Some Reflections on the Statistical Postulate: Typicality, Probability and Explanation between Deterministic and Indeterministic Theories[1]

Valia Allori
Department of Philosophy
Northern Illinois University
vallori@niu.edu

Abstract

A common way of characterizing Boltzmann's explanation of thermodynamics in term of statistical mechanics is with reference to three ingredients: the dynamics, the past hypothesis, and the statistical postulate. In this paper I focus on the statistical postulate, and I have three aims. First, I wish to argue that regarding the statistical postulate as a probability postulate may be too strong: a postulate about typicality would be enough. Second, I wish to show that there is no need to postulate anything, for the typicality postulate can be suitably derived from the dynamics. Finally, I discuss how the attempts to give preference to certain stochastic quantum theories (such as the spontaneous collapse theory) over deterministic alternatives on the basis that they do not need the statistical postulate fail.

1. Introduction

Boltzmann's statistical mechanics provides an explanation of the macroscopic laws of thermodynamics, such as 'entropy always increases', in terms of the microscopic Newtonian laws. In his seminal book "Time and Chance" (2001) David Albert has made especially clear how this is done, and what problems this account faces. As a consequence, his book has been extremely influential in the discussions about the foundation and the philosophy of statistical mechanics (see also Loewer 2020). Very briefly, the view is that Boltzmann's derivation of the macroscopic laws only needs three ingredients:

1. The Newtonian law of motion,
2. The statistical postulate,
3. The past hypothesis.

Roughly put, (1) describes the dynamics of the microscopic components of the thermodynamics objects; (2) specifies the probability measure grounding the meaning of the probabilities arising from the statistical derivation of thermodynamic laws; and (3) guarantees that entropy will likely increase in the future but not in the past by postulating an initial low entropy state for the universe.

In addition, Albert argues that a corresponding quantum statistical explanation based on a universal quantum theory (such as possibly the pilot-wave theory, the spontaneous localization theory, and Everettian mechanics) would look very similar to the

[1] Forthcoming in V. Allori (ed.) "Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature," World Scientific (2020).

classical one. However, he maintains that in case of the spontaneous localization theory the ingredients would reduce to the following two:

1. The stochastic law of motion,
2. The past hypothesis.

That is, contrarily to what happens in any deterministic framework, in the spontaneous localization schema there is no need for an additional statistical postulate. The reason is that the theory, being indeterministic, has probabilities already 'built into' it. Hence, statistical mechanical probabilities in the spontaneous localization framework are just the quantum probabilities. Because of this, Albert concludes, the spontaneous collapse theory should be preferred to the other alternative quantum theories.

In this paper, I focus my attention on the statistical postulate, and I wish to make three points. First, I argue that the postulate need not be understood as a postulate about probability; second, I argue that it is not a postulate after all; and third that, given the first two points, there is no argument remaining to prefer the spontaneous localization theory over the deterministic alternatives. Here is how I plan to accomplish these tasks. After having reviewed this account of Boltzmann's work in Section 2, I argue that the statistical postulate is not needed in two steps. In the first step, developed in Section 3, I assume for the sake of the argument that one indeed has a postulate. However, I argue that it is not necessarily a postulate about probabilities; rather, one can appeal to the weaker notion of typicality. Therefore one is led to discuss the notion of typicality (rather than probability) measure. In the second step, discussed in Section 4, I reconstruct and use (what I take to be) Sheldon Goldstein's argument (2001, 2011) to show that the typicality measure is not postulated but can be suitably extracted from the dynamics under certain constraints, such as stationarity and generality, which are regarded 'natural'.[2] Moreover, as I elaborate in Section 5, I reconstruct the argument developed by Nino Zanghí (2005) to show that the difference between deterministic and indeterministic theories is not as substantial as one may think. Elaborating on the concepts introduced so far, one can provide a unifying account of the statistical mechanical explanation for both kinds of theories, which therefore rely on the same ingredients. As a consequence, one cannot conclude that one type of theory or the other will be better off in explaining the macroscopic laws from a microscopic dynamics, therefore diffusing the argument to prefer spontaneous localization theories to deterministic quantum theories.

2. Boltzmann's Approach to Statistical Mechanics

The evolution of macroscopic objects is generally very complicated. Nonetheless, their behaviour is governed by simple and general physical laws such as the laws of thermodynamics. These laws are phenomenological: they merely describe regularities on the macroscopic level. A natural question is then whether they can be derived by a more fundamental theory. Suppose such a theory is Newtonian mechanics, according to which the world is described in terms of point-like particles moving according to Newton's

[2] See also Dürr (2009), and Zanghí (2005).

second law of motion. If the theory is complete then every physical system is describable by it, and one could (at least in principle) reproduce the behaviour of macroscopic objects thought as composed of microscopic Newtonian particles. However, to do this one needs to solve Newton's equation for macroscopic bodies, and there are at least two obvious problems. First, one would need to know exactly which forces act between the particles. Moreover, since macroscopic objects are composed by an incredibly large number of particles, it is practically impossible to compute their trajectories. Nonetheless, exploiting the fact that there are so many particles, one can use statistical methods to have enough information on macroscopic systems even under these conditions. The resulting theory is called statistical mechanics and it has been primarily developed by Ludwig Boltzmann and Josiah Willard Gibbs. In this paper I will focus on the work of Boltzmann, which is the basis of the model I aim to discuss.[3]

There is another challenge to recover the laws of thermodynamics from the underlying microscopic dynamics. It is connected to the time-reversibility of Newtonian mechanics, which is in stark contrast with the irreversibility of the macroscopic laws. Arguably, to say that Newtonian mechanics is time reversible is to say that if we flip the sign of the time variable in Newton's equation, the solutions of the new equation are still solutions of the original equation. That means that we cannot tell from the behaviour of an object whether it is moving forward or backward in time. Since we are assuming that macroscopic bodies are a collection of microscopic Newtonian particles, this time-reversibility should be observed also at the macroscopic level. However, this is empirically false: the macroscopic phenomena of our everyday experience are all time-directed: a perfume sprayed from the corner will spread out in the whole room. The opposite, namely the perfume getting back to the initial corner, spontaneously never happens. This macroscopic irreversibility is captured by the second law of thermodynamics, according to which a quantity called 'entropy' (or 'thermodynamic entropy') always increases. Hence, macroscopic phenomena always happen in the direction of increasing entropy. The reason why the perfume spreads in the room is that the state of spread-out perfume has a higher-entropy than the perfume-in-the-corner state. This is the explanation why this phenomenon – perfume spreading out – happens, while the opposite does not. The problem now is that if we want to derive thermodynamics from Newtonian mechanics then we need to derive macroscopic irreversibility from microscopic reversibility. So, we need to get rid of the second solution in which the perfume comes back into the corner. But how? This is the challenge Boltzmann faced and that we will review in the next sections along the lines discussed in Albert (2001) and Loewer (2020).

2.1 Microstates, Macrostates and Entropy

Assuming Newtonian mechanics, the complete description at a time of any single particle is given by the pair (r, v) of its position and velocity at that time. Therefore, the complete

[3] For a comparison between the Boltzmaniann and Gibbsian approaches to statistical mechanics, see Goldstein et al. (2020), Wallace (2020), and Werndl and Frigg (2020) and references therein.

dynamical description at a fixed time t of a body composed of N particles is given by the so-called *microstate* $X = (r_1, \ldots . r_N, v_1, \ldots . v_N)$, namely the set of the positions and velocities of all particles at time t. The set of all possible microstates constitutes *phase space*, which is in a sense the space of all possible ways the world can be. Phase space can be partitioned into sets, called *macrostates*, to describe the state of a system of particles from a macroscopic point of view. This coarse-graining is done specifying some macroscopic thermodynamically relevant properties such as temperature, pressure and volume: each macrostate is such that all of its microstates have the same thermodynamic properties. Thus, macroscopic properties are functions on the microstate on phase space which vary very slowly on a microscopic scale, so that each macrostate has the same macroscopic property.[4] Therefore, there are different ways in which the same macrostate (with some given macroscopic properties) can 'come from' different microstates.[5] In other words, then, a macrostate is the collection of all the possible ways the system can microscopically give rise to the macroscopic properties of the macrostate. In this sense, knowing the macrostate of a system does not tell us its actual microstate: all the microstates in a macrostate are macroscopically identical. And the bigger the macrostate is, the less information one has about the microscopic composition.

Different macrostates in general are composed by a different number of microstates. One can define the size of a macrostate in phase space by (suitably) 'counting' how many microstates it is composed of. For a fixed energy, there is a particular macrostate which has the following empirical feature: bodies are 'drawn to it' during their motion, and when a body finally reaches it, it will not leave it afterwards (spontaneously, i.e. if nothing takes it away). For these reasons, this macrostate is called *equilibrium* state. For instance, two bodies at different temperatures put into contact will reach an equilibrium temperature, given by the mean of the initial temperatures of the two bodies, and afterwards the temperature will not change.[6] This is connected with the second law of thermodynamics

[4] For instance, temperature is defined as the mean kinetic energy of the particles: $T = T(X) = \frac{1}{K_B} \sum_{i=1}^{N} \frac{1}{2} m v_i^2$, where $K_B = 1.38 \times 10^{-23}\ m^2\ kg\ s^{-1}\ K^{-1}$ is the Boltzmann constant, which provides the macroscopic-microscopic scaling (one can think that the number of particles in a macroscopic body should be at least of the order of Avogadro number, that is $N \sim 10^{23}$).

[5] For example, a macrostate of a gas in a room with volume V is composed of all the ways in which the molecules are arranged such that one finds a gas filling the room. The gas can be in macrostate $X = (r_1, r_2, \ldots, v_1, v_2, \ldots)$. However, If we swap some positions, or change a little some of their velocities such that, say, the microstate is now $X' = (r_2, r_4, \ldots, r_3, \ldots,\ r_1, \ldots, v'_1, v_2, \ldots)$, one still get a gas filling the room of volume V. That is, both are microstates in the same macrostate. Also, consider a generic system in microstate $X = (r_1, r_2, \ldots, v_1, v_2, \ldots)$. Now swap the velocities of particle 1 and particle 2 such that the microstate is now $X'' = (r_1, r_2, \ldots, v_2, v_1, \ldots)$.. This change makes no difference to the value of macroscopic properties such as temperature, hence both microstates belong to the same macrostate defined by temperature T. Therefore, a gas filling a room of volume V and a generic macroscopic body at temperature T , say, are each made of particles whose microstate (unknown to us) belongs to the macrostate characterized respectively by that that volume and that temperature value.

