Ehrenfest and Ehrenfest-Afanassjewa on the Ergodic Hypothesis

Patricia Palacios

Department of Philosophy (KGW), University of Salzburg

Abstract

Ehrenfest and Ehrenfest-Afanassjewa's seminal article on statistical mechanics highlighted a crucial assumption at the heart of Boltzmann's statistical mechanics: the ergodic hypothesis. The importance of this article for transmitting the problems related with the ergodic hypothesis has been widely recognized, but Ehrenfest and Ehrenfest-Afanassjewa have been strongly criticized for not having provided a fair account of Boltzmann's statistical mechanics. In this chapter, I outline Ehrenfest and Ehrenfest-Afanassjewa's treatment of the ergodic hypothesis and I evaluate the role of this discussion for the development of the ergodic theory in the 20th century. I will conclude that the major contribution of Ehrenfest and Ehrenfest-Afanassjewa comes precisely from what has been regarded by some historians of science as historical inaccuracies of the article.

Contents

1 Introduction

In the *Conceptual Foundations of the Statistical Approach to Mechanics* (Ehrenfest and Ehrenfest-Afanassjewa 1959) – also known as the "Encyklopädie article" – Ehrenfest and Ehrenfest-Afanassjewa gave a prominent role to what they dubbed as "the ergodic hypothesis", suggesting that Boltzmann's entire program lacks a firm foundation because it relies on this hypothesis of questionable validity. The importance of Ehrenfest and Ehrenfest-Afanassjewa's article for transmitting the problems related with the ergodic hypothesis has been widely recognized, but they have been strongly criticized for not having provided a fair account of Boltzmann's statistical mechanics. It has been argued that they exaggerated the role of the ergodic hypothesis in Boltzmann's program (Brush, 1967, 1971; von Plato, 1991); that they exaggerated the importance of the equivalence between phase averages and time averages (Brush, 1967; Uffink, 2007); and that they misunderstood the meaning of the ergodic hypothesis as conceived by Boltzmann (von Plato, 1991; Brush, 1967). In this chapter, I will analyse the discussion of Ehrenfest and Ehrenfest-Afanassjewa around the ergodic hypothesis and I will evaluate the role of this discussion for the development of the ergodic theory in the 20th century. I will conclude that the major contribution of the *Encyklopädie* article comes precisely from what has been regarded by some historians of science as historical inaccuracies of the article. In particular, I will argue that they advanced Boltzmann's interpretation of probabilities as time averages by emphasizing the role of the ergodic hypothesis and by highlighting the importance of the equivalence between phase and time averages.

This chapter is organized as follows. In Section 2, I review the origins of the ergodic hypothesis in Boltzmann's statistical mechanics, by highlighting the connection between the ergodic hypothesis and Boltzmann's time average interpretation of probabilities. In Section 3, I discuss Ehrenfest and Ehrenfest-Afanassjewa's criticism of the ergodic hypothesis and I point out that one their major contributions was to pose a new puzzle in the foundations of statistical mechanics, which I call "the Ehrenfest and Ehrenfest-Afanassjewa's puzzle". Subsequently, in Section 4, I illustrate how this puzzle encouraged the development of the impossibility theorems in 1913. In Section 5, I argue that this puzzle also played a role in the introduction of the notion of metric transitivity, which led to the establishment of the ergodic theorems by Birkhoff and von Neumann. I point out that although these theorems solved the Ehrenfest and Ehrenfest-Afanassjewa's puzzle, they transformed the problem of ergodicity into the problem of proving that the systems of interest are metrically transitive. Finally, in Section 6, I review the recent discussion in the foundations of statistical mechanics around the problem of metric transitivity.

2 The origin of the ergodic hypothesis

Consider a typical situation of a dilute gas enclosed in a finite container with N identical polyatomic molecules, each with r degrees of freedom. The molecules collide with each other and with the walls of the container and the collisions are governed by short-range repelling potentials. The possible states of this system are represented by points in a 2rN-dimensional phase space Γ, with q position coordinates and p momentum coordinates. Assume that the energy of the system is E, so that the state must lie on the energy surface Γ_E, which is $2rN - 1$ dimensional. At time t, the state of the system will be determined exactly by the simultaneous position and momentum coordinates of the N molecules:

$$q_1^1, ..., q_r^1; q_1^2, ..., q_r^2; ..., ; q_1^N, ..., q_r^N$$

,

$$p_1^1, ..., p_r^1; p_1^2, ..., p_r^2; ..., ; p_1^N, ..., p_r^N$$

The corresponding changes in the states of the gas model are expressed, following Ehrenfest and Ehrenfest-Afanassjewa (1959)'s notation, by the following Hamiltonian equations of motion:

$$\frac{dq_s^k}{dt} = \frac{\delta E}{\delta p_s^k} \qquad \frac{dp_s^k}{dt} = \frac{\delta E}{\delta q_s^k}.$$

One can easily observe that the number of variables is enormous, so even if the system is deterministic, it is not possible to know the exact initial conditions of the system and there is little chance of integrating these equations to find the exact solutions. Such pragmatic difficulties motivated a statistical approach to the study of these kinds of systems, which began with Maxwell and Boltzmann in the second half of the 19th century.

Maxwell (1860) was the first to characterize the equilibrium state of a gas by a probability distribution function f. Almost a decade later, Boltzmann

(1868) derived this probability distribution in the presence of external forces suggesting that if the system is left alone, the probability of the molecular velocities will *always* assume Maxwellian distribution.[1] To derive this result he uses, for the very first time, a time average interpretation of these probabilities, whereby the probability of the equilibrium state is identified as the relative time in which the system is in that state when left alone for "a very long time". In a short communication about Maxwell's work in 1879, he refers to his own interpretation of probabilities in the following terms:

> There is a difference in the conceptions of Maxwell and Boltzmann in that the latter characterizes the probability of a state by the average time in which the system is in this state, whereas the former assumes an infinity of equal systems with all possible initial states. (quoted in von Plato 1991, p.71)

But what exactly is a time average and how does it help interpret probabilities in statistical mechanics? Let us recall that Boltzmann (1868) had adopted Maxwell's characterization of equilibrium in terms of stationary probability distributions, where macroscopic observables correspond to phase averages over the phase space with respect to a specific probability measure, i.e., the microcanonical measure. In a modernized notation, this can be written as follows:

$$\langle f \rangle = \int_{\Gamma_E} f(x)p(x)dx, \tag{1}$$

where $f(x)$ is the phase function that gives the value of the observable for each microscopic state in the phase space, $p(x)$ is the probability distribution density and x represents the state of the system written in local coordinates. Since this average is independent of the dynamics of the system there is no immediate interpretation for this probability, in other words it is not clear why an arbitrarily chosen system should have this distribution. In order to establish a connection with the dynamics of the system that is robust upon the initial conditions, Boltzmann postulated that the probabilities are time averages.[2] In modern terms this can be defined as follows. Let $T(t, x)$ represent the dynamical evolution of the system, if x is the state at time 0,

