

Chapter 6

Measures over Initial Conditions

Meir Hemmo and Orly Shenker

Abstract This paper concerns the meaning of the idea of typicality in classical statistical mechanics and how typicality is related to the notion of probability.

6.1 Introduction

This paper concerns the meaning of the idea of typicality in classical statistical mechanics and how typicality is related to the notion of probability. Our thoughts about these issues have been greatly influenced along the years by numerous conversations with Itamar Pitowsky. In his last paper [1] which he devoted to the issue of typicality, he writes:

Consider a finite but large collection of marbles. When one says that a vast majority of the marbles are white one usually means that all the marbles except possibly very few are white. And when one says that half the marbles are white, one makes a statement about counting, and not about the probability of drawing a white marble from the collection.

Here Itamar is making a sharp distinction between the *size* of a set of outcomes of an experiment and the probability of these outcomes. The size of a set of outcomes is fixed by a measure defined on the event space. In the discrete case, the size of the set is fixed by counting the number of outcomes that belong to it. Itamar thought that in the discrete case the measure obtained by counting is natural, and therefore he thought that it is worthwhile to generalize this measure to the continuous case. In his paper (ibid.) he argues that the Lebesgue measure in the continuous case is

M. Hemmo (✉)

Department of Philosophy, University of Haifa, Haifa, Israel

e-mail: meir@research.haifa.ac.il

O. Shenker

Department of Philosophy, The Hebrew University of Jerusalem, Jerusalem, Israel

e-mail: orlyshenker@pluto.mscc.huji.ac.il

the natural extension of the counting measure in the discrete case, and he takes this result to establish a preference for the Lebesgue measure in the continuous case. This means that in classical statistical mechanics, for example, the Lebesgue measure is the natural measure to determine sizes of sets in the state space. If this is right, the problem of justifying the choice of measure in classical statistical mechanics is partially solved. The reason why it is only partially solved is that on the standard way of thinking about statistical mechanics, the problem concerns the justification of the statistical mechanical probabilities, and as Itamar himself stresses (in the quotation above) the measure of sets is not enough to determine probability.

Despite the distinction between measure and probability, Itamar thought (see *ibid.*) that the Lebesgue measure in statistical mechanics plays some role, admittedly weak, in the explanation of thermodynamic behavior. In this paper we examine this question. Our starting point is similar to Itamar's that measure is indeed different from probability, but while Itamar thought that the Lebesgue measure is natural in some *a priori* sense, it seems to us that the choice of measure in physics is guided by experience, which in turn guides our choice of probabilistic laws.

The structure of the paper is as follows. We begin in Sect. 6.2 by describing the so-called typicality approach (as it is usually framed in the context of deterministic theories in physics). In Sect. 6.3 we describe the way in which probabilistic statements in classical statistical mechanics ought to be understood. In Sect. 6.4 we examine arguments based on the classical dynamics to the effect that the Lebesgue measure is natural in statistical mechanics. In Sect. 6.5 we analyze the significance of Lanford's theorem in classical statistical mechanics, and we explain how the theorem ought to be understood without appealing to typicality. Section 6.6 is the conclusion.

6.2 Typicality

In classical statistical mechanics the standard way of understanding the thermodynamic behavior of systems around us appeals to a *probability* distribution over the initial microstates of the systems (compatible with the initial thermodynamic macrostate). On the standard way of thinking one says that given the uniform probability distribution (relative to the Lebesgue measure) over the initial macrostate, it is highly *probable* that the system will, for example, approach equilibrium after some designated time. In this way, the behavior of the system is explained by the fact that its actual microstate is highly likely to sit on a trajectory, which will take it to equilibrium at the time in question. Here the high probability pertains to subsystems of the universe, and it is assumed further that the trajectory of the whole universe that gives rise to this high probability itself sits on an initial condition which has high probability. Note that here there are two notions of probability: a probability distribution over the initial macrostate (i.e. the microstates compatible with the macrostate at some present time) of subsystems of the universe, and a probability distribution over the initial conditions of the universe.

Another important example of the central role played by the measure in explaining physical behavior in statistical mechanics is in Einstein's [2] account of Brownian motion, as developed by Wiener (see [3]). As is well known, Wiener has proved that the so-called Wiener measure of trajectories in the phase space of a Brownian particle which are continuous but nowhere differentiable is one. The explanation of the actual behavior of Brownian particles is based on the assumption that their actual trajectories belong to this measure one set. Avogadro's number is derived from this assumption.