[6] Likewise, consider a box divided in two regions separated by a wall, and a gas in one of the regions. When the separation wall is removed, the gas starts expanding, and in time it will occupy the whole container. This is the equilibrium state of the gas, since after having occupied the whole box the gas will not change anymore.

according to which entropy always increases, given that the (thermodynamic) entropy of equilibrium is maximal. From the point of view of statistical mechanics, we need to suitably define entropy and to define what the equilibrium state is, from a microscopic point of view. The equilibrium state is just a macrostate like any other. With one crucial difference: it contains incredibly many more microstates than the other macrostates. In fact, as we saw before, the number of microstates in a macrostate depends on how many ways there are to obtain the same macroscopic features. And there are so many more ways for the microstates to be distributed in order to correspond to, say, a uniform temperature macrostate than there are in order to correspond to a macrostate with non-uniform temperature.[7] That means that the size of the equilibrium macrostate is larger than any other macrostate. One can therefore understand why microstates 'tend' to equilibrium and (almost) never leave it afterwards: a microstate, in its wandering through phase space, will sooner or later fall into such a big state, and afterwards it will stay there simply because there are so many ways for a microstate to be in equilibrium.[8]

Boltzmann defined the *entropy* of a given microstate as a measure of the volume of the phase space occupied by the macrostate in which the microstate is located. It can be shown that it is equivalent to the thermodynamic entropy, so Boltzmann was left to show that his entropy also increases in time. That is, he had to show that the volume in phase space of the macrostates a microstate will cross during its motion will increase in time. Indeed, we have just see that it does: microstates move from regions of phase space of smaller volume to ones of bigger volume, the biggest of which is the equilibrium state. Thus, entropy increases in the future. Notice however that, contrarily to the case of thermodynamic entropy, it is *possible* for Boltzmann entropy to decrease. In fact, there could be some microstates that behave 'anti-thermodynamically' in the sense that they accurately manage to avoid equilibrium and go into macrostates which are smaller than the one they are currently in. However, even if these microstates are possible, the fact that the volume of the equilibrium state is so big guarantees that such states are very few. That is, the microstates in a given macrostate that manage to avoid such a large state as equilibrium are very few. Thus, more cautiously, one can conclude that the overwhelmingly vast majority (rather than all of them) of microstates in a given macrostate will evolve toward a bigger macrostate, hence larger entropy, in the future.

---

[7] Perhaps this is best seen considering the example of the gas again: initially the gas is in a given macrostate, say M_1, and later it is in equilibrium state M_E. When in M_1, each microscopic particle could be anywhere in the first half of the box. That is, its microstate could be anything between (x, y, z, v_x, v_y, v_z) and $(x + L, y + L, z + L, v_x, v_y, v_z)$, where L is such that $V = L^3$ is the volume of half of the box (assuming velocities do not change). Which means that the volume in phase space of each particle (namely the number of ways the particles can be) is, initially, V. At equilibrium, each particle will have double the volume to roam through, so that the volume in phase space for each particle is $2V$. This is true for all particles, and therefore, assuming that there are N particles in the gas, their volume in phase space is initially $Vol\,(M_1) = V^N$ and at equilibrium is $Vol(M_E) = (2V)^N$. Accordingly, $\frac{Vol(M_E)}{Vol(M_1)} = \frac{(2V)^N}{V^N} = 2^N \approx 2^{10^{23}}$, assuming that N is of the order of the Avogadro number.

[8] The view I just described has been challenged by some commentators (see Frigg and Werndl 2012, Werndl 2013, Werndl and Frigg 2015 ad references therein), who have argued that the approach to equilibrium is explained only if the dynamics play a crucial role. For more on this, see Section 3.4.

2.2 The Past Hypothesis

Unfortunately, as soon as this strategy is understood, one also realizes that, because the underlying Newtonian dynamics is time-reversal invariant, not only the majority of microstates will go into a macrostate with higher entropy, but also most of them will *come from* a state of higher entropy as well. In fact, consider a half-melted ice cube on a kitchen table. The ice cube will likely fully melt, because the state corresponding to a fully melted ice cube has a higher entropy than the present one. However, where is the half-melted cube is most likely to come from? If we forget our memories of the past (that half melted ice cubes usually come from the freezer), and we reason merely in terms of sizes of macrostates, we should immediately see that the present state is likely to come from a past macrostate of larger size. That is, the half-melted ice cube is most likely to come from a fully-melted ice cube. This is because the laws are invariant under time reversal. Therefore, for the vast majority of microstates, entropy increases both in the past and in the future, which is not empirically adequate, assuming we do not want to say that our memories are unreliable. In fact, empirically one needs to find that while the vast majority of microstates in a given macrostate (half-melted ice cube) will evolve into another macrostate with higher entropy (fully melted ice cube), only a very tiny minority of the microstates in the present macrostate (half-melted ice cube) comes from the initial macrostate (ice cubes in the freezer).

One needs to break this symmetry between past and future, and there are various ways to do so. One which is considered among the most promising is the so-called past hypothesis, which postulates that the universe begun with a microstate of extremely low entropy.[9] I will not discuss the reasons why the past hypothesis is considered to be better than the other proposed solutions, since whether the past hypothesis is true or not does not affect my arguments in this paper. However, it should be clear that if one postulates that the universe begun with an extremely low entropy, one effectively 'cancels out' the possibility of it moving toward an even lower entropy, guaranteeing that it is overwhelmingly likely for the microstate to go into a macrostate of higher entropy.[10]

2.3 The Statistical Postulate

Let us now consider the last element in Boltzmann's account as presented by Albert and Loewer. As we have just seen, Boltzmann's account of the second law, supplemented by the past hypothesis, does not tell us how all states will evolve. Rather, it will tell us *how the vast majority* of them will move, namely toward an increasing entropy. However, if we want to have a theory which makes predictions and provides explanations, Albert and Loewer point out, we need to talk about probabilities, not about the number of

[9] Whether the past hypothesis is also able to account for our experience of time is discussed in Hemmo and Shenker (2020).

[10] For a discussion of the past hypothesis and the proposals on how to eliminate it, see Lazarovici and Reichert (2020) and references therein.

microstates. In other words, we want to be able to say not only that the great majority of microstates goes toward a higher entropy state. We also want to be able to say that the *probability* of entropy increasing is high (or that entropy increase is likely). Similarly, one would not only want to say that there are very few 'abnormal' or exceptional (or anti-thermodynamical) states for which entropy decreases. One would also like to say that it is *unlikely* that entropy decreases (or that the probability of entropy decreasing is low). In this way, one can say that Boltzmann's account provides a probabilistic explanation of the second law of thermodynamics, even if the underlying theory is deterministic.

To put things differently, one could have definite results (i.e. with probability 1) if one could solve exactly all the equations for all the particles. However, since one cannot do that for practical reasons, the results obtained are only probabilistic. Thus, the second law of *thermodynamics* is that (thermodynamic) entropy *always* increases, while the *statistical mechanics* version of the second law is that (Boltzmann) entropy *almost always* increases. This should be the same as saying that the *probability* of entropy increasing is extremely high. Here is where probabilities enter the picture. In this way one sees that a statistical mechanical explanation is characterized by providing the prediction not of what will certainly happen, but of what will probably happen. Nevertheless, as we just pointed out, in Boltzmann's theory we do not find probabilities. Rather, there are statements about what the great majority of microstates will do. Hence, we want that something like: 'for the great majority of situations entropy will increase,' to be equivalent to: 'entropy has high probability of increasing,' or to 'entropy almost always increases.' One can move from one locution to the other if all the microstates are 'equiprobable,' in the sense that neither one of them is special in some way or other. That amounts to defining over phase space a 'measure' which does not privilege one microstate over another. A measure is, roughly put, a way of counting how many ways microscopic particles (i.e. microstates) can be arranged in order to give rise to the same macroscopic properties (i.e. macrostates).

The statistical postulate amounts to the assertion that the measure to be used in making inferences about the past and the future (respectively explanations and predictions) is the uniform measure (over the suitable regions of phase space), also known as the Lebesgue-Liouville measure. According to Albert and Loewer, the choice of this measure has to be introduced as a postulate because, they argue, there is no acceptable justification for the choice of such measure. In fact, there could be an infinite number of measures: the uniform measure counts all states equally, but other measures might not. So why is the uniform measure special? The usual answer found in physics books is that the measure is a reflection of our state of ignorance with respect to the microstate the system is in. That is, the measure reflects the fact that an observer is uniformly ignorant of the true microstate the system is in: the observer will assign the same probability to each microstate because she does not know which microstate the system is in. However, this cannot be right, Albert and Loewer argue, since one should not use epistemology to guide our metaphysics: the behavior of macroscopic bodies cannot possibly be a function of how ignorant we are about their composition. For one thing, the former is objective, the latter is not.[11] Thus, since no other justification is provided, the only option is to make this assumption a postulate.

In the next section, I do not dispute that the uniform measure is postulated. I instead argue that this measure need not be a probability measure. That is, I argue that one does not need to invoke the notion of probability in order to explain the macroscopic appearances. Rather, the weaker notion of typicality is sufficient. If so, then the statistical postulate should be understood as a typicality postulate. I then argue in Section 4 that one does not even need to add a postulate after all, given that the notion of typicality which explains the macroscopic appearances can be suitably derived from the microscopic dynamics.

3. How to Avoid Probability Talk: The Notion of Typicality

Let's discuss the statistical postulate more in detail. As we just saw in Section 2.3, the statistical postulate was introduced to bridge the gap between the talk in terms of number of microstates that we use when describing a system in terms of the microscopic dynamics, and the probability talk that we use when discussing in the predictions and explanations of macroscopic phenomena. For instance, when we see an ice cube in a glass of water we want to be able to predict that it is probable that the ice will melt in the next five minutes. Likewise, we want to be able to say that we can explain that an ice cube has melted in the last five minutes by saying that its probability of melting was high. Using statistical mechanics we have 'merely' shown respectively that we can predict that the vast majority of microstates corresponding to an ice cube now will evolve in the next five minutes into a macrostate with higher volume corresponding to a melted ice cube, and that we can explain why the vast majority of microstates corresponding to an ice cube five minutes ago has evolved into the macrostate corresponding to a melted ice cube now. So, we need to connect the locution 'the vast majority of microstates go toward an increasing entropy,' which is given by Boltzmann's account, with the locution 'the probability of entropy increasing is high,' otherwise we would not have explained the regularities. The translation-rule can be provided if we count the microstates equally, using a uniform probability measure, so that we can say that 'vast majority' means 'high probability' as just discussed. Then, the question arises as to why the uniform measure is the correct probability measure. As we have seen in the previous subsection, the usual 'ignorance' justification is not tenable because the microstate will move however it will, independently of whether we know anything about it or not. Thus, in absence of an alternative justification, Albert and Loewer think that this measure should be postulated.
Notice that two claims are being made here:

1) The statistical postulate is needed because *probabilities* are needed to explain/predict;
2) The statistical postulate cannot be inferred from the theory but needs to be *postulated*.

[11] See also Hemmo and Shenker (2012) for additional comments on the (lack of) justification of the uniformity measure.