[1] Boltzmann noted that there might be exceptions to his derivation, for example, when the trajectory is periodic. However, he believed that such behavior would be destroyed by the slightest disturbances from outside (Uffink, 2007, p.39)

[2] This use of time averages for interpreting the expectation values w.r.t the Maxwellian stationary probability distribution should be distinguished from the related use of time

then the future state at time t is $T(t, x)$. The time average of a function of state f is:

$$f^* = lim_{t\to\infty}(1/t)\int_0^t f(T(t, x))dt. \qquad (2)$$

Although Boltzmann did not attempt to demonstrate the equivalence between phase averages and time averages (we will come back to this point in the next section), he did explicitly use time averages to derive the Maxwellian distribution (Brush, 1967; Uffink, 2007; von Plato, 1991). The main strength of Boltzmann's result is its generality, indeed it is robust upon any particular assumption about collisions or any other detail of the mechanical model involved, with the only requirement that the system must obey the constancy of total energy. The main weakness of this result is that it depends on what Ehrenfest and Ehrenfest-Afanassjewa (1959) baptized as the *ergodic hypothesis*, i.e. the assumption that the trajectory of the system will eventually pass through all points of the phase space.[3] As Maxwell (1879, p. 713) puts it:

> The only assumption which is necessary for the direct proof [of the distribution law of energy] is that the system, if left to itself in its actual state of motion, will, sooner or later, pass through every phase which is consistent with the equation of energy.

Boltzmann (1868) was not unaware of the importance of this assumption for his derivation and argues: "If all initial states lead to periodic motions

averages to explain the empirical success of the microcanonical ensemble, which leads to the further problem of explaining that time averages are equal to the results of a macroscopic measurement. The latter is usually justified by assuming that measurements take an amount of time that is long compared to microscopic relaxation times. Since this chapter focuses on the contribution of Ehrenfest and Ehrefest-Afanassjewa on the ergodic hypothesis, I will mostly refer to the use of time averages in the interpretation of probabilities and not to the explanation of the success of the microcanonical ensemble. For an analysis of this second problem see e.g. van Lith 2001; Uffink 2007; Palacios 2018.

[3]The name "ergodic hypothesis" was coined by Ehrenfest and Ehrenfest-Afanassjewa (1959). In fact, Boltzmann introduced the word *Ergode* only in 1884 to denote an ensemble of systems with a certain probability distribution in phase space. This use of the word "ergodic" by Ehrenfest and Ehrenfest-Afanassjewa has led to some controversy among commentators. For Brush (1967, p. 169), Ehrenfest and Ehrenfest-Afanassjewa gave an entire different meaning to ergodicity than the one given by Boltzmann. For Uffink (2007, p.39) instead, they were justified in using the term "ergodic" to denote this hypothesis (see further discussion in Section 3).

not running through all possible states compatible with the total energy, there would be an infinity of different possible temperature equilibria"[p.96]. Furthermore, he was not unaware of the controversial character of this assumption. In fact, he recognized the possibility of the existence of periodic motions that fail to be ergodic. However, by studying the behavior of so-called Lissajous figures, he observed that small irregularities would destroy this regular behavior obliging the system to pass through every phase consistent with the equation of energy. He then justifies the ergodic hypothesis more generally by appealing to irregular distortions of external forces or surrounding gas molecules:

> The great irregularity of thermal motion and the manifold forces affecting bodies from the outside make it probable that the atoms of the warm body, through the motion we call heat, run through all the positions and velocities compatible with the equation of kinetic energy, so that we can use the equations developed above on the coordinates and component velocities of the atoms of warm bodies. (Boltzmann, 1871, p. 679)

It is useful to illustrate this idea with a simple example. Consider a case of a hard-sphere system in a box where every particle moves on the same straight line being reflected at each end from a perfectly smooth parallel wall. Such systems would remain on a region of the phase space without visiting the entire phase space. However, those systems would be extremely unlikely since the slightest perturbation would destroy the perfect alignment.

In spite of this, Boltzmann recognized the hypothetical character of this justification and remained sceptic about the validity of this hypothesis. This is clear by the fact that in later works (e.g. Boltzmann 1877) he attempted to characterized thermal equilibrium differently making explicit that these alternative approaches avoid the hypothesis of ergodicity (See also Uffink 2007, p.42).

3 Ehrenfest and Ehrenfest-Afanassjewa's critique of the ergodic hypothesis

In their *Encyklopädie* article, Paul Ehrenfest and Tatjana Ehrenfest-Afanassjewa (1959) wrote an extensive and influential critique of Boltzmann's statistical mechanics accusing this approach of relying on what they regarded as a doubtful hypothesis, i.e. the ergodic hypothesis.

> The fundamental assumption underlying [Boltzmann's] investigation is the hypothesis that the gas models are ergodic systems [...]. With the help of this hypothesis Boltzmann computed the time average of, for instance, the kinetic energy of each atom (the same value is obtained for all atoms!). (Ehrenfest and Ehrenfest-Afanassjewa, 1959, p. 24)

> However, the existence of ergodic systems (i.e. the consistency of their definition) is doubtful. So far, not even one example is known of a mechanical system for which the single G-path approaches arbitrarily closely each point of the corresponding energy surface. Moreover, no example is known where the single path actually traverses all points of the corresponding energy surface (Ibid, p. 22).[4]

However, the emphasis in the doubtful character of what they called "the ergodic hypothesis" was not the main contribution of the *Encyklopädie* article. As mentioned above, Boltzmann (1871) had already recognized the controversial character of this assumption. The main contribution of the *Encyklopädie* article, as I see it, comes precisely from what has been regarded by some historians of science as historical inaccuracies of the article (Brush, 1967, 1971; von Plato, 1991). In particular, that they exaggerated the role that ergodicity played in Boltzmann's statistical mechanics and that they attributed to Boltzmann an assumption that he probably never believed in. I will argue next that these aspects, which go beyond Boltzmann's investigations, are precisely the ones that really advanced the development of the ergodic theory in the 20th century.

Although there is no consensus among modern commentators about the role and status of the ergodic hypothesis in Boltzmann's approach, most of them agree that the ergodic hypothesis was a justification for the time average interpretation of probabilities (von Plato, 1991; Uffink, 2007; Brush, 1967). Based on what has been said in the previous section, a straightforward way

[4] They define a single G-path as the trajectory of the moving image point corresponding to the phase changes of a gas model (Ehrenfest and Ehrenfest-Afanassjewa 1959, footnote 73).

of giving a time average interpretation of probabilities is by equating time averages f^* and phase averages $\langle f \rangle$:

$$\int_{\Gamma_E} f(x)p(x)dx = lim_{t\to\infty}(1/t)\int_0^t f(T(t,x))dt \tag{3}$$

However, as modern commentators have pointed out (von Plato, 1991; Uffink, 2007), this particular motivation for assuming ergodicity is not to be found anywhere in Boltzmann's writings and seemed to have been introduced for the first time by Ehrenfest and Eherefest-Afanassjewa in their review of Boltzmann's statistical mechanics. von Plato (1991, p. 78) expresses this as follows:

> This simplified reading of Boltzmann is due to the review in 1911 of foundations of statistical mechanics by Paul and Tatiana Ehrenfest. They called the above justification for assuming a single trajectory [the equivalence between phase and time averages] the 'Boltzmann-Maxwell justification', and it has since then been accepted as standard. But that particular motivation for assuming ergodicity is not used by Boltzmann.