A question that immediately arises concerning this understanding is what could a probability distribution over the *initial* conditions of the universe possibly mean. A probability distribution suggests some sort of a random sampling of an initial condition out of the set of all possible conditions. But with respect to the initial conditions of the universe any such sampling (if it is to be physical) would be external to the universe, and therefore this seems to suggest an empirically meaningless fairy tale. This problem does not arise with respect to subsystems of the universe, since one can ground a probability distribution over initial conditions in experience (as we show in Sect. 6.3). Moreover, probability in physical theories is usually conceived as involving (or as being tested by) repetitions of experiments, which in the case of the initial conditions of the universe are trivially impossible.

We understand the typicality approach¹ as an attempt to solve these problems by appealing to a certain natural measure over initial conditions, where the measure is *not* understood as a probability measure (see [5] for a similar construal).

Here is an example of how the distinction between typicality and a probability distribution over initial conditions is made:

When employing the method of appeal to typicality, one usually uses the language of probability theory. When we do so we do not mean to imply that any of the objects considered is random in reality. What we mean is that certain sets (of wave functions, of orthonormal bases, etc.) have certain sizes (e.g., close to one) in terms of certain natural measures of size. That is, we describe the behavior that is typical of wave functions, orthonormal bases, etc. However, since the mathematics is equivalent to that of probability theory, it is convenient to adopt that language. For this reason, we do not mean, when using a normalized measure μ , to make an "assumption of *a priori* probabilities," even if we use the word "probability." Rather, we have in mind that, if a condition is true of most D , or most H , this fact may suggest that the condition is also true of a concrete given system, unless we have reasons to expect otherwise. [7].

And in another place [8], they say:

When we express that something is true for most H or most ψ relative to some normalized measure μ , it is often convenient to use the language of probability theory and speak of a random H or ψ chosen with distribution μ . However, by this we do not mean to imply that the actual H or ψ in a concrete physical situation is random, nor that one would obtain, in repetitions of the experiment or in a class of similar experiments, different H 's or ψ 's whose empirical distribution is close to μ . That would be a misinterpretation of the measure μ , one

¹For various formulations and extensive discussions of the typicality approach, see Dürr et al. [4], Maudlin [5], Callender [6].

that suggests the question whether perhaps the actual distribution in reality could be non-uniform. This question misses the point, as there need not be any actual distribution in reality. Rather, Theorem 1 means that the set of “bad” Hamiltonians has very small measure μ .

There are three different statements made here about the idea of typicality:

- (1) The set of initial conditions compatible with the initial macrostate of the universe is divided into two subsets, T1 and T2 such that all the microstates in T1 but not in T2 give rise to some property F. The property F may be for example the approach to equilibrium in statistical mechanics, or the Born rule in Bohmian mechanics.
- (2) There is some natural (normalized) measure μ over the initial conditions such that $\mu(T1)$ is close to one (and $\mu(T2)$ is close to zero). In this sense, *most* initial conditions, as determined by μ , are in T1 (and are called typical).
- (3) In a given experiment, the actual initial microstate of the universe belongs to T1.

Let us explain these three statements in turn. The statement in (1) above expresses a contingent fact about the dynamics, namely a fact about how the initial conditions are mapped by the equations of motion into microstates at later times. There are various theorems in classical statistical mechanics that demonstrate that special cases of (1) hold under some conditions with some appropriate property F. Examples are Lanford’s theorem in which F is (roughly) entropy increase and the Birkhoff-von Neumann theorem in which F is the so-called pointwise ergodic theorem, which we discuss below. Statement (1) is not controversial in our discussion.

The notion of *most* in statement (2) above requires a measure over the phase space. That is, there are infinitely many ways to determine the size of subsets of a continuous set of points. The question is on what grounds one can justify the choice of measure, or the choice of some class of measures. Usually, in classical statistical mechanics the measure chosen is the Lebesgue measure (or the class of measures absolutely continuous with the Lebesgue measure), and in quantum mechanics the measure is given by the absolute square of the wavefunction. The grounds for these choices are that each of these measures has a preferred dynamical status in the theory.