In this section and the next I argue that none of them is necessarily the case: one does not need probabilities in order to get a satisfactory scientific explanation of the macroscopic phenomena (current section); and even if one did, the probabilities could be grounded in something which is not postulated but derived from the theory using symmetry considerations (Section 4).

3.1 The Typicality Measure

To being with, let me reiterate that the current discussion is not an argument usually present in physics book, in which the uniform measure is justified on epistemic grounds. Moreover, in most physics books there is no mention of the statistical postulate in these terms. In the work of physicists, such as Jean Bricmont (1995, 2001, 2020), Detlef Dürr (2009), Sheldon Goldstein (2001, 2011), Joel Lebowitz (for a first example, see his 1981 paper), and Nino Zanghí (2005), however the postulate and the probabilities appear only indirectly as mediated by another notion, namely the notion of typicality. What is the relation between typicality and probability? In this subsection, I discuss how why typicality, rather than probability, has a more fundamental role in the explanation of macroscopic laws. In the next subsection, I focus on the relations between the two concepts.

To proceed, let us assume for now that the statistical postulate holds: the uniformity measure is postulated. However, I do not assume it is a probability measure. As just mentioned, in the physics literature mentioned above, one finds that 'the vast majority of microstates approach equilibrium,' say, is translated into 'the approach to equilibrium is typical' rather than into 'the approach to equilibrium is probable.' The notion of 'typicality' is notoriously controversial[12] but I think it boils down to the following: a given behavior is typical for an object of a given type if and only if one can show that the vast majority of the systems, suitably similar to the original one, would display that behavior.[13] For instance, consider approach to equilibrium. This behavior is typical, in the sense that the vast majority of thermodynamic bodies display this behavior. Technically, as already seen, the notion of 'vast majority' is defined in terms of a measure μ, which allows 'counting' (i.e. evaluating their number) the exceptions E to a given behavior, fixed a given tolerance. 'Great majority' means therefore that the exceptions to the given behavior, given a tolerance, are few when counted using the measure μ. This is equivalent to say that the size of the set of exceptions, as counted by μ, is small. That is, the set of thermodynamically exceptional states (the ones for which entropy does not increase and thus they do not go to equilibrium) is small. One could be tempted to think that this means that if the set of exception to the behavior 'go toward equilibrium' is very small, then it is highly *probable* that systems will go toward equilibrium. And thus that it's only natural to interpret μ as a probability measure. However, this may be too big a step. In fact, up to this point the measure merely does this:

[12] See Badino (2019) and references therein for a discussion.

[13] Formally, given an object x in a set S, a property P is typical for objects in S if and only if almost all elements of S possess P. See also Wilhelm (2019) on typicality in scientific explanation.

1. It counts the microstates,
2. It allows to define what it means that the size of the set of exceptions E is small.

These two conditions define a *typicality* measure. In contrast, a probability measure is much richer. For once, it also has to be additive (that is, the probability of the sum of two sets is the sum of the probabilities of each set), in contrast to the typicality measure. In addition, while there is a difference between the probability being ½, say, and being 1/3., nothing of the sort is required by the typicality measure, whose role is to make sense of claims that describe/account for/explain phenomena as holding with some very rare exception, without specifically quantify them. Similarly, the typicality measure does not have to satisfy the probability axioms.[14]

Let us now turn to the question of whether we need something more, namely whether we also need a probability measure. I think not: a typicality measure, which provides us with a rough guidance about the relative size of sets in phase space, is enough for the current purposes. In fact, the *only* reason we need a measure for the size of the macrostates is to count the number of microstates in them. The details do not count: it does not matter what precise number the measure gives us. What matters instead is that the size of the set of exceptions (i.e. abnormal or anti-thermodynamic states) is extremely small, regardless of how much it precisely is. That is, we need a typicality measure, not a probability one. Thus, postulating that the typicality measure is the uniform measure (as dictated by the statistical postulate), the second law of thermodynamics holds, *typically*. That is, the set of the exceptional (i.e. anti-thermodynamic) microstates is very small.

Let us see in a bit more detail why typicality is able to explain everything that we seek an explanation for in recovering macroscopic laws from microscopic ones.[15]

1) We want to explain why the 'probability' that a state is in equilibrium is very high. However, we do not need a precise probability estimate about how likely it is; we do not need to distinguish between the equilibrium state being reached with 0.95 probability and 0.99 probability. All we need is a rough estimate, which is what typicality gives us. Thus, we 'merely' have to explain why the equilibrium state is typical. And to do that we need to show that it occupies the vast majority of phase space. That is, we need to show that if A is the set of microstates that do not belong to the equilibrium macrostate, then the size of A is much smaller than the size of the equilibrium macrostate.
2) We also want to explain why systems starting from equilibrium remain there with 'high probability.' Again, details do not matter as we want to explain why the departure from equilibrium is atypical. That is, we need to show that the size of the set of microstates starting from equilibrium and remaining there, as measured by the typicality measure, is overwhelmingly larger that the size of the exceptions, namely the set of microstates whose temporal evolution brings them outside of equilibrium.
3) Similarly, we want to explain why systems outside of equilibrium evolve towards it with 'high probability.' Like before, this is actually the request for the reason why equilibrium-evolving behavior is typical for states outside of equilibrium. That is, we

[14] For more differences between typicality and probability, see Wilhelm (2019).

[15] For more on the connection between typicality and explanation, see Section 3.3.

want to show that the size of the set of exceptions, namely the set of microstates starting outside of equilibrium which will not go toward equilibrium, is very small with respect to the size of the set of microstates outside of equilibrium which will go to equilibrium.

Whit these qualifications, one can answer to the one-million-dollar-question about entropy increase:

4) We want to explain why entropy has an extremely 'high probability' of increasing. That is, since all that is needed is an estimate of the entropy rather than a precise computation, we want to explain why typical states increase their entropy during their temporal evolution. That is, we want to show that typical states in equilibrium will remain there (maintaining their entropy constant) and typical states outside of equilibrium will evolve towards it (increasing their entropy).

So, since in Boltzmann's account using the uniform measure as a typicality measure one can show all of the above, then Boltzmann's account explains the macroscopic laws as well as the macroscopic asymmetry in terms of time-symmetric microscopic laws, even without invoking the notion of probability.

3.2 Typicality and Probability

However, one could object that we still have not responded to the original question about the second law of thermodynamics. In fact, one could complain that, since typicality is not probability, one can only say that the second law is typical, but not that it is probable. And this seems wrong. Thus, one should further investigate the connection between the two notions: more often than not, we express probabilistic statements and not typicality statements when describing a macroscopic phenomenon in this way.

Indeed, one may argue that the notion of typicality grounds the one of probability. That is, probability theory is a mathematical idealization developed in order to make precise the meaning of 'vast majority' through theorems like the law of large numbers. It is that notion which has to satisfy the axioms of probability, not the typicality measure from which it is derived. Be that as it may, one could also outline the connections between the various senses of probability, typicality and their explanatory role. The situation is not simple and presumably one should write another paper on this.[16] Hoverer, here are some simple considerations. First, there is a straightforward connection with the notion of subjective probabilities as degrees of belief. In fact, the measure $\mu(A)$, where A is for example the set of microstates for which the entropy grows, can be interpreted as the degree of belief of an observer making an inference on the probability that the entropy will increase. Thus, the measure plays a role in the justification of why we have a reasonable degree of belief about the growth of entropy. Nonetheless, since this is someone's degree of belief, it cannot help in explaining physical phenomena, given that only an objective notion has the potentiality of doing that (recall Albert's objection to epistemic accounts of the probability measure).

[16] See also Maudlin (2020).

Perhaps more interesting is the connection with probabilities as relative frequencies. When an experimenter prepares a set of repeated experiments with substantially identical initial macroscopic preparations, she will obtain empirical regularities. These empirical distributions are the relative frequencies of the various outcomes. Take a set of N gases concentrated one the corner of N similar boxes; let them evolve freely; check what has happened after 2 hours, say; record the N results: the first gas spreads out in the first box; the second spreads out in the second box; and so on. In general, the empirical distributions present statistical patterns: for instance, the vast majority of the gases spreads out in their container. Now, how do we explain the empirical distributions? We show that the distribution predicted by the theory describing the phenomena in question matches the one observed. In this case, we need to show that the observed distribution ρ_{emp} and the theoretical distribution predicted by statistical mechanics ρ_{theo} agree. That is, we need to show that, theoretically, the vast majority of gases expands when evolving freely. That is, one would then have to show that, with the right measure of typicality, the theoretical distribution and the empirical one are very close. Formally, one needs to show that $|\rho_{emp} - \rho_{theo}| < \varepsilon$., with some positive constant ε small at will and with the distance measured by the typicality measure. Having said that, then, the connection between typicality and probability as empirical frequencies is that typicality provides the measure with which one can compare the empirical and the theoretical frequencies.[17]

3.3 Explanation Based on Typicality

To sum up the results of the previous subsections, in this account to explain why a given regularity occurs is to explain why the regularity is typical. This can be done specifying the laws of nature and the typicality measure (in addition to the past hypothesis).[18] Formally, as discussed by Isaac Wilhelm (2019), this is the explanatory schema based on typicality: if an object x has property A (or belonging to a given set), and a property B which is typical for A-type objects, then the explanation of why x has B is given by this argument schema [typ]:

$$\frac{\begin{array}{c} x \text{ is } A \\ \text{Typically, all } As \text{ are } Bs \end{array}}{\therefore \text{ x is } B}$$

That is, the explanation of the fact that a given system has a given property is given by showing that this property is typical for objects of that type. In other words, if one has shown that B is typical of A -objects (the objects that belong to the set of objects having the property A), the one has explained why x has the property B too. For instance, the explanation of the fact that this gas expands is given by the fact that one can prove that free expansion is typical for gases. That is, for the vast majority of initial conditions one can show that gases expand.

[17] For more on the relation between typicality and probability, see Volchan (2007), Goldstein (2011), Pitowsky (2012), Hemmo and Shenker (2012), Wilhelm (2019), Maudlin (2020).

[18] This is similar to what Albert and Loewer conclude, with the only difference that here we have typicality rather than probability.