Similarly, Uffink (2007, p. 42) claims:

> There is however *no* evidence that Boltzmann ever followed this line of reasoning neither in the 1870s, nor later. He simply never gave any justification for equating time and particle averages, or phase averages, at all. Presumably, he thought nothing much depended on this issue and that it was a matter of taste.

Now since Boltzmann attributes himself a time average interpretation of probabilities, one may wonder why he did not explicitly attempt to equate phase and time averages. One can speculate that since he was aware of the possibility of periodic motions, he did not believed that this equivalence holds exactly (von Plato, 1991, p.78). One can also believe, as Uffink (2007) does, that for him nothing much depended on this equivalence. If this latter interpretation is correct, then one should give Ehrenfest and Ehrenfest-Afanassjewa all the credit for having emphasized the important role of this equivalence for a time average interpretation of probabilities. Indeed, for them this equivalence was not just "a matter of taste" but the only way of warranting the uniqueness of the stationary probability distribution, which means that all motions associated with the same total energy yield the same value for the time average of any function (Ehrenfest and Ehrenfest-Afanassjewa, 1959, p.22). Their reasoning can be summarized as follows. If one identifies the ergodic hypothesis as the assumption that if a system that is left to itself will pass through all the phase points compatible with its total energy. Then, given that a point in phase space cannot lie on more than one trajectory, all systems with the same value of the total energy will follow the same trajectory and their averages over infinite time intervals will be equal. The equivalence between phase and time averages gives in this way a neat interpretation of probabilities in equilibrium statistical mechanics and a clear connection to the dynamics of the system that does not depend on the initial conditions. Another significant advantage of equating time and phase averages is that phase averages can be calculated in many cases, whereas the time averages cannot (Moore, 2015). The problem, as Ehrenfest and Ehrenfest-Afanassjewa stated it, is that this equivalence seems to rely crucially on the ergodic hypothesis, which they intuitively believed was not only doubtful but mathematically impossible, but more on this later.

From what has been said here, one can see that the *Encyklopädie* article, perhaps by exaggerating the role that ergodicity played in Boltzmann's approach advanced a time average interpretation of probabilities and prompted at least three important challenges that were later conceived as "the ergodic problem". The first challenge was to demonstrate that the limit involved in the definition of time averages f^* exists. The second challenge was to prove that this limit is independent of x and equal to the phase average. The third challenge was to determine the validity of the ergodic assumption or any analogous assumption required to derive this equivalence. As it will be seen later, the first two problems had to await the ergodic theorems of 1931 and 1932 and the new concept of metric transitivity (that replaced ergodicity) to be given a definite solution (See Section 5).[5] The third problem does not have an unequivocal answer yet, but the impossibility theorems of 1913 demonstrated that at least the strict ergodic hypothesis was not a valid assumption (See Section 4).

Another aspect of the *Encyklopädie* article that has been regarded as

[5]Physicists frequently identify the term "ergodicity" with "metric transitivity", however, I will argue below that these are two well-defined different concepts that should not be confused.

a historical error concerns the definition of the ergodic hypothesis (Brush, 1967). It is important to note that Boltzmann never (at least not explicitly) associated what Ehrenfest and Ehrenfest-Afanassjewa defined as ergodic hypothesis with the word *Ergode*, which was actually introduced only in Boltzmann (1884), and was used to denoted a stationary ensemble (i.e. the micro-canonical ensemble) with only one integral of motion, its total energy. However, Boltzmann (1884) did assume that every element of such an ensemble traverses every phase point with the given energy. According to Uffink (2007, Footnote 20) this would have justified Ehrenfest and Ehrenfest-Afanassjewa in using this term in their formulation of the hypothesis. But beyond the discussion of whether or not the "ergodic hypothesis" should receive that name, Ehrenfest and Ehrenfest-Afanassjewa have been strongly criticized for attributing to Boltzmann an hypothesis that he probably never believed in. More to the point, Ehrenfest and Ehrenfenst-Afanassjewa defined the ergodic hypothesis as follows:

Definition 3.1. *Ergodic hypothesis* The single, undisturbed motion of the system, if pursued without limit in time, will finally traverse "every phase point" which is compatible with its given energy. (Ehrenfest and Ehrenfest-Afanassjewa, 1959, p. 21)

For Brush (1967) and von Plato (1991), this definition is misleading for at least two reasons. First, because it refers to one single trajectory instead of an ensemble in which there is a continuous number of different trajectories. Second, because it uses the terms "every phase point", which is something that most probably Boltzmann never believed in. Regarding the first point, it is true that Boltzmann introduced the word *Ergode* to denote ensembles instead of single trajectories. However, Boltzmann did not introduce the notion of ensemble until 1884, whereas the hypothesis that the system will pass through "every phase" consistent with its given energy was used much earlier. For instance, in 1871 he explicitly refers to the motion of a point-mass in a plane under the influence of an attractive force. He says that if the force is described by a potential function $1/2(ax^2 + by^2)$ (the compound harmonic motion which results in the so-called Lissajous figures) and the ratio of the periods of the two motions is irrational, then the point-mass goes through all possible positions within a certain rectangle (Brush 1967, p. 169). Therefore, the formulation of Boltzmann's hypothesis in terms of single trajectories seems to reflect at least the first uses of this hypothesis by Boltzmann. One should also note that Ehrenfest and Ehrenfest-Afanassjewa had a further reason to formulate this hypothesis in terms of single trajectories, since, as mentioned above, by the uniqueness of mechanical trajectories, there would be essentially one trajectory, so that one can replace a long trajectory of a single system by an average over all points on the energy surface.

The second objection to Ehrenfest and Ehrenfest-Afanassjewa's formulation of the ergodic hypothesis was that it is stronger than the hypothesis that Boltzmann probably had in mind. Although Boltzmann explicitly used the terms "through every point", Brush (1967) pointed out that what he probably meant was something close to what Ehrenfest and Ehrenfest-Afanassjewa termed as the "quasi-ergodic" hypothesis, which they define as follows (footnote 98):

Definition 3.2. *Quasi-Ergodic hypothesis* The single, undisturbed motion of the system, if pursued without limit in time, will approach "arbitrarily closely each point" which is compatible with its given energy, which means that the trajectory is dense.