Statement (3), as stated above, seems as expressing the brute fact, without further reasoning, that the microstate of the universe invariably (in every experimental set up) belongs to T1. But since there are microstates of the universe that don’t belong to T1 this fact calls for a justification. It is evident that (2) is taken in the typicality approach to completely justify (3), that is if T1 were to contain only a small fraction of the microstates of the universe, one would not see (3) as justified. It is important to stress that in this approach the justification of (3) makes no appeal to probability. Rather, it is the measure of T1 that is supposed to do the whole work. This implies that, lacking reasons to expect otherwise, microstates of the universe that belong to T2 are not realized.

In short, there are two questions that need be answered in the context of typicality: what justifies the choice of measure in (2), and what justifies the passage

from (2) to (3). In particular, the question we consider is whether there are grounds that justify the choice of measures in a way that *explains* the observed behavior of physical systems. If such grounds could be spelled out the problems concerning the meaning of probability distributions over the initial conditions of the universe would obviously evaporate together with the probability distribution itself. In the subsequent sections we attempt to answer these two questions. We will see that statements (2) and (3) are both wanting. Again, statement (1) is not controversial in the context of typicality. Our analysis will lead us to reject the typicality approach.

6.3 Probability in Classical Statistical Mechanics

In order to set the stage we need to go into some detail concerning the way in which probability statements arise in classical statistical mechanics and how precisely the choice of measure over the state space is carried out.

Consider the paradigmatic case of an ideal gas S , which is initially confined by a partition to the left half of a container, and then, by removing the partition, is allowed to expand. Finally, the gas fills out the entire container. Suppose that we set up a very large number of such gases $S_1 \dots S_k$, all of which are prepared in the same initial macrostate M_0 in which the gas is confined to the left half of the container by a partition. We then remove the partitions and follow the spontaneous macroscopic evolution of these gases for a certain time interval Δt , and we see by simple counting that the overwhelming majority of the gases $S_1 \dots S_k$ quickly reach and then remain in macrostate M_1 in which they fill up the entire container. We now wish to predict the evolution of another system, call it S_{k+1} , which is prepared in the same initial macrostate as $S_1 \dots S_k$. We know that the dynamical equation of motion that governs the evolution of S_{k+1} is the same as the ones governing $S_1 \dots S_k$, but we do not know the details of this dynamics, nor do we know the exact initial microscopic conditions of S_{k+1} and therefore all we can rely on in this prediction is the above experiment.

Can we infer from the experiment with $S_1 \dots S_k$ that S_{k+1} is highly likely to end up in macrostate M_1 ? That is, can we use the experiment with $S_1 \dots S_k$ in order to come up with a probabilistic law, on which we can base our bets regarding the evolution of S_{k+1} ? The answer is, of course, yes, we can infer the probabilities from the finite observed relative frequencies.² This inference is valid just to the extent that we can infer from experience any other physical law or prediction, such as $F = ma$. However, the way in which our probabilistic predictions can be *justified*, and the *extent* to which they can be justified – are not always clear in the literature, as we show later.

To see how to understand probabilistic statements in statistical mechanics let us describe the above experiment in the phase space of the gas. Classical mechanics

²This inference is a subtle issue which depends on how probability is understood. We don't address this question here.

tells us that the universe consists of microscopic particles, and that our experience is an *effect* of the *microstate* of the universe, which is the state of those particles. However, it is a physical fact that our senses are too coarse to reflect the full details of the microscopic structure of the universe; we can only perceive some of its general features. In this sense our experience is *macroscopic*. In the above experiment, we can only observe relative frequencies of *transitions* between macrostates of the gas. Let us see how these transitions are described in the phase space, and then how these relative frequencies are accounted for in the phase space.

The phase space of a system (in our example, of any of the systems S_i) is partitioned into sets of microstates, which are indistinguishable by an observer; these sets are called macrostates. The phase space regions corresponding to the macrostates express the observer's maximal observational capability, and therefore while the observer can tell which macrostate contains the actual microstate of the system at the time of observation, it cannot tell which part of the macrostate contains that microstate.

