# Typicality, Irreversibility and the Status of Macroscopic Laws

Dustin Lazarovici[*], Paula Reichert[†]

*Ludwig-Maximilians-Universität München*
*Mathematisches Institut*
*Theresienstr. 39, D-80333 Munich, Germany.*
Tel.: +49 (0)89 2180 4623

**Abstract**

We discuss Boltzmann's probabilistic explanation of the second law of thermodynamics providing a comprehensive presentation of what is called today the *typicality account*. Countering its misconception as an alternative explanation, we examine the relation between Boltzmann's H-theorem and the general typicality argument demonstrating the conceptual continuity between the two. We then discuss the philosophical dimensions of the concept of typicality and its relevance for scientific reasoning in general, in particular for understanding the reduction of macroscopic laws to microscopic laws. Finally, we reply to various common criticisms of the typicality account.

[*]lazarovici@math.lmu.de
[†]reichert@math.lmu.de

# Contents

# 1 Introduction

Over the last two decades, a series of papers by various distinguished mathematical physicists stressed the importance of the concept of *typicality* as a basis for probabilistic reasoning in physics, in particular as a basis for the explanation of the second law of thermodynamics in statistical mechanics (Lebowitz, 1993; Bricmont, 1995; Penrose, 1999; Goldstein, 2001). None of the authors took much credit for the ideas he presented, each of them rather stressed that he was recapturing or reformulating the groundbreaking insights of Ludwig Boltzmann who, more than one century ago, had shown how to explain and derive macroscopic regularities from the underlying laws governing the motion of the microscopic constituents of matter. However, reintroducing these ideas to physicists, mathematicians and philosophers proved to be highly necessary as their relevance is rarely appreciated today and the response to the papers of Lebowitz, Goldstein and others shows that they are still subject to widespread misconceptions and misunderstandings.

In this paper, we provide a comprehensive summary of the typicality account, spelling out some of the details and subtleties that have remained unspoken in the before-mentioned presentations, thus leaving room for objections and misunderstandings that we hope to eliminate (Section 2). We then demonstrate the conceptual continuity between this typicality account and Boltzmann's famous H-theorem, showing that it is false that they are often viewed as competing accounts of macroscopic irreversibility (Section 3). Putting things in wider perspective, we discuss the relevance of typicality for scientific reasoning in general, in particular for understanding the reduction of macroscopic laws to microscopic laws (Section 4). Finally, we address some of the most common objections against the typicality account that have been raised in the contemporary literature (Section 5).

# 2 The typicality account

## 2.1 Boltzmann's statistical mechanics

Our discussion is concerned with the explanation of the irreversible thermodynamic behavior of macroscopic systems. The term "thermodynamic behavior" thereby refers to the ubiquitous phenomenon that physical systems, prepared or created in a non-equilibrium state and then suitably isolated from the environment, tend to evolve to and then stay in a distinguished macroscopic configuration called the *equilibrium state*. Familiar examples are the spreading of a gas, the mixing of milk and coffee, the disappearance of temperature gradients, and so on.

Historically, this empirical regularity was captured by the *second law of thermodynamics*, positing the monotonous increase of a macroscopic variable of state called *entropy*, which attains its maximum value in equilibrium. The main task of *statistical mechanics* is to explain this macroscopic regularity on the basis of the underlying laws guiding the behavior of the system's micro-constituents.

A crucial ingredient to the understanding of this issue is the distinction between macro- and microstate of a system. Whereas the microstate $X(t)$ of a system is given by the complete specification of all its microscopic degrees of freedom, its macrostate $M(t)$ is specified in terms of physical variables that characterize the system on macroscopic scales (like its volume, pressure, temperature, and so on). The macroscopic state of a system is completely determined by its microscopic configuration, that is $M(t) = M(X(t))$, but one and the same macrostate can be realized by a large number of different microstates all of which "look macroscopically the same". The partitioning of the set of microstates into sets corresponding to macroscopically distinct states is therefore called a *coarse-graining*. Turning to the phase-space picture of Hamiltonian mechanics for an $N$-particle system, a microstate corresponds to one point $X = (q, p)$ in phase space $\Omega \cong \mathbb{R}^{3N} \times \mathbb{R}^{3N}$, $q = (q_1, q_2, ..., q_N)$ being the position- and $p = (p_1, p_2, ..., p_N)$ the momentum-coordinates of the $N$ particles, whereas a macrostate $M$ corresponds to an entire region $\Gamma_M \subseteq \Omega$ of phase space, namely the set of all microstates that realize $M$. The microscopic laws of motion are such that any initial microstate $X_0$ determines the complete microevolution $X(t) = \phi_t(X_0)$ of the system, represented by a unique trajectory in phase space going through $X_0$, thereby also determining its complete macro-evolution $M(X(t))$ as the microstate passes through different macro-regions.

These concepts are pretty much forced on us if we accept the supervenience of macroscopic facts on microscopic facts and they are essential to appreciating the nature of the problem. The second law of thermodynamics describes an empirical regularity about the *macro-evolution* $M(t)$ of a physical system. This macro-evolution, however, supervenes on the evolution of the system's microscopic configuration which is determined by precise and unambiguous laws of motion. The aspiration of statistical mechanics is thus to explain or justify the empirical regularity expressed in the macroscopic law on the basis of the underlying microscopic theory. This seems like a quite formidable task, though, as it requires us to reconcile the *irreversibility* of thermodynamic behavior with the *time-reversal symmetry* of the microscopic laws of motion. The task was nevertheless accomplished by Ludwig Boltzmann at the end of the 19th century. His account, we recall, was crucially based on two profound insights:

1. The identification of the (Clausius) entropy with the (logarithm of) the phase-space volume corresponding to its current macrostate. Formally:

$$S = k_B \ln |\Gamma_{M(X)}|, \tag{1}$$

   where $k_B$ is the Boltzmann constant and $|\Gamma_M|$ denotes the volume (the Lebesgue or Liouville measure) of the phase-space region $\Gamma_M$. The Boltzmann entropy is thus de facto a logarithmic measure of the phase-space volume corresponding to the system's macrostate.

2. The understanding that the *separation of scales* between the microscopic and the macroscopic level leads to enormous differences in the phase-space volume corresponding to states with different values of entropy. In particular, we will generally find that the equilibrium region – by definition the region of maximum entropy – is vastly larger than any other macro-region, so large, in fact, that it exhausts almost the entire phase space. In other words: almost every microstate *is* an equilibrium state.

The two points are actually related in the following sense: Note that the logarithm in equation (1) makes it that substantial differences in $S$ correspond to huge differences in the respective phase-space measure. And note that the separation of scales between micro- and macro-level, that is, in particular, the number of microscopic degrees of freedom in a macroscopic system, is characterized by Avogadro's constant which is of the order of $10^{23}$. And since the entropy associated with specific values of the macroscopic observables will in general grow with $N$ (the thermodynamic entropy is a so-called extensive variable of state), this means that the differences in phase-space volume corresponding to different entropy levels will depend exponentially on this already huge number. In other words, we will generally find that for macroscopic systems, i.e. for systems with a very large number of microscopic degrees of freedom, the partitioning of microstates into macrostates does not correspond to a partitioning of phase space into regions of roughly the same size, but into regions whose sizes vary by a great many orders of magnitude, with the equilibrium region being by far the largest.

These insights are the key ingredients in Boltzmann's account of the second law of thermodynamics. What we learn from them is, first and foremost, that the thermodynamic behavior that we want to explain is not a feature of certain *special* micro-evolutions, but rather the kind of macro-behavior that would correspond to almost any conceivable trajectory that the configuration of a macroscopic system, starting in a non-equilibrium region, could follow through phase space. Indeed, the dynamics of a system whose microscopic configuration starts out in a tiny non-equilibrium macro-region would have to be very peculiar to *avoid* carrying the microstate into larger and larger phase-space volumes – corresponding to gradually increasing entropy – and finally into the equilibrium region, where it will remain for the foreseeable future. This is why Boltzmann's arguments are extremely robust against the details of the microscopic theory, giving us an understanding of thermodynamic behavior as a virtually universal feature of macroscopic systems (see also the insightful remarks in Einstein's 'autobiographical notes', 1949).

On the other hand, recent publications have devoted many pages to pointing out the obvious fact that if we consider a system in an initial non-equilibrium microstate, we cannot plausibly conclude that it *must* evolve into equilibrium solely on the basis that the equilibrium-region is vastly larger than the non-equilibrium region (e.g. Uffink 2008; Frigg 2009, 2011, see also the discussion in Section 5). This is, as such, a correct observation though one cannot stress enough that, as a point of criticism, it doesn't pertain in any way to what the typicality account actually claims – or to what Boltzmann had claimed at least after 1876.

Indeed, we even know for a fact that, given a low-entropy macrostate $M$, there exist microscopic configurations realizing $M$ that will *not* evolve into equilibrium but follow a trajectory of *decreasing* entropy instead. And this is not, in the first place, due to any involved mathematical or philosophical argument, but a straightforward consequence of the *time-reversal symmetry* of the microscopic laws, as was famously pointed out by Johann Loschmidt in his "reversibility objection". Hence, Lebowitz rightly warned us, quoting Ruelle, that the ideas of Boltzmann are "at the same time simple and rather subtle" (1993, p. 7). We will elaborate on these subtleties in the following section.

### 2.2 The typicality account

To build on the basic principles of Boltzmann's statistical mechanics and go into the details of the typicality account, let's discuss the paradigmatic example of a gas in a box. We thus consider a system of $N \approx 10^{23}$ particles, interacting by a repelling short-range potential, which are confined to a finite volume within a box with reflecting walls. Now assume that we find or prepare the gas in the macrostate $M_2$ sketched below (Fig. 1), that is, we consider a particle configuration that looks, macroscopically, like a gas filling out about half of the accessible volume. What kind of macroscopic evolution should we predict for this system?