The reason that Brush (1967) and other historians of science (e.g. Borel 1915; von Plato 1991) offer for concluding that Boltzmann interpreted "the ergodic hypothesis" in the sense of the "quasi-ergodic hypothesis" is that in some passages, he actually added the qualification "approximation" to the description of ergodic behavior without pointing out a distinction between these behaviors and what we could understand as strict ergodic behaviors. For example, in a discussion of Kelvin's test-cases in 1892 he claims: "all possible sets of values of x, y, and θ which are consistent with the equation of *vis viva* are obtained with any required degree of approximation [*mit beliebiger Annherung erreicht werden*]" (quoted in Brush 1967, p. 173). They suppose then that the use of the terms "through every point" were just used by Boltzmann as an approximation of trajectories going "arbitrarily close to every point" (Brush 1967, p. 174). We could express this idea in terms of contemporary philosophy of science and say that Boltzmann believed strict ergodic trajectories to be a *straightforward idealization* of the trajectories going arbitrarily close to every point, i.e. an idealization that constitutes an approximation of realistic behavior and that can therefore be de-idealized without loosing explanatory power (See Butterfield (2011)).

If the previous interpretation is correct, then it seems then that Ehrenfest and Ehrenfest-Afanassjewa were wrong in taking the idea of trajectories going literally "through every point" in Boltzmann's writings too seriously, after all Boltzmann just meant trajectories going "arbitrarily close to every point" (Borel, 1915; von Plato, 1991; Brush, 1967). However, this aspect, which appears to be a historical misconception of the *Encyklopädie* article and the result of "some careless statements made by the Ehrenfests" (Brush, 1967, p. 169), is at the same time one of the main contributions of the article about the ergodic hypothesis. As I see it, Ehrenfest and Ehrenfest-Afanassjewa advanced Boltzmann's ideas by pointing out that the difference between "ergodic behavior" and "quasi-ergodic behavior" is mathematically essential, which implies that the ergodic hypothesis cannot be de-idealized and replaced by the quasi-ergodic hypothesis without loosing explanatory power. In footnote (98) they illustrate this essential difference between "ergodic" and "quasi-ergodic" behavior by help of the following example. Consider a geodesic line of a torus for which the ratio of the two numbers of turnings in the two directions is irrational. Such a geodesic intersects the meridian at infinitely many points P_h, which are densely distributed everywhere over the circumference. No matter how many times one turns around the torus along the geodesic line, one will never get from a point P_h to the diametrically opposite point Q on the meridian. This is because if that were the case, then twice the same number of revolutions would bring us back to P_h and the system would stay in a particular region of the phase space, thus failing to describe a dense trajectory. This means that the points that can be visited by a dense trajectory constitute a subset of all points of the phase space. "From this one can easily see that the set of all those points P_h which can be reached by a given geodesic line form a denumerable subset in the continuum of all those points on the circumference which the geodesic line approaches arbitrarily closely" (Ehrenfest and Ehrenfest-Afanassjewa, 1959, Footnote 98). Here Ehrefest and Ehrenfest-Afanassjewa were not only pointing out an essential difference between "ergodic" and "quasi-ergodic" trajectories, but also suggesting that the notion of strict ergodic behavior was not consistent and therefore mathematically impossible (See Section 4).

But there was a further and more important reason to distinguish between "ergodic" and "quasi-ergodic" behavior, namely that they suspected that quasi-ergodic behavior does not entail the desired conclusion that Maxwell's distribution is the only stationary distribution over the energy surface and therefore that it does not warrant the equivalence between phase and time averages: "[W]e must say that for a "quasi-ergodic" system on each surface $E(q,p)$ there will be a continuum of $\infty^{(2rN-2)}$ different G-paths with different values of the constants $c_2, ..., c_{2rN_1}$. Hence one cannot extend the Boltzmann-Maxwell justification [...] to quasi-ergodic systems" (footnote 99). "The time average in question can change quite discontinuously from path to path for a quasi-ergodic system, because we obtain it by averaging over an infinite time interval" (footnote 102).

Although they did not offer a careful proof of the previous statements, their distinction between ergodic and quasi-ergodic hypothesis suggests that the ergodic hypothesis corresponds to what in the contemporary philosophy of science (Fletcher et al., 2019) is called an "essential idealization", i.e. an idealization that cannot be de-idealized without loosing explanatory power. In this case the idealized assumption was the ergodic hypothesis, which Ehrenfest and Ehrenfest-Afanassjewa suggested cannot be de-idealized by the weaker and more plausible "quasi-ergodic hypothesis". The latter ideas led them to the following puzzle:

> *Ehrenfest and Ehrenfest-Afanassjewa's puzzle:* On the one hand, by assuming ergodicity, one can demonstrate the equivalence between phase and time averages. Yet the existence of ergodic systems is doubtful. On the other hand, if one de-idealizes this assumption by the weaker hypothesis of quasi-ergodicity, which is probably true for some systems, one does not obtain the desired equivalence.

This clear way of presenting the issues surrounding the ergodicity assumption is in my view the second major contribution of the *Encyklopädie* article on this topic, apart from the above mentioned emphasis on the equivalence between time and phase averages. We will see in the next sections that in order to solve the Ehrenfest and Ehrenfest-Afanassjewa's puzzle, one needs to introduce elements of modern topology and the new notion of metric transitivity, which, as I will argue below, corresponds neither to the ergodic hypothesis nor to the quasi-ergodic hypothesis.

4 Proof of the impossibility of ergodic systems

We have seen above that the *Encyklopädie* article not only raised suspicion on the actual existence of ergodic systems but also on the mathematical possibility of ergodic systems. The challenge was therefore not only for physicists, who were asked to justify the validity of this assumption but also for mathematicians, who were now challenged to demonstrate that ergodic systems cannot exist in principle. Mathematicians Arthur Rosenthal and Michel Plancherel accepted the challenge and independently demonstrated in 1913, soon after the publication of the *Encyklopädie* article, that Ehrenfest and Ehrenfest-Afanassjewa's intuition was correct and that a mechanical system represented by a phase space with more than one dimension cannot pass through every point on the energy surface. Rosenthal (1913) explicitly recognizes the direct influence of Ehrenfest and Ehrenfest-Afanassjewa in his "Proof of the impossibility of ergodic gas" and begins his article by saying: "In view of the fact that no example of such an ergodic system has been demonstrated with certainty P. and T. Ehrenfest doubted the existence of ergodic systems (i.e. they doubted that their definition is not contradictory). In the following it will be shown that this doubt was correct; i.e., *it will be shown* that not only *no gas* is an ergodic system, but also that in general such systems cannot exist." (p. 796)

