuity. In the definition of translation-continuity, the perturbations of the set A are ones such that every point of A undergoes the same displacement. But this does not seem essential to the rationale offered by Malament and Zabell, and it seems that the rationale should hold equally well for other transformations of our measure space, whether or not each point gets the same displacement, as long as the displacements are small. With this in mind, consider the stronger condition of displacement-continuity, which we'll define as follows.

> A measure υ is *displacement-continuous* if and only if, for every
> measurable set A, for any $\epsilon > 0$, there exists $\delta > 0$ such that
> $|\upsilon(H(A)) - \upsilon(A)| < \epsilon$ for all homeomorphisms H of the measure
> space such that $\|H(x) - x\| < 0$ for all x.

This seems to be an intuitively reasonable condition to place on our probability measures. Unfortunately, it's not satisfied by Lebesgue measure, and, indeed, not satisfiable![43]

While Malamant and Zabell's remarks do not quite have the consequence that's needed, if one could motivate, on the basis of physical considerations, translation-continuity, or some other continuity condition that's equivalent to being absolutely continuous with respect to Lebesgue measure, then something that resembles the reasoning Werndl employs seems likely to go through.

3.3.2   Extending Typicality Results

Typicality accounts attempt to extend results that apply to Boltzmann-like gases to other macroscopic systems. Their success depends on the plausibility of their extension. Notably, these accounts assume that relevant systems have energy hypersurfaces that are dominated, like a Boltzmann-like gas, by their respective equilibrium macrostate. They also assume that the microstates that constitute the equilibrium macrostate are, like a Boltzmann-like gas, typical. Both of these assumptions are dubious.

First, as Frigg (2011: p.89) notes, despite it often being stated as if it were

---

[43]See Luczak and Myrvold (unpublished) for more on this point, and for a discussion of translation-continuity, its rationale, and other continuity conditions.

a general truism, the result that the equilibrium macrostate dominates a system's energy hypersurface has only been proven for an ideal gas. That is, for a system of non-interacting particles. What's more, as Uffink (2007: p.976) explains, it's central to the proof that we're dealing with an ideal gas. Since most systems are not ideal gases, not even approximately, one may worry about using this assumption to account for their behaviour. Moreover, as Frigg (2011: p.90) also notes, it seems likely that this assumption is false, at least for some thermodynamically well-behaved systems.[44]

Second, even if we grant that systems have energy hypersurfaces that are dominated by their respective equilibrium macrostate, it doesn't follow that the measure of this set is one or close to one. In fact, it's possible that there are systems with distinct nonequilibrium macrostates that share the same Boltzmann entropy value whose union has a measure greater than the equilibrium macrostate. David Lavis has shown that this possibility is realised by the Baker's gas (a gas whose particles move according to the Baker's transformation) and the Kac ring model.[45]

Of course, most supporters of typicality are under no illusions about the results they actually possess or about the ones they don't. They are also not under any illusions about what they can logically infer about the behaviour of other systems on the basis of the results they actually possess. These claims are especially true of those who support the unspecified dynamical view. Many supporters of typicality seem to think of themselves as offering arguments that are of heuristic value. They see themselves as offering a

---

[44]See also Callender (2010).

[45]See Lavis (2005: pp.255-258) for a more detailed description of the Baker's gas and the relevant result, and see Lavis (2008: Sec.2) for the Kac ring model result.

general framework that highlights the form rigorous results would likely take. Because of this, supporters are usually untroubled by these concerns. They are also unmoved by those who demand more rigorous results; they often see them as unnecessary.

3.4　The Limitations of Typicality Accounts

Suppose for the moment that we can treat Lebesgue measure as the appropriate and unique measure of typicality. Suppose also that every isolated thermodynamic system has an energy hypersurface that is dominated by its equilibrium macrostate. Or even just that many of the systems we're interested in have this feature. Suppose further that this state has measure one (or is close to one) and that the set of microstates that evolve under the dynamics of the system that spend incredibly long periods of time in the equilibrium macrostate also has measure one (or is close to one). Suppose that by ordinary preparation procedures systems are placed into one of these microstates with high probability. That is, with probability one or with a probability close to one. And suppose that every isolated thermodynamic system can be shown either to be ergodic or epsilon ergodic. Or even just that many of the systems we're interested in can. An aim of this section is to highlight that typicality accounts have severe limitations, even if we grant all of these things.

Typicality accounts do not provide us with the resources to answer important questions such as: what will this system do in the next five minutes? Ten minutes? Hour? Year? What states will it pass through on its way to

equilibrium? If it reaches equilibrium, how long will it stay there before it moves out of equilibrium? How likely is it that we'll see the system fluctuate out of equilibrium in the next few minutes? If it does, how large a fluctuation should we expect? Etc. They do not underpin facts about the rates in which systems approach equilibrium, or about the kinds of states they pass through on their way to equilibrium, or about fluctuation phenomena. Interestingly, it's typically these other quantitative facts, and not the qualitative ones the literature has focused on, that we care about most. Moreover, none of these views help us form expectations about any of these quantitative facts or help us justify the expectations we may already have about the behaviour of systems that begin away from equilibrium, having formed them on the basis of experience. They also do not underpin quantitative relations such as the fluctuation-dissipation theorem (see Sec.4.1).

Each view fails to underpin these facts, and fails to answer these questions, for the same reason: they do not incorporate enough dynamical information. The dominance view is unable to underpin these facts and answer these questions because it does not incorporate *any* dynamical information. We have no way of accounting for what a system will do without *some* account of its dynamics. The unspecified dynamical view is unable to underpin these facts, and answer these questions, because it, as its name suggests, leaves a system's dynamics, for the most part, unspecified. And both the ergodic and epsilon ergodic views fail because they are tied to dynamical properties that only hold in the long term time limit. Consequently, neither of these properties can be used to validly conclude anything about a system's state

at a particular time, or about its behaviour over finite time intervals.[46]

Of course, none of this should come as much of a surprise once we remember what typicality accounts hope to achieve. These accounts attempt to show, in a way that satisfies goals such as **G1-4**, that isolated macroscopic systems that begin away from equilibrium spontaneously approach equilibrium and that they remain in equilibrium for incredibly long periods of time. Importantly, none of these goals makes any reference to the rates in which systems approach equilibrium, or about the states they pass through on their way to equilibrium, or about fluctuation phenomena. And why would they? Not only would this undermine the generality of these claims, which is surely one of their virtues, but it would also betray the macroscopic theory they're intended to be compatible with: thermodynamics. That is, a theory that does not speak to any of these things. In fact, a theory no less, that despite it's somewhat misleading name, does not contain any dynamics.

In light of all of these shortfalls, it seems natural to turn some of our attention towards understanding and accounting for the success of the techniques physicists use to model the behaviour of systems that begin away from equilibrium. These techniques lead to equations that track a large and important set of facts that concern a system's behaviour. These equations also answer the kinds of questions we're often most interested in. Understanding and accounting for the success of these techniques and equations would provide us with an underpinning of a large and important set of quantitative facts that typicality accounts cannot. Accounting for their success would also provide

---

[46]In fact, these criticisms also apply to views which hold that relevant systems are mixing. Some of these criticisms, however, do not apply to views which hold that relevant systems are K-systems or Bernoulli. Again, for more on the ergodic hierarchy, see Berkovitz et al. (2006) and Frigg et al. (2014).

us with an underpinning, albeit a mosaic one, of the general qualitative facts that are presently at the centre of foundational discussions. Here's another way of making these points. While there are certain qualitative facts we want to understand about the behaviour of systems that begin away from equilibrium, we *also* want to understand an associated set of quantitative facts, and an account of how systems approach equilibrium quantitatively will *a fortiori* tell us *that* they approach equilibrium and remain in equilibrium for incredibly long periods of time.

What's more, as an added bonus, this alternative line of research is more in line with actual scientific practise. It's worth noting that when physicists model the behaviour of systems that begin away from equilibrium, they are typically not, at the same time, attempting to satisfy goals such as the ones that drive typicality accounts. While they are sensitive to the tension between the existence of irreversible macroscopic behaviour and an underlying reversible dynamics, and while they are concerned about getting their models to agree with experience, they are not concerned about expressing themselves in ways that would underpin very general facts that concern the behaviour of a wide variety of macroscopic systems. They are happy for their models, which incorporate more specific dynamical information than the typicality accounts presented above, to simply describe and predict the behaviour of particular systems, or to describe and predict the behaviour of particular classes of systems.

3.5 Summary

This chapter outlined and examined a collection of typicality accounts. These accounts aim at explaining why isolated macroscopic systems spontaneously approach equilibrium and why they remain in equilibrium for incredibly long periods of time. They attempt to underpin these qualitative facts and satisfy goals such as **G1-4**, thereby reconciling the existence of irreversible macroscopic processes with underlying reversible dynamics.