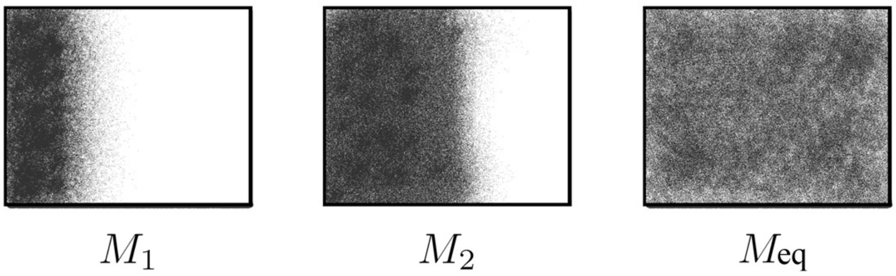


Figure 1: Thermodynamic evolution of a gas

Well, a simple combinatorial argument due to Boltzmann shows that *the overwhelming majority* of microstates that the system could possibly evolve into will look, macroscopically, like $M_{eq}$, i.e. like a gas that is homogeneously distributed over the entire volume of the box. In fact, one can readily conclude that the phase-space volume corresponding to this *equilibrium macrostate* $M_{eq}$ is about $2^N \approx 10^{10^{23}}$ times (!) larger than the phase-space volume occupied by configurations with substantially lower entropy (in agreement with our general reasoning above). Hence, as the particles move with different speeds in different directions, scattering from each other and occasionally from the walls, the system's microstate wanders around on an erratic path in the high-dimensional phase space and we would expect, by all reasonable means, that this path will soon end up in the equilibrium region $\Gamma_{M_{eq}}$ and then leave $\Gamma_{M_{eq}}$ only very rarely, corresponding to small fluctuations of the entropy about its maximal value. (Larger fluctuations, e.g. from $M_{eq}$ back into $M_2$, are *possible* as well, however, as Boltzmann

(1896b) already noted, the time-scales on which substantial fluctuations are to be expected are so astronomical that they have no empirical relevance.)

However, it is clear – and it was clear to Boltzmann – that there are initial conditions in $\Gamma_{M_2}$ for which the system will not exhibit this "expected" (we will later say *typical*) macro-behavior, but follow an anti-thermodynamic trajectory of decreasing entropy. For if we consider a macrostate of even lower entropy, $M_1$, the time-reversal symmetry of the equations of motions implies that for every solution corresponding to a macro-evolution from $M_1$ to $M_2$, there exists another solution carrying an initial microstate in $\Gamma_{M_2}$ into the lower-entropy macro-region $\Gamma_{M_1}$. (Indeed, we only have to take the solutions that have evolved from $\Gamma_{M_1}$ into $\Gamma_{M_2}$ and reverse the momenta of every single particle.) And yet, as Boltzmann understood, the microscopic initial conditions in $\Gamma_{M_2}$ that lead to such an anti-thermodynamic evolution are *extremely special ones* relative to all possible microstates realizing $M_2$. The correct assertion is thus that *almost all* initial configurations in $\Gamma_{M_2}$ will evolve into the equilibrium-region $\Gamma_{M_{eq}}$, while only a very small set of "bad" initial conditions will show the anti-thermodynamic evolution from $\Gamma_{M_2}$ into $\Gamma_{M_1}$. We will make these arguments more precise in a minute.

For now, let us note that it's more appropriate, in fact, not to consider any individual trajectory but the set of all solutions with initial condition in $\Gamma_{M_2}$. The dynamics of a system of about $N \approx 10^{23}$ particles is very chaotic, in the sense that even small variations in the initial configuration can lead to considerable differences in the time-evolution. Under the Hamiltonian dynamics, the set of microstates realizing $M_2$ at the initial time will thus quickly spread over phase space (respectively a submanifold compatible with the constants of motion) with the overwhelming majority of microstates ending up in the equilibrium-region and only a small fraction of special initial configurations evolving into the comparably tiny macro-regions of equal or lower entropy.

(A side-note on the last point: "chaos" is one of those notions that are notoriously difficult to exhaust by precise mathematical definitions, though it's clear that some form of dynamical instability is characteristic of thermodynamic systems with many degrees of freedom. Various attempts to capture this characteristic by rigorous – and usually very idealized – mathematical concepts, as well as the fruitfulness of some of these concepts for the respective fields of mathematics, have often created the impression that any one them in particular must play a central role in the foundations of statistical mechanics. However, as stressed before, the explanation of thermodynamic behavior is much more robust against the details of the microscopic model and doesn't hinge on any narrowly conceived property of the underlying dynamics. In particular, the relevant systems might easily fail to be ergodic, or mixing, or have everywhere positive Lyapunov exponents – to throw around some jargon –, and they probably do, though their overall behavior would have to be completely qualitatively different from what it's generally understood to be in order to render the typicality account irrelevant.)

To summarize our discussion, Boltzmann's analysis tells us that it cannot be true that *every* non-equilibrium configuration will follow the second law of thermodynamics and undergo an evolution of increasing entropy. The appropriate conclusion, however, is that *typical* microscopic configurations realizing a low-entropy initial macrostate will evolve into equilibrium and then stay in equilibrium for most of the time.

It should go without saying, though we have to emphasize nonetheless, that this typicality account is an explanation or an explanatory scheme – not a proof. As plausible as the conclusion may be, proving it in a rigorous fashion for any particular (reasonably complex) model remains an extremely difficult and largely unresolved problem in mathematical physics.

### 2.3 The measure of typicality

Throughout the above argument, the intuitive notions of *almost all* and *extremely special*, that we used synonymously to *typical/atypical*, were understood in terms of the stationary *Liouville measure*, i.e. in terms of the *phase-space volume* corresponding to the set of microstates that have the respective property. More precisely, for a perfectly isolated system with total energy $E$, we would have to consider instead of the Liouville measure the induced *microcanonical measure* $\mu_E$ on the hypersurface $\Gamma_E \subset \Omega$, to which the motion of the system is confined in virtue of the energy conservation. For simplicity, we will usually omit this distinction and merely refer to "phase space" and the "measure" or "size" of macro-regions.

In any case, a crucial property of the Liouville measure as well as the microcanonical measure is their *stationarity* under the microscopic time-evolution. Intuitively, this is to say that the Hamiltonian flow $\phi_t$ behaves like an incompressible fluid on phase space. Formally, it means that for all measurable sets $A \subseteq \Omega$ and all times $t \in \mathbb{R}$, we have $|\phi_t(A)| = |A|$. This is such an essential feature because it means that

a) the notion of typicality is *timeless*, i.e. a typicality statement does not depend on a reference to any external time-parameter.

b) the Hamiltonian dynamics "care about" the size of the macro-regions that play such a central role in the argument, in the sense that the stationary measure as a measure on *initial conditions* carries over to a well-defined measure on *solution trajectories*, which is such that the number of trajectories passing through a phase-space region at any given time is proportional to the size of that region.

Turning back to Boltzmann's explanation of the second law, we note that the Liouville measure (respectively the microcanonical measure) as a *typicality measure* serves two purposes in the argument:

1. To establish that the region of phase space corresponding to the macrostate $M_2$ is very much larger than the region of phase space corresponding to the macrostate $M_1$, and that the region of phase space corresponding to the equilibrium macrostate $M_{eq}$ is very much larger than the region of phase space corresponding to the macrostate $M_2$, so large, in fact, that it occupies almost the entire phase-space volume.

   It is easy to learn about this "dominance of the equilibrium state" (Frigg, 2009) and yet hard to appreciate what it is really saying, since the scale of the proportions expressed by the innocuous term "almost entirely" are beyond anything that we could intuitively grasp (just think of the ratio $10^{10^{23}} : 1$ for the gas-model).

   Note that together with the stationarity of the phase-space measure, the dominance of the equilibrium state already implies that (by far) most solution trajectories are in equilibrium (by far) most of the time – which is not quite what we need, but it's not too far away, either.

2. To define a notion of typicality *relative to the current macrostate of the system*, allowing us to assert, for instance, that *almost all* microstates in the non-equilibrium region $\Gamma_{M_2}$ will evolve into equilibrium.

   Regarding the meaning of "almost all", one should note that it's only in the idealized situation of a *thermodynamic limit* (where the number of microscopic degrees of freedom goes to infinity) that one can expect the exception set of "bad" configurations to be of measure *zero*, while if we argue about a realistic system, the *atypicality* of such configurations is substantiated by the fact that they have very very small (though positive) measure compared to that of all microstates realizing $M_2$.

In fact, stationarity of the Liouville measure allows us to estimate the measure of the good microstates relative to the bad microstates in $\Gamma_{M_2}$ by the ratio of phase-space volume occupied by $M_2$ to the phase-space volume corresponding to lower-entropy states. For let $\Gamma_{M_{low}}$ be the region of phase space corresponding to states of (substantially) lower entropy and let $B \subset \Gamma_{M_2}$ be the set of initial conditions in $\Gamma_{M_2}$ that will have evolved into $\Gamma_{M_{low}}$ after a time $\Delta t$, then $\Phi_{\Delta t}(B) \subseteq \Gamma_{M_{low}}$ and thus $|B| = |\Phi_{\Delta t}(B)| \leq |\Gamma_{M_{low}}|$, so that $|B| : |\Gamma_{M_2}| \approx |\Gamma_{M_{low}}| : |\Gamma_{M_2}| \approx 1 : 10^{10^{23}}$.