In order to understand the proofs offered by Rosenthal (1913) and Plancherel (1913), it is necessary to review some of the results obtained in pure mathematics during the second half of the nineteenth century, in which these proofs were based. The ergodic hypothesis, as formulated by Ehrenfest and Ehrenfest-Afanassjewa, implies that a certain curve, which can be placed in correspondence with a straight line (the time axis), eventually visits all points of the phase space. This means that such a curve would appear one-dimensional from the time axis point of view and multidimensional from the phase space point of view, if the phase space has more than one dimension. This obliged to introduce a method for comparing the sizes of infinite classes of different dimensionality. Cantor, who developed his work on set theory during 1871 and 1897, furnished such a method by using the criterion of one-to-one correspondence. The consequence of Cantor's theory that was relevant for the understanding of ergodicity was the proof that any n-dimensional manifold can be put into one-to-one correspondence with any m-dimensional manifold, where n and m may have different dimensionality (Cantor 1878).[6] However, this proof lacks the property of continuity, which means that points that are close together in the n-dimensional manifold may be mapped into points that are far apart in the m-dimensional manifold. Cantor suspected then that bicontinuous one-to-one mapping may serve as a criterion for proving that two sets have the same dimensionality, which was

[6]See Brush 1967 for a historical review of these results

finally demonstrated by Brouwer (1911).

Other developments needed to establish the impossibility of ergodic systems in 1913 were the results offer by Borel (1898) and Lebesgue (1902), who completed Cantor's method for comparing infinite sets of points by providing a method for determining the length, area, volume, and more generally, the "measure" of sets of points. In these approaches the measure of a point is defined to be zero whereas the measure of all real numbers in a finite interval is defined as the length of that interval. Based on these results, Plancherel (1913) and Rosenthal (1913) published their celebrated proofs. Without going into technical details, Rosenthal (1913)'s proof can be summarized as follows. Consider a gas system of N particles and r degrees of freedom, where all states of the system are represented by a $2rN$-dimensional phase space Γ. One can choose a small region G of the energy surface Γ_E in which all the partial derivatives of the energy with respect to the coordinates and momenta are continuous functions of the coordinates and momenta, such that at least one of these derivatives is different from zero everywhere in that region. One can then map this region G into a $2rN-1$-dimensional cube. According to the ergodic hypothesis, the representative point of the system must pass through every point in that region G. This can happen in two ways: i) In a finite time interval (one-dimensional time axis), the representative point enters G, visits all points, and comes out again, which means that the $2rN-1$-dimensional region G should be mapped onto a line of finite length, continuously and one-to-one. ii) Or the representative point passes into and out of G infinitely many times. Rosenthal's proof shows that neither of these alternatives is possible, except in the trivial case when the phase space is one-dimensional. For both alternatives imply that there is a bicontinuous one-to-one mapping in sets of different dimensionality, which according to Brouwer's proof is impossible. Plancherel's proof follows a similar reasoning but using Lebesgue theory of measure. In particular, the proof is based on the result that the time-axis, a line, is a set of measure zero with respect to a region of two or more dimensions.[7]

These proofs demonstrated that the first part of what has been called here "Ehrenfest and Ehrenfest-Afanassjewa's puzzle" was correct: a mechanical system cannot be ergodic, except in cases when the phase space is one-dimensional. It remained to be demonstrated the truth of the second part of this puzzle, namely that the weaker "quasi-ergodic" hypothesis was

[7]See Brush 1967 for more details about this proof

not sufficient to derive the equivalence between time and phase averages. An indirect proof of this statement had to await the establishment of the ergodic theorems of Birkhoff (1931) and von Neumann (1932), which stated the necessary and sufficient conditions for the equivalence between time and phase averages. We review these results in the following section.

5 The ergodic theorems and the notion of metric transitivity

In the 1930's, George D. Birkhoff (1931) and John von Neumann (1932) published two separated papers containing different versions of what is now known as "the ergodic theorem". This theorem provided a key insight into the problem presented in a clear way for the first time by Ehrenfest and Ehrenfest-Afanassjewa, namely the justification of the hypothesis that time averages equal phase averages. At the same time, it initiated an entire new field of mathematical research called ergodic theory, which has thrived more than 80 years.[8]

The basic concept that allowed von Neumann and Birkhoff to arrive at the celebrated result was the notion of "metric transitivity", introduced for the first time by Birkhoff and Smith (1928). In order to understand these results and the notion of metric transitivity, one needs to become familiarized with elements of measure theory. Let (X, ϕ_t, μ) be a dynamical system given by a metric space (phase space) X, a continuous map $\phi : X \to X$, and equipped with a normalized Lebesgue-measure μ (i.e. $\mu = 1$) restricted to X. A point $x \in X$, which represents a particular state of the system, "moves" in time, generating a "flow" that we denote as $\phi_t(x)$. $\phi_t(x)$ is the position to which the system moves after time t so that $\phi_t(x)$ is the solution of the differential equation with initial value x at time $t = 0$. One should also add that $\phi_t(x)$ is a homeomorphism of X onto itself, which satisfies the group property $\phi_t \phi_s = \phi_{t+s}$. It also preserves the Liouville measure μ. The notion of "metric transitivity", which replaced the old notion of "ergodicity" introduced by Boltzmann (1868), was then defined as follows:

Definition 5.1. *Metric transitivity* The dynamical system (X, ϕ_t, μ) is *metrically transitive* iff for any measurable set of nonzero measure V and for *almost every point* $x \in X$ it holds that $\{\phi_t(x)\} \cap V \neq \emptyset$ for some time t.

[8] See Mackey (1974) for an excellent historical review of the ergodic theory.

This means that eventually almost every point $x \in X$ (i.e. all points except for a set of measure 0) visits every measurable set V in X. If a dynamical system is metrically transitive, it follows that it cannot be decomposed into two (or more) invariant regions of non-zero measure.

By using this notion of metric transitivity, the theorems of Birkhoff (1931) and von Neumann (1932) established that the time limit $t \to \infty$ used in the following definition of time averages f^* exists for almost all x, and is independent of x when it exists:

$$f^* = lim_{t\to\infty}(1/t)\int_0^t f(\phi_t(x))dt, \tag{4}$$

where f is an integrable function on the phase space X, representing a physical measurement on a system that is in state $x \in X$. [9]

In short, they demonstrated that the following theorem (the Ergodic Theorem) holds:

Theorem 5.1. *Ergodic Theorem* If the dynamical system (X, ϕ_t, μ) is metrically transitive, then the limit of f^* exists and coincides with the phase average $\langle f \rangle = \int_X f(x)d\mu(x)$, for *almost all* $x \in X$.

The proof of theorem 5.1 constituted the first crucial step towards the solution of the long-standing problem of the equivalence between phase and time averages posed in a clear way by Ehrenfest and Ehrenfest-Afanassjewa. However, as one can observe, this theorem depends essentially on the assumption that systems are "metrically transitive", an assumption that is not quite easy to justify. In this sense, one can say that the ergodic theorem transformed the question of equivalence between time and phase averages into the question of whether the flow ϕ representing the time evolution of the system is metrically transitive. I will return to this problem in the next section, but first I will put these results in the context of the previous discussion around the ergodic hypothesis.