This chapter began by discussing Boltzmann's H-theorem—which can be seen as the first serious attempt to account for the behaviour of systems that begin away from equilibrium—and highlighted that it gives rise to one of the very issues that is at the centre of contemporary philosophical discussions of statistical mechanics: reconciling the existence of irreversible macroscopic processes with underlying reversible dynamics. It then outlined and discussed several typicality accounts. As was noted, these accounts are inspired by, and build on, one of Boltzmann's attempts to resolve the issues that were sparked by his original derivation of the H-theorem. This chapter considered some of the most pressing problems facing typicality accounts and, where possible, discussed their most promising solutions. It then explained that these accounts have severe limitations, even when we ignore the usual concerns. This chapter highlighted that they do not underpin a large and important set of facts that concern the behaviour of systems that begin away from equilibrium. This chapter suggested that this was a consequence of what they aim at: underpinning two very general qualitative facts, while satisfying goals such as **G1-4**. While typicality accounts may underpin some important aspects of thermodynamics, they do not underpin many of the im-

portant quantitative facts that concern the behaviour of systems that begin away from equilibrium. They are also of no help when it comes to forming expectations about the rates in which systems equilibrate, or about the kinds of states they pass through on their way to equilibrium, or about fluctuation phenomena.

To remedy these shortfalls, this chapter ended by suggesting that we pursue an alternative line of research. That we attempt to understand and ground the success of the techniques and equations physicists commonly use to model the behaviour of systems that begin away from equilibrium. This line of research, if successful, has the potential to underpin not just some important qualitative facts but also a large set of important quantitative facts.

Chapter 4

Another Way to Approach the Approach to Equilibrium

The previous chapter ended by highlighting the limitations of typicality accounts. To remedy this shortfall, it promoted pursuing a different line of research. That we attempt to understand and account for the success of the techniques and equations physicists use to model the behaviour of systems that begin away from equilibrium. The previous chapter indicated that by understanding and grounding their success, we will not only be able to underpin the qualitative facts the literature has focused on but we will also be able to underpin many of the important quantitative facts that typicality accounts cannot.

This chapter takes steps in this promising direction. It outlines and examines a technique commonly used to model the behaviour of an important kind of system. More accurately, this chapter outlines and examines a technique commonly used to model the approach to equilibrium, and subsequent

behaviour, of a system in which a Brownian particle is introduced to an isolated homogeneous fluid at equilibrium. As this chapter highlights, the technique generates a collection of quantitatively accurate equations that track important aspects of the system's behaviour. This chapter also attempts to account for the success of the model, by identifying and grounding the technique's key assumptions.

This chapter is composed of four sections. The first section outlines a way of modelling the Brownian particle's approach to equilibrium and subsequent behaviour. As one might have expected, the approach takes its cue from the theory of Brownian motion. It appeals to Langevin dynamics. The aim of this section is to present the technique in a manner that reflects the way physicists standardly model this kind of system.[1] As this section highlights, the approach generates a collection of interesting equations that track aspects of the particle's behaviour. These equations provide us with the resources to answer the kinds of questions commonly asked about this kind of system. The second section narrows in on, and discusses, the approach's key assumptions. The third section attempts to motivate a particular microphysical claim that would ground the success of the technique, justify its key assumptions, and underpin the facts the equations it leads to track. The fourth, concluding

---

[1]It should be noted that the system and technique discussed in this chapter draw on a range of assumptions supported by equilibrium statistical mechanics. While there are benefits to considering systems and techniques that do not appeal to such assumptions, such as those that are described by, and that generate, Boltzmann's equation, there are also costs. For one thing, these free-of-equilibrium-statistical-mechanical-assumption systems and techniques are conceptually more difficult than the ones discussed in this chapter. Another reason to focus on the systems and techniques of Langevin dynamics are their novelty. There is already a wealth of philosophical literature that discusses Boltzmann's equation. As far as the author can tell, no one, at least at the time of publication, has discussed Langevin dynamics in the philosophical literature on statistical mechanics.

section, includes a summary and a suggestion about future research.

4.1 Modelling a Brownian Particle's Behaviour

The theory of Brownian motion is quite possibly the simplest approximate way to treat the dynamics of nonequilibrium systems. The theory arose from investigations into the irregular behaviour of objects such as pollen grains and dust particles when they are placed into various kinds of fluids. The phenomenon became widely known through the work of Robert Brown, a botanist, who, in 1827, reported on the strange behaviour of pollen grains floating on water. Albert Einstein (1905) and Marian Smoluchowski (1906) provided the first theoretical analyses of Brownian motion. Since then, a number of authors have helped develop the theory. Paul Langevin, a prominent French physicist during the early parts of the 20th century, was one such contributor.[2] What's particularly interesting about the theory of Brownian motion is that it can be applied successfully to many other phenomena (e.g. the motions of ions in water and the reorientation of dipolar molecules). In fact, the theory has been extended to situations in which the "Brownian particle" is not really a particle at all, but is instead some collective property of a macroscopic system. An example is the instantaneous concentration of any component of a chemically reacting system near thermal equilibrium. In this situation, the irregular fluctuations of the concentration in time correspond to the irregular motion of the Brownian particle. What's more, Brownian

---

[2]See, for example, Langevin (1908).

motion continues to be a topic of research nearly 200 years after Brown's work. As Pathria and Beale (2011: p.599) explain, much of the current interest is due to the growth in the technological importance of colloids across a wide range of fields, and the development of digital video and computer image analysis.

The *Langevin equation* is the theory's fundamental dynamical equation. It contains both frictional forces and random forces. It's a linear, first-order, inhomogeneous differential equation. The Langevin equation is often used to construct expressions that track interesting macroscopic variables. We often have detailed quantitative information about these variables and they capture what we're often most interested in tracking. This section examines the evolution of a system in which a Brownian particle is introduced to an isolated homogeneous fluid at equilibrium. It also highlights the Langevin equation's use in the construction of equations that track the behaviour of two interesting macroscopic variables—the Brownian particle's mean squared displacement and its mean squared velocity. The theory also possesses a fluctuation-dissipation theorem. It relates the forces that appear in the Langevin equation to each other. This theorem has many important and far-reaching generalisations. This section includes an elementary version of the theorem.

When a single Brownian particle is immersed in an isolated homogeneous fluid at equilibrium it typically does one of three things, depending on its initial velocity. If the Brownian particle enters the fluid with a mean squared velocity (an expectation value around which the actual value fluctuates) that's greater than the value given by the equipartition theorem, then

it typically slows down to that value. If the Brownian particle enters the fluid with a mean squared velocity that's less than the value given by the equipartition theorem, then it typically speeds up. The particle's speeding up and slowing down is the result of collisions with the molecules that comprise the fluid. If the Brownian particle enters the fluid with a mean squared velocity that's close to the value given by the equipartition theorem, then it typically remains at this value. When the Brownian particle reaches equilibrium, it jostles about, on macroscopic time scales, with a mean displacement of zero. Slight deviations from this behaviour can occur, but large deviations are extremely rare. From the standpoint of an ordinary observer, the Brownian particle's short-term behaviour seems to be random, both as it approaches equilibrium and once it has reached equilibrium. As such, there isn't much more an unaided observer can say about its motion.

While the motion of a Brownian particle appears to be random, it's nonetheless describable by the same equations of motion that model the behaviour of other classical systems.[3]

We'll consider the one-dimensional motion of a spherical particle with radius a, mass m, position x, and velocity V, in a fluid medium with viscosity η.

Newton's equation of motion for the particle is

---

[3]One might argue that it's inappropriate to be talking about classical mechanics in connection with the Langevin equation, since the wave packets of molecules will spread too quickly for a classical approximation to be appropriate.

However, in an appropriate regime, even if we can't treat the molecules as approximately localised, the Wigner function will, to a good approximation, obey the same Liouville equation that the classical probability density function obeys, and hence thinking about what classical trajectories will do is a useful way of visualising the behaviour of the quantum probability density. The expectation values we consider can be taken to be expectation values of quantum observables, and much the same things will be said in that context.
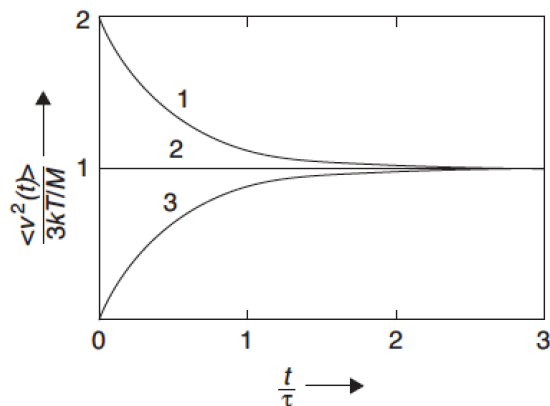
Figure 4.1: The mean squared velocity of a Brownian particle as a function of time. Curves 1,2, and 3 correspond, respectively, to the initial conditions $v^2 = 6KT/m, 3KT/m$, and 0. τ is the characteristic time. (Pathria and Beale 2011: p.591)

$$m\frac{dv}{dt} = F_{total}(t). \tag{4.1}$$

$F_{total}(t)$ is the total instantaneous force on the particle at time t. This force is the result of the particle's interaction with the fluid molecules. If the positions of the molecules are known as functions of time, then, in principle, this force is a known function of time. In this sense, the particle is not subject to a "random force" at all.