### 2.4 Irreversibility

By incorporating into our discussion what is essentially Boltzmann's answer to Loschmidt's reversibility objection, we have already seen how the typicality account solves the greatest challenge to our reductionist enterprise: the *prima facie* contradiction between the irreversibility of thermodynamic processes and the reversibility of the underlying mechanical laws. To emphasize how this apparent contradiction is resolved, we recall that it was crucial to the typicality argument that it referred to (typical or atypical) initial conditions *relative to the initial macrostate.* Of course, in terms of overall phase-space volume, a non-equilibrium macrostate occupies a vanishingly small fraction of phase space to begin with, corresponding (if you will) to a very low *a priori* probability. The relevant notion of typicality when discussing convergence to equilibrium from a non-equilibrium macrostate $M_2$ is thus defined by the phase-space measure conditioned on the fact that the initial microstate of the system is in the respective (low-entropy) region $\Gamma_{M_2}$.

Now, as we already observed, the time-symmetry of the microscopic laws is manifested in the fact that the phase-space volume occupied by the bad initial conditions in $\Gamma_{eq}$, for which the system will fluctuate out of equilibrium into the macrostate $M_2$ (let's say), is just as large as the phase-space volume occupied by the good initial conditions in $\Gamma_{M_2}$ for which the system will relax into equilibrium. In other words, over any given period of time, there are just as many solutions that evolve *into* equilibrium, as there are solutions evolving *out* of equilibrium into a lower entropy state – but the first case is nevertheless *typical* for systems in non-equilibrium, whereas the second case is *atypical* with respect to all possible equilibrium configurations in $\Gamma_{eq}$. It is this fact and this fact alone that establishes the irreversibility of thermodynamic behavior.

## 2.5 The Past Hypothesis and the thermodynamic arrow

By asserting that the origin of the thermodynamic asymmetry lies only in the specialness of the initial low-entropy macrostates, the typicality account is shifting the explanatory burden from why it is that a system in non-equilibrium relaxes to equilibrium (once macroscopic constraints are removed), to why it is that we find systems in such special states in the first place. Note that a typical configuration *simpliciter*, i.e. a typical configuration with respect to *all possible microstates*, is a state for which the system is in equilibrium, will be in equilibrium for most of its future, and has been in equilibrium for most of its past – which thus describes a time-symmetric situation.

Of course, as long as we are preoccupied with boxes of gas or melting ice-cubes or other confined systems, their low-entropy states will always be attributable to influences from outside, i.e. to the fact that these systems are actually part of some larger system (usually containing a physicist, or a freezer, or the like) from which they "branched off" to undergo a (more or less) autonomous evolution as (more or less) isolated subsystems. This presupposes, however, that those larger systems are out of equilibrium *themselves*, otherwise they couldn't give rise to subsystems with less then maximal entropy without violating the second law. Hence, if we think this through to the end, we arrive at the question why it is that we find *our universe* in such a very special state, far away from equilibrium (and how we justify our believe that its state was even more special the farther we go back in the past). This is what Goldstein calls the "hard part of the problem [of irreversibility]" (Goldstein, 2001, p. 49) and it concerns, broadly speaking, the origin of irreversibility and the thermodynamic arrow of time in our universe. Dealing with the "hard problem" would require us to discuss the meaning and the status of the *Past Hypothesis*[1] stipulating a very-low-entropy beginning of our universe. This issue, however, is beyond the scope of the present paper and we shall return to it only briefly in the course of our discussion.

[1] The term "Past Hypothesis" is due to Albert (2000), though the necessity of such an assumption was already noted by Boltzmann (1896a, pp. 252–53). See also Feynman (1967) and Carroll (2010) for a very nice discussion.

## 3 Typicality and the H-theorem

Although the formula engraved on Boltzmann's tombstone is equation (1), connecting the entropy of a microstate with the "probability" of the corresponding macrostate, his name is at least as intimately associated with the Boltzmann equation and the H-theorem, describing, in a more quantitative manner, convergence to equilibrium for a low-density gas. This H-theorem is of great interest in the light of our previous discussion, first, because it illustrates very clearly the need for a typicality argument and second, because it can be viewed as a concrete implementation of the general scheme that we've just presented as the "typicality account". By expanding on these points, we also want to counter two misconceptions regarding the H-theorem that seem to be wide-spread in the contemporary literature. These misconceptions have most likely arisen from Boltzmann's first presentation of the H-theorem but persisted despite his more refined argumentation in later writings. The first misconception is manifested in the charge that the explanation of macroscopic irreversibility in the context of the H-theorem is begging the question, because the derivation of the Boltzmann equation is grounded in an explicitly time-asymmetric assumption about the microscopic dynamics (see e.g. Uffink, 2008; Price, 1996, 2002).

The second, more basic misunderstanding is that the H-theorem and the typicality account are somehow *competing* accounts of macroscopic irreversibility and the convergence to equilibrium. Witness, for instance[2], Huw Price who writes with respect to the latter:

> *In essence, I think – although he himself does not present it in these terms – what Boltzmann offers is an alternative to his own famous H-Theorem. The H-theorem offers a dynamical argument that the entropy of a non-equilibrium system must increase over time, as a result of collisions between its constituent particles. [...] The statistical approach does away with this dynamical argument altogether.* (Price, 2002, p. 27)

We are convinced that the reason why Boltzmann did not present the "statistical approach" as an alternative to the H-theorem is that, in fact, it isn't. Understood correctly, there is a clear conceptual continuity between the H-theorem and the typicality account, so that the latter does not appear as

[2]Similarly, the pertinent entry in the Stanford Encyclopedia of Philosophy (Uffink, 2008) presents Boltzmann's work as a series of rather incoherent (and ultimately wanting) attempts to explain the second law.

a break with Boltzmann's earlier work but as a distillation of its essence (cf. Goldstein, 2012; Goldstein and Lebowitz, 2004). To make this case, we shall first review what the H-theorem is about and how it's grounded in the microscopic theory.[3]

### 3.1 Boltzmann's equation and the H-theorem

Recall that the microstate of an $N$-particle system is represented by a point $X = (q_1, ..., q_N; p_1, ..., p_N) \in \Omega$ in $6N$-dimensional phase space, comprising the position and momenta of all particles. The same state (modulo permutations of the particles) can also be represented as $N$ points in the 6-dimensional *$\mu$-space*, whose coordinates correspond to position and velocity of a *single particle*, i.e. $X \to \{(q_1, v_1), ..., (q_N, v_N)\}$, with $v_i := p_i/m$. The H-theorem is concerned with the evolution of a function $f_X(q, v)$ on $\mu$-space, that is supposed to provide an efficient description of the most important (macroscopic) characteristics of the gas in the microstate $X$. This function is defined as the *empirical distribution* or *coarse-grained density* of points in $\mu$-space. In principle, one can think of dividing $\mu$-space into little cells whose dimension is large enough to contain a great number of particles, yet very small compared to the resolution of macroscopic observations, and counting the number of particles in each cell. For fixed $q$ and $v$, $f_X(q, v)$ thus corresponds to the proportion of particles located *near* $q$ with velocity *approximately* $v$. In the limit where the size of the cells goes to zero, the empirical distribution becomes the actual distribution

$$f_{X(t)}(q, v) := \frac{1}{N} \sum_{i=1}^{N} \delta(q - q_i(t))\, \delta(v - \frac{1}{m} p_i(t)).$$

We give this formula to emphasize that, although $f_X(q, v)$ is technically a *probability measure*, there's is nothing *random* about it. In fact, it's more adequate to think of it as a *macroscopic variable*, determined, as it always is, by the microscopic configuration of the system. In particular, the distribution function does not describe a random system or an ensemble of systems but pertains to a *coarse-grained description* of an *individual* system, so that every microstate $X$ determines a unique $f_X(q, v)$, while many different microscopic configurations will coarse-grain to the same $\mu$-space density.

Now the first crucial result is that although the empirical distribution can be different for different microscopic configurations $X$, it is in fact (more

[3]For a good introduction, see, for instance, Davies (1977). For a detailed mathematical treatment, see Spohn (1991), Villani (2002), Lebowitz (1981).

or less) the same for an *overwhelming majority* of possible $X$. That is, one can show that for typical $X \in \Gamma$, the distribution function is of the form

$$f_X(q,v) \propto e^{-\frac{1}{2}m\beta v^2},$$

for some constant $\beta$ that is later identified with the inverse temperature of the system. This is the famous *Maxwell* or *Maxwell-Boltzmann distribution*, which is hence the *equilibrium distribution* of the gas. The distribution having no $q$-dependence means that the gas is homogeneously distributed over the entire volume with no correlations between position and velocities, i.e. with *uniform temperature.*

The goal of Boltzmann's famous H-theorem is thus to show the convergence of an initial *non-equilibrium* distribution $f_0(q,v)$ to the Maxwell-distribution $f_{eq}(q,v)$. The result is thereby based on three claims:

1) For a low-density gas, the time-evolution of $f_{X(t)}(q,v)$ is well described by an effective equation now known as the *Boltzmann equation.*

   Starting with an initial distribution $f_0(q,v) = f_{X_0}(q,v)$, it's important to distinguish the function $f_{X(t)}(q,v)$ – whose time-evolution is always determined by that of the microstate $X(t)$ – from the solution $f(t,q,v)$ of the Boltzmann equation with initial condition $f(0,q,v) = f_0(q,v)$ (respectively a smooth approximation thereof). The relevant claim is then that *for typical initial conditions*, $f_{X(t)}(q,v)$ will be (in a precisely specified way) *close* to $f(t,q,v)$ for a sufficiently long period of time, thus providing an *effective* description of the system's time-evolution.