A longer discussion would be appropriate, but let me observe that the explanatory schema sketched above shares some similarities with Hempel's covering law model (1965), whose main idea was that explanations are arguments with laws of nature as premises. The idea is roughly that if one finds the relevant law that made the phenomenon happen, one has found the explanation for why it happened. Here, explanations are also arguments, but the explanation is given not by nomological facts but by typicality facts. However, the difference between the two types of facts is not in kind but merely of degree: they are both nomological facts, but while laws of nature are exception-less, typicality facts are not. That is, one can think of typicality facts as nomological facts which allow for rare exceptions. Notice that this is compatible with typicality explanations being used in a macroscopic context in which it seems fine to allow for exceptions, while it is not used at the fundamental level, in which exceptions are seen as problematical. If so, the parallel with the deductive-nomological (DM) model is striking:

$$\frac{\begin{array}{c} x \text{ is A} \\ \text{All } As \text{ are } Bs \end{array}}{\therefore\ x \text{ is } B}$$

That is, if one can show that there is an exception-less regularity such that all As are Bs, this explains why this x, which is an A, is also a B. Similarly, in the typicality schema, if one can show that B is typical of A-objects, this explain why this x, who is an A, has also property B. Nonetheless, the schemas are not identical: while in the DN one can deduce that x is a B from the other premises, this is not so for the typicality schema because there are exceptions. That is, no deductive derivation can be provided, not even in principle: in the case of Boltzmann, for instance, there will always be anti-thermodynamics states. Thus Boltzmann proved 'merely' that most systems have increasing entropy, not that all of them did. Accordingly, the DN model could not be used to explain the laws of thermodynamics, but the typicality schema could.[19]

In addition, there is a sense in which the typicality explanation is similar to Hempel's inductive statistical (IS) model. The original idea was that if one can prove that two features occur together with high probability then one explain why something possessing one also possesses the other. Formally, an explanation is provided by the following inductive argument:

$$x \text{ is } A$$

[19] Moreover, as in the DN model, also in the typicality schema every explanation is a prediction: a given gas (x) is in non-equilibrium (A); it's typical for non-equilibrium gas to expand freely (B); so this gas expands freely. This is an explanation of why the gas in this box has expanded, or a prediction that it will expand, if the gas has not been observed yet. However, unlike the DN model, the typicality schema has not the problem of asymmetry. Let's recall what this problem is. Given the height of a flagpole, the height of the sun on the horizon, and the laws of optics (and trigonometry), one can deduce the length of the shadow cast by the pole. This is an adequate explanation as well as a prediction of the shadow's length of a given flagpole of that height, observed or not. However, one could also use the length of the shadow, together with the height of the sun on the horizon, and the laws of optics (and trigonometry), to deduce the height of the flagpole. This is an accurate prediction, but not an explanation of the height of the pole. Critics track this down to the idea that explanation is asymmetric while deduction is not. This is not true for the typicality explanation, because the explanation is not deductive but follows the schema above.

$$\frac{Prob(A,B) = r}{\therefore x \text{ is } B}[r]$$

where r is the strength of the induction. For instance, if this apple (x) is red (A), and if one can show that there is a high probability (r) that red apples are also sweet (B), then one has explained why this apple is sweet.[20] Compare this with the typicality schema [typ]. According to the IS model, a phenomenon of type A is explained to be also a B by showing that the probability of some$A$ being also a B is high. In typicality account a phenomenon of type A is explained of being also a B by showing that being a B is typical of being an A. So these schema are very similar, if not formally identical. The difference is in the logic operators (Typ vs. $Prob$), and we have seen that typicality is not probability.[21] Some probabilistic explanations are not typicality explanations: consider a nucleus of the element X having the probability of decaying being less than 0.5. Then, one can explain why this nucleus has decayed using the IS model, providing therefore a probabilistic explanation. However, it would not be typical for X to decay, since r is small.[22]

Be that as it may, I think that the reason why the apple argument in terms of probabilities appears to be explanatory (to those who think it is explanatory) is because what one has in mind is a typicality argument instead. That is, we could have said: science has shown that red apples are typically sweet, and that is why this red apple is sweet. Details on the strength of the induction (whether it is 0.98 or 0.97) do not matter much, as long as r is large. This is typical of typicality reasoning (pun intended), not of probabilistic reasoning. This is compatible with Hempel's original idea that the IS model works only for large probabilities.[23] Also, to make the point that typicality explanations are not probability ones, one could point out that probabilistic reasoning can be counterintuitive, while typicality reasoning is not. For example, think of the probability of having a rare (one case over 10,000 people) but terrible disease when testing positive for it, assuming the testing methods has 99% of accuracy. Most people will be terrified, but they would be wrong: using Bayes theorem your probability of having the disease is 0.01. So, testing positive but not having the disease is, arguably, explained by the fact that Bayes theorem show that the probability of having the disease was low. However, this is extremely counterintuitive. In contrast typicality arguments are not like that at all. They are used all the time in everyday reasoning arguably because they are very intuitive: typically, dogs are affectionate, and that is why my dog is affectionate. Interestingly, this is connected to the use of stereotypes: we immediately nod in agreement when someone says that our husband's bad behavior is explained by the fact that will all probability all men are jerks. We use probability talk, but since details do not matter, we are actually

[20]I am assuming that these claims are explanatory. People have objected to this, but pointing at this misses the point: all I am saying is that as long as the IS model is explanatory, so is the typicality model.

[21] See Crane and Wilhem (2020) for two proposals for a formal logic for typicality arguments, one based on propositional modal logic, and the other on intuitionistic logic.

[22] See Wilhelm (2019) for more examples.

[23] Later, critics of the model argued that also small probabilities can be explanatory. See Strevens (2000) for a nice summary and an argument that the strength of the induction doesn't necessarily correlate with the strength of the explanation. This does not affect my point because the argument is not that all explanations are typicality explanations.

using a typicality explanation: typically, men are jerks.[24] Similarly, I think, we may use probability talk to present Boltzmann's explanation of the macroscopic laws, but we actually have in mind typicality: typically, entropy increases.

On another note, as pointed out by Wilhelm (2019), notice that there seems to be the right correlation between typicality and causation. The claim that 'woodpeckers typically have a pointy beak' may be further explained citing some genetic factor, which may play a causal role in beaks having a pointy shape. Notice however, that this is itself a typicality explanation: those genes typically give rise to such beaks. This shows that the typicality schema is not wedded to the covering law model, even if it shares with it some important features.

Aside from the comparison with the covering laws model, let us explore whether the explanation based on typicality is adequate. Among the desiderata for a satisfactory scientific explanation, one can list the following: informativeness, predictive power, and expectability. A phenomenon is suitably explained if the account is able to provide an informative and concise description of the phenomenon in question, and typicality is able to do that, as also pointed out by Wilhelm (2019). It is informative because it tells us about the behavior of the vast majority of systems: typical systems approach equilibrium. And it does so very succinctly, in a single sentence, merely using the uniform measure as typicality measure to count the number of microstates. Notice that it is informative only because it tells us about the typical behavior, namely the behavior that the vast majority of systems will display. Otherwise, one would not know, by merely accessing the macrostate, whether the system's microstate will evolve thermodynamically or not.

Moreover, as also emphasized by Dustin Lazarovici and Paula Reichert (2015), typicality can provide us with predictive power, which is not surprising, given what we observed above in relation with the connection with the covering law model. In fact, if one shows that a given property is typical, then one has reasons to expect to see this property in other systems similar to the original one. Given that one has proven that a typical system will approach equilibrium, this is what one should expect a system to generally do. Notice how, obviously, this predictability would not be possible if one had shown that only *some* states, not the great majority of them, approach equilibrium. Consider a gas in the corner of a box. If one could only show that some state will approach equilibrium while other will not, what can one conclude about the behavior of a given gas which hasn't been observed yet? Not much, since the size of equilibrium-approaching states and the size of non-equilibrium-approaching states are comparable (in our pretend-example). It is only when we can rule out the possibility of non-thermodynamic states because the size of the corresponding states is incredibly small that we can predict what is likely to happen in similar circumstances.

[24] Notice that I am not saying that stereotypes, by themselves, can explain: in contrast, stereotypes do not really explain unless someone actually shows that they hold typically. The point was merely that stereotypes are commonly used in accounting for certain phenomena happening, in contrast with 'true' probabilistic explanations, in which the details mater, which are notoriously difficult to grasp (see the above mentioned use of Bayes theorem).

In addition, typicality provides an explanation which holds for most initial conditions: a phenomenon has been explained if it holds for typical initial conditions. That is, with rare exceptions as defined by a suitable measure of typicality. As Bricmont (2001, 2020) has suggested, if something is typical, no further explanation seems to be required. For instance, the fact that this particular gas expands is not surprising, once Boltzmann's account has been provided. What would be surprising is if it did not expand. A less satisfactory explanation would be, on the contrary, one which is true only for very peculiar initial conditions. In fact, too many things could be explained by appealing to special initial conditions, for one can always find an initial condition that will account for the phenomenon. If we accept this type of explanations, allowing for fine-tuned initial conditions, then one could 'explain' everything. Let me elaborate. There is a sense in which to explain a phenomenon is to remove the surprise in seeing this phenomenon happening. This is related to what Hempel had in mind with his notion of 'expectability.' We are not surprised that a piece of salt will dissolve in water because we know the reason why it happens: the positive ions in water (H+) attract the negative chloride ions, and the negative ions in water (O--) attract the positive sodium ions. However, consider the case of a monkey who, by randomly hitting computer keys, ends up writing the "Divine Comedy." One can account for this fact by cherry picking an initial condition for which this actually happened. So, there is a sense in which the phenomenon is 'explained.' However, the fact that the monkey ended up writing that book is extremely surprising, and by pointing to a special, perhaps unique, initial condition that made it true does not help much to remove the surprise. So, there is a sense in which we are not completely explaining the phenomenon if we rely on special initial conditions. In other words, monkey writing books is not something that we expect. This is not something that monkeys typically do, because monkeys, typically, cannot read or write, for starters. It is not impossible that they write a wonderful book: indeed, it could just be a very lucky set of keyboard strokes. However, if they end up writing this wonderful book, we find it surprising. That is why I used the world 'lucky' in the previous sentence: the event has happened because of a ridiculously special initial condition. A slightly different initial condition would not have brought about a similar event. Instead, if one were to point out that most initial conditions would have the same outcome, the surprise will cease. In other words, a satisfactory explanation is one which, for the majority of initial conditions, monkeys would indeed write books like the "Divine Comedy." However, this typically does not happen: monkey randomly hitting computer keys would typically write gibberish. So, when asking for an explanation of a given phenomenon what we are actually asking for is a reason *why we should not find the phenomenon surprising*, and the response is that the phenomenon happens for most of the initial conditions. Moreover, in this framework, if we rely on a special initial condition to account of why a phenomenon has happened, we are not truly providing an explanation for it because we are not removing the surprise element. Or, in other words, the 'explanation' lacks expectability. Something similar happens in statistical mechanics: for the great majority of initial conditions, entropy will increase, not just for a special one. Because of this, the surprise is removed and the phenomenon explained. Moreover, one can account for why anti-thermodynamic states are not observed by pointing out that they are atypical, but Boltzmann's statistical mechanics does not provide an explanation of *why* atypical states happen, other than that they typically do not happen.[25]

One could object that this is something missing, namely we need to be able to explain *all* phenomena, including the atypical ones. That is fair enough; however let me just notice that the situation remains unchanged if we use probabilities instead, and explain improbable events. In any case, one could always rely on initial conditions: the only explanation of why *this* gas is not expanding, as opposed to that other one, is that this is what the initial conditions, together with the dynamics, bring about for it. Indeed, some have argued that the requirements for a satisfactory scientific explanation I discussed above are too strict: one would provide a satisfactory explanation of the phenomenon even if one relies on special initial conditions.[26] All that is required is that there is one such condition that would bring the phenomenon about. That may be so. However, as already seen, in this way one would not be able to account for the link between expectability and explanation: it would be difficult to understand why one would find monkey writing books surprising while one would not find surprising that gases expand when evolving freely.