It's usually not practical or even desirable to look for an exact expression of $F_{total}(t)$. Experience teaches us that in ordinary cases this force is dominated by a frictional force, $-\zeta V$, that is proportional to the velocity of the Brownian particle. The friction coefficient is given by Stokes' law, $\zeta = 6\pi\eta a$.

If this were the whole story, the equation of motion for the Brownian particle would be

$$m\frac{dV}{dt} \cong -\zeta V, \tag{4.2}$$

whose solution is

$$V(t) = e^{-\zeta t/m}V(0). \tag{4.3}$$

According to (4.3), the velocity of the Brownian particle decays, over time, to zero. This result, however, cannot be quite right. It follows from the equipartition theorem that the mean squared velocity of the particle at thermal equilibrium is $\langle V^2 \rangle_{eq} = KT/m$. Obviously, the assumption that $F_{total}(t)$ is dominated by the frictional force has to be modified.

Standardly, an additional force, $\delta F(t)$, is added to the frictional force. For reasons that will become clear shortly, this force is referred to as the fluctuating force. The equation of motion then becomes

$$m\frac{dV}{dt} = -\zeta V + \delta F(t). \tag{4.4}$$

This is the Langevin equation for a Brownian particle. It partitions the total force in two. Because both parts are the result of the Brownian particle's interaction with the fluid, it's reasonable to think that a fundamental relation exists between them.

The additional force, $\delta F(t)$, is the result of continual collisions between the Brownian particle and molecules in the fluid. It's typically assumed, given the observed randomness of an individual trajectory, that this force

has the following properties.

$$\langle \delta F(t) \rangle = 0, \tag{4.5}$$

and

$$\langle \delta F(t) \delta F(t') \rangle = 2B\delta(t - t'). \tag{4.6}$$

Because of these properties, $\delta F(t)$ is standardly referred to as the fluctuating force. B is a measure of the strength of the fluctuating force. The delta function in time indicates that there is no correlation between collisions in distinct time intervals dt and dt'.

The Langevin equation can be solved to give

$$V(t) = e^{-\zeta t/m}V(0) + \int_0^t dt' e^{-\zeta(t-t')/m}\delta F(t')/m. \tag{4.7}$$

The first term describes the exponential decay of the particle's initial velocity. The second term describes the extra velocity produced by the fluctuating force.

This expression can be used, in conjunction with (4.5) and (4.6), to model the equilibration of the Brownian particle's mean squared velocity. There are three contributions to $V(t)^2$. The first one is

$$e^{-2\zeta t/m}V(0)^2. \tag{4.8}$$

It decays to zero at long times. Next are two cross terms, each first order in the noise,

$$2V(0)e^{-\zeta t/m} \int_0^t dt' e^{-\zeta(t-t')/m} \delta F(t')/m. \tag{4.9}$$

When (4.5) is used to average over the noise, these terms vanish. The final term is second order in the noise:

$$\int_0^t dt' e^{-\zeta(t-t')/m} \delta F(t') \int_0^t dt'' e^{-\zeta(t-t'')/m} \delta F(t'')/m^2. \tag{4.10}$$

Now the product of the two noise factors is averaged, according to (4.6). This leads to

$$\int_0^t dt' e^{-\zeta(t-t')/m} \int_0^t dt'' e^{-\zeta(t-t'')/m} 2B\delta(t'-t'')/m^2. \tag{4.11}$$

The delta function removes one time integration, and the other can be done directly. The resulting mean squared velocity is

$$\langle V(t)^2 \rangle = e^{-2\zeta t/m} V(0)^2 + \frac{B}{\zeta m}(1 - e^{-2\zeta t/m}). \tag{4.12}$$

In the long time limit, the exponentials fall out, and this quantity approaches $B/\zeta m$. At long times the mean squared velocity obtains the value KT/m, as given by the equipartition theorem. As a consequence, we find that

$$B = \zeta kT. \tag{4.13}$$

This result is known as the *fluctuation-dissipation theorem.* It relates the strength B of the fluctuating force to the magnitude ζ of the friction.

The evolution of one of the system's macrovariables is described by (4.12).

It models the system's spontaneous approach to equilibrium. It provides qualitative and quantitative information. It tells us to expect the system to approach equilibrium reasonably quickly and for it to remain in equilibrium for incredibly long periods of time. It helps us form expectations about the kinds of macrostates the system will be in at various stages of its evolution and it helps us form expectations about how quickly the system will reach equilibrium. It also helps us form expectations about fluctuation phenomena. What's more, as the fluctuation-dissipation theorem reveals, (4.12) also leads to quantitative information about the strength of the fluctuating force.

The Langevin equation can also be used, again with the help of (4.5) and (4.6), to derive an expression for the mean squared displacement of a Brownian particle that's been in the fluid for a long time.

Multiplying both sides of the Langevin equation by x gives

$$mx\frac{dV}{dt} = m\left[\frac{d}{dt}(xV) - V^2\right] = -\zeta xV + x\delta F(t). \tag{4.14}$$

If we take the average of both sides and use (4.5), we find that

$$m\langle\frac{d}{dt}(xV)\rangle - m\langle V^2\rangle = -\zeta\langle xV\rangle. \tag{4.15}$$

From the equipartition theorem we have $\frac{1}{2}m\langle V^2\rangle = \frac{1}{2}KT$, so (4.15) becomes

$$m\langle\frac{d}{dt}(xV)\rangle = KT - \zeta\langle xV\rangle. \tag{4.16}$$

This is a first-order differential equation that can be solved using the integrating factor technique. Its solution is

$$\langle xV \rangle = Ce^{-\frac{\zeta}{m}t} + \frac{KT}{\zeta}, \tag{4.17}$$

where C is a constant of integration. Let $\tau \equiv \frac{\zeta}{m}$, so that τ^{-1} is a characteristic time constant of the system. If we take as our initial condition that $x = 0$ at $t = 0$, then (4.17) entails that $C = -\frac{KT}{\zeta}$. So we have

$$\langle xV \rangle = \frac{KT}{\zeta}(1 - e^{-\tau t}). \tag{4.18}$$

If we replace $\langle xV \rangle$ by $\frac{1}{2}\frac{d}{dt}\langle x^2 \rangle$ and integrate (4.18) with respect to t, we obtain

$$\langle x^2 \rangle = \frac{2KT}{\zeta}[t - \frac{1}{\tau}(1 - e^{-\tau t})]. \tag{4.19}$$

(4.19) gives the *mean squared displacement* of the Brownian particle.

Note two interesting limiting cases. *Case 1*: $t << \tau^{-1}$. If $t < \tau^{-1}$, then using a Taylor series expansion,

$$e^{-\tau t} = 1 - \tau t + \frac{1}{2}\tau^2 t^2 - \dots. \tag{4.20}$$

So for $t << \tau^{-1}$ and $O(t^3) = 0$,

$$\langle x^2 \rangle = \frac{KT}{m}t^2. \tag{4.21}$$

This means that on very short time scales the particle behaves as if it were a free particle moving along with constant velocity $V = (KT/m)^{\frac{1}{2}}$.

Case 2: $t >> \tau^{-1}$. If $t >> \tau^{-1}$, then $e^{-\tau t} \to 0$. So (4.19) simply becomes

$$\langle x^2 \rangle = \frac{2KT}{\zeta} t. \tag{4.22}$$

(4.19) provides both interesting qualitative information and interesting quantitative information about one of the system's macrovariables. It models the Brownian particle's behaviour. The Brownian particle jostles about with a mean displacement of zero and with a mean squared displacement given by (4.19). It tells us to expect it to behave as if it were a free particle moving along with constant velocity on very short time scales, and for it to behave like a diffusing particle executing a random walk on longer time scales. It helps us form expectations about the kinds of macrostates the system will be in at various stages of its evolution, and it helps us form expectations about the Brownian particle's fluctuations. And this does not exhaust the quantitative information that can be extracted from the model. Since we're able to measure $\langle x^2 \rangle$ experimentally, then, if we know the size and density of the particles, as well as the viscosity of the medium, we can deduce from these observations a good approximation of K, Boltzmann's constant. Moreover, if we have knowledge of the gas constant, then we can also calculate the value of Avogadro's number.