We now formulate what we take to be the essential way for calculating transition probabilities in statistical mechanics. Suppose that at time t_0 an observer O finds the system S in macrostate M_0 (as for example in our experiment above; see Fig. 6.1). Suppose also that O knows the laws of classical mechanics, which govern S 's evolution in time. If O knows the Hamiltonian of S , that is: if O knows the equations of motion of S , then O can (in principle) calculate the evolution of all the trajectory segments that start out in the microstates contained in M_0 and find out the end points of these trajectory segments after the time interval Δt . These end points make up a set, which we call the *dynamical blob* $B(t_0 + \Delta t)$ of S at $t_0 + \Delta t$ given that it was in M_0 at t_0 . In general, the region covered by $B(t_0 + \Delta t)$ overlaps with several macrostate regions; for instance, it may partially overlap with M_1 (in which the gas fills out the entire container), and with some other macrostates, such as M_2 or M_3 in which the macrostate of the gas is different. If the system S , which started out in M_0 at t_0 , is observed to be (say) in macrostate M_1 at $t_0 + \Delta t$, then this means that the microstate of S is actually in the region of *overlap* between the region of macrostate M_1 and the region of the dynamical blob $B(t_0 + \Delta t)$. Now, in our above experiment, O carries out the experiment k times (or on k identical systems). In some of these experiments – actually in most of them (in our story) – at $t_0 + \Delta t$

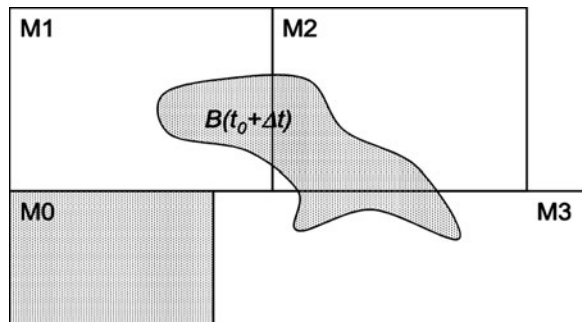


Fig. 6.1 The time evolved blob $B(t)$ spreads over different macrostate

the system S is observed to be in M_1 and in other fewer experiments it is found in M_2 or M_3 , or more precisely in the regions of *overlap* of the dynamical blob $B(t_0 + \Delta t)$ with these macrostates, with some relative frequencies F_1, F_2 and F_3 respectively. These relative frequencies are the empirical basis on which the probabilistic statements of the theory can be based, and on the basis of which these statements can be tested or justified.

The next step towards constructing or justifying the probabilistic theory is as follows. Given the above experimental outcomes, we have the relative frequencies with which systems of type S that start out in M_0 at t_0 are found in the macrostates M_1, M_2 or M_3 . We conclude that the phase points of our k systems evolved into the regions of overlap of the dynamical blob $B(t_0 + \Delta t)$ with the macrostates M_1, M_2 or M_3 . We then conjecture on the basis of our experience that this statistical behavior will be repeated (more or less) in the future. Since any of the microstates in M_0 is a possible initial condition of S_{k+1} and since the phase space is continuous, such a generalization of our experience requires that we impose a measure on the phase space. We identify the set of probability measures that, if applied to the continuous phase space of S , yield a measure of the regions of overlap of the blob $B(t_0 + \Delta t)$ with the macrostates M_1, M_2 or M_3 that are (to a satisfactory approximation) identical with the relative frequencies F_1, F_2 and F_3 , respectively. There are many – possibly infinitely many – such measures, and all of them are *empirically* adequate. Among them we *choose* one measure, using *pragmatic* criteria such as simplicity, convenience, meshing with other theories, etc. Call this measure μ . The (normalized) measures of the regions of overlap are then given by $\mu(B(t) \cap M_1) \approx F_1$, $\mu(B(t) \cap M_2) \approx F_2$, $\mu(B(t) \cap M_3) \approx F_3$. This measure μ is imposed *over the blob* $B(t_0 + \Delta t)$ and provides the basis for predicting the evolution of system S_{k+1} in terms of transition probability (roughly) as follows:

(*) The transition probability that S_{k+1} will evolve to macrostate M_i at $t_0 + \Delta t$ given that it was in macrostate M_0 at t_0 , is equal to

$$\mu(B(t_0 + \Delta t) \cap M_i | B(t_0) = M_0) \approx F_i.$$

That is, the transition probability from the macrostate M_0 at t_0 to M_i at $t_0 + \Delta t$ is equal to the (normalized) measure of the region of overlap of the blob $B(t_0 + \Delta t)$ with the macrostates M_i . This is the basis of our probabilistic theory.