Moreover, let me emphasizes that the idea of a satisfactory explanation being an explanation for most initial conditions is compatible with scientific practice in physics. For example, one of the reasons why Alan Guth (1981) proposed his theory of inflationary cosmology is that the big bang model requires strict assumptions on initial condition. In contrast, inflation would explain all the phenomena without relying on these special assumptions, and for this reason is considered a better theory: "The equations that describe the period of inflation have a very attractive feature: from almost any initial conditions the universe evolves to precisely the state that had to be assumed as the initial one in the standard model" (Guth and Steinhardt 1984).

In this respect and to conclude, let me add a remark regarding the past hypothesis. The past hypothesis has been introduced in order to break the past-future symmetry of the microscopic laws, by postulating that the universe had an initial very low entropy. One may think that this is a problem that undermines the whole account based on typicality: haven't we just said that a satisfactory explanation should hold for the typical initial condition? In contrast, a low entropy state is a very atypical state. So, have we ended up explaining what is typical by postulating something atypical? Isn't that bad? One thing that can be said to mitigate the problem is that, with respect to this low entropy initial macrostate, the initial microstate of the universe is typical in regard to its future evolution, which accounts for the entropy increase.[27] However, I honestly do not think this is helping

[25] The notion of expectability and the one of typicality go together as long as the world is typical. In fact in a typical world, entropy increases, and expectability goes with typicality. However, in an atypical world, one would expect something different than what it is typical, as in that world entropy would decrease. This is because what is expected comes from what is 'usual' in our world, and only in typical worlds what is usual and what is typical are the same. Thank you to Katie Elliot, Barry Loewer and Tim Maudin to make me be explicit on that.

[26] See, e.g., Valentini (2020). Also, see Myrvold (2020).

[27] See Lazarovici and Reichert (2015).

a lot, given that the initial macrostate is incredibly small. Indeed, many think this is a serious problem and propose mechanisms to make the initial state typical.[28] In contrast, Humeans such as Craig Callender (2004) argue that there is no need to explain the past hypothesis because, in a Lewisian fashion, it is simply one of the axioms of the best system of the world.[29] Notice, however, that from the point of view of the typicality account, the situation gets better, rather than worse: add the typicality, rather than the probability, postulate, and everything follows, without any additional need for explanation.

3.4 Objections to the Typicality Account

The typicality approach as applied in statistical mechanics described so far is what Massimiliano Badino (2020) dubs the "simple typicality account," or STA. In this view, the approach to equilibrium is explained entirely in terms of the size of the macrostates, by showing that the vast majority of microstates will fall into the equilibrium state. However, according to some critics, this approach dismisses the dynamics, which apparently plays no role in explaining the approach to equilibrium. Accordingly, Roman Frigg and Charlotte Werndl propose the one Badino calls "combined typicality approach," or CTA (Frigg and Werndl 2012, Werndl 2013, Werndl and Frigg 2015). The idea is that one should show that equilibrium is approached for the typical *dynamics*, in addition to showing it is approached for the typical initial condition. Frigg and Werndl prove that typical Hamiltonians produce systems which are epsilon-ergodic, namely they are such that the time spent in a macrostate is proportional to the size of the macrostate. Because the equilibrium state is the largest of the macrostates, an epsilon-ergodic system will spend in it most of the time, explaining in this way the approach to equilibrium of typical dynamics as well as for typical initial conditions.

In this regard, let me enter into some details about the STA and the role of the dynamics: it is not true that, strictly speaking, in the STA the dynamics is ignored. Frigg and Werndl prove that phenomena such as the approach to equilibrium are to be explained for *most* (typical) initial conditions and for *most* (typical) dynamics. Instead, Frigg and Werndl complain, the STA merely does that for most initial conditions, forgetting about typical Hamiltonians. However, this is not so. Indeed, the STA manages to do something *more*, rather than less, *general*. That is, in the STA one is able to account for phenomena such as the approach to equilibrium dynamics for most initial conditions, *without* being specific about any feature the dynamics needs to have. That is, the phenomenon is explained for *most* (typical) initial conditions and for *all* dynamics. This is why the dynamics is never mentioned: not because it is irrelevant, but because for most initial conditions the details of the dynamics do not matter, and the system will reach

[28] To this end, Penrose (1999) for instance proposes his 'Weyl curvature hypothesis' as an additional law in order to explain how the low entropy initial state is not atypical. In addition, Carrol and Chen (2005) put forward a model whose purpose is to completely eliminate the past hypothesis. See Lazarovici and Reichert (2019) for a proposal built on Chen and Carrol's model.

[29] For more on this approach and its challenges, see also Olsen and Meacham (2020).

equilibrium *regardless* of the Hamiltonian. For more on this 'genericity' of the Hamiltonian, see the next section.

Finally, to counter the idea that the dynamics plays no role in the STA, in the next section I show that the statistical postulate can be derived from the dynamics. If so, the dynamics plays a big role in the STA, namely the role of selecting the typicality measure, henceforth reducing the gap between the STA and CTA.

4. How to Dispense of the Statistical Postulate: The Stationarity Argument

Up to now we have simply assumed that the choice of the measure, typicality or probability, had to be postulated. In this section I wish to explore what I take to be the proposal put forward by Goldstein (2001, 2011), Dürr (2009), and Zanghí (2005). That is, the proposal that the typicality measure is derived from the dynamics introducing suitable symmetry constraints. I find it extremely surprising that this argument has received very little attention in the literature,[30] because not only it provides a non-epistemic justification for the uniform measure but also shows how the dynamics plays a crucial role in the typicality account.[31]

The main idea, I take it, is that the typicality measure is the uniform measure (the Lebesgue-Liouville measure) not because it is uniform, but because of these two features: 1) it is *time-translations invariant*; and 2) it is *generic* with respect to the Hamiltonian of the system. Let's see what these features amount to starting with the first. A measure is time-translation in variant when the volume it defines in phase space is conserved. That is, if A is any set in phase space, and A_{-t} is the set of points in phase space that evolve into A after a time t, then A and A_{-t} have the same volume (this is Liouville's theorem). The reason why time-translation invariance, also called stationarity, is a requirement for the typicality measure is connected with the idea that no temporal instant needs to be privileged. In fact, as Goldstein, Dürr and Zanghí notice, the measure counts the space-time histories of the universe, while the phase space point is just a convenient way of representing them. So, when counting the histories, one needs to regard the initial time merely as conventional, by not privileging any particular time, and a time-translation invariance measure would guarantee that. The uniform measure is not the only time-translation invariant measure. In fact, given a conservative force, there could be other measures, whose explicit form depends on the particular law of the force. However, it is argued, one can single out a

[30] Aside from the comments of Bricmont (2001, 2020), discussed in Section 4.1.

[31] Itamar Pitowsky (2012) proposes a justification of the uniform measure as the typicality measure which has been later criticized by Werndl (2013). She proposes that the choice of the typicality measure should be done using symmetry consideration, in particular the typicality measure should be invariant with respect to the dynamics. Maudlin (p.c.) maintains that there are other ways of justifying the use of the uniform measure. For instance, one could say that the uniform measure is privileged because it is the one that phase space inherits from the spatial measure. However, considerations like this will have less weight when one moves to the quantum domain, especially in the context of the pilot-wave theory. In that context in fact the choice of the typicality measure is justified by an argument which is a direct analog of the stationarity argument discussed in this section, and this is one of the reasons I think it is worth exploring.

unique measure by requiring an invariant measure that is also *generic* with respect to the dynamics: that is, it is *independent* of the particular law of the force. Tying the measure to the dynamics arguably makes the choice of the uniform measure *natural*: the uniform measure is the typicality measure because it is the only stationary generic measure.

This argument, which I will call from now the 'stationarity argument,' can be therefore summarized as follows:

P1: The purpose of the typicality measure is to count microstates (definition);

P2: One can do this at different times (definition/construction);

P3: One should not privilege one time over any other (the initial time is conventionally chosen)

subC : Thus the typicality measure should be time-translation invariant or stationary;

P4: If the dynamics is Hamiltonian, the typicality measure should be independent from the specific form of the Hamiltonian (genericity);

P5: The only measure satisfying these two requirements is the uniform measure (mathematical proof);

C: Thus the uniform measure is the typicality measure.

4.1 Objections to the Stationarity Argument

The stationary argument for the uniform measure as typicality measure has been criticized most notably by Bricmont (2001), who argues that there are several problems.

First, the argument applies only to Hamiltonian systems. As we just saw, Hamiltonian systems are volume preserving: they are such that the phase-space volume, defined by the uniform measure, is preserved under time evolution. Instead, Bricmont urges us to consider certain dissipative systems like dynamical systems with a chaotic attractor. These systems usually give rise to solution flows which contract volumes in phase space. This volume contraction gives rise to a set in phase space called an *attractor*, toward which solutions ultimately evolve. Certain dissipative systems are chaotic, that is they show sensitive dependence of initial conditions. For chaotic systems the attractor is special because merely studying the dynamics on the attractor is sufficient to have information about the overall dynamics. As Bricmont points out, the uniform measure is not time-translation invariant for these systems (because the volume is not conserved) but it is still the typicality measure: for most initial conditions in the basin of attraction, counted using the uniform measure, the relevant empirical frequencies will be correctly reproduced. How do we justify this choice? According to Bricmont the uniform measure is chosen as typicality measure not because it is stationary (which it is not). Rather, the justification is more generally based on what Bricmont calls "Bayesian grounds" (Bricmont 2020): *the typicality measure is whichever measure that reproduces the relevant empirical frequencies*, regardless of whether it is stationary or not. This argument, according to Bricmont, provides a more general account of the reasons why the uniform measure is the typicality measure which holds for non-Hamiltonian systems. So, why should one need another argument for Hamiltonian systems?

Another objection raised by Bricmont against the stationary account of the typicality measure is that it is too sophisticated to genuinely be part of the scientific explanation of thermodynamic phenomena. In fact, he notices that we are trying to provide an account of scientific explanation, and as such the proposed account should be not too distant from our intuitive understanding of the notion of explanation. Because of this, time-reversal invariance does not seem the right notion, since when we give our folk explanations we never use anything that remotely resembles stationarity.

To conclude, let me add an objection which challenges the premise requiring genericity for the measure. Indeed, I find this to be the weakest point of the stationarity argument as nowhere it has been argued why such a constraint is required: why would one want the measure to be independent of the Hamiltonian, hence the potential? If the whole point of caring about Hamiltonian systems was to ensure that they have the chance of describing the universe, since there's only one universe, there's also only one Hamiltonian. So why one would care about other possible dynamics?