As a final point, consider the results of the following real study. In the study, experimenters observed the behaviour of a spherical particle immersed in water and recorded aspects of its Brownian motion. The study highlights the predictive accuracy of (4.19) and (4.22). For the complete details of the study, see Lee, Sears, and Turcotte (1973). In the study, $T \simeq 300K$, $\eta \simeq 10^{-2}$ poise, and $a \simeq 4 \times 10^{-5}$cm. The study found that 403 values of the

net displacement, Δx, of the particle, observed after successive intervals of 2 seconds each, were distributed as follows:

Displacement Δx, in units of $\mu (= 10^{-4}$cm)	Frequency of occurrence n
less than -5.5	0
between -5.5 and -4.5	1
between -4.5 and -3.5	2
between -3.5 and -2.5	15
between -2.5 and -1.5	32
between -1.5 and -0.5	95
between -0.5 and $+0.5$	111
between $+0.5$ and $+1.5$	87
between $+1.5$ and $+2.5$	47
between $+2.5$ and $+3.5$	8
between $+3.5$ and $+4.5$	5
greater than $+4.5$	0

Table 4.1: Observed displacements of a spherical particle immersed in water. (Pathria and Beale 2011: p.591)

The data indicates that $\langle x^2 \rangle = 2.09 \times 10^{-8}$cm^2. This observationally calculated result is remarkably close to the value given by (4.22).

4.2   Some Comments on the Approach

The purpose of this section is to narrow in on the approach's key assumptions and to set aside aspects of the approach that may appear mysterious. Justifying these assumptions is central to underpinning the qualitative and quantitative facts the technique tracks.

4.2.1 Constructing a Langevin Equation

First, it may seem as though the Langevin equation came out of nowhere. Its form was motivated by aspects of ordinary experience. The discussion began by noting that experience teaches us that in ordinary cases the total force acting on the particle is dominated by a frictional force, $-\zeta V$, proportional to the velocity of the Brownian particle. It was then indicated that this couldn't be the whole story, since it's at odds with what we observe over longer periods of time. This led to adding a force to the frictional force that would account for the difference. We were thus led to (4.4), the Langevin equation for a Brownian particle. This kind of motivation is common.[4] It is, however, "top down". Since it's motivated by phenomenal considerations, one may worry about whether and how it connects to the microphysics. Happily, it can be constructed from a combination of physical and mathematical considerations.

Consider a similar but simpler system than the one discussed in the previous section, a one-dimensional system that's comprised of a relatively large Brownian particle with mass M, which is hit from both sides by molecules of mass m.[5] We'll also assume that $M \gg m$ and that collisions are elastic. The velocity of the Brownian particle before and after a single collision will be denoted by V and V', respectively, and the velocity of a molecule before and after a collision will be denoted by v and v', respectively. If we combine the equations for conservation of momentum and energy, then we can write the velocities after the collision in terms of the velocities before the collision. That is,

---

[4]See Reif (1965: Sec.15.5), Zwanzig (2001: Ch.1), Mazenko (2006: Ch.1), and Pathria and Beale (2011: Ch.15), for example.

[5]This system is also discussed by de Grooth (1999).

$$V' = \frac{M-m}{M+m}V + \frac{2m}{M+m}v \qquad (4.23)$$

and

$$v' = \frac{m-M}{M+m}v + \frac{2M}{M+m}V. \qquad (4.24)$$

Using the assumption that $M \gg m$, and the following approximations,

$$\frac{M-m}{M+m} \approx 1 - 2\frac{m}{M} + O\left(\left(\frac{m}{M}\right)^2\right), \qquad (4.25)$$

$$\frac{M}{M+m} \approx 1 - \frac{m}{M} + O\left(\left(\frac{m}{M}\right)^2\right), \text{ and} \qquad (4.26)$$

$$\frac{m}{M+m} \approx \frac{m}{M} + O\left(\left(\frac{m}{M}\right)^2\right), \qquad (4.27)$$

(4.23) can be written as

$$V' = \left(1 - \frac{2m}{M}\right)V + \frac{2m}{M}v. \qquad (4.28)$$

So the change in momentum of the Brownian particle due to a single collision is

$$\Delta P = 2mv - 2mV. \qquad (4.29)$$

And the momentum change of the Brownian particle due to N collisions is

$$\Delta P_N = 2m \sum_{i=0}^{N-1} v_i - 2m \sum_{i=0}^{N-1} V_i. \tag{4.30}$$

If we consider a time interval Δt that's small enough that the velocity of the Brownian particle does not change appreciably, but because $M \gg m$, we still have a large number of collisions, then the second sum in (4.30) can be approximated by $2mNV = 2mnV(t)\Delta t$, where $V(t)$ is the velocity of the Brownian particle at time t, and n is the mean number of collisions per second so that $N = n\Delta t$. So, we have,

$$\Delta P_N = 2m \sum_{i=0}^{n\Delta t-1} v_i - 2mnV(t)\Delta t. \tag{4.31}$$

Dividing both sides by Δt yields an expression for the time derivative of the Brownian particle's velocity.

$$M\frac{dV}{dt} = -\gamma V + F_s, \tag{4.32}$$

where the damping constant $\gamma = 2mn$ and

$$F_s = \frac{1}{\Delta t} \sum_{i=0}^{n\Delta t-1} 2mv_i. \tag{4.33}$$

F_s is a fluctuating force, if, like before, it's assumed to have the following properties:

$$\langle F_s(t) \rangle = 0, \tag{4.34}$$

and

$$\langle F_s(t)F_s(t') \rangle = 2B\delta(t - t'). \tag{4.35}$$

(4.34) entails that the incoming velocities of colliding molecules have an expectation value of zero. (4.35) entails that the incoming velocities of colliding molecules at distinct times are uncorrelated.

(4.32) has the same form as the Langevin equation. It, however, contains explicit expressions for both the damping force and the fluctuating force. Much like (4.4), (4.32) can be used together with (4.34), (4.35), and the equipartition theorem, to arrive at important results such as (4.22).

4.2.2   The Time-Reversal Non-Invariance of the Langevin Equation

Something else that may have seemed mysterious about the approach described in the previous section is the time-reversal non-invariance of the Langevin equation. The one-dimensional version, (4.32), is also asymmetric under time-reversal. This too may seem mysterious. Each equation describes the evolution of a system whose underlying dynamics are symmetric under time-reversal.

To see that (4.32) is asymmetric under time-reversal, first note that the expectation value of the fluctuating force is zero. Then, taking the expectation values of both sides of (4.32), we get the conclusion that the expectation value of V decays exponentially to zero. The temporal reverse of this would have an exponentially increasing expectation value of V. This reasoning also applies to the Langevin equation. Again, the expectation value of the fluctuating force is taken to be zero. Since the frictional force term is temporally asymmetric, we are led, when we take the expectation values of both sides of (4.4), to the conclusion that the expectation value of V decays exponentially

to zero. And similarly, the temporal reverse would have an exponentially increasing expectation value of V.

The time-reversal non-invariance of the Langevin equation is the result of the temporal asymmetry of its frictional force term and the properties attributed to its second, fluctuating force term. The same is true of (4.32). Interestingly, both equations, understood simply as equations that divide the total force acting on the Brownian particle in two, are compatible with time-reversal invariance. What's more, this compatibility holds even after the choice is made to split the total force into a frictional force and what is left over. What ensures the time-reversal non-invariance of these equations, given the time-reversal non-invariance of their respective frictional force terms, are the properties attributed to their second terms.

The time-reversal non-invariance of the Langevin equation prompts further questions. First, why do physicists standardly assume that the Langevin equation's second term has the properties identified in (4.5) and (4.6)? And second, what does their rationale reveal about the microphysics of those systems whose behaviour is accurately modelled by the Langevin equation, and the equations derived from it (e.g. (4.12) and (4.19))? Otherwise said, what would have to be true, or at least be approximately true, at the microlevel to account for the success of equations such as (4.4), (4.12), and (4.19), given that the underlying dynamics is symmetric under time-reversal?

First, it's standardly assumed that the second force term has an expectation value of zero because the Brownian particle, observed at ordinary scales, seems just as likely to move in any direction at any moment as any other. In the one-dimensional model, the analogous thought is that the Brownian

particle seems just as likely to move, at any moment, to the left as it is to the right. What's more, if observation provided us with reason to think that the Brownian particle was more likely to move in one direction than another—say, because we thought it was more likely to be struck from one side rather than another—then we would describe this thought using some force expression and extract it from the second term.

Interestingly, (4.5) is standardly presented alongside an assumption analogous to (4.6).[6] Notice too that (4.34) was presented along with (4.35). Both (4.6) and its analogue, (4.35), say that no correlation exists between the forces acting on the Brownian particle at any two distinct times.[7]

Importantly, what standardly underlies and motivates both (4.5) and (4.6), or, analogously, (4.34) and (4.35), is something that closely resembles the following microphysical assumption.