2) For a solution $f(t,q,v)$ of the Boltzmann-equation, the *H-function*

   $$H(f(t,q,v)) := \int f(t,q,v)\log f(t,q,v)\mathrm{dqdv}$$

   is monotonously decreasing in $t$.[4]

3) The H-functional reaches its *minimum* for the Maxwell-distribution $f_{eq}(q,v)$.

   Together with 2) this implies, in particular, that the Maxwell-distribution is a *stationary* solution of the Boltzmann-equation.

Statements 2) and 3) are fairly standard mathematical results. The crux of the matter lies in statement 1). When Boltzmann first presented the H-theorem in 1872, he argued that a diluted gas *must* evolve in accord with

[4]While the "true" microscopic $H(f_{X(t)}(q,v))$ fluctuates and only decreases "on average".

his equation; he later had to mitigate this statement claiming, in effect, only that it would *typically* do so. Indeed, we will see that 1), and therefore the H-theorem, are genuinely typicality statements.

### 3.2 Molecular chaos

Boltzmann's derivation of what is now known as the Boltzmann equation is famously based on the *Stoßzahlansatz* or the assumption of *molecular chaos*.[5] This is an assumption about the *relative frequencies* of collisions between the particles in the gas. Denoting by $\mathcal{N}(t,q;v_1,v_2)$ the number of collisions happening near $q$ in a small time-interval around $t$ between particles with velocity (approximately) $v_1$ and $v_2$, the Stoßzahlansatz is:

$$\mathcal{N}(t,q\,;v_1,v_2) \propto N^2\, f(t,q,v_1) f(t,q,v_2)\, |v_1 - v_2|\; \mathrm{d}t\, \mathrm{d}q\, \mathrm{d}v_1 \mathrm{d}v_2, \tag{2}$$

i.e. the relative frequency of scattering events between particles of different velocities happening in the cell around $q$ is assumed to be proportional to the density of particles with the respective velocities near the respective position. The scattering probability being proportional to the product of $f(t,q,v_1)$ and $f(t,q,v_2)$ means that particles of different velocities are assumed to be *statistically independent* as they contribute to collisions. This is, more specifically, the meaning of *molecular chaos*.

Boltzmann's derivation, although a brilliant physical argument, was far from a rigorous proof. There are many mathematical subtleties involved in statement 1), concerning, for instance, the existence and uniqueness of solutions to the Boltzmann equation. However, if we can generously overlook these technical points, it is true that *if* and *as long as* the assumption of molecular chaos and equation (2) are valid, statement 1) is correct. Hence, we have to ask: What is the status of molecular chaos and how is it justified?

First and foremost, we have to keep in mind that there is nothing *random* about the interactions in a gas. Which particles are going to collide and how they are going to collide is completely determined by the initial conditions of the system and the microscopic laws of motion. For the purpose of illustration, let's imagine that we could freeze the system at time $t = 0$ and arrange the position and momentum of every single particle before letting the clock run and the system evolve according to the deterministic laws of Newtonian

[5] *Assumption*, unfortunately, is not a perfectly accurate translation of the German word *Ansatz*. Whereas the first is sometimes used synonymously with a logical *premise*, the later has a distinctly pragmatic element and can refer to something more akin to an "approximation" or a "working hypothesis".

mechanics. (Note that there is no issue here as to whether we let the clock run "forwards" or "backwards" – the problem is symmetric with respect to the time-evolution in both directions.) We could then, for instance, arrange the initial state in such a way that "slow" particles will almost exclusively scatter with other "slow" particles and "fast" particles with other "fast" particles. But such initial configurations are obviously very special ones. For *typical* microscopic configurations, coarse-graining to the initial distribution $f_0(q, v)$, we will find that the relative frequency with which particles of different velocities meet for the first collision is roughly proportional to the density of particles with the respective velocities near the respective position, i.e. given by eq. (2). This is nothing more and nothing less than the *law of large numbers*, based, in effect, on simple combinatorics. The validity of (2) *at the initial time* is thus, as all law-of-large-number statements, a *typicality statement* and as such another mathematical fact.

We observe here the fundamental difference between the probability density $f(t, q, v)$ and the typicality measure. The "scattering probability" at time $t$ is defined in terms of $f(t, q, v)$, though it's only for typical initial conditions that the relative frequency of scatterings is actually close to the expectation value. And typical initial conditions are defined, as usual, by the microcanonical measure restricted to the initial macro-region $\Gamma_0 := \{X \in \Gamma_E \mid f_X(q, v) = f_0(q, v)\}$.

This brings us, finally, to the critical part of the H-theorem. For assume that after an (infinitesimal) time-interval $\Delta t$ for which the validity of the Boltzmann-equation is established, the distribution function has evolved into $f(\Delta t, q, v)$. How do we know that (2) is still a good approximation for all but a small set of initial conditions? It is still true that eq. (2) is satisfied for typical microscopic configurations realizing the *current* distribution, i.e. counting all possible configurations that coarse-grain to $f(\Delta t, q, v)$. But we cannot count all these configurations, since the microstates relevant to our considerations are constraint by the condition that they have evolved from the macro-region corresponding to the initial distribution $f_0(q, v)$. Mathematically, these dynamical constraints on the "combinatorics" translate into the fact that the $\mu$-space coordinates of the particles at time $t > 0$ are no longer *statistically independent*, making it prima facie questionable whether a law-or-large-number statement for the relative frequencies of particle collisions, i.e. molecular chaos, still holds. This is, notably, the only meaningful way in which interactions *build up correlations* and it should be distinguished from more naive "causal" intuitions to the effect that the physical states of any two individual particles are independent before – but not after – they collide. Note, in particular, that the described situation is still completely symmetric between the time evolution towards the "future" and the time evolution towards the "past" of the distinguished initial state.

Boltzmann's *Stoßzahlansatz* now can be understood as the assumption that statistical independence is (sufficiently well) preserved by the microscopic time-evolution, or, in other words, that the relative frequency of collisions is always the typical one with respect to the current empirical distribution ($\approx$ the current macrostate) of the system. The mathematician refers to such a proposition as *propagation of molecular chaos*. Deriving the Boltzmann equation from a microscopic model in a rigorous mathematical sense is thus to validate this *ansatz*, i.e. to show that for typical initial conditions equation (2) remains *approximately* satisfied on sufficiently long time-scales. Sufficiently long, that is, to describe the thermodynamic evolution of a gas into equilibrium.

So, does molecular chaos propagate? That is, do the dynamics of a gas preserve statistical independence well enough to justify the *Stoßzahlansatz* over relevant time-scales? Based on physical intuition and various encouraging results, there is little doubt that the answer is affirmative. Given that the microscopic dynamics are very chaotic, that the number of particles in a gas is huge and the gas, by assumption, very diluted so that re-collisions (collisions between particles that have already collided in the past) are very rare, it is highly plausible that the relative frequency of collisions shouldn't become too special – in the sense of deviating significantly from the expectation value (2) – unless the initial configuration itself was very special. And yet, this is extremely difficult to *prove*; so difficult, in fact, that, as of to date, the best mathematical results available are valid only for very short times and a very restricted class of particle-interactions.[6]

Moreover, it is important to understand that, unless one considers the thermodynamic limit of an infinitely large system, equation (2) will hold at best *approximately* for *all but a small* set of "bad" initial conditions, that this approximation will get worse with time, and that the approximation is only good enough until it isn't. Eventually, a typical system will exhibit sizable fluctuations out of equilibrium at which point its evolution is no longer adequately described by the Boltzmann equation.

[6] See Lanford (1975) and King (1975) for the landmark results and Gallagher et al. (2012), Pulvirenti et al. (2013) for recent extensions to more general potentials.

### 3.3 The H-theorem as a typicality statement

With all that said, we can now appreciate the fact that the H-theorem is not an alternative way to explain convergence to equilibrium and the irreversibility of thermodynamic behavior, but rather a concrete exemplification of the explanatory scheme that we have presented before, in more general terms, as the typicality account. While the micro/macro distinction does not appear as prominently in the formulation of the H-theorem, it is essential that the empirical distribution $f(q, v)$ pertains to a *coarse-grained* description of the system, hence distinguishing a macro-region in phase space consisting of all microscopic configurations coarse-graining to the same $\mu$-space density. Convergence to equilibrium is then established for *typical initial conditions* with respect to that initial non-equilibrium macro-region. And the equilibrium state – characterized by the Maxwell-distribution to which a non-equilibrium distribution typically converges according to the H-theorem – is, as always, distinguished by the fact that it's the one realized by an *overwhelming majority* of all microscopic configurations. As Boltzmann himself beautifully explained:

> *The ensuing, most likely state [...] which we call that of the Maxwellian velocity-distribution, since it was Maxwell who first found the mathematical expression in a special case, is not an outstanding singular state, opposite to which there are infinitely many more non-Maxwellian velocity-distributions, but it is, to the contrary, distinguished by the fact that by far the largest number of all possible states have the characteristic properties of the Maxwellian distribution, and that compared to this number the amount of possible velocity-distributions that deviate significantly from Maxwell's is vanishingly small.* (Boltzmann, 1896a, p. 252, translation by the authors)

Despite the common focus on the *Stoßzahlansatz*, there is a compelling case to make that the tendency to equilibrium is by all means *explained* by the dominance of the equilibrium state. (Although it will not appear among the premises of the H-theorem, nor necessarily as an explicit part of the proof!) The explanatory role of the Stoßzahlansatz is then somewhat subsidiary to this insight, namely to express the fact that it's thus the "most likely" micro-evolutions that will carry a non-equilibrium distribution into equilibrium.