4.2 Replies to Objections to the Stationarity Argument

First, consider the objection based on chaotic dynamics: if for non-Hamiltonian systems the uniformity measure is chosen as the typicality measure without using stationarity considerations, why use them for Hamiltonian systems? This is, in my view, a well-thought objection to the stationarity account of the typicality measures, and as such should be carefully considered. The best response to this, I think, is the following. Upon reflection, one should not find it surprising that for non-Hamiltonian systems the uniform measure is the typicality measure even if it is not stationary. Indeed, I think it would be unreasonable to expect that stationarity would pick out the correct typicality measure in these cases. In fact, while in Hamiltonian systems the volume in phase space is preserved, and the natural weighting of all points in phase space is equal, this is not so for non-Hamiltonian systems because the system is dissipative and the attractor carry more weight. And this in turn is not surprising because non-Hamiltonian systems are open systems. They are dissipative system, so there has to be somewhere they dissipate to. Because of this, they are *less general* than Hamiltonian systems, in the sense that while Hamiltonian systems may reproduce the behaviour of the whole universe, non–Hamiltonian systems at best can reproduce the behaviour of open subsystems. To put it in another way, the universe is a Hamiltonian system, while non-Hamiltonian systems may be descriptions of non-isolated subsystems. If so, the stationarity constraint needs to be fulfilled to select the typicality measure for the universe, not its subsystems. Then, when describing dissipative system, which are open subsystems of the universe, the choice of the typicality measure has been *already made*. Then one looks for other 'natural invariant' or 'physical' measures, such as the Sinai-Ruelle-Bowen (SRB) measure, defined on the attractor, which is time-translation invariant. However, when considering that these systems are merely subsystem of a more general Hamiltonian system, the apparent paradox (why the SRB measure, which is stationary, is not the typicality measure, while the uniform measure is, even if it's not stationary?) is resolved: the uniform measure is selected as the typicality measure because it is the stationary measure for the universe, and whatever is stationary in a subsystem does not really matter. Bricmont's question ('why one needs to look for an additional justification for Hamiltonian systems based on stationarity if one already accepts the Bayesian justification for non-Hamiltonian systems?') seems compelling only if one considers non-Hamiltonian systems to be more general than Hamiltonian ones. However, in this context, this is not the case, and the logic is reversed: first one looks at the universe (which is described by a Hamiltonian) to find the typicality measure using stationarity (and this is the uniform measure); and then at its subsystems (which are not). When dealing with them, stationarity is no longer relevant and the typicality measure is the uniformity measure because it has been inherited by the one for the universe.

Now, let us move on to the second objection that the stationarity argument is too sophisticated to genuinely capture what we mean by explanation. In response, one could say that the ingredients of a scientific explanation need not to be familiar to us to make the theory an adequate account of scientific explanation. One judges a theory of explanation to be adequate if it is able to reproduce our intuitions regarding which truly are explanations and which are not, not necessarily using notions that we are already familiar with. For instance, as we have seen, in the DN model of explanation, an explanation is a valid deductive argument in which (at least) one of the premises is a law of nature. The adequacy of this model is judged by considering whether every explanation as given from the model is also intuitively an explanation, and *vice versa*. Indeed, one of the counterexamples of the model points to asymmetry in the account that contradicts our intuition. One in fact can derive the length S of the shadow cast by a flagpole from the height H of the pole and the angle θ of the sun above the horizon and laws about the rectilinear propagation of light. This derivation is thus an explanation according to the deductive nomological model, and that seems right. However, the 'backward' derivation of H from S and θ, which is also an explanation according to the model, intuitively does not seem explanatory. While it makes sense to say that the shadow of a flagpole being a particular length is explained by the flagpole having a particular height, we do not explain the flagpole having a certain height in terms of the shadow being of a particular length. Rather, the flagpole has that height because it was constructed that way, for other reasons. The model of explanation has therefore to pass the test of intuition. Nonetheless, this is a test for the *outcome* of the model, namely what counts as an explanation, rather than for the *ingredients* used to arrive to the model's outcome. In this case, one may notice, also the ingredients used in the model to generate explanations are familiar: deductive arguments and laws of nature. However, it does not seem to me that using other unfamiliar notions would be problematic, unless they appear in the explanation itself. In fact, consider Boltzmann's account of thermodynamics. We have said that the second law, say, is explained by claiming that it can be shown that 'entropy-increasing behaviour is typical.' The sentence within quotes is the explanation. The notions in it are 'entropy' and 'typicality,' which have an intuitive meaning that is qualified and made precise in the process of working out the explanation. Moreover, even if the notions involved to derive such an explanation are far from being intuitive, they are built bottom-up from intuitive notions into sophisticated mathematical notions. Think of the notions of macrostate, microstate, measure, and so on. Notice that also stationarity is like that. It is a notion which is far from intuitions, which however is connected with the intuitive idea that no temporal instant should be privileged. The explanation of the choice of the typicality measure uses intuitive notions, through its implementation with the sophisticated notion of stationarity, to explain physical phenomena. In other words, it is this intuition that provides the reason why stationarity enters the explanatory machinery. As Einstein (1936) said: "the whole of science is nothing more than a refinement of our everyday thinking". However, this refinement can lead us far from intuition without losing its legitimacy. Be that as it may, let me conclude that, interestingly, Bricmont agrees that 'entropy-increasing behaviour is typical' is an explanation of the phenomena, even if it includes the notions above (which are built from intuition into mathematical notions without being themselves intuitive). However, he maintains that stationarity is not explanatory because it is not a notion we commonly use. Nonetheless, it seems inconsistent to complain about stationarity being 'counterintuitive' and thus not explanatory, if one agrees that typicality is 'counterintuitive' but explanatory.

As far as Bricmont's own view (Bricmont 2020), he thinks that the best way to justify the uniform measure is in a Bayesian framework. In this account, the probabilities used in statistical mechanics are seen as epistemic: they express our ignorance and are used to update one's probabilities estimate when new information becomes available. In this way, the uniform measure is taken to be a generalization of the principle of indifference, according to which one should not introduce any bias, or information that one does not have. Since this account resembles the accounts that Albert criticized in his book, there may be problems as to how is it that our ignorance can *explain* the behavior of objects. In particular, in such a Bayesian account, two people may disagree about what counts as 'vast majority', and thus they may disagree about what counts as typical. Since typicality is used to explain, they may arrive to different explanation for the same physical phenomena (or may fail to explain some phenomena), and this is far from being desirable because we want explanations to be objective. Perhaps, a response to this kind of arguments would be that explanation based on typicality only requires 'coarse-grained' constraints (such as that the number of non-thermodynamic states is overwhelmingly smaller than the number of thermodynamic states) and therefore no two people may actually disagree on what is typical and what is not. Otherwise, one may want to link this notion of entropy to the notion of rationality, which is what Bricmont (2020) suggests. However, this account seems to share very similar objections as the epistemic view, as well as new ones: what is rationality? More discussion on this is needed, but here I will merely recall that the debate over rationality and rational decision making in Everettian mechanics, even if the context is different, is still wide open.[32]

Finally, let me discuss the objection that it is unclear why the measure needs to be generic under the dynamics. Here is a possible answer. While it is true there is a unique true Hamiltonian H, when looking at subsystems within the universe this may be different: one can use some effective Hamiltonian H_{eff} (if the subsystem of the universe is still Hamiltonian), which is an approximation of the true one, H, for that subsystem under

[32] See Wallace (2012) and references therein.

consideration. If we use H to find the typicality measure, we will find many stationary ones; likewise, if we use H_{eff}, we find another bunch of stationary measures. Which one should we choose? Since both H and H_{eff} should give rise to the same empirical results (otherwise we have done a bad approximation), then one should require that the stationary measure they find be the same. In addition, one could argue[33] that in order to provide a genuine explanation one would have to require genericity: it is not the specific form of the Hamiltonian that gets the fact explained, it is not because of some special fact of the Hamiltonian that the fact is explained; rather it is explained independently of what kind of Hamiltonian we have. Moreover, and perhaps more importantly, notice that the request for genericity is what guarantees that the dynamics is 'irrelevant' in the sense discussed in Section 4.2: the phenomena are explained for typical initial conditions and for all (rather than for typical) Hamiltonians. That is, the genericity of the Hamiltonian is what guarantees that explanation is so general that the details of the dynamics do not matter at all.

4.3 Boltzmann's Ingredients in the Typicality Account

To conclude this section, let me summarize the situation. We started from the characterization of Boltzmann's explanation of macroscopic laws in terms of the classical dynamical laws, the statistical postulate and the past hypothesis. Some reflections on the statistical postulate lead us to the conclusions that:

1) Typicality (not necessarily probability) is enough to explain;
2) The correct typicality measure (namely the one that proves empirically adequate) may be inferred from the dynamical laws using symmetries considerations, and therefore not postulated.

If the arguments presented here are sound, one could conclude that the statistical postulate is not needed because the typicality measure is suitably derivable from the dynamics. As a consequence, the ingredients of Boltzmann's explanation are now reduced to the following:

1) The laws of motion (which determine the typicality measure on phase space);
2) The past hypothesis.[34]

5. Quantum Statistical Mechanics

Now that we have discussed Boltzmann's account in the classical domain, let's discuss about its possible generalizations. If one wishes to generalize Boltzmann's explanatory schema to the quantum domain, *prima facie*, one should not expect some fundamental differences, especially in the case of deterministic theories like for instance the pilot-wave

[33] Maudlin (p.c.).

[34] Interestingly, but from a very different perspective, Eddy K. Chen (2020) has recently argued that one can dispense of the statistical postulate in a quantum extension of statistical mechanics by assuming that the ontology is given by the density matrix rather than by the wave-function.

theory.[35] However, the spontaneous localization theory provides something new, namely intrinsic stochasticity. In fact, this is a theory in which the wave-function does not always evolve according to the Schrödinger equations. Rather, it does for some time, then at a random time the wave-function 'collapses' into a random localization, then it continues to evolve according to the Schrödinger dynamics, and so on. Albert (2001) argued that this theory can provide a *dynamical* explanation for the statistical postulate. If so, the statistical mechanical probabilities would just be the quantum mechanical probabilities. That is, Albert argues that one can dynamically derive the statistical postulate if the spontaneous localization theory is true. In the last section I also argued that one can dynamically derive the statistical postulate, even if the arguments are very different. In this section I plan to show that this approach can be extended to all theories, including indeterministic quantum ones. Before entering into this, let us present Albert's argument for his thesis.