Collision Assumption: That v is, at any time, independent of V.

Spelled out, the underlying claim is that the velocity of any incoming colliding fluid molecule (drawn from the distribution of velocities of molecules of the fluid), is, at any time, probabilistically independent of the incoming velocity of the Brownian particle.

Interestingly, the collision assumption is temporally asymmetric. To see this, consider the one-dimensional model. Next, note that the collision as-

---

[6]See Zwanzig (2001: p.5), Kadanoff (2000: p.120), and Mazenko (2006: p.8), for example.

[7]It's worth noting that concerns about the conflict between the time-reversal non-invariance of the Langevin equation and the time-reversal invariance of the underlying dynamics do not standardly play a role in the determination of properties attributed to the Langevin equation's second term.

sumption has to do with the probability distribution of the velocities of molecules that are about to hit the Brownian particle. Now consider (4.24). In particular, consider the probability distribution of v'. That is, the probability distribution of the velocities of molecules that have just collided with the Brownian particle. If the expectation value of v is zero, as (4.34) entails, then the expectation value of v' is proportional to V. So, if V is greater than its equilibrium value, then, on average, colliding molecules that are in the path of the Brownian particle's motion will have a greater velocity after the collision than before. Meanwhile, colliding molecules that are travelling on the Brownian particle's path will, on average, have a post-collision velocity that's less than their pre-collision velocity. When we velocity-reverse the situation, we get a state of affairs in which the distribution of velocities of incoming colliding molecules are, contrary to the assumption, not independent of the Brownian particle's velocity. They appear conspiratorial. The ones travelling on its path will, on average, have a pre-collision velocity that's greater than the average molecule and the ones traveling in its path will, on average, have a pre-collision velocity that's less than average.

4.2.3   Why Does the Langevin Equation Work?

Since (4.4), (4.12), and, in particular, (4.19) have consequences that are confirmed by experiment, it's worth considering what would have to be true, or be at least approximately true, at the microlevel, to justify the use of the collision assumption, and to ensure that its consequences—(4.5) and (4.6)—are at least approximately true. Answering this helps explain the success of equations such as (4.4), (4.12), and (4.19). It also helps to underpin the

quantitative and qualitative facts these equations track. As one might have expected, the answer is suggested by the assumption itself. The natural suggestion is that it's a convenient approximation of the following microphysical fact.

> **Microphysical Fact:** That v is at most independent of V at some initial time and that at any other time v is *effectively* independent of V.

The claim here is that, at all times (except, perhaps, initially), the velocity of any incoming colliding fluid molecule and the incoming velocity of the Brownian particle are approximately probabilistically independent.[8]

Notice that we don't claim that the velocities of incoming colliding molecules are, at all times, *in fact* independent of the incoming velocity of the Brownian particle, as the collision assumption has it. The reason for this more modest claim is that these velocities are at least correlated for very short times after collision. This can be seen by reflecting on what we think's going on at the microlevel. Once the Brownian particle is introduced to the fluid, it begins to collide with fluid molecules. The subsequent positions and momenta of the molecules, like the Brownian particle, are the result of collisions. Because the forces acting on the particle can be expressed by smooth functions—at least to a good approximation—the forces resulting from these collisions will be correlated.

But it should be remembered that the motion of the Brownian particle appears to be random on macroscopic time scales. So it seems reasonable to

---

[8]The idea then is that, for practical purposes, one can treat the velocities of these colliding particles *as if* they are probabilistically independent. That is, one may be justified in using the collision assumption.

think, as the microphysical fact intends to suggest, that any correlations that form between the velocities of colliding particles wash out incredibly quickly. Happily, weakening (4.6) (or (4.35)) to accommodate these considerations—say by instead assuming that correlations between fluctuating forces at distinct times exponentially decay on scales that are incredibly short compared to the system's relaxation to equilibrium—yields results that closely approximate (4.19) and (4.13). Similar results also hold if we assume that correlations between fluctuating forces at distinct times are Gaussian on scales that are extremely short compared to the system's relaxation to equilibrium.

Having identified the approach's crucial assumption, the collision assumption, it's important to ask whether its use is justified. Or, more crucially, what reason do we have to think that the microphysical fact is, in fact, true? This question is important, since we can hardly ground the success of the technique, or the facts its resulting equations track, on an unsupported microphysical claim.

4.3 Motivating the Collision Assumption

The microphysical fact, if true, provides reason to utilise the collision assumption. Support for the microphysical fact comes from at least two sources: from a scientifically informed reflection on the microphysical behaviour of relevant systems, and from results that emerge from the study of an idealised system, a hard sphere gas.

First, when thinking about the behaviour of a Brownian particle that's been placed into an isolated homogeneous fluid at equilibrium, it seems rea-

sonable to have the following picture in mind. Once the Brownian particle is placed into the fluid, it begins to collide with fluid molecules as it approaches equilibrium. Since the fluid is composed of a very large number of molecules, which are involved in very many collisions, it's reasonable to think that whenever two of them collide it's overwhelmingly likely that a very large number of collisions will occur before they collide again—if, in fact, they ever collide again. This suggests that by the time the Brownian particle were to collide with some molecule it had already encountered neither would, in effect, carry a memory of their past collision. That is, any correlation that resulted from their previous interaction would have effectively been washed out. Because these collisions happen incredibly quickly, so too would the washing out of correlations. So then, we can reasonably expect the forces acting on the Brownian particle to be, at any moment, effectively random, and we can expect the velocities of incoming colliding molecules to be, at any moment, effectively independent of the velocity of the Brownian particle. The reasonableness of this picture supports the truth of the microphysical fact.

Notice too how bizarre the temporal reverse of a situation in which a Brownian particle slowed to equilibrium would be. That is, a situation in which the microphysical fact did not hold. If we reversed the velocities of all of the molecules that comprise the system, then we would witness a conspiratorial situation similar to the one described at the end of the last section. The Brownian particle would miraculously speed up and the velocities of the incoming colliding molecules would be correlated with it. The molecules that are about to hit the Brownian particle from behind would, on average, have a higher than average velocity, and the ones that are in its path would, on

average, have a lower than average velocity.

What makes the first picture reasonable and the second situation bizarre is that only in the second situation are the velocities of incoming colliding molecules correlated *before* they interact. What's peculiar about this is that in the absence of some earlier event that would account for this correlation, their correlation is to be explained by some event that lies in their future. And this flies in the face of the fundamental idea that causes precede their effects.[9]

If we assume that the microphysical fact holds (or that something close to it holds), then we are led to equations that accurately model and predict the behaviour of systems we can and do observe. The models outlined in earlier sections attest to this fact. So we have support for the truth of the microphysical fact. As the discussions of the temporal reversals of the ordinary behaviour of these systems suggest, deviations from such an assumption lead us to expect systems to behave in ways we simply do not ever witness. So we have further support for the truth of the microphysical fact.

The study of idealised models also supports the truth of the microphysical fact. Oscar Lanford III has shown that an instance of the microphysical fact is true, under certain conditions, for a hard sphere gas model.[10] Lanford has shown that when the initial state of the gas satisfies a certain condition—that, in effect, amounts to there being a lack of correlation between the momenta

---

[9]This claim might be controversial to some. Huw Price (1996), for example, disputes its truth at the microphysical level. The following things, however, are not controversial. First, we take this idea for granted in our ordinary dealings with the world. That it's universally true is the default position. Second, even on reflection, it's not disputed at the macroscopic level. Third, scientists standardly develop successful physical models on the basis that causes precede their effects.

[10]See Lanford III (1975, 1976, & 1981).

and positions of gas molecules—the system effectively sustains this property, as it approaches equilibrium.[11] The result, which follows from Lanford's theorem, applies to systems whose behaviour is described by Boltzmann's equation. These systems also have an underlying dynamics that is invariant under time-reversal. While the result holds only for a short period of time, there are good reasons to think that it can be extended.[12]

On the basis of these considerations, it seems reasonable to think that the microphysical fact is true. Its truth would support the use of the collision assumption and its consequences, (4.5) and (4.6) ((4.35) and (4.35)). In particular, it would justify the use of (4.5) and (4.6) in the derivations of (4.12) and (4.19). Its truth would also help explain the success of these equations, and help underpin the quantitative and qualitative facts these equations track.

4.4   A Summary and a Suggestion

The last chapter suggested pursuing a particular line of research. More accurately, it promoted turning some of our attention towards understanding and grounding the success of the techniques and equations physicists use to model the behaviour of systems that begin away from equilibrium. This chapter took steps in this direction by outlining, examining, and attempting to ground the success of a technique and equation used to model the behaviour of a Brownian particle that's been immersed in a homogeneous

---

[11]See Uffink (2007: Sec.6.4), Uffink and Valente (2015), and Valente (2014) for discussions of Lanford's results.