Finally, we understand that the *irreversibility* of the Boltzmann equation (as an effective description of a system's macro-evolution) is – as it cannot be otherwise – a consequence of the fact that non-equilibrium configurations converging to equilibrium are *typical* with respect to the corresponding "macrostate", whereas microscopic configurations leading to the time-reversed evolution are *atypical* with respect to all equilibrium configurations, i.e. all microstates coarse-graining to $f_{eq}(q, v)$. In particular, the origin of the asymmetry is, as always, the assumption of a non-equilibrium (and in this sense special) initial distribution.

As mentioned before, one will often encounter the claim that the irreversibility of the Boltzmann equation is a result of the *Stoßzahlansatz* being an explicitly time-asymmetric assumption. This is not correct. The assumption of molecular chaos breaks the time-symmetry only in the obvious (and necessary) sense that it applies to the thermodynamic evolution but not to the reversed motion; but this does not mean that any time-asymmetry is smuggled into the derivation of the H-theorem in addition to the one introduced by the assumption of a non-equilibrium initial distribution.

This misunderstanding, we believe, is mostly based on the failure to recognize molecular chaos, respectively the *Stoßzahlansatz*, as a *typicality statement*. For *typical* initial conditions, eq. (2) is equally valid for the time-evolution in *both temporal directions*. However, the microscopic configurations that have evolved from a state of lower entropy are *ipso facto* atypical with respect to their evolution in the reversed (past) time direction.

To put it differently, if the assumption of molecular chaos is justified in the sense explained before, it will hold for typical initial configurations realizing a non-equilibrium distribution, for which the H-theorem thus asserts convergence of the distribution function to a Maxwellian distribution (towards the future as well as towards the past) and it will also hold for typical equilibrium configurations, for which the H-theorem thus asserts that the equilibrium distribution is stationary. There is no reason, however, why it must hold for those equilibrium configurations that are the time-reversal of states that have just evolved from non-equilibrium, which are, after all, a vanishingly small subset of the equilibrium region. And we know, of course, that it doesn't, that those states are precisely contained in the set of bad configurations for which the particles are correlated in such a way as to undergo a macro-evolution of decreasing entropy (increasing H) that cannot be described by the Boltzmann equation. And we also know that the atypicality of these states (with respect to their evolution in one temporal direction) is explained by, or at least a necessary consequence of, the fact that the system is assumed or constrained or observed to be in a special (i.e.

non-equilibrium) state at one particular moment in time.

The only deeper question that may be left is why the Boltzmann equation is in fact *relevant*, i.e. why it is a good description of an *actual* gas in our *actual* world. To understand the answer to this question is thus to appreciate the meaning and relevance of typicality statements in general.

# 4 Typicality and the status of macroscopic laws

## 4.1 The 'logic' of typicality statements

One of the hurdles that may have stood in the way of appreciating Boltzmann's contribution and the relevance of typicality is the fact that Nagelian schemes of reduction and the related *deductive-nomological models* of physical explanation did not quite capture the subtleties of Boltzmann's arguments.[7] According to these often criticized yet very persistent theories, a microscopic explanation of the second law of thermodynamics – respectively a reduction by the microscopic theory – must be a *derivation* of the macroscopic law from the microscopic laws plus suitably specified "auxiliary assumptions" or "circumstances" in which the macroscopic law is supposed to hold. There is a sense in which this characterization is correct, although to get a grip on what sense this is, we'll have to say more about what we mean by "derive" and by a "macroscopic law". First, we want to emphasize one of the more problematic aspects of this view, which is that an understanding of the relationship between the macroscopic regularity and the underlying microscopic laws in purely logical terms misses the crucial role that *initial conditions* play in the explanation of a macroscopic phenomenon.

For what is it to *derive* the thermodynamic behavior of, let's say, a gas from the microscopic laws of motion? Is it to show that there exists at least one microscopic configuration for which the gas will relax to equilibrium? Is it to show that it will happen for *all* possible (non-equilibrium) configurations? The insufficiency of the first statement and the falsehood of the second must severely question the adequacy of purely deductive schemes of explanation. For suppose we wanted to account for the thermodynamic behavior of a certain type of physical system by a scheme of the form $\forall x\big(F(x) \Rightarrow G(x)\big)$, where $x$ ranges of all possible realizations of the corresponding microscopic model and the predicate $G$ is a suitable formulation

[7] See Dizadji-Bahmani et al. (2010) for a recent defense of Nagelian reduction. On typicality, see, e.g., Maudlin (2007), Bricmont (1995), Dürr (2009), Goldstein (2012), and Zanghì (2005).

of "showing effectively/approximately thermodynamic behavior". Then the antecedent $F(x)$ would have to contain a clause more or less equivalent to the statement "The initial conditions of the system $x$ are such that $G(x)$". But then the deduction becomes too trivial to be relevant. *Of course* there exist initial conditions for which the gas will expand. There are also initial conditions for which the gas will contract. And initial conditions for which the gas will transform into a banana. In other words, for a system $x$ with sufficiently many degrees of freedom and sufficiently non-trivial dynamics it is practically *always* possible to maintain that it has the (macroscopic) property $G$ because the initial conditions were such that $G(x)$. The only thing that can provide explanatory value in this context is the assertion of *typicality*, i.e. the assertion that $G$ is not a feature of certain *special* initial conditions, but a physical fact that would arise from *almost any* initial condition. This is also to assure that the explanatory work is done, as much as possible, by the fundamental physical *laws*, rather than by some fine-tuned arrangement of microscopic degrees of freedom.

Note however that the relevant statement is now, logically and syntactically, a proposition about $G$ rather than a proposition about any particular $x$. The 'logic' of the statistical explanation of the second law is thus not to state a set of (statistical) assumptions about an *individual* system from which to infer its thermodynamic behavior, but to spell out a physical account that *grounds* the explanation of thermodynamic behavior in the notion of typicality.

### 4.2 Typicality vs. probability

Indisputably, the common way of speaking is not to assert that a macroscopic feature $G$ is typical, but to say $G(x)$ is *very likely* or that we infer $G(x)$ *with high probability.* Such a probabilistic statement must, however, raise two additional questions: a) what is it supposed to mean? and b) how do we justify a probabilistic result on the basis of deterministic microscopic laws? We cannot discuss here in detail how the different "interpretations" of the concept of probability (subjectivist, frequentist, Humean best system, etc.) fare in this context, but we want to emphasize a few general points to capture the intricacy of the issue.

First, it would seem rather odd (and detached from scientific practice) if, in order to account for the second law of thermodynamics, we would have to add to the mechanical laws a quantitative assumption about the distribution of initial conditions of boxes of gas, or the like, that we find in our universe.

Second, the fact that we are generally ignorant about the exact microstate of a system is true, but largely irrelevant. It is absurd to think that the validity of the second law of thermodynamics could in any way depend on what we know or believe or are able to observe.

Finally, if we are serious about our commitment to argue within the paradigm of a particular deterministic theory, we have to take it to the conclusion that there is nothing more random about the physical processes that give rise to subsystems in non-equilibrium states than about the entropy-increasing processes going on within these subsystems, once they are suitably isolated from their environment. Eventually one has to wonder why it is true *as a matter of fact* that whenever someone prepares a gas in a low-entropy state, it never ends up in one of the "bad" microscopic configurations for which the gas would contract rather than expand. And then one has to take seriously the fact that an act of "preparation" is itself a physical process, following the same set of physical laws, with its outcome determined by suitably specified initial conditions. Why are *these* initial conditions always good ones, then? To defer the source of randomness to the outside, from the box of gas to the shaky hands of the experimentalist or to exterior perturbations preventing the subsystem from being perfectly isolated, is just to pass the buck. But the buck must stop, eventually, with the universe itself. For the universe is what it is, it exists once and only once, there is nothing before and nothing outside. And we either live in a universe that obeys the second law of thermodynamics (on cosmological scales and, with the possibility of very rare exceptions, in its branching sub-systems) or we don't.

All that said, what is the difference between a statement of probability and a typicality statement, and why is typicality the more appropriate concept in this context?

For one thing, contrary to the conventional use of probabilities, typicality is not a *quantitative* concept. The role of the typicality measure is only to realize and give precise meaning to the notion of "almost all" or "the overwhelming majority of" initial conditions and although it is common and convenient and natural to use the Liouville measure, at least in the context of classical mechanics, many different measures would yield the same notion of typicality.[8] In particular, we are not committed to giving meaning to the

[8]On the other hand, many measures would yield a different notion of typicality. One can think, for instance, of singular measures, concentrated on a single point in phase space. Such a measure may even turn out to be stationary, in case that this particular microstate

exact number that a typicality measure assigns to every (measurable) subset of phase space. The only values that are meaningful in this context are 1 (or those close to 1) and 0 (or those close to 0) indicating what Bernoulli called *moral certainty* and *moral impossibility*.[9] Probability statements – as an expression of *statistical regularities* – can then be *grounded* in such typicality statements, e.g. by asserting that the probability ($\sim$ relative frequency) of 'heads' or 'tails' in a *typical* series of coin tosses is 1/2.

Second, in making a typicality statement, we do not commit ourselves to talking about actual or hypothetical *ensembles* of systems, nor do we use probabilistic concepts to express our "guess" – in terms of *information* or *knowledge* or *believe* – about a system's actual microstate. A typicality statement refers to nothing more and nothing less than the fact that a certain (coarse-grained/macroscopic) property or behavior of a physical system is typical given the respective microscopic laws, i.e. that it's the kind of property or behavior that our fundamental theory predicts for an overwhelming majority of microscopic configurations compatible with appropriately specified (macroscopic) boundary conditions: Typically, a coin tossed repeatedly (and fairly) for a large number of times will land about as often on *heads* as on *tails*. Typically, an ice cube at room temperate will melt. According to the laws of quantum mechanics, a series of particles shot through a double-slit will typically (though not necessarily) display an interference pattern when registered on a screen behind the slits. According to classical mechanics, it typically won't (though it possibly might).