In Section 3, we said that Ehrenfest and Ehrenfest-Afanassjewa posed the following puzzle: On the one hand, if we assume ergodicity, then one can derive the equivalence between time averages and phase averages. However,

[9]von Neumann (1932) demonstrated that the functions of x on the time average converge and Birkhoff (1931) proved further that this convergence was pointwise almost everywhere. See Moore (2015) for more details on the difference between von Neumann and Birkhoff's results.

real systems cannot be ergodic. On the other hand, if we assume the weaker hypothesis of quasi-ergodicity, which may be true of some systems, then one cannot derive the equivalence between phase and time averages. In Section 4, we saw that the first part of this puzzle was correct, indeed Rosenthal (1913) and Plancherel (1913) demonstrated that real systems represented by a phase space with more than one dimension cannot be ergodic. This naturally raises the question of whether metric transitivity is equivalent to the original ergodic hypothesis. If they were equivalent, then theorem 5.1 would loose interest, since it would not be applicable to almost any real system of interest. Fortunately, the notion of metric transitivity has been proven to be weaker than the original ergodic hypothesis and therefore immune to the impossibility proofs of ergodic systems elaborated by Rosenthal and Plancherel (Moore, 2015). The key aspect that weakens the definition of metric transitivity is the introduction of the expression "for almost every x" in the definition of metric transitivity, which means for all x except for a set of measure zero.[10] Furstenberg (1961) demonstrated by considering an r-dimensional torus, that convergence "almost everywhere" can be replaced by "convergence everywhere" only in cases of $r = 1$, which is consistent with Rosenthal and Plancherel's results. In all other cases, i.e, for $r > 1$, one needs to impose further restrictions on the transformation ϕ.

Another question that arises in light of the Ehrenfest and Ehrenfest-Afanassjewa's puzzle is whether the notion of metric transitivity corresponds to what they dubbed as the "quasi-ergodic" hypothesis. If this were the case, then the second part of the puzzle would have been proven to be false, since "quasi-ergodicity" would have been demonstrated to be sufficient for the equivalence between phase and time averages. Interestingly, although metric transitivity is frequently taken as equivalent to the quasi-ergodic hypothesis (e.g. Lebowitz and Penrose (1973)), the former can be demonstrated to be stronger than the later.[11] More specifically, as Ehrenfest and Eherenfest-Afanassjewa define it, the quasi-ergodic hypothesis corresponds

[10] There is an interesting foundational problem associated with the definition of metrical transitivity for "almost every x", which is called "the measure zero problem". The issue is that it is hard to demonstrate that states with probability measure zero can be neglected without begging the question, i.e. without presupposing that phase averages equal time averages. I will not discuss this problem further for lack of space, but the reader can see Uffink (2007); van Lith (2001); Frigg (2016) for a detailed discussion around this issue.

[11] Perhaps the misleading title of von Neumann's paper (1932) "Proof of the quasi-ergodic hypothesis contributed to this confusion

to the hypothesis that the orbits (trajectories) are topologically dense in the phase space (i.e. they pass arbitrarily close to every point of phase space). It can be demonstrated that metric transitivity implies quasi-ergodicity, when X is a compact metric, which is generally assumed in the ergodic theorems (Moore, 2015; Kolyada and Snoha, 1997). However, the converse is not valid, since it requires the further assumption that X has no isolated point, which does not necessarily hold in the systems of interest (Kolyada and Snoha, 1997).[12] In fact, it is not even true that a minimal flow with an invariant measure, in which every orbit is dense, is metrically transitive (Moore, 2015). Since metric transitivity is a necessary and sufficient condition for the validity of theorem 5.1., then one can conclude that the "quasi-ergodic hypothesis" is too weak to establish the equivalence between phase and time average, which means that the second part of the Ehrenfest and Ehrenfest-Afanassjewa's puzzle was also true.

Let us now summarize the contribution of Ehrenfest and Ehrenfest- Afanassjewa to the development of the ergodic theorems in the 1930's. The first direct contribution was to point out for the very first time the importance of the hypothesis that time averages equal phase averages in a time average interpretation of probabilities. In absence of this emphasis on the role of this hypothesis, there may have not been enough motivation to provide such careful mathematical proofs of this equivalence. Another direct contribution of Ehrenfest and Ehrenfest-Afanassjewa was to suggest that neither the ergodic hypothesis nor the quasi-ergodic hypothesis can serve to derive the desired equivalence between time and phase averages for the systems of interest, which we have called here the Ehrenfest and Ehrenfest-Afanassjewa's puzzle. This motivated mathematicians and physicists to find a different hypothesis, i.e. metric transitivity, that can serve as a basis to derive the equivalence between phase and time averages. The proof of this "ergodic" theorem based on the notion of metric transitivity finally solved the puzzle prompted two decades before by Ehrenfest and Ehrenfest-Afanassjewa. But, as said above, it led to a different, yet more specific problem: the problem of demonstrating that the real physical systems of interest are in fact metrically transitive. In the next section, I will discuss this issue further.

[12] Systems in which X has no isolated point is said to be a standard dynamical system (Kolyada and Snoha, 1997)

6 The problem of metric transitivity

How can one prove that the flow of a dynamical system is metrically transitive? And how can we be sure that metrically transitive systems exist? These have proved to be very challenging questions that have motivated an enormous amount of research (e.g. Oxtoby and Ulam 1941; Markus and Meyer 1974; Sinai 1970). A promising result to the existence of metrically transitive systems was offered Oxtoby and Ulam (1941), who showed that on a compact polyhedron equipped with a finite Lebesgue measure, all measure preserving homeomorphisms are metrically transitive in a topological sense. A more concrete example of metrically transitive systems was examined by Sinai (1970), who considered a model of dynamical systems, where the molecules were contained in a cubical enclosure and moved with periodic boundary conditions. Although the model was not entirely realistic (it allowed for collisions between the molecules but not with the walls of the container), he proved that the system was (approximately) metrically transitive. These results contrasted with the ones offered by Markus and Meyer (1974), who showed that for Hamiltonian dynamical systems, almost all systems fail to be metrically transitive. The latter was reinforced by the so called KAM (Kolmogorov, Arnold, Moser) Theorem, which stated that when the interactions among the molecules are non-singular, the phase space will contain islands of stability where the flow is not metrically transitive (see Lichtenberg and Lieberman 2013; Earman and Rédei 1996). To be more specific, the theorem shows that if one starts by a Hamiltonian system with quasi-periodic trajectories and adds perturbation terms that are intended to eliminate this periodic behavior, there will still remain "islands" of periodic behavior so that the system fails to be metrically transitive. Based on these results one can conclude that most systems of interest in statistical mechanics are very probably not metrically transitive (Wightman 1985; Earman and Rédei 1996; van Lith 2001).