[12]For a discussion of its extension, see Valente (2014: Sec.7.2) and Lanford III (1976: p.14).

fluid. This chapter highlighted that the technique generates a collection of interesting and quantitatively accurate equations whose predictions are confirmed by experiment. This chapter attempted to account for the success of the model, by identifying and motivating the technique's key assumptions. It noted that support for them can be traced back to an endorsement of the collision assumption, which is a convenient approximation of the microphysical fact. The body of the chapter ended with several reasons to think that the microphysical fact is, in fact, true.

Of course, the technique and equation discussed in this chapter are but the first in a collection of more elaborate and predictively accurate techniques and equations.[13] There are also limitations on the model's applicability—despite its generality. Naturally, it would be good for future investigations to consider and examine more elaborate techniques and equations, and for them to also examine more complicated systems (e.g. systems that are influenced by external forces). It would be good for these investigations to uncover and discuss the presuppositions these techniques and equations trade on (if any), and, where possible, to discuss the grounds we have for believing that these conditions are satisfied.

A more concrete suggestion is to examine and motivate the steps that lead to the Fokker-Plank equation—a probabilistic version of the Langevin equation. The Fokker-Plank equation is the starting point for many useful calculations in nonequilibrium statistical mechanics. For example, it's used to determine the rate at which a Brownian particle crosses a potential barrier (Zwanzig 2001: p.40). The Fokker-Plank equation is often called upon

---

[13]See, for example, Zwanzig (2001) and Mazenko (2006).

in situations in which it's suitable to model a system's behaviour using a *non-linear* Langevin equation. Linear Langevin equations are easy to solve analytically, non-linear Langevin equations are not. Physicists standardly side-step this difficulty by constructing a Fokker-Plank equation that corresponds to a given Langevin equation. It would be good to highlight whatever presuppositions are required to make these steps legitimate, and to discuss the grounds we have for believing that these conditions are met.

More obviously though, it would be good for future investigations to return to Boltzmann's equation (and the many versions of it) and to examine and motivate the steps that lead to it. Boltzmann's equation accurately and quantitatively models the behaviour of dilute gases that begin away from equilibrium. This fact is often lost amidst talk of its problems.[14] As the discussion of Boltzmann's equation in Section 3.1 and the underpinning of the Langevin equation offered in Section 4.3 suggest, this research project would involve an attempt to motivate an assumption similar to the *Stoßzahlansatz*.

The successful grounding of these, and other techniques and equations, would likely underpin many of the important and interesting quantitative facts that typicality accounts cannot. Accounting for the success of these techniques and equations would also lead, in a mosaic fashion, to an underpinning of the very concepts that are currently at the centre of foundational discussions of statistical mechanics.

---

[14]See also Wallace (2013: p.14).

Chapter 5

Conclusion

This thesis commented on aspects of foundational debates that arise in connection with statistical mechanics and thermodynamics.

One of its overarching goals was to promote a particular line of research: understanding and accounting for the success of the techniques and equations physicists use to model the behaviour of systems that begin away from equilibrium.

This thesis helped explain what philosophers mean when they say that an aim of nonequilibrium statistical mechanics is to account for certain aspects of thermodynamics, by making the issue of reconciling the existence of thermodynamically irreversible processes with underlying reversible dynamics more clear. This contributed to the overarching goal of promoting a particular line of research by making clear the goals philosophers have traditionally set themselves. This enabled us, in later chapters, to better appreciate and evaluate their attempts to achieve them, and to better situate and appreciate how the proposed line of research fits with the contemporary

literature.

As was noted, many of the leading attempts to reconcile the existence of thermodynamically irreversible processes with underlying reversible dynamics proceed by way of, and are typically part and parcel of, discussions that attempt to underpin the following qualitative facts: (i) that isolated macroscopic systems that begin away from equilibrium spontaneously approach equilibrium, and (ii) that they remain in equilibrium for incredibly long periods of time. As we saw, these attempts appeal to phase space considerations and notions of typicality. This thesis examined leading typicality accounts and highlighted their severe limitations. For one thing, they do not underpin a multitude of interesting quantitative facts that arise in connection with the behaviour of systems that begin away from equilibrium. To remedy this and other shortfalls, this thesis promoted working on the alternative project, since accounting for the success of the techniques and equations physicists use to model the behaviour of systems that begin away from equilibrium helps underpin not just the qualitative facts mentioned above, which the literature has focused on, but also many of the important quantitative facts that typicality accounts cannot.

This thesis also took steps in this promising direction. The fourth chapter outlined, examined, and attempted to ground the success of a technique commonly used to model the approach to equilibrium and subsequent behaviour of a Brownian particle that's been introduced to an isolated homogeneous fluid at equilibrium. It attempted to account for the success of the model, by identifying and grounding the technique's key assumptions. The chapter ended by offering some suggestions about where this research may be taken

next. It suggested examining more elaborate techniques and to ground their success along with the predictively accurate equations they help generate.

Bibliography

Albert, David Z. 2000. *Time and Chance*. Harvard University Press.

Berkovitz, Joseph, Roman Frigg, and Fred Kronz. 2006. "The Ergodic Hierarchy, Randomness and Hamiltonian Chaos". *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 37: pp. 661–691.

Birkhoff, G. D. 1931a. "Proof of a Recurrence Theorem for Strong Transitive Systems". *Proceedings of the National Academy of Sciences* 17: pp. 650–655.

———. 1931b. "Proof of the Ergodic Theorem". *Proceedings of the National Academy of Sciences* 17: pp. 656–660.

Blackmore, John. 1999. "Boltzmann and Epistemology". *Synthese* 119: pp. 157–189.

Boltzmann, Ludwig. 1872. "Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen". *Wiener Berichte* 66: pp. 275–370.

———. 1877. "Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung resp.

den Sätzen über das Wärmegleichgewicht." *Wiener Berichte* 76: pp. 373–435.

———. 1909. *Wissenschaftliche Abhandlungen,* vol. I, II, and III. Leipzig: Barth.

Boyling, J. B. 1972. "An Axiomatic Approach to Classical Thermodynamics". *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* 329: pp. 35–70.

Bricmont, J. 1995. "Science of Chaos or Chaos in Science?" *Annals of the New York Academy of Sciences* 775: pp. 131–175.

———. 2001. "Bayes, Boltzmann and Bohm: Probabilities in Physics". In *Chance in Physics*, edited by J. Bricmont, G. Ghirardi, D. Dürr, F. Petruccione, M-C. Galavotti, and N. Zanghì, vol. 574 of *Lecture Notes in Physics*. Springer Berlin Heidelberg, pp. 3–21.

Brown, Harvey, Wayne Myrvold, and Jos Uffink. 2009. "Boltzmann's H-theorem, its Discontents, and the Birth of Statistical Mechanics". *Studies in History and Philosophy of Science Part B* 40: pp. 174–191.

Brown, Harvey R., and Jos Uffink. 2001. "The Origins of Time-Asymmetry in Thermodynamics: The Minus First Law". *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 32: pp. 525–538.

Brush, S.G. 1986. *The Kind of Motion We Call Heat*, vol. 2. North-Holland Publishing Company.

Callender, Craig. 1999. "Reducing Thermodynamics to Statistical Mechanics: The Case of Entropy". *The Journal of Philosophy* 96: pp. 348–373.

———. 2001. "Taking Thermodynamics Too Seriously". *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 32: pp. 539–553.

———. 2010. "The Past Hypothesis Meets Gravity". In *Time, Chance and Reduction*, edited by Gerhard Ernst and Andreas Hüttemann. Cambridge University Press, pp. 34–58.

Carathéodory, C. 1909. "Untersuchungen über die Grundlagen der Thermodynamik". *Mathematische Annalen* 67: pp. 355–386.

Cercignani, C. 1998. *Ludwig Boltzmann. The Man who Trusted Atoms.* Oxford: Oxford University Press.

Clausius, Rudolf. 1856. "On a Modified Form of the Second Fundamental Theorem in the Mechanical Theory of Heat". *London, Edinburgh and Dublin Philosophical Magazine and Journal of Science.* Vol. XII: pp.81–98.

de Grooth, Bart G. 1999. "A Simple Model for Brownian Motion Leading to the Langevin Equation". *American Journal of Physics* 67: pp. 1248–1252.

de Regt, Henk W. 1996. "Philosophy and the Kinetic Theory of Gases". *The British Journal for the Philosophy of Science* 47: pp. 31–62.

Dorfman, Jay Robert. 1999. *An Introduction to Chaos in Nonequilibrium Statistical Mechanics*, vol. 14. Cambridge University Press.

Earman, John, and Miklós Rédei. 1996. "Why Ergodic Theory Does Not Explain the Success of Equilibrium Statistical Mechanics". *The British Journal for the Philosophy of Science* 47: pp. 63–78.

Ehrenfest, P., and T. Ehrenfest. 2002. *The Conceptual Foundations of the Statistical Approach in Mechanics*. Dover Publications.