Note that in general $\mu(M_i)/\mu(M_j)$ need not be equal to $\mu(B(t) \cap M_i)/\mu(B(t) \cap M_j)$.³ Note further that despite the deterministic dynamics these transition probabilities between *macrostates* are physically objective provided the partition to the macrostates is objective.⁴

What is the significance of taking the μ (normalized) measure over the blob $B(t_0 + \Delta t)$ as underwriting our probabilities for measurement outcomes? It is crucial

³ This implies that the transition to a given macrostate need not be equal to the entropy of that macrostate, even if both are measured by the same measure μ .

⁴ This last condition needs to be flashed out; we skip this here.

to see that the probabilistic statements are about *transitions* from M_0 at t_0 to any one of the macrostates M_i . We don't distribute probabilities relative to the μ measure over the initial macrostate M_0 at t_0 . Of course, if the measure μ is invariant under the classical dynamics, e.g. if it happens to be the Lebesgue measure, then one can map, backwards (as it were), the measure of regions over the blob at later times to the corresponding regions over the *initial* macrostate. That is, in this case the measure of a set of points in M_0 is equal to the measure of the time evolved set of points to which it is mapped by the dynamics. Once the (normalized) measure is fixed (by the probabilities) one can distribute uniform probabilities relative to the Lebesgue measure over the initial macrostate. But note that this interpretative move is derivative. In general, whether or not the measure that best fits our observations is the Lebesgue measure, or more generally a measure that is invariant under the dynamics, is a contingent matter.

We can now see what justifies the choice of measure and what justifies probabilistic statements in classical statistical mechanics, and moreover how these two issues are related. First, probabilistic statements are grounded in the experience of relative frequencies in the way stated above. Second, the choice of measure is dictated inductively (not uniquely) by the observed relative frequencies. That is, the measure is implied by the probabilities rather than the other way around. We can only justify empirically transition probabilities as sketched in (*) above rather than distributions over initial conditions.

The implications of this analysis for the typicality approach are as follows.

1. The probability measure μ is applicable only to subsystems of the universe. Of course, if the dynamics is deterministic, each microstate of all the subsystems of the universe can be mapped backwards to the initial conditions and the measure over the initial conditions will depend on the measure at the later times. But in this way the justification of the choice of the measure over the initial conditions is grounded in experience, and therefore it cannot be taken to explain (non-circularly) experience. Note that this argument applies to the question of the choice of the measure regardless of whether the measure is understood as determining the typical set of initial conditions (as in the typicality approach) or as a probability measure over the initial conditions of the universe (as in standard approaches to statistical mechanics).
2. This strategy of grounding the measure over the initial conditions of the universe in experience can hold only with respect to a fraction of all possible initial conditions of the universe (compatible with the initial macrostate). It excludes by construction initial conditions that lead to a universe at the later times which is macroscopically different from what we see.
3. Our ignorance about the initial microstate of S_{k+1} is often illustrated by appealing to some random sampling of a point out of M_0 . Of course, this idea need not be taken too seriously (as describing a fairy tale about some mechanism of selection). However, the point to be stressed here is the following. A random sampling is a sampling that depends only on the measure. The measure with respect to which the sampling is random need only be the measure that fits the observed relative frequencies in experience. In particular the measure need not

be the Lebesgue measure, and may not even be conserved under the dynamics. By appealing to the *probability* measure we can now justify statements about the probability of randomly sampling initial conditions for subsystems of the universe. Here unlike the statement (3) of the typicality approach, the sampling is described in terms of probability rather than typicality. The role of the measure in our approach is derivative rather than fundamental and is patently probabilistic.

6.4 Are There Natural Measures?

In the literature there are attempts to justify the choice of the measure (in the typicality approach and in other approaches) on the basis of dynamical considerations.

An argument sometime given for preferring the Lebesgue measure as ‘natural’ on the basis of the classical dynamics is the invariance of the Lebesgue measure under the dynamics as expressed by Liouville’s theorem. If a measure is invariant under the dynamics it means that the measure of a given set of points in the state space is equal to the measure of the set to which it is mapped by the time evolution equations for all times. Of course this feature has very attractive properties (simplicity, elegance, etc.) but it is unclear why this fact is relevant at all to the issue at stake, namely the explanation and prediction of physical behavior.