Different reactions can be found among philosophers of science on the consequences of these results. For some (e.g. Earman and Rédei 1996; van Lith 2001) these results lead at the conclusion that the traditional ergodic (or metrically trasitivity) program for interpreting probabilities and explaining the success of phase averaging should be abandoned. Others have suggested instead (Vranas 1998; Frigg and Werndl 2011) that the ergodic program can be rescued by appealing to what they call "epsilon-ergodicity" (or more precisely "epsilon-metrical transitivity"). This latter view is motivated by the possibility that most systems that fail to be metrically transitive have a probability measure that is *close enough* to the microcanonical. More specifically, Vranas (1998) examines computational evidence for the existence of systems that he calls "epsilon-ergodic", which are systems that have an invariant subset B of measure $1 - \epsilon$, such that i) ϵ is tiny or zero and ii) for *almost every point* $x \in X$ it holds that $\{\phi_t(x)\} \cap B \neq \emptyset$ for some time t. Then, by generalizing the notion of absolute continuity, he proved that if a system is "epsilon-ergodic" (which here can be understood as "epsilon-metrically transitive"), the probability measure associated to that system will be close to the microcanonical. One can build an intereresting parallel with the previous discussion around the ergodic hypothesis and interpret this "epsilon-ergodicity hypothesis" as an analogue of the quasi-ergodic hypothesis. The hope is then that one can de-idealize the hypothesis of metrical transitivity by the hypothesis of "epsilon-ergodicity" or "epsilon metrical transitivity" without loosing explanatory power, since it is expected that in the latter case the probability measure will be *close enough* to the microcanonical measure.

Even accepting Vranas' results, there still remains the question of whether most of systems of interest are "epsilon ergodic" or more precisely "epsilon metrically transitive". There is important evidence suggesting the existence of systems that display thermodynamic behavior and yet are not metrically transitive or even "epsilon-metrically transitive" (Frigg 1999, Uffink 2007). For example, in a solid the molecules can oscillate around fixed positions so that the phase trajectory of the system can only access a small part of the energy hypersurface (Uffink, 2007; Frigg, 2016). Frigg and Werndl (2011) have argued that this is not as problematic as it seems, since one can still use the ergodic theory for the restricted set of cases that are proven to be "epsilon ergodic" (epsilon metrically transitive) such as gases. However, Earman and Redei (1996) are skeptic about this line of reasoning, since they claim that if there are non-metrically transitive systems that display thermodynamic-like behavior, then it is likely that the same mechanisms that explain this behavior in non metrically transitive systems also explain the behavior of metrically transitive systems (or epsilon-ergodic) systems.[13]

The question about the validity of the ergodic theory based on the notion of metric transitivity that followed the discussion started by Ehrenfest and

[13] Werndl and Frigg (2015) prove a theorem that establishes that for equilibrium to exist three factors need to cooperate: the choice of macro-variables, the dynamics of the system, and the choice of the state space. For them a consequence of this theorem is that focusing on ergodicity as the crucial property for the existence of an equilibrium state is misleading.

Ehrenfest-Afanassjewa remains open in the foundations of statistical mechanics. One of the reasons why this theory has not been given up despite the complications to demonstrate that systems are in fact metrically transitive and other problems associated with this approach is that this theory gives a solid foundation to the time average interpretation of probabilities and a neat mechanical explanation of thermodynamic equilibrium (Frigg, 2016).[14] Giving up the time average interpretation of probabilities in order to get rid of the problems associated with the ergodic hypothesis is a high price to pay, since the alternatives face similar if not more dramatic problems. Indeed, the frequentist interpretation of probabilities violates the requirement of von Mises's theory (van Lith, 2001; Frigg, 2016), and the propensity interpretation (Popper, 1959) is inconsistent with the assumption of a deterministic underlying micro theory (Clark and Butterfield, 1987; Frigg, 2016). A different strategy to deal with the problem of interpreting probabilities consists in avoiding probabilities all in all. This approach is known as the "typicality approach" and is based on the distinction between 'typical states', which correspond to equilibrium states and 'atypical' states, which are non-equilibrium states (Lebowitz, 1993; Goldstein, 2001). Although this program solves some problems associated with the ergodic theory, it has been criticized for not establishing a clear connection with the dynamics of the system (Frigg, 2009, 2010). Very recently, Wallace (2016) has suggested a novel interpretation of statistical mechanical probabilities based on quantum-mechanical probabilities, yet the empirical validity of this approach is still to be seen.

[14] There are other important problems associated with the ergodic approach that have not been mentioned here because they are not directly related with the problems discussed by Ehrenfest and Ehrenfest-Afanassjewa. One of these problems is the measure zero problem, which I mentioned on footnote 9. The other problem concerns the justification of infinite-time limits in the definition of time averages. The worry here is that if one wants to explain the empirical success of the microcanonical distribution it is not clear why one should interpret measurements as time averages, even less as infinite time averages (Uffink, 2007; Frigg, 2016; Palacios, 2018). Finally, there is the problem that time average interpretation cannot easily be generalised to time-dependent phenomena and therefore this theory seems to be restricted to the explanation of equilibrium and cannot be used as a general non-equilibrium theory (See Uffink 2007; van Lith 2001).

7 Conclusion

We have traced the history of the ergodic hypothesis from its origins to recent discussions in the foundations of statistical mechanics highlighting the contribution of Ehrenfest and Ehrenfest-Afanassjewa in this debate. We have seen that apart from pointing out the difficulties associated to the demonstration of the existence of ergodic systems, they motivated a distinction between "ergodic" and "quasi-ergodic" systems and, more importantly, they emphasized the role of the hypothesis that time averages equal phase averages in a time average interpretation of probabilities. We have seen that the latter, which has been sometimes regarded as a historical error of their analysis, served to the postulation of the ergodic theorem in the 1930's and encouraged the development of the concept of metric transitivity, which continues playing a role in the foundations of statistical mechanics.

References

Birkhoff, G. D. (1931). Proof of the ergodic theorem. *Proceedings of the National Academy of Sciences*, 17(12):656–660.

Birkhoff, G. D. and Smith, P. (1928). Structure analysis of surface transformations. *Journal de Mathématiques pures et appliquées*, 7:345–380.

Boltzmann, L. (1868). Studien über das Gleichgewicht der lebenden Kraft zwischen bewegten materiellen Punkten. *Wiener Berichte*, 58:517–560.

Boltzmann, L. (1871). Einige allgemeine Sätze über Wärmegleichgewicht. *Wiener Berichte*, 63:679–711.

Boltzmann, L. (1877). Über die beziehung zwisschen dem zweiten Haubtsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung resp. dem Sätzen über das Wärmegleichgewicht. *Wiener Berichte*, 76:373–435.