A typicality statement is thus an objective physical fact, *in principle* derivable from the fundamental (microscopic) laws that we take as the basis of our considerations. (It is a fact that even Laplace's demon should care about, to the degree that he cares about physics.) But what exactly is it a fact about? Well, typicality is, first and foremost, the answer to the question that stood at the very beginning of our discussion, namely: what is the connection between the*macroscopic regularities* that physics is supposed to account for and the underlying microscopic laws. Another way to put this

happens to be a stationary point of the dynamics. So why not take such a measure to define "typicality", meaning that a property is typical if and only if it is instantiated by this one particular configuration? We trust the reader to answer this question for himself.

[9]See Bernoulli (1713). Such typicality statements can be understood in the sense of *Cournot's principle*, which is one of the basic principles underlying the philosophy of Kolmogorov's *Grundbegriffe*, but also stands in the philosophical tradition of great mathematicians such as Emile Borel, Maurice Fréchet or Paul Lévy. See Shafer and Volk (2006) for a beautiful essay on this topic.

is to ask: What is the nomological status of the “macroscopic laws”?

### 4.3 The status of macroscopic laws

Philosophically, the truly remarkable yet often unacknowledged aspect about the probabilistic character of thermodynamic laws is not the way in which laws that once have been thought to be exact turn out to be merely approximately true, but the way in which the regularities expressed by these laws turn out to be *contingent* rather than *necessary* truths. In other words, if we accept the microscopic laws as fundamental, we have to concede that the so-called macroscopic laws, even in an approximate or statistical sense, are in fact *no laws at all* in that they lack the status of nomological necessity. For all we know, the initial condition of our universe (conceived as a Newtonian universe) could have been such that systems, prepared or created in a low-entropy state, would regularly end up on one of the “bad” trajectories that undergo an anti-thermodynamic evolution. That is to say that there are possible Newtonian universes in which gases are regularly found to contract rather than expand, in which heat does sometimes flow from a colder to a hotter body and in which macroscopic objects such as balls and chairs and tables occasionally jump up in the air (while cooling off accordingly to account for the conservation of energy) simply because a large number of particles happen to move in the same direction at the same time. In these counterfactual yet (nomologically) possible universes, it is simply not true that such events are very unlikely because they happen all the time.[10]

And yet, we would insist, it’s more than a mere contingency, more than a *factum brutum* that our universe is not like that. And indeed, our physical theory has more to say here – notabene without assigning us the impossible task of determining the actual boundary conditions of our universe – for it tells us that the initial conditions of a Newtonian universe would have to be *exceedingly special* to give rise to subsystems violating thermodynamic laws as more than astronomically rare exceptions. Thermodynamic laws, in other words, are *statistical regularities of typical universes.* And it is this characterization, we suggest, that specifies their connection to the underlying microscopic laws and grounds their own law-like status.

Kripke (1980) famously explained the difference between logical and

[10] Of course, among all possible Newtonian universes there will be some with no thermodynamic arrow and no interesting structures at all. But here, to make a point, we consider universes that are hospitable to intelligent life, while the second law of thermodynamics fails to hold in branching systems just so often as to make a fool out of physicists.

nomological supervenience by the following metaphor: B-properties supervene logically on A-properties if, after fixing the A-properties of the world, there was nothing else God could (or needed to) do for fixing the B-properties. The A-properties, we say, logically entail the B-properties. In case of a nomological supervenience, however, God, after making sure of the A-facts, still had some work to do for making sure of the B-facts by determining *laws of nature* relating B-properties to A-properties. Going one step further we could say: a property of our world that is *typical* for these laws is a fact or regularity for which God, after fixing the laws of nature and the fundamental ontology of the world, still had a *little bit* of work to do in choosing appropriate initial conditions for our universe. However, while almost any possible choice (compatible with the relevant macroscopic constraints) would have been fine to make sure that this property is instantiated, God would've had to be utterly meticulous – and maybe somewhat malicious – to arrange the initial configuration of the universe in such a way that isn't.

Turning back to the 'second law', we have to note one subtlety in connection with the Past Hypothesis (see Section 2). According to the Past Hypothesis, the initial *macrostate* of our universe was a very *special* one, marking the low-entropy end of the thermodynamic arrow of time. However, with respect to this macrostate, the initial *microstate* of the universe was *typical* (in regard to its future evolution), thus explaining the increase of entropy in the universe as a whole and in any of its branching subsystems. All in all, there is no contradiction, but a clear *tension* between the typicality account and the Past Hypothesis. The resolution of this tension is considered by many as one of the most profound problems of modern physics.[11]

What else is left to say? Not much, we believe. To understand that a certain regularity is typical and yet to wonder why it is that we observe this regularity in nature (and why we should expect this regularity to persist in the future), is to ask why our universe is typical, i.e. why it is, in this particular respect, like the overwhelming majority of all possible universes instantiating the same set of fundamental laws. And while we don't know how to answer, except maybe with Einstein's bon mot that "God is subtle, but he is not malicious", the very question seems to us utterly uncompelling.

[11]See, for instance, Penrose (1999) and his "Weyl curvature hypothesis" as a proposal for an additional law restricting the initial state of the universe, but also Callender (2004) arguing from a Humean perspective *against* the need for further explanation of the Past Hypothesis. See Carroll (2010) for a very nice discussion of the problem and ibid. as well as Carroll and Chen (2004) for an attempt to dispose of the Past Hypothesis altogether.

Explanations have to end somewhere. If we can establish that a certain property is typical for a particular kind of system, this should elevate any sense of mystery or puzzlement as to why we find such systems instantiating the respective property. Hence, we should consider the phenomenon to be reasonably and conclusively *explained* on the basis of the microscopic theory. Similarly, if we can establish that a macroscopic feature or regularity is typical for a certain kind of system, we should by all reasonable means expect to find this feature realized in a given system of said kind. In this sense, it constitutes a *prediction* of the microscopic theory.

In this fashion, typicality statements figure in a *way of reasoning* about nature. In fact, since the situation in which we find ourselves towards the world is necessarily one in which all we can ever hope to know about the world's state is compatible with a plurality of fundamental (microscopic) matters of fact, the relevant *explanatory* and *behavior guiding* statements that we can extract from the fundamental laws of physics are virtually always results about typical solutions of the equations of motion.

Finally, we shall emphasize again that a typicality reasoning is a non-deductive reasoning. Logically, the fact that something has been shown to be typical doesn't imply anything about any *particular* instance. In other words, it is always possible for a particular system – and ultimately our universe – to be atypical in the relevant respect. But facts that strike us as atypical are usually the kind of facts that cry out for *further* explanation. This is why a Casino manager has not just economic interest but reasonable grounds to suspect cheating if a player hits three jackpots in a single night. And this is why good scientific practice would eventually require us to revise our theory and look for a different set of laws, rather than endorsing an explanation of empirical data based on special initial conditions or, if you will, a streak of bad luck. In the end, it is not logically but epistemically inconsistent to accept a certain physical theory and accept at the same time that our universe is (in any relevant respect) an atypical model of that theory, for this would undermine any reasons to endorse the theory in the first place.

## 5 Reply to critics

In this final section, we are going to address the most common objections to the typicality account that have been raised in the contemporary philosophical literature. We will argue that these objections are unfounded and often based on unnecessary misunderstandings of the actual argument.

### 5.1 Missing the point of typicality

One of the most common mistakes in the debate about Boltzmann's statistical mechanics is the failure to appreciate the difference between a typicality statement and an inference about particular instances. Consider for example the objection of Roman Frigg in reply to Goldstein (2001):

> *Goldstein suggests that a system approaches equilibrium simply because the overwhelming majority of states in $\Gamma_E$ are equilibrium microstates [...]. This is wrong. If a system is in an atypical microstate [...], it does not evolve into a equilibrium microstate just because the latter are typical; typical states do not automatically function as attractors. (Uffink, 2007, 979–980) provides the following example. Consider a trajectory $x(t)$, i.e. the set $\{x(t) = \phi_t(x(t_0)) \mid t \in [t_0, \infty)\}$, a set of measure zero in $\Gamma_E$. Its complement, the set $\Gamma_E \setminus x(t)$ of points not laying on $x(t)$, has measure one. Hence the points on $x(t)$ are atypical while the ones not on $x(t)$ are typical (with respect to $\Gamma_E$, $\mu$, and the property 'being on $x(t)$'). But from this we cannot conclude that a point on $x(t)$ eventually has to move away from $x(t)$ and end up in $\Gamma \setminus x(t)$; in fact the uniqueness theorem for solutions tells us that it does not. The moral is that non-equilibrium states do not evolve into equilibrium states simply because there are overwhelmingly more of the latter than of the former, i.e. because the former are atypical and the latter are typical.* (Frigg, 2009, pp. 8–9).

Of course, no one suggests, in the naive sense implied by Frigg, that any *specific trajectory* must move to equilibrium "simply because" the overwhelming majority of states are equilibrium states – just as no one suggests that any *specific* lottery ticket must lose "simply because" the overwhelming majority of possible combinations are losing combinations. In the alluded sense, a lottery ticket loses simply because someone picked the wrong numbers and a system converges to equilibrium simply because its microscopic evolution carries the configuration into equilibrium. What adds explanatory value in these cases is not a statement that identifies further cause or sufficient reason for why the respective result must obtain in any single instance, but a statement which asserts that the instances in which it does obtain are the typical ones. In particular, Goldstein's argument – which is the same as our argument, which is the same as Boltzmann's argument – is not about what every single solution trajectory must do, but about what the great majority of them will (i.e. for typical initial conditions).