A similar argument is sometimes given in the case of ergodic dynamics. Obviously, the ergodic theorem gives a preferred status to the Lebesgue measure (or to any measure absolutely continuous with the Lebesgue measure) since it shows that the relative frequency of any macrostate M along an infinite trajectory is equal to the Lebesgue measure of M for a Lebesgue measure one of points in the phase space of the system. There are various senses in which the preferred status of the Lebesgue measure here is irrelevant for the issue at stake. First, the ergodic theorem yields no predictions concerning finite times, and therefore strictly speaking the theorem is not empirically testable. For example, it is extremely difficult to distinguish empirically between an ergodic system and a system with KAM dynamics (see [9]). Second, even if the dynamics of the universe is granted to be ergodic and even if one accepts the fairy tale about an initial random sampling, this does not imply that the sampling is random relative to the Lebesgue measure. One can say metaphorically that God could have used a non-Lebesgue sort of die in sampling at random the initial condition of the universe even if the universe were ergodic. Third, and with respect to the typicality approach. Consider again statement (3) in Sect. 6.1. Here the idea is that the fact that T_1 has measure (close to) one suggests that the initial condition of the universe belongs to T_1 . Since the measure is not to be understood as a probability measure, this seems to mean that the measure zero set is excluded as impossible in some sense. But the measure zero set belongs to the initial macrostate of the universe and we don’t see what justifies this exclusion.

Finally, it is important to stress in this context that in understanding the ergodic theorem as a theorem about probability one must identify from the outset that a set

of Lebesgue measure zero (one) has zero (one) probability. Although the theorem is usually understood in probabilistic terms, it should be stressed that this identification is *not* part of von Neumann's and Birkhoff's ergodic theorem. Whether or not the Lebesgue measure may be interpreted as the right *probability* measure for thermodynamic systems depends on whether it satisfies our probability rule (*).

Another argument sometimes given for taking the Lebesgue measure as the natural measure in statistical mechanics is that the Lebesgue measure of a macrostate corresponds to the thermodynamic entropy of that macrostate. However, this correspondence is true *only if* the Second Law of thermodynamics (even in its probabilistic version) is true. But as we argued elsewhere (see our [7, 8, 10], Chap. 5) the Second Law of thermodynamics is not universally true in statistical mechanics.

6.5 Lanford's Theorem

The above conclusion has implications for the significance of measure one theorems in statistical mechanics. We focus here as an example on Lanford's theorem.⁵ Lanford proved on the basis of the classical equations of motion, that, roughly, given some specific initial macrostate, and some specific kind of Hamiltonian, a *Lebesgue* measure one of the microstates in that macrostate will evolve to a macrostate with larger entropy, after a certain short time.⁶ Can such a theorem endow the Lebesgue measure with a status that is stronger than that of an empirical generalization (as sketched in (*) above)?

In terms of our transition probabilities Lanford's theorem proves that the Lebesgue measure of the overlap between the blob $B(t_0 + \Delta t)$ and the macrostate E of equilibrium (or some other high entropy macrostate) is 1. Of course, as we said above, since the Lebesgue measure is conserved under the dynamics, one may interpret Lanford's theorem as referring to the Lebesgue measure of subsets of the *initial* macrostate M_0 at t_0 . However, inferring anything about the measure of subsets of the initial macrostate is an artifact of the contingent fact that the Lebesgue measure matches the observed relative frequencies.

Another crucial point in this context is the following. There are two different and logically independent ways of understanding the role of the Lebesgue measure in Lanford's theorem. **(A)** The size of the overlap between the blob $B(t_0 + \Delta t)$ and the macrostate E , as determined by the Lebesgue measure, is 1; **(B)** Upon a random sampling of a point out of the blob $B(t_0 + \Delta t)$, one is highly likely to pick out a point

⁵ For details concerning Lanford's theorem see Uffink [11].

⁶ The fact that a Maxwellian Demon is compatible with classical statistical mechanics demonstrates that there can be no theorem in mechanics that implies a universal entropy increase. See Albert [10, Chap. 5] and Hemmo and Shenker [12, 13].

from the overlap of the blob with E . The distinction between (A)-type statements about sizes of sets and (B)-type statements about probabilities is general.

Lanford's theorem is about the size of the overlap with E , that is, it is only an (A)-type theorem, whereas in order to make predictions about the future behavior of S-type systems (such as our S_{k+1} in the above example) one needs to add a (B)-type statement, *which is not proven by Lanford's theorem*. In other words, assuming that we already know from experience that the Lebesgue measure of the overlap regions (of the blob with the macrostates) matches the relative frequencies of the macrostates, Lanford's theorem provides possible mechanical conditions, which underwrite these observations.