Boltzmann, L. (1884). Über die Eigenshaften monozyklischer und anderer damit verwandter Systeme. *Crelle's Journal für die reine und angewandte Mathematik*, 98:68–94.

Borel, E. (1898). *Leçons sur la théorie des fonctions.* Paris: Gauthier-Villars.

Borel, E. (1915). *Mécanique Statistique. Exposé d'aprés l'article allemand de P. Ehrenfest, T. Ehrenfest.* Paris : Gauthier-Villars.

Brouwer, L. E. (1911). Beweis der invarianz der Dimensionenzahl. *Mathematische Annalen*, 70(2):161–165.

Brush, S. G. (1967). Foundations of statistical mechanics 1845–1915. *Archive for History of Exact Sciences*, 4(3):145–183.

Brush, S. G. (1971). Proof of the impossibility of ergodic systems: The 1913 papers of Rosenthal and Plancherel. *Transport Theory and Statistical Physics*, 1(4):287–298.

Butterfield, J. (2011). Less is different: Emergence and reduction reconciled. *Foundations of physics*, 41(6):1065–1135.

Cantor, G. (1878). Ein beitrag zur Mannigfaltigkeitslehre. *Journal für die reine und angewandte Mathematik*, 84:242–258.

Clark, P. and Butterfield, J. (1987). Determinism and probability in physics. *Proceedings of the Aristotelian Society, Supplementary Volumes*, 61:185–243.

Earman, J. and Rédei, M. (1996). Why ergodic theory does not explain the success of equilibrium statistical mechanics. *The British Journal for the Philosophy of Science*, 47(1):63–78.

Ehrenfest, P. and Ehrenfest-Afanassjewa, T. (1959). *The conceptual foundations of the statistical approach to mechanics,* Cornell University Press, Ithaka. English translation by M.J. Moravcsik of "Begriffliche Grundlagen der statistischen Auffassung in der Mechanik". *Encyklopädie der mathematischen Wissenschaften* IV-32: F. Klein (ed.) 1-90. 1911). Re-used by Dover, Minneola (2015).

Fletcher, S. C., Palacios, P., Ruetsche, L., and Shech, E. (2019). Infinite idealizations in science: an introduction. *Synthese*, 196(5):1657–1669.

Frigg, R. (2009). Typicality and the approach to equilibrium in boltzmannian statistical mechanics. *Philosophy of Science*, 76(5):997–1008.

Frigg, R. (2010). Why typicality does not explain the approach to equilibrium. In *Probabilities, causes and propensities in physics.* Springer.

Frigg, R. (2016). A field guide to recent work on the foundations of statistical mechanics. In *The Ashgate companion to contemporary philosophy of physics*. Routledge.

Frigg, R. and Werndl, C. (2011). Explaining thermodynamic-like behavior in terms of epsilon-ergodicity. *Philosophy of Science*, 78(4):628–652.

Furstenberg, H. (1961). Strict ergodicity and transformation of the torus. *American Journal of Mathematics*, 83(4):573–601.

Goldstein, S. (2001). Boltzmanns approach to statistical mechanics. In *Chance in physics*. Springer.

Kolyada, S. and Snoha, L. (1997). Some aspects of topological transitivitya survey. *Grazer Math. Ber.*, 334:3–35.

Lebesgue, H. (1902). Intégrale, longueur, aire. *Annali di Matematica Pura ed Applicata (1898-1922)*, 7(1):231–359.

Lebowitz, J. L. (1993). Macroscopic laws, microscopic dynamics, time's arrow and boltzmann's entropy. *Physica A: Statistical Mechanics and its Applications*, 194(1-4):1–27.

Lebowitz, J. L. and Penrose, O. (1973). Modern ergodic theory. *Physics Today*, 26(2):23–29.

Lichtenberg, A. J. and Lieberman, M. A. (2013). *Regular and stochastic motion*. Springer Science & Business Media.

Mackey, G. W. (1974). Ergodic theory and its significance for statistical mechanics and probability theory. *Advances in Mathematics*, 12(2):178–268.

Markus, L. and Meyer, K. R. (1974). *Generic Hamiltonian dynamical systems are neither integrable nor ergodic*. Memoirs of the American Mathematical Society 144.

Maxwell, J. C. (1860). Illustrations of the dynamical theory of gases. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 19(124):19–32.

Maxwell, J. C. (1879). On Boltzmann's theorem on the average distribution of energy in a system of material points. *Transactions of the Cambridge Philosophical Society*, 12:547–570.

Moore, C. C. (2015). Ergodic theorem, ergodic theory, and statistical mechanics. *Proceedings of the National Academy of Sciences*, 112(7):1907–1911.

Neumann, J, v. (1932). Proof of the quasi-ergodic hypothesis. *Proceedings of the National Academy of Sciences*, 18(1):70–82.

Oxtoby, J. C. and Ulam, S. M. (1941). Measure-preserving homeomorphisms and metrical transitivity. *Annals of Mathematics*, pages 874–920.

Palacios, P. (2018). Had we but world enough, and time... but we dont!: Justifying the thermodynamic and infinite-time limits in statistical mechanics. *Foundations of Physics*, 48(5):526–541.

Plancherel, M. (1913). Beweis der Unmöglichkeit ergodischer mechanischer Systeme. *Annalen der Physik*, 347(15):1061–1063.

Popper, K. R. (1959). The propensity interpretation of probability. *The British journal for the philosophy of science*, 10(37):25–42.

Rosenthal, A. (1913). Beweis der Unmöglichkeit ergodischer Gassysteme. *Annalen der Physik*, 347(14):796–806.

Sinai, Y. G. (1970). Dynamical systems with elastic reflections. ergodic properties of dispersing billiards. *Uspehi Mat Nauk*, 25(2):141–192.

Uffink, J. (2007). Compendium to the foundations of classical statistical physics in handbook for the philosophy of physics. In *Handbook for the Philosophy of Physics*. Elsevier, Amsterdam.

van Lith, J. (2001). Ergodic theory, interpretations of probability and the foundations of statistical mechanics. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 32(4):581–594.

von Plato, J. (1991). Boltzmann's ergodic hypothesis. *Archive for History of Exact Sciences*, 42(1):71–89.

Vranas, P. B. (1998). Epsilon-ergodicity and the success of equilibrium statistical mechanics. *Philosophy of Science*, 65(4):688–708.

Wallace, D. (2016). Probability and irreversibility in modern statistical mechanics: Classical and quantum. In *Quantum Foundations of Statistical Mechanics.* Oxford University Press, forthcoming.

Werndl, C. and Frigg, R. (2015). Reconceptualising equilibrium in Boltzmannian statistical mechanics and characterising its existence. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 49:19–31.

Wightman, A. S. (1985). Regular and chaotic motions in dynamical systems: Introduction to the problems. In *Regular and Chaotic Motions in Dynamic Systems*. New York: Plenum.