So what is the point of the "counterexample" that was formulated by Jos Uffink and that made such an impression on Frigg? Evidently, it is correct that a solution $x(t)$ will never enter the phase-space region $\Gamma_E \setminus \{x(t)\}$ despite the fact that this set has measure one. *Typical* solutions, however, will. In fact, it follows from the "uniqueness theorem" that every other solution (with the same total energy) lies entirely in $\Gamma_E \setminus \{x(t)\}$. Hence, leaving aside the fact that this artificially crafted phase-space region is of no physical interest to begin with, it remains unclear what the example is supposed to demonstrate and how it's supposed to hit the typicality account. With all due respect, the debate seems a bit like some people trying to explain that typical lottery tickets will fail to win the jackpot because of the huge number of possible combinations, while others are running around with a winning lottery ticket in order to disprove them.

### 5.2 The 'measure zero problem'

If Uffink's example works at all, then as another instance of the notorious *measure zero problem* which is basically the observation that, as soon as we go to a more *fine-grade* description, any physical system is found to be atypical with respect to some (more or less natural) properties. In particular, for a continuous state-space and a nonsingular measure, the *actual* microscopic configuration and, as we just noted, even the entire trajectory of a system will constitute a set of measure zero. While this observation is almost traditionally presented as a serious challenge to typicality arguments (not least by Frigg himself, 2009, p. 23, but see also, e.g., Sklar, 1973) we don't believe it to cause much of an embarrassment for the kind of reasoning we have defended so far.

There are facts and regularities that can be explained on the basis of the microscopic laws by virtue of being typical (like ice-cubes melting at high temperature). There are contingent facts about physical systems that are not typical, but can be explained in a different sense – usually by tracing them back to other (even more) special states of affair. For instance, the current state of our office is certainly atypical with respect to the exact distribution of objects on the desk, but we can tell some sort of causal story about how a used coffee mug ended up near the keyboard and how the blue book came to lie on top of the heavier red one. Finally there are facts like the one that a trajectory through some state-space will never cross its complement – which do not require further explanation but are well-suited for creating confusion where none is due.

### 5.3 Misidentifying the macrostates

One objection to the typicality account that we have encountered on various occasions goes back to Lavis (2005) and attacks the very foundation of Boltzmann's argument by claiming that the equilibrium macrostate (the state of highest entropy) will not generally make up the largest part of phase space. Allegedly, this is because states of lower entropy might be "degenerate" with the sum of their measures exceeding the measure of the equilibrium region.

The premise of Lavis' argument follows Boltzmann's combinatorial analysis of the *gas in a box*, in which the volume of the box is coarse-grained into finitely many cells while counting the number of particles in each cell. Lavis observes – considering, for instance, the simple case of $N = 8$ particles distributed over $m = 4$ cells – that while the most likely occupation $(2, 2, 2, 2)$ (meaning: every cell contains exactly two particles) corresponds to larger phase-space measure than, say, $(3, 2, 2, 1)$, there are actually 12 different permutations of $(3, 2, 2, 1)$, all describing possible macroscopic configurations. Hence, he continues, the sum of the measures of such degenerate states exceed that of the largest "macrostate" $(2, 2, 2, 2)$, which Boltzmann – according to Lavis – would identify as thermodynamic equilibrium.

It is hard to understand, from our point of view, how this could have caught on as a serious issue in Boltzmann's statistical mechanics. Of course, while having *exactly* $N/m$ particles in each of the $m$ cells is more likely (corresponding to larger phase-space volume) than any other specific occupation of cells, this is overall a very *special* situation. In fact, the probability of this exact equidistribution goes to zero as $N$ becomes large. However, for large (macroscopic) $N$ and small (microscopic) particles, having precisely $N/m$ particles in each of the $m$ cells is *macroscopically indistinguishable* from configurations in which some cells contain one or two or ten or even a million particles more than others. To suggest that a gas is already out of equilibrium if there are, say, a few more molecules on the left side of the box than on the right side of the box is to miss the whole point of the micro/macro distinction in the first place (and to attack a bad caricature of Boltzmann's argument).

In other words, the exact equidistribution of the particles over the cells (that Lavis falsely identifies with thermodynamic equilibrium) and the small deviations from this exact equidistribution (that he wants to weigh against the former) actually coarse-grain to *one and the same* macrostate, all corresponding to thermodynamic equilibrium.

More precisely, it is an elementary result of probability theory that, for large $N$, the phase-space measure is concentrated on configurations for which the number of particles in each cell deviates at most $\sim \sqrt{\frac{N}{m}}$ from the mean value. For $N \approx 10^{23}$ and $m \ll N$, this means that microstates corresponding to local density-fluctuations of less than a billionth of a percent exhaust almost the entire phase-space volume. In all or most circumstances, this would correspond to the relevant and appropriate notion of *thermodynamic equilibrium.*

In conclusion, the case presented by Lavis has very litte to do with the degeneracy of lower-entropy states and everything with considering reasonable macrostates in the first place. In general, while one shouldn't discount the possibility to come up with some pathological counterexample (rather than a wrong one), the degeneracy of entropy levels is not a real issue. As repeatedly emphasized, the phase-space measures corresponding to different entropy levels do not vary only by a little, but by a great many orders of magnitude, so that even all lower-entropy states taken together will in general correspond to much smaller phase-space volume than a single macrostate of larger entropy. In particular, the *entire* non-equilibrium region of phase space is minuscule compared to the phase-space region corresponding to thermodynamic equilibrium.

## 5.4 The role of the dynamics

Recently, the typicality account has come under attack in the philosophical literature for its lack of mathematical rigor and the alleged failure to make precise the dynamical assumptions on which the argument is supposed to rest (Uffink, 2008; Frigg, 2009, 2011; Frigg and Werndl, 2011, 2012). Frigg and Werndl (2012) even go as far as declaring that the typicality account is "mysterious" because the "connection with the dynamics" is unclear (p. 918). Jos Uffink writes on a similar note (as a conclusion to his "counterexample" discussed in 5.1):

> *[I]n order to obtain any satisfactory argument why the system should tend to evolve from non-equilibrium states to the equilib-*

> *rium state, we should make some assumptions about its dynamics. In any case, judgments like 'reasonable' or 'ridiculous' remain partly a matter of taste. The reversibility objection is a request for mathematical proof (which, as the saying goes, is something that even convinces an unreasonable person).* (Uffink, 2007, p. 61)

We have already seen that these objections (in particular the ones by Frigg and Uffink, but see also Frigg and Werndl, 2012, who refer explicitly to Frigg, 2009, 2011) are based to a large degree on a simple misunderstanding of what the typicality account actually argues for. Other than that, the dissatisfaction expressed by these authors is clearly based on a certain expectation about what an explanation of thermodynamic behavior (or maybe a physical explanation in general) is supposed to be. Upon the most reasonable reading of their queries, the authors insist that an account of the 'second law' would have to involve a precise mathematical assumption about a system's microscopic dynamics that *logically implies* its thermodynamic behavior (see also Frigg and Werndl, 2011, p. 632). Admittedly, the typicality account doesn't quite live up to this expectation. However, we are convinced that this sort of naive deductive explanation is neither possible nor necessary nor, in fact, desirable.

To begin with, we should emphasize again that the typicality account is – by design – an explanation or an explanatory scheme, not a proof. And while everyone is entitled to his epistemic standards, there are good reasons why (to paraphrase Uffink) a reasonable person is sometimes willing to settle for less than a mathematical theorem. If as an explanation of macroscopic phenomena we accepted nothing short of rigorous proof, the atomic hypothesis would yet have to earn its merits.

When it comes to the subtle issue of macro-to-micro reduction, it lies in the nature of the problem that rigorous results are rare and difficult to come by. Obviously, we cannot just solve the equations of motion for $N \approx 10^{23}$ particles to check if the typicality account tells us the right story about a systems thermodynamic behavior. Instead, a mathematical physicist has to look for appropriate models and tools, simplifications and idealizations to bridge the gap between the beauty and simplicity of the fundamental microscopic laws and the complexity of the macroscopic world. This is neither a straightforward nor a purely deductive process. But rather than acknowledging that life – and physics – is sometimes hard, certain approaches to the foundations of statistical mechanics (in particular, the various ergodic programs, old and new) have succumbed to the idea that if we just find the right mathematical language (e.g. if we study microscopic models on the level of *dynamical systems*) the dynamics of a trillion trillion interacting particles, that are just extremely complicated and complex and difficult to handle, will somehow reduce to a simple and concise mathematical property that can serve as an *explenans* for thermodynamic behavior. But these programs have not only been (so far) unsuccessful, their whole premise seems to us utterly naive and there is simply no good reason to expect that it should even be possible to identify such a specific dynamical property as some authors have in mind.

Maybe more importantly, there is also no compelling reason to insist on such a property. Some critics may have missed the fact that it is not by omission but a result of Boltzmann's analysis that his explanation of thermodynamic behavior doesn't rely on any special feature of the microscopic time-evolution. What we learn from this analysis, is, simply put, that the role of the dynamics is basically restricted to carrying a great majority of microstates, starting out in a designated, vanishingly small region of phase space that corresponds to thermodynamic non-equilibrium, reasonably quickly into the rest of phase space that corresponds to thermodynamic equilibrium. And this is so much weaker and so much more evident as an "assumption" about the dynamics of thermodynamic systems that it's hard to see how it could be further explained or made more plausible by reducing it to a more formal and abstract mathematical concept.

Finally, we have to make the general point that – contrary to what certain philosophical theories have claimed and contrary to what certain authors have come to take for granted – there is a difference between a physical explanation and a logical deduction starting from a complete set of precise axioms. For while it lies in the nature of mathematical proof and logical deduction that the truth of the conclusion depends rigidly on the truth of the premises, it is essential for a good physical argument to be reasonably stable against small deviations from its underlying assumptions, in particular when they are themselves the result of approximations and idealizations (see Schwartz, 1992, for a beautiful elaboration on this point.)