Ehrenfest-Afanassjewa, T. 1925. "Zur Axiomatisierung des zweiten Hauptsatzes der Thermodynamik" 33: pp. 933–945.

Einstein, Albert. 1905. "On the Motion—Required by the Molecular Kinetic Theory of Heat—of Small Particules Suspended in a Stationary Liquid". *Annalen der Physik* 17: pp. 549–560.

Emch, Gérard, and Chuang Liu. 2002. *The Logic of Thermo-Statistical Physics*. Springer Berlin Heidelberg.

Fowler, Ralph, and Edward Guggenheim. 1939. *Statistical Thermodynamics*. Cambridge: Cambridge University Press.

Friedman, K. S. 1976. "A Partial Vindication of Ergodic Theory". *Philosophy of Science* 43: pp. 151–162.

Frigg, Roman. 2008a. "Chance in Boltzmannian Statistical Mechanics". *Philosophy of Science* 75: pp. 670–681.

———. 2008b. "A Field Guide to Recent Work on the Foundations of Statistical Mechanics." In *The Ashgate Companion to Contemporary Philosophy of Physics*, edited by Dean Rickles. Ashgate Publishing.

———. 2009. "Typicality and the Approach to Equilibrium in Boltzmannian Statistical Mechanics". *Philosophy of Science* 76: pp. 997–1008.

———. 2011. "Why Typicality Does Not Explain the Approach to Equilibrium". In *Probabilities, Causes and Propensities in Physics*, edited by M. Suárez, vol. 347. Springer Netherlands, pp. 77–93.

Frigg, Roman, Joseph Berkovitz, and Fred Kronz. 2014. "The Ergodic Hierarchy". In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, summer 2014 edn.

Frigg, Roman, and Charlotte Werndl. 2011. "Explaining Thermodynamic-Like Behavior in Terms of Epsilon-Ergodicity". *Philosophy of Science* 78: pp. 628–652.

———. 2012a. "Demystifying Typicality". *Philosophy of Science* 79: pp. 917–929.

———. 2012b. "A New Approach to the Approach to Equilibrium". In *Probability in Physics*, edited by Y. Ben-Menahem and M. Hemmo. The Frontiers Collection, Springer Berlin Heidelberg, pp. 99–113.

Giles, Robin. 1964. *Mathematical Foundations of Thermodynaics*. Oxford: Pergamon Press.

Goldstein, S. 2001. "Boltzmann's Approach to Statistical Mechanics". In *Chance in Physics*, edited by J. Bricmont, G. Ghirardi, D. Dürr, F. Petruccione, M-C. Galavotti, and N. Zanghì, vol. 574 of *Lecture Notes in Physics*. Springer Berlin Heidelberg, pp. 39–54.

Goldstein, S., and J. L. Lebowitz. 2004. "On the (Boltzmann) entropy of non-equilibrium systems". *Physica D: Nonlinear Phenomena* 193: pp. 53–66.

Gottwald, Georg A, and Marcel Oliver. 2009. "Boltzmann's Dilemma: An Introduction to Statistical Mechanics via the Kac Ring". *SIAM review* 51: 613–635.

Jauch, J. M. 1972. "On a New Foundation of Equilibrium Thermodynamics". *Foundations of Physics* 2: pp. 327–332.

———. 1975. "Analytical Thermodynamics. Part I. Thermostatics—General theory". *Foundations of Physics* 5: pp. 111–132.

Kac, Mark. 1959. *Probability and Related Topics in Physical Sciences*, vol. 1 of *American Mathematical Society*. Interscience Publishers.

Kadanoff, Leo. 2000. *Statistical Physics: Statics, Dynamics and Renormalization*. World Scientific.

Kardar, M. 2007. *Statistical Physics of Particles*. Cambridge University Press.

Kelvin. 1851. "On the Dynamical Theory of Heat, with Numerical Results Deduced from Mr Joule's Equivalent of a Thermal Unit, and M. Regnault's Observations on Steam". *Transactions of the Royal Society of Edinburgh XX* : pp 48–53.

Kestin, J. 1979. *A Course in Thermodynamics*, vol. Vol. I. New York: McGraw-Hill.

Klein, M. J. 1973. "The Development of Boltzmann's Statistical Ideas". In *The Boltzmann Equation: Theory and Applications*, edited by D. Cohen and W. Thirring. Vienna: Springer, pp. 53–106.

Lanford III, O. E. 1975. "Time Evolution of Large Classical Systems". In *Dynamical Systems, Theory and Applications*, edited by J. Moser, vol. 38 of *Lecture Notes in Physics*. Springer Berlin Heidelberg, pp. 1–111.

———. 1976. "On a Derivation of the Boltzmann Equation". In *Nonequilibrium Phenomena I: The Boltzmann Equation*, edited by J. L. Lebowitz and E. W. Montroll, vol. X of *Studies In Statistical Mechanics*. North-Holland Publishing Company, pp. 1–17.

———. 1981. "The Hard Sphere Gas in the Boltzmann-Grad Limit". *Physica A: Statistical Mechanics and its Applications* 106: pp. 70–76.

Langevin, Paul. 1908. "Sur la Théorie du Mouvement Brownien". *Comptes Rendus de l'Acadmie des Sciences* 146: pp. 530–533.

Lavis, David A. 2005. "Boltzmann and Gibbs: An Attempted Reconciliation". *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 36: pp. 245–273.

———. 2008. "Boltzmann, Gibbs, and the Concept of Equilibrium". *Philosophy of Science* 75: pp. 682–696.

Lazarovici, Dustin, and Paula Reichert. 2014. "Typicality, Irreversibility and the Status of Macroscopic Laws". http://philsci-archive.pitt.edu/10895/.

Lebowitz, J. L. 1993a. "Boltzmann's Entropy and Time's Arrow". *Physics Today* 46: pp. 32–38.

———. 1993b. "Macroscopic Laws, Microscopic Dynamics, Time's Arrow and Boltzmann's Entropy". *Physica A: Statistical Mechanics and its Applications* 194: pp. 1–27.

———. 1994. "Time's Arrow and Boltzmann's Entropy". In *Physical Orgins of Time Asymmetry*, edited by W. H. Zurzek J. J Halliwell, J. Pèrez-Mercador. Cambridge University Press.

———. 1999. "Statistical Mechanics: A Selective Review of Two Central Issues". *Reviews of Modern Physics* 71: pp. 346–357.

Lee, John Francis, Francis Weston Sears, and Donald Lawson Turcotte. 1973. *Statistical Mechanics.* 2nd edn. Addison Wesley.

Lieb, Elliott H., and Jakob Yngvason. 1999. "The Physics and Mathematics of the Second Law of Thermodynamics". *Physics Reports* 310: pp. 1–96.

Luczak, Joshua, and Wayne C. Myrvold. "Getting Clear About Absolute Continuity and Translation Continuity".

Malament, David B., and Sandy L. Zabell. 1980. "Why Gibbs Phase Averages Work–The Role of Ergodic Theory". *Philosophy of Science* 47: pp. 339–349.

Maxwell, J. C. 1871. *Theory of Heat.* Longmans, Green, and Company.

Mazenko, Gene. 2006. *Nonequilibrium Statistical Mechanics.* Wiley.

Mendoza, Eric, and Sadi Carnot. 1960. *Reflections on the Motive Power of Fire.* Dover Publications.

Myrvold, Wayne C. 2012. "Probabilities in Statistical Mechanics: What are they?" http://philsci-archive.pitt.edu/9236/.

———. Forthcoming. "Probabilities in Statistical Mechanics". In *Oxford Handbook of Probability and Philosophy*, edited by Christopher Hitchcock and Alan Hájek. Oxford University Press.

Pathria, Raj, and Paul Beale. 2011. *Statistical Mechanics.* 3rd edn. Academic Press.

Pauli, W. 1973. *Pauli Lectures on Physics*, vol. Volume 3. Thermodynamics and the Kinetic Theory of Gases. Cambridge, MA: The MIT Press.

Penrose, Roger. 1989. *The Emperor's New Mind.* Oxford University Press.

Pitowsky, Itamar. 2012. "Typicality and the Role of the Lebesgue Measure in Statistical Mechanics". In *Probability in Physics*, edited by Yemima Ben-Menahem and Meir Hemmo. Springer Berlin Heidelberg, pp. 41–58.

Price, Huw. 1996. *Time's Arrow and Archimedes' Point.* Oxford University Press.

Ramsey, Norman. 1956. "Thermodynamics and Statistical Mechanics at Negative Absolute Temperatures". *Physical Review* 103: pp. 20–28.

Reif, Frederick. 1965. *Fundamentals of Statistical and Thermal Physics.* New York: McGraw-Hill.

Ridderbos, Katinka. 2002. "The Coarse-Graining Approach to Statistical Mechanics: How Blissful is Our Ignorance?" *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 33: pp. 65–77.