To appreciate this point, note that if the measure μ that matches our experience were not the Lebesgue measure, but some other measure (that may not be absolutely continuous with Lebesgue) then Lanford's theorem would have a completely different significance: for instance, it could happen that by the measure μ the number of systems that go to equilibrium given Lanford's Hamiltonian would be small. The theorem that a set of Lebesgue measure one of points has a certain property (such as approaching equilibrium after some finite time interval) would be empirically insignificant – unless this fact is supplemented by the additional fact that the Lebesgue measure happens to correspond (to a useful approximation) to the observed relative frequencies.

The general structure of Lanford's theorem is that it proves a certain statement about the dynamics of the form of (1) in the typicality approach (see Sect. 6.2). That is, Lanford's theorem shows that a certain subset of microstates $T1$ share some property F (entropy increase, for example), such that all the points in $T1$ are mapped by the dynamics to points in $T1^*$. Moreover, the theorem shows that the subset $T1$ has Lebesgue measure one. But nothing in this theorem justifies the choice of the measure. In particular, the fact that $T1$ has Lebesgue measure one does not constitute such a justification. What's important in Lanford's theorem is that it identifies two sets $T1$ and $T1^*$ and proves that $T1$ evolves to $T1^*$ under the dynamics. That is, the theorem is about the structure of trajectories. The fact that $T1$ has Lebesgue measure one is important only if there are independent reasons for preferring the Lebesgue measure. As we saw in Sect. 6.3 such reasons can be grounded essentially only in experience.

6.6 Conclusion

In this paper we showed that one can understand the full scope of classical statistical mechanics by appealing to the notion of *transition* probabilities between macrostates, without resorting to probability distributions over initial conditions or to typicality considerations.

Acknowledgement This research is supported by the Israel Science Foundation, grant numbers 240/06 and 713/10.

References

1. Pitowsky, I.: Typicality and the role of the Lebesgue measure in statistical mechanics, In Y. Ben-Menahem and M. Hemmo (eds.), *Probability in Physics*, pp. 41–58. The Frontiers Collection, Springer-Verlag Berlin Heidelberg (2011)
2. Einstein, A.: Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annal. Phys.* **17**, 549–560 (1905). English translation in: Furth, R. (ed.) *Einstein, A., Investigations on the Theory of Brownian Motion*. Dover, New York (1926)
3. Pitowsky, I.: Why does physics need mathematics? A comment. In: Ulmann-Margalit, E. (ed.) *The Scientific Enterprise*, pp. 163–167. Kluwer, Dordrecht (1992)
4. Dürr, D., Goldstein, S., Zanghi, N.: Quantum equilibrium and the origin of absolute uncertainty. *J. Stat. Phys.* **67**(5/6), 843–907 (1992)
5. Mauldin, T.: What could be objective about probabilities? *Stud. Hist. Philos. Mod. Phys.* **38**, 275–291 (2007)
6. Callender, C.: The emergence and interpretation of probability in Bohmian mechanics. *Stud. Hist. Philos. Mod. Phys.* **38**, 351–370 (2007)
7. Goldstein, S., Lebowitz, J., Mastrodonato, C., Tumulka, R., Zanghi, N.: Normal typicality and von Neumann’s quantum ergodic theorem. *Proc. R. Soc. A* **466**(2123), 3203–3224 (2010)
8. Goldstein, S., Lebowitz, J., Mastrodonato, C., Tumulka, R., Zanghi, N.: Approach to thermal equilibrium of macroscopic quantum systems. *Phys. Rev. E* **81**, 011109 (2010)
9. Earman, J., Redei, M.: Why ergodic theory does not explain the success of equilibrium statistical mechanics. *Br. J. Philos. Sci.* **47**, 63–78 (1996)
10. Albert, D.: *Time and Chance*. Harvard University Press, Cambridge (2000)
11. Uffink, J.: Compendium to the foundations of classical statistical physics. In: Butterfield, J., Earman, J. (eds.) *Handbook for the Philosophy of Physics, Part B*, pp. 923–1074. (2007)
12. Hemmo, M., Shenker, O.: Maxwell’s demon. *J. Philos.* **107**(8), 389–411 (2010)
13. Hemmo, M., Shenker, O.: Szilard’s perpetual mobile. *Philosophy of Science* **78**, 264–283 (2011)