Against this backdrop, one reason why it's not useful to tie the explanation of thermodynamic irreversibility to a very specific and precise characteristic of the dynamical system is that any such characteristic is bound to be very sensitive to small changes in the microscopic model or boundary conditions. For instance, it is often claimed that ergodic properties are essential to the explanation of thermodynamic behavior and many philosophers such as Frigg and Werndl have engaged in the debate whether the assumption of ergodicity – or some variant thereof – actually applies to the gas models most commonly studied by physicists and mathematicians. However, regardless of the mathematical interest in this question, the debate is rather pointless from a foundational point of view. For while it may or may not be the case that a hard-sphere gas in an ellipsoidal box with perfectly reflecting walls is ergodic[12], the slightest perturbation of the interaction potential or the slightest inhomogeneity in the walls would most likely destroy this special feature anyhow. In particular, the very question whether, say, *the air in your living room* has good ergodic properties (or is accurately modeled by a system with good ergodic properties) is not just hopeless but ridiculous.

The details of ergodicity (or its variant promoted by Frigg and Werndl) need not concern us here, for the moral of this story is that, regardless of the merits of certain formal concepts, our physical and philosophical understanding of thermodynamic behavior should better *not* depend too rigidly on any such specific and narrowly defined mathematical premise if it's supposed to apply to anything more than highly idealized models.

None of this is to say that we shouldn't continue the efforts to corroborate Boltzmann's insights with rigorous mathematical results. However, it is important to understand that, contrary to what some commentators have suggested, the difference between the explanatory scheme that we have presented and a more rigorous proof of thermodynamic behavior is not some secret ingredient like a dynamical assumption that proponents of the typicality account have missed to specify, but a heap of very hard, very technical work in mathematical physics. Good physics and good philosophy of physics, on the other hand, is also about appreciating where our understanding of an issue depends on rigorous formalization and technical proof and where it doesn't.

**Acknowledgements** We are grateful to Detlef Dürr, Sheldon Goldstein, Tim Maudlin and Nino Zanghì for teaching us almost everything we know about the subject of this paper. Thanks to Jean Bricmont, Mathias Frisch and Jenann Ismael for insightful remarks on various occasions.

[12]Ergodicity is probably true for the hard-sphere gas, and almost certainly failing for any more realistic model. For why ergodicity is largely irrelevant in the first place, see e.g. Schwartz (1992), Bricmont (1995) and Goldstein (2001).

## References


Albert, D. (2000). Time and Chance. (Cambridge: Harvard University Press)

Bernoulli, J. (1713). Ars conjectandi, opus posthumum. Accedit Tractatus de seriebus infinitis, et epistola gallicé scripta de ludo pilae reticularis. (Basel: Thurneysen Brothers). Reprinted: Bernoulli, J. (2006). The art of conjecturing. (Baltimore: The John Hopkins University Press)

Boltzmann, L. (1896). Vorlesungen über Gastheorie. (Leipzig: Verlag v. J. A. Barth, Leipzig). Nabu Public Domain Reprints.

Boltzmann, L. (1896). Entgegnung auf die wärmetheoretischen Betrachtungen des Hrn. E. Zermelo. Wiedemann's Annalen, 57, 773-784.

Bricmont, J. (1995). Science of Chaos or Chaos in Science? Physicalia Magazine, 17, 159-208.

Bricmont, J. (2001). Bayes, Boltzmann and Bohm: Probabilities in Physics. (In J. Bricmont, D. Dürr, et al. (Eds.), Chance in Physics. Foundations and Perspectives (pp. 3-21). Berlin: Springer.)

Callender, C. (2004). There is No Puzzle about the Low Entropy Past. (In C. Hitchcock (Ed.). Contemporary Debates in Philosophy of Science (pp. 240-255). London: Blackwell.)

Carroll, S., & Chen, J. (2004). Spontaneous Inflation and the Origin of the Arrow of Time. ArXiv: hep-th/0410270.

Carroll, S. (2010). From Eternity to Here. The Quest for the Ultimate Theory of Time. (USA: Dutton, Penguin Group)

Davies, P. C. W. (1977). The Physics of Time Asymmetry. (Berkeley: University of California Press)

Dizadji-Bahmani, F., Frigg, R. & Hartmann, S. (2010). Who's Afraid of Nagelian Reduction? Erkenntnis, 73, 393-412.

Dürr, D. (2009). Bohmian Mechanics. (Berlin: Springer)

Einstein, A. (1949). Autobiographical Notes. (In P. A. Schilpp (Ed.), Albert Einstein: Philosopher Scientist, The Library of Living Philosophers (p. 43), sixth printing 1995. La Salle (IL): Open Court.)

Ehrenfest, P. und T. (1911). Begriffliche Grundlagen der statistischen Auffassung in der Mechanik. Enzyklopädie der mathematischen Wissenschaften mit Einschluss ihrer Anwendungen, Band 4, 3-90.

Feynman, R. (1967). The Character of Physical Law. (Cambridge: The M.I.T. Press)

Frigg, R. (2009). Typicality and the Approach to Equilibrium in Boltzmannian Statistical Mechanics. Philosophy of Science, 76, 997-1008.

Frigg, R. (2011). Why Typicality Does Not Explain the Approach to Equilibrium. (In M. Suárez (Ed.), Probabilities, Causes and Propensities in Physics (pp. 77-93). Dordrecht: Springer.)

Frigg, R., & Werndl, C. (2011). Explaining Thermodynamic-Like Behaviour In Terms of Epsilon-Ergodicity. Philosophy of Science, 78, 628-652 .

Frigg, R., & Werndl, C. (2012). Demystifying Typicality. Philosophy of Science, 79, 917-929.

Gallagher, I., Saint Raymond, L., & Texier, B. (2012). From Newton to Boltzmann: the case of short-range potentials. Preprint: ArXiv: 1208.5753v1 [math.AP].

Goldstein, S. (2001). Boltzmann's Approach to Statistical Mechanics. (In J. Bricmont, D. Dürr, et al., Chance in Physics. Foundations and Perspectives (pp. 39-54). Berlin: Springer.)

Goldstein, S. (2012). Typicality and Notions of Probability in Physics. (In Y. Ben-Menahem, & M. Hemmo (Eds.), Probability in Physics. The Frontiers Collection (pp. 59-71). Berlin: Springer.)

Goldstein, S. & Lebowitz, J. (2004). On the (Boltzmann) entropy of non-equilibrium systems. Physica D: Nonlinear Phenomena, 193, Issues 1-4, 53-66.

King, F. (1975). BBGKY hierarchy for positive potentials. Dissertation, University of California at Berkeley.

Kripke, S. (1880). Naming and Necessity. (Oxford: Blackwell)

Lanford, O. E. (1975). Time Evolution of Large Classical Systems. (In J. Moser (Ed.), Lecture Notes in Physics, Vol. 38 (pp. 1-111), Berlin: Springer.)

Lavis, D. (2005). Boltzmann and Gibbs: An Attempted Reconciliation. Studies in History and Philosophy of Modern Physics, 36, 245-273.

Lebowitz, J. (1981). Microscopic dynamics and macroscopic laws. Annals New York Academy of Sciences, 220-233.

Lebowitz, J. (1993). Macroscopic laws, microscopic dynamics, time's arrow and Boltzmann's entropy. Physica A, 194, 1-27.

Maudlin, T. (2007). What could be objective about probabilities? Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics, 38, Issue 2, 275-291.

Penrose, R. (1999). The Emperor's New Mind. (Oxford: Oxford University Press)

Price, H. (1996). Time's Arrow & Archimedes' Point. New Directions for the Physics of Time. (New York: Oxford University Press)

Price, H. (2002). Burbury's Last Case: The Mystery of the Entropic Arrow. (In C. Callender (Ed.), Time, Reality & Experience (pp. 19-56). Cambridge: Cambridge University Press.)

Pulvirenti, M., Saffiro, C., & Simonella, S. (2013). On the validity of the Boltzmann equation for short range potentials. Preprint: ArXiv: 1301.2514v1 [math-ph].

Schwartz, J.: (1992) The pernicious influence of mathematics on science. (In M. Kac, G.-C. Rota, J. and Schwartz (Eds.), Discrete Thoughts (pp. 19-25). Boston: Birkhäuser.)

Shafer, G., & Volk, V. (2006). The Sources of Kolmogorov's Grundbegriffe. Statistical Science, 21, No. 1, 70-98.

Sklar, L. (1973). Statistical Explanation and Ergodic Theory. Philosophy of Science, 40, No. 2, 194-212.

Spohn, H. (1991). Large Scale Dynamics of Interacting Particles. (Berlin: Springer)

Uffink, J. (2007). Compendium of the foundations of classical statistical physics. (In J. Butterfield, J., & J. Earman (Eds.), Handbook for the Philosophy of Physics (pp. 923-1047). Amsterdam: Elsevier.)

Uffink, J. (2008). Boltzmann's Work in Statistical Physics. The Stanford Encyclopedia of Philosophy.

Villani, C. (2002). A review of mathematical topics in collisional kinetic theory. (In S. Friedlander, & D. Serre (Eds.), Handbook of mathematical fluid dynamics, Vol. 1 (pp. 71-305). Amsterdam: Elsevier.)

Zanghì, N. (2005) I fondamenti concettuali dell'approccio statistico in fisica. (In V. Allori, M. Dorato, F. Laudisa, & N. Zanghì (Eds.), La Natura Delle Cose. Introduzione ai Fundamenti e alla Filosofia della Fisica (pp. 139-228). Roma: Carocci.