Since the overwhelming majority of microstates is thermodynamically normal (that is, entropy-increasing), they are stable. In fact everything close to them is also likely to be normal. In contrast, abnormal microstates (that is, entropy-decreasing) are very unstable, since they are surrounded almost always by normal microstates. Because of this, any abnormal system is extremely close to being in a normal state. Albert's idea is that the effect of a wave-function collapse like those happening in the spontaneous localization theory, with overwhelming likelihood, will keep a normal microstate normal, and will make an abnormal microstate 'jump' into a normal one. In the theory, to technically implement the collapse the wave-function is multiplied by a Gaussian, which effectively restricts the support of the wave-function to a random and very small region of space. In this sense, there is a set of random and small macrostates (the regions after the collapse) to which the wave function can go to at any time, each of which with its own probability distribution, given by the quantum rules. That is, there is *automatically* a probability distribution on each macrostate, *in contrast* with deterministic theories, where one has to add it by hand. The region after the collapse is, by construction, smaller than any region possibly representing a macrostate. However, the size of the set of abnormal states is much smaller, so one could still say that the vast majority of microstates in the collapsed region will go toward a higher entropy state, in agreement with the statistical mechanical predictions. Therefore, Albert claims that the spontaneous collapse theory can do away with the statistical postulate: it is the dynamics itself, being open toward the future and assigning a probability to each possible collapsed region, that fills in the gap between the microstate number talk and the probability talk.

5.1 The Indeterministic Case

Let's now explore how my approach may translate into the quantum domain and to indeterministic theories. Given the conclusions drawn in Section 4, namely that the statistical postulate is not needed and the relevant explanatory notion is derivable from the dynamics, one naturally wonders whether this extends to indeterministic theories as well. In Section 4 we considered Boltzmann's explanation when the fundamental theory of

[35] See Goldstein et al (2020).

the world is deterministic. As already discussed, it seems that in an indeterministic theory the dynamical laws are time-directed in the sense that while the past is determined the future is open, and each possible future has its own probabilities of happening. Albert argued that this is likely to help, getting rid of the statistical postulate. However, I argue, following Zanghí's suggestions (2005), that the differences are not as striking as one may think, and do not majorly affect the structure of Boltzmann's explanatory schema (with qualifications).

Indeterministic theories are notoriously difficult to get a grip on, so one may use the simple model proposed by Ehrenfest (Baldovin et al 2019, Ehrenfest 2015). This model can be taken as the prototypical example of indeterministic dynamics and can be used to study the differences and the similarities between deterministic and indeterministic theories (Zanghì 2005). Suppose there are N numbered balls in two boxes, one on the right and on the left. The Ehrenfest process is a discrete-time process which describes the indeterministic dynamics defined by a series of random jumps of the balls from one box to the other, one at a time. To describe this dynamics, somewhat oddly but ultimately usefully, one can think in terms of macrostates and microstates also in this context. The microstate at a given time is the list of ball locations in one or the other box at that time. That is, $X = (a_1, \dots, a_N)$, such that the i-th entry is 0 if the i-th ball is in the right box and is 1 if it is in the left one. At another time, take a ball at random from a box and put it in the other (or: let one ball at random 'jump' from one box to the other). Consequently, the microstate changes. For instance, assuming for simplicity there are only two balls, there are only four possible initial states: (1,1), both balls on the left, (0,0), both balls on the right, (1,0) and (0,1), respectively the first ball on the left and the second on the right, and the other way around. If the actual initial state is, say, (1,1), then the system can evolve at time $t = 1$ into (1,0) or (0,1) with probability ½ each. Assuming the state at $t = 1$ is (1,0), then, it can evolve at time $t = 2$ into either (0,0) or (1,1), again with ½ probability each. And so on. Now drop the simplification that there are just two balls. If at the beginning all balls are in the left container, that is $X_0 = (1,1,1,1 \dots)$, it can be shown that, as the number of jumps increases, the number of balls contained in each box will tend to be the same. If so, an example of microstate after a sufficiently long time T is $X_T = (1,0, \dots ,0,1 \dots)$, with an equal number of randomly distributed zeros and ones. This is so for combinatorial reasons, given that there are many more ways in which N balls can be distributed half in the left and half in the right box than any way. Thereby, one can define an 'equilibrium' macrostate as the state of all the microstates in which the balls are half in one box, and half in the other. We can call it equilibrium because after this configuration is reached, the systems will tend to stay there, just like in statistical mechanics.

Can we write a law for this behavior? A first possibility is to describe the law in pure probabilistic terms: if the system is at a given initial time $t = 0$ in the state X_0, at the time t the system will have certain probability (namely $1/2^N$) to evolve in one of the 2^N possible states. In this picture the past is fixed and the future open, and accordingly there is a fundamental difference between indeterministic and deterministic dynamics, in which both the past and the future are fixed. Nevertheless, Zanghí points out, there is another mathematically equivalent way of reformulating the law avoiding the postulation of a fundamental difference between past and future. Let us start from the microstate X_0 at $t = 0$. At time $t = 1$, there are N possible states; at $t = 2$, the different states are N^2; and so on, so that at $t = n$ they are N^n. In other words, from the initial microstate X_0, which describes how the balls are arranged in the two boxes, the microstate, namely the balls distribution, evolves in time as described by a branching graph (See figure 1).

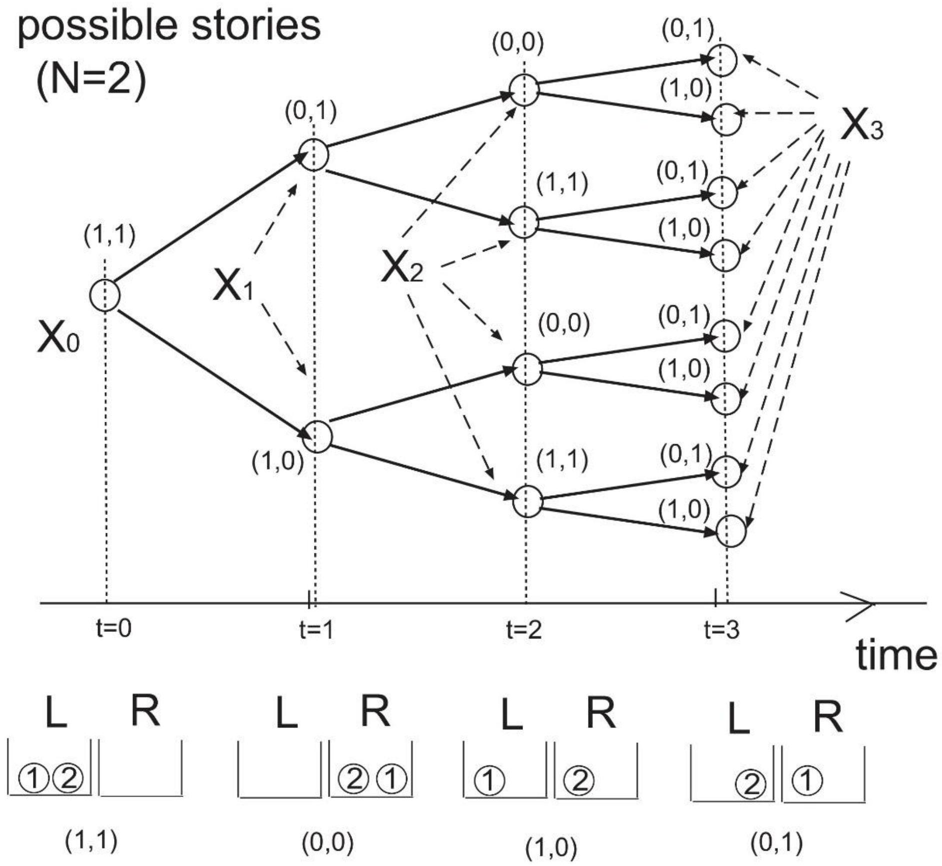

Figure 1: Possible histories in an Ehrenfest process. For graphical simplicity I have assumed that there are only two balls (N=2).

In a deterministic dynamics, one microstate evolves in time into another microstate, and therefore, the complete history of the world is a single curve in phase space. In contrast, in an indeterministic theory described in this way, given that multiple states become available to the incoming microstates, the history of the world involves a continuous ramification in phase space. Accordingly, the set of all the histories of the world for all the 2^N possible initial microstates is the union Ω of the graphs corresponding to the different initial microstates, which becomes the space of all the possible state of affairs of the world in the case of an indeterministic theory. This branching scenario is a pictorial representation of what having multiple future means. However, even if there are multiple future available, the system will invariably evolve into only one of them. That is, if we assume a timeless view and look at the space of possible histories, there will *always* be single histories, namely single sequences of states at different times: one sequence for each path the system at one time will have taken at a later time.

Now, let us go back to Boltzmann's explanation. One of the ingredients was the typicality measure, which is a measure in phase space, given that phase space represents the space of all the possible histories of the world. However, in this context, the space of the possible histories of the world is Ω, namely the union of the graphs corresponding to each possible initial microstate. Accordingly then, the typicality measure will have to be taken on such space and thus applied to all the possible histories. If the stationary account of typicality is correct, then, the indeterministic dynamics will determine as typicality measure the one which is stationary on the space of possible histories. Namely, the one does not privilege any temporal instant.

One can define the notion of time-translation invariance in this context as follows: assuming that A is a set of possible histories and that A_{-t} is the set of these histories translated back in time, then the measure is time-translation invariance if it assigns the same size to both sets: $\mu(A) = \mu(A_{-t})$. In fact, since there is no privileged time, a possible history is a sequence of instantaneous states with both endings open, of the form: $X = [\ldots, X_{-1}, X_0, X_1, \ldots]$, which is infinite on both the right and the left hand sides. Translating this sequence back in time of, say, a single time unit, means considering another way the world can be, one in which: the state at time $t = 1$ is identical with the state at time $t = 0$ of the original sequence, namely X_0; the state at time $t = 2$ is identical with the state at time $t = 1$ of the original sequence, namely X_1; and so on.[36]

How does this apply to the spontaneous collapse theory? The situation is more complicated because there are not only two states to jump between, but infinitely many, since the positions in which the wave-function can be localized after a collapse form a continuum.[37] Moreover, the accessible states are constrained by the quantum probability rules rather than being accessible with the same probability. However, it seems that given the initial wave-function, the set of possible accessible states at a later time will still be a graph. Similarly, therefore, the space of possible microstates is the union of these graphs, one for each initial possible microstate. Hence, the typicality measure is the stationary, generic measure on this space as defined by the dynamics.

If this is so, there are more similarities than differences between deterministic and indeterministic theories: in deterministic theories, given an initial state, the dynamical laws determine the typicality measure on the space of possible histories of the world, namely phase space, and in indeterministic theories, given an initial state, the dynamical laws determine the typicality measure on the space of possible histories, namely Ω . Thus, in both deterministic and indeterministic cases, given an initial condition and the dynamics, 'everything follows.' That is, the set of possible states and the typicality measure are specified by the dynamics.

6. Conclusions

Let us finally go back to the originally discussed Boltzmann's account of macroscopic regularities. According to Albert and Loewer's view discussed here, three ingredients are needed:

---

[36] See also Bedingham and Maroney (2017) and Allori (2019) for a similar view to define time reversal for indeterminist theories.