Schroeder, D.V. 2000. *An Introduction to Thermal Physics.* Addison Wesley.

Schulman, L. S. 1997. *Time's Arrow and Quantum Measurement.* Cambridge, U.K.: Cambridge University Press.

Sklar, L. 1995. *Physics and Chance: Philosophical Issues in the Foundations of Statistical Mechanics.* Cambridge University Press.

Smoluchowski, Marian. 1906. "Zur kinetischen Theorie der Brownschen Molekularbewegung und der Suspensionen". *Annalen der Physik* 21: pp. 756–780.

Thompson, Colin J. 1972. *Mathematical Statistical Mechanics.* The Macmillan Company: New York.

Uffink, Jos. 1996. "Nought But Molecules in Motion". *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics* 27: pp. 373–387.

———. 2001. "Bluff Your Way in the Second Law of Thermodynamics". *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 32: pp. 305–394.

———. 2007. "Compendium of the Foundations of Classical Statistical

Physics". In *Philosophy of Phyics*, edited by J. Butterfield and J. Earman. Amsterdam: North Holland.

———. 2014. "Boltzmann's Work in Statistical Physics". In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, fall 2014 edn.

Uffink, Jos, and Giovanni Valente. 2015. "Lanford's Theorem and the Emergence of Irreversibility". *Foundations of Physics* 45: pp. 404–438.

Uhlenbeck, G. E., and G. W. Ford. 1963. *Lectures in Statistical Mechanics*. Providence: American Mathematical Society.

Valente, Giovanni. 2014. "The Approach Towards Equilibrium in Lanford's Theorem". *European Journal for Philosophy of Science* : pp. 1–27.

van Lith, J. 2001. "Ergodic Theory, Interpretations of Probability and the Foundations of Statistical Mechanics". *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 32: pp. 581–594.

Visser, Henk. 1999. "Boltzmann and Wittgenstein or How Pictures Became Linguistic". *Synthese* 119: pp. 135–156.

von Plato, J. 1994. *Creating Modern Probability: Its Mathematics, Physics and Philosophy in Historical Perspective*. Cambridge University Press.

Vranas, Peter B. M. 1998. "Epsilon-Ergodicity and the Success of Equilibrium Statistical Mechanics". *Philosophy of Science* 65: pp. 688–708.

Wallace, David. 2013. "What Statistical Mechanics Actually Does". http://philsci-archive.pitt.edu/9846/.

Werndl, Charlotte. 2013. "Justifying Typicality Measures of Boltzmannian Statistical Mechanics and Dynamical Systems". *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 44: pp. 470–479.

Werndl, Charlotte, and Roman Frigg. 2014. "Rethinking Boltzmannian Equilibrium". Forthcoming in: Philosophy of Science. Available at: http://philsci-archive.pitt.edu/10756/.

———. 2015. "Reconceptualising Equilibrium in Boltzmannian Statistical Mechanics and Characterising its Existence". *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 49: pp. 19–31.

Zanghì, N. 2005. "I Fondamenti Concettuali dell'Approccio Statistico in Fisica". In *La Natura Delle Cose. Introduzione ai Fundamenti e alla Filosofia della Fisica*, edited by F. Laudisa V. Allori, M. Dorato and N. Zanghí. Carocci: Roma.

Zwanzig, Robert. 2001. *Nonequilibrium Statistical Mechanics*. Oxford University Press.

Joshua Luczak

Education

Ph.D., Philosophy   2015
The University of Western Ontario

M.A. (First Class Honours), Philosophy   2010
Monash University

B.A.\B.Sc. (Hons.), Philosophy, Mathematics   2007
Monash University

Publications

2015. "Reflections on Lazare and Sadi Carnot." *Metascience*, Vol. 24, Issue 1: 87–89.

2013. "On Chance." *Metascience*, Vol. 22, Issue 1: 85–87.

Selected Presentations

"On How to Approach the Approach to Equilibrium."

15th Congress of Logic, Methodology, and Philosophy of Science. University of Helsinki, Finland. August 3-8, 2015.

"Toy Models: What They Are, What They're Not, and What They're Good For."

Second Irvine-Pittsburgh-Princeton Conference on the Mathematical and Conceptual Foundations of Physics. University of California, Irvine, USA. March 20-21, 2014.

17th Oxford Philosophy Graduate Conference. University of Oxford, UK. Nov. 15-17, 2013.

"Toy Models: Interesting? Yes. Approximations? Idealisations? No."

Models and Simulations 6. University of Notre Dame, USA. May 9-11, 2014.

Congress on Logic and Philosophy of Science. Ghent University, Belgium. Sept. 16-18, 2013.

Philosophy Graduate Student Association (PGSA) Colloquium, The University of Western Ontario, Canada, March 20, 2013.

"Fatalism, Bivalence, and You"

Canadian Philosophical Association Annual Meeting. Wilfrid Laurier University/University of Waterloo, Canada. May 27-30, 2012.

"A *Good* Defence of Abortion"

Centre on Values and Ethics (COVE) Graduate Student Conference. Carleton University, Canada. March 10-11, 2012.

Teaching

Philosophy Instructor

Georgetown University
Beginning Logic   Fall 2015
Contemporary Moral Issues   Fall 2015

Royal Melbourne Institute of Technology
Business Ethics   Spring 2010, Fall 2011

Singapore Institute of Management University
Business Ethics   Fall 2011

Philosophy Guest Lecturer

The University of Western Ontario
Introduction to Logic   Summer 2013

Monash University

Life, Death, and Morality	Summer 2011
Reason and Argument	Summer 2010

Philosophy Teaching Assistant[1]

The University of Western Ontario

Introduction to Philosophy	2013–2014F
Reasoning and Critical Thinking	2012–2013F
Advanced Introduction to Philosophy	2011–2012F

Monash University

Human Rights Theory	Fall 2011
Philosophy of Film	Summer 2010, Summer 2011
Time, Self and Mind	Spring 2010
Life, Death and Morality	Fall 2008, Fall 2009, Fall 2010
Philosophical Explorations of World Religions	Summer 2010
Time, Self and Freedom	Spring 2008, Spring 2009
God, Freedom and Evil	Spring 2009
Reason and Argument	Fall 2008, Fall 2009
Philosophy of Sex	Spring 2008

Philosophy Grader

Monash University

Ethics	Summer 2008, Summer 2009
Life, Death and Morality	Summer 2008
Time, Self and Freedom	Spring 2007
Primary Logic	Spring 2007

Mathematics Teaching Assistant

Monash University

Analysis of Change	Fall 2006, Fall 2007, Fall 2008
Techniques of Modelling	Spring 2006, Spring 2008
Linear Algebra with Applications	Fall 2007
Multivariable Calculus	Spring 2006

---

[1]F = Full year course.

Grants, Scholarships, and Awards

Munich Center for Mathematical Philosophy Visiting Fellowship, Ludwig-Maximilians-Universität	2016
Graduate Thesis Research Award, The University of Western Ontario	2015
Rotman Graduate Research Assistantship, The University of Western Ontario	2013, 2014, 2015
USC Teaching Honour Roll Award of Excellence, The University of Western Ontario	2012-13
British Society for the Philosophy of Science Travel Award, British Society for the Philosophy of Science	2013
Graduate Student Teaching Award Nominee, The University of Western Ontario	2012, 2013
Western Graduate Research Scholarship, The University of Western Ontario	2011–Present
Graduate Teaching Assistantship, The University of Western Ontario	2011–Present
Faculty of Arts and Humanities Chancellor's Entrance Scholarship, The University of Western Ontario	2011
Dean's Commendation of Excellence in Teaching, Monash University	2008, 2009, 2010
Dean's Sessional Teaching Commendation, Monash University	2009

School of Mathematical Science Scholarship, 2006
Monash University

Academic Service and Professional Affiliations

PhilPapers Area and Subject Editor, 2014–Present
Philosophy of Physical Science, Philosophy of Probability
Thermodynamics and Statistical Mechanics,
Models and Idealization, Probabilistic Reasoning
Philpapers.org

Centre for Digital Philosophy 2014–Present
Editorial and Administrative Assistant,
The University of Western Ontario

Rotman Institute of Philosophy 2012–Present
Student Member,
The University of Western Ontario

Rotman External Public Outreach 2013–Present
Committee Member,
The University of Western Ontario

Rotman K-12 Outreach 2013–Present
Committee Member,
The University of Western Ontario

Logic, Mathematics, and Physics Graduate 2012, 2013, 2014
Philosophy Conference Co-Organiser,
The University of Western Ontario

Logic, Mathematics, and Physics Graduate 2013, 2014
Philosophy Conference Reviewer,
The University of Western Ontario

Philosophy Graduate Student Association Mentor, 2012, 2013
The University of Western Ontario