[37] I am leaving aside considerations about the so-called primitive ontology approaches to GRW theory (see Allori, Goldstien Tumulka and Zanghí 2008) but a generalization seems straightforward.

1. The laws of motion,
2. The past hypothesis,
3. The statistical postulate.

From this, one can construct an argument that the spontaneous localization theory does not need to introduce the statistical postulate, for probability appears once and, so to speak, it is 'in the right place' within the dynamics. This, if this theory were true, one would need two, rather than three, ingredients, and on this basis such a theory should be preferred over the deterministic alternatives.

In this paper I have done the following:

1) I have shown, building on the work of Goldstein and collaborators, that, assuming there is a statistical postulate that specifies how to count states, it is not necessary for it to be about probabilities. In fact the notion of typicality is enough.
2) I have also shown that the statistical postulate, now understood as a typicality postulate, can be derived from the dynamics under symmetry constraints (stationarity, genericity).
3) Finally I have shown, following Zanghí, that both indeterministic and deterministic theory can ground a typicality-based understanding of macroscopic phenomena. In fact, when appropriately reformulated, they both require the same two ingredients:
    a. The specification of the dynamical laws, which determine the space of possible histories as well as the typicality measure;
    b. The past hypothesis.

If that is so, then both indeterministic and deterministic theories have basically the same relevant structure, and therefore there cannot be any fundamental difference in how satisfactorily they explain the macroscopic laws.

Acknowledgements

I am particularly grateful to Jean Bricmont and Barry Loewer for their comments on an early version of this paper. Also, I thank the participants of the 11th MUST (Models of Explanation) conference in Turin, Italy (June 11-13, 2018), as well as the participants of the Rutgers Philosophy of Probability Workshop (October 24-26, 2019), in particular David Albert, Katie Elliot, Carl Hoefer, Tim Maudlin, Wayne Myrvold, Isaac Wilhelm.

References

- Albert, David, Z. 2001. *Time and Chance*. Cambridge: Harvard University Press
- Allori, Valia. 2019. "Quantum Mechanics, Time and Ontology." *Studies in History and Philosophy of Modern Physics* 66: 145-154.
- Allori, Valia, Sheldon Goldstein, Roderich Tumulka, and Nino Zanghí. 2008. "On the Common Structure of Bohmian Mechanics and the Ghirardi-Rimini-Weber Theory." *The British Journal for the Philosophy of Science* 59 (3): 353-389.
- Baldovin, Marco, Lorenzo Caprini, Angelo Vulpiani. 2019. "Irreversibility and Typicality: A Simple Analytical Result for the Ehrenfest Model" *Physica* A 524: 422-429.
- Bedingham, Daniel J., and Owen J.E. Maroney. 2017. "Time Reversal Symmetry and Collapse Models." *Foundations of Physics* 47 (5): 670–696.
- Bricmont, Jean. 1995. "Science of Chaos or Chaos in Science?" *Annals of the New York Academy of Sciences* 775 (1): 131–175.
- Bricmont, Jean. 2001. "Bayes, Boltzmann and Bohm: Probabilities in Physics In: Jean Bricmont, GianCarlo Ghirardi, Detlef Dürr, Francesco Petruccione, Maria Carla Galavotti, and Nino Zanghí (eds) *Chance in Physics: Foundations and Perspectives*. *Lecture Notes in Physics*, 574: 3-21. Berlin: Springer.
- Bricmont, Jean. 2020. "Probabilistic Explanations and the Derivation of Macroscopic Laws". In V. Allori (ed.) "Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature," World Scientific.
- Callender, Craig. 2004. "There is No Puzzle about the Low Entropy Past". In Christopher Hitchcock (ed.). *Contemporary Debates in Philosophy of Science:* 240-255. London: Blackwell.
- Carroll, Sean, & Chen, Jennifer. 2005. "Spontaneous Inflation and the Origin of the Arrow of Time." *General Relativity and Gravitation* 37: 1671–1674.
- Chen, Eddy K. 2020. "Time's Arrow in a Quantum Universe: On the Status of the Statistical Mechanical Probabilities." In V. Allori (ed.) "Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature," World Scientific.
- Crane, Harry and Isaac Wilhelm. 2020. "The Logic of Typicality." In V. Allori (ed.) "Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature," World Scientific.
- Dürr, Detlef. 2009. *Bohmian Mechanics*. Berlin: Springer.
- Ehrenfest, Paul and Tatiana Ehrenfest. 2015. The Conceptual Foundations of the Statistical Approach in Mechanics. Dover Publications, New York, reprint edition.

- Einstein, Albert. 1936. "Physics and Reality." *Journal of the Franklin Institute* 221: 349-382.
- Frigg, Roman. 2009. "Typicality and the Approach to Equilibrium in Boltzmannian Statistical Mechanics". *Philosophy of Science* 76: 997-1008.
- Frigg, Roman. 2011. "Why Typicality Does Not Explain the Approach to Equilibrium". In M. Suárez (ed.): *Probabilities, Causes and Propensities in Physics:* 77-93. Dordrecht: Springer.
- Frigg, Roman and Werndl, Charlotte. 2011. "Explaining Thermodynamic-like Behaviour in terms of Epsilon-ergodicity." *Philosophy of Science* 78: 628-652.
- Frigg, Roman and Werndl, Charlotte. 2012. "Demystifying Typicality." *Philosophy of Science* 5: 917-929.
- Ghirardi, GianCarlo, Rimini, Alberto, and Weber, Tulio. 1986. "Unified Dynamics for Microscopic and Macroscopic Systems". *Physical Review* D. 34: 470.
- Goldstein, Sheldon. 2001. "Boltzmann's Approach to Statistical Mechanics," In: Jean Bricmont, GianCarlo Ghirardi, Detlef Dürr, Francesco Petruccione, Maria Carla Galavotti, and Nino Zanghí (eds) *Chance in Physics: Foundations and Perspectives*. *Lecture Notes in Physics*, 574: 39-54. Berlin: Springer.
- Goldstein, Sheldon. 2011. "Typicality and Notions of Probability in Physics". In: Yemima Ben-Menahem, and Meir Hemmo (eds) *Probability in Physics*: 59-71. Heidelberg: Springer, The Frontiers Collection.
- Goldstein, Sheldon, Joel L. Lebowitz, Roderich Tumulka, and Nino Zanghì. 2019. "Gibbs and Boltzmann Entropy in Classical and Quantum Mechanics". In V. Allori (ed.) "Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature," World Scientific.
- Guth, Alan H. 1981. "Inflationary Universe: A Possible Solution for the Horizon and Flatness Problems." *Physical Review* D 23: 347-356.
- Guth, Alan H., and Paul J. Steinhardt. 1984. "The Inflationary Universe." *Scientific American*: 116-128.
- Hemmo, Meir, and Orly R. Shenker. 2012. *The Road to Maxwell's Demon. Conceptual Foundations of Statistical Mechanics.* Cambridge: Cambridge University Press.

- Hemmo, Meir, and Orly R. Shenker. 2012. *The Road to Maxwell's Demon. Conceptual Foundations of Statistical Mechanics.* Cambridge: Cambridge University Press.

- Hemmo, Meir, and Orly R. Shenker. 2020. *Can the Past Hypothesis Explain the Psychological Arrow of Time?* In V. Allori (ed.) "Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature," World Scientific.
- Hempel, Carl G. 1965. *Aspects of Scientific Explanations and Other Essays in the Philosophy of Science*. New York: The Free Press.
- Lazarovici, Dustin, & Reichert, Paula. 2015. "Typicality, Irreversibility and the Status of Macroscopic Laws". *Erkenntnis* 80 (4): 689-716.
- Lazarovici, Dustin, & Reichert, Paula. 2020. "Arrow(s) of Time without a Past Hypothesis." In V. Allori (ed.) "Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature," World Scientific.

- Lebowitz, Joel. 1981. “Microscopic Dynamics and Macroscopic Laws.” *Annals New York Academy of Sciences*: 220-233.
- Loewer, Barry. 2020. “The Mentaculus Vision.” In V. Allori (ed.) “Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature,” World Scientific.
- Maudlin, Tim. 2020. “The Grammar of Typicality.” In V. Allori (ed.) “Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature,” World Scientific.
- Myrvold, Wayne C. 2020. “Explaining Thermodynamics: What Remains to be Done?” In V. Allori (ed.) “Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature,” World Scientific.
- Olsen, Ryan A., and Christopher Meacham. 2020. *Eternal Worlds and the Best System Account of Laws*. In V. Allori (ed.) “Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature,” World Scientific.
- Penrose, Roger. 1999. *The Emperor’s New Mind.* Oxford: Oxford University Press.
- Pitowsky, Itamar. 2012. Typicality and the Role of the Lebesgue Measure in Statistical Mechanics. In J. Ben-Menahem and M. Hemmo (eds.): *Probability in Physics*, 51-58. Berlin and New York: Springer.
- Strevens, Michael. 2000. ‘Do Large Probabilities Explain Better?’ *Philosophy of Science* 67: 366–90.
- Uffink, Jos. 2007. “Compendium of the Foundations of Classical Statistical Physics”. In J. Butterfield and J. Earman (eds.): *Handbook for the Philosophy of Physics,* 923-1047. Amsterdam: Elsevier.
- Uffink, Jos. 2008. “Boltzmann’s Work in Statistical Physics”. The Stanford Encyclopedia of Philosophy. , http://plato.stanford.edu
- Valentini, Antony. 2020. “Foundations of Statistical Mechanics and the Status of the Born Rule in the de Broglie-Bohm pilot-wave theory.” This volume.
- Volchan, Sergio B. 2007. “Probability as typicality”. *Studies in History and Philosophy of Modern Physics* 38: 801-814.
- Wallace, David M. 2012. *The Emergent Universe.* New York: Oxford University Press.
- Wallace, David M. 2020. “The Necessity of Gibbsian Statistical Mechanics”. In V. Allori (ed.) “Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature,” World Scientific.
- Werndl, Charlotte. 2013. “Justifying Typicality Measures of Boltzmannian Statistical Mechanics and Dynamical Systems. *Studies in History and Philosophy of Modern Physics* 44(4): 470-479.
- Werndl, Charlotte, and Roman Frigg. 2020. “Taming Abundance: on the Relation between Boltzmannian and Gibbsian Statistical Mechanics.” In V. Allori (ed.) “Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism and Laws of Nature,” World Scientific.
- Wilhelm, Isaac. 2019. “Typical: A Theory of Typicality and Typicality Explanation”. *The British Journal for the Philosophy of Science*. https://doi.org/10.1093/bjps/axz016

- Zanghì, Nino. 2005. "I Fondamenti Concettuali dell'Approccio Statistico in Fisica". In: Valia Allori, Mauro Dorato, Federico Laudisa, and Nino Zanghì (eds.): *La Natura Delle Cose. Introduzione ai Fundamenti e alla Filosofia della Fisica*: 139-228. Roma: Carocci editore.