Irreversibility and randomness

Nino Dekkers* and Klaas Landsman†

Abstract

We make precise sense of the idea of "molecular chaos" through algorithmic randomness of microscopic trajectories, and ground macroscopic irreversibility in the lack of symmetry under time reversal of this property. This concept of randomness is defined relative to an underlying probability measure $\mathbb{P}$ on the space of trajectories. In deterministic models like Newtonian N-particle flow in dilute gases of hard spheres (as considered by Boltzmann) or the Kac ring model these may be reduced to their initial conditions, in which case $\mathbb{P}$ makes the particles i.i.d. at $t = 0$. In the (stochastic) Ehrenfest urn model, on the other hand, the importance of *trajectories* as the decisive random objects comes out more clearly. We consider each of these models from this point of view, including a conceptual analysis of the recent (post-Lanford) microscopic derivation of the full Boltzmann equation for long times. We also show to which extent algorithmic randomness is stronger than necessary for the derivation of Boltzmann-like equations, in giving rise to an infinite number of other macroscopic properties. In the light of Chaitin's incompleteness theorems for algorithmic randomness, the price for this scenario is the impossibility of explicitly displaying algorithmically random microscopic trajectories.

Contents

*Department of Mathematics and Computer Science, Technical University Eindhoven, n.dekkers@tue.nl
†Radboud Center for Natural Philosophy, Radboud University, Nijmegen, landsman@math.ru.nl

1 Introduction

Despite approximately 150 years of effort and progress, starting with Boltzmann's great papers from 1872 and 1877, the problem of explaining irreversible physical processes from an underlying reversible dynamics remains. We restrict ourselves to Boltzmann's program, which in our view (and that of many others) contains the main physical ideas but requires some (very considerable!) additional mathematical sophistication and conceptual clarification.[1] In particular, our focus is on the assumptions behind the derivation of irreversible macroscopic evolution equations (like the Boltzmann equation) from reversible (Newtonian) microscopic dynamics.[2]

In §2 we review the recent derivation of the Boltzmann equation for long times by Deng, Hani, and Ma (2024), who completed the work of Lanford (1975) and others. We will emphasize:

- What they in fact prove and how this relates to Boltzmann's own ideas from 1872 and 1877;
- What remains desirable to prove in order to understand the actual physics of dilute gases;
- The origin of irreversibility in their approach.[3]

The one-particle distribution function $f(\mathbf{z},t)$ which Deng, Hani, and Ma (2024) prove to satisfy the Boltzmann equation under suitable assumptions is an *average* (or expectation value) with respect to a certain probability measure on the space of microscopic degrees of freedom of a gas, which in this case is simply a Newtonian N-particle system, where $N \to \infty$. Though their result is extremely difficult and significant, it does not explain why the (initial) N-particle configuration of some *given* gas induces a one-particle distribution function which in the limit $N \to \infty$ satisfies the Boltzmann equation.[4] That would require *pointwise* results, initially to the effect that the Boltzmann equation holds *almost surely* (or "typically", as the neo-Boltzmannians have it).[5] But even this would fail to identify *which* microscopic configurations support the Boltzmann equation. This is the main problem we wish to address in this paper via the theory of algorithmic randomness.

As a warm-up, let us review the strong law of large numbers for a fair coin toss, in which similar issues are encountered. Let $2^{\mathbb{N}}$ be the set of infinite binary sequences, equipped with the (uniform or fair) Bernoulli measure $f^{\mathbb{N}}$ (see Appendix A for details). For each $N \in \mathbb{N}$, define

$$S_N : 2^{\mathbb{N}} \to [0,1]; \qquad S_N(s) := \frac{1}{N}\sum_{n=1}^{N} s(n). \tag{1.1}$$

For later use and conceptual clarity, we note that S_N is equivalent to the empirical measure

$$\mu_N : 2^{\mathbb{N}} \to \mathrm{Prob}(\{0,1\}); \qquad \mu_N(s) := \frac{1}{N}\sum_{n=1}^{N} \delta_{s(n)}, \tag{1.2}$$

[1] See Boltzmann (1872, 1877) and the commentaries by Darrigol (2018), Uffink (2007, 2022), Brown, Myrvold, and Uffink (2009), and Badino (2011). For general historical perspectives on irreversibility in the 19th century also see Brush (1974) and van Strien (2013). For later developments see Mackey (1992), Sklar (1993), and Zuchowski (2024).

[2] We agree with Grad (1958), Uffink (2001), and Roberts (2022) that the possible connection between irreversible macroscopic equations that supervene on reversible microscopic dynamics and a possible arrow of time is doubtful.

[3] This updates the analysis of the proof for short times by Lanford (1975, 1976) by Uffink and Valente (2015).

[4] Compare interpretations of quantum mechanics that fail to explain why *individual* measurements have outcomes.

[5] What we identify as the neo-Boltzmannian (or "typicality") program is expounded by e.g. Lebowitz (1999), Lazarovici and Reichert (2015), and Bricmont (2020). The textbook by Spohn (2012) also follows this program. The core idea, which indeed may be attributed to Boltzmann (1877), is that macroscopic quantities are coarse-grained functions of microstates and follow macroscopic equations for all "typical" microstates. In the thermodynamic limit, these and other "typical" properties hold almost surely with respect to some prior probability distribution on the microstates.

where $\mathrm{Prob}(\{0,1\})$ is the set of probability measures on $\{0,1\}$, and, for $a \in \{0,1\}$, the symbol $\delta_a \in \mathrm{Prob}(\{0,1\})$ is the point measure $\delta_a(U) = 1$ if $a \in U$ and $\delta_a(U) = 0$ if $a \notin U$ ($U \subseteq \{0,1\}$). For a finite set like $\{0,1\}$ it is easier to work with probability *distributions* p defined on $\{0,1\}$ than with probability *measures* defined on subsets of $\{0,1\}$; in that case, abusing notation, δ_a in (1.2) is a function on $\{0,1\}$ namely the Kronecker delta $\delta_a(b) = \delta_{ab}$. Since $a = \delta_a(1)$, we have

$$S_N(s) = \mu_N(s)(1). \tag{1.3}$$

The most basic result is $\langle S_N \rangle_{f^{\mathbb{N}}} \to 1/2$, or, equivalently, $\langle \mu_N \rangle_{f^{\mathbb{N}}} \to f$. Since, as we shall see in §2, the empirical measure μ_N is a toy version of Boltzmann's one-particle distribution function, this kind of convergence of averages is conceptually similar to what has been proved for the Boltzmann equation. But averages contain no information about *single* outcome sequences. The law of large numbers is a step forward in this direction: in its strong form, it states that $f^{\mathbb{N}}$-almost surely,

$$S_N(s) \to 1/2; \qquad\qquad \mu_N(s) \to f. \tag{1.4}$$

That is, one may remove a subset $\mathscr{N}$ from $2^{\mathbb{N}}$ such that $f^{\mathbb{N}}(\mathscr{N}) = 0$ and (1.4) holds for all $s \in 2^{\mathbb{N}} \backslash \mathscr{N}$ (i.e., $s \notin \mathscr{N}$). But $\mathscr{N}$ is not given explicitly and one would like to know *which* $s \in 2^{\mathbb{N}}$ satisfy (1.4). This question is addressed by the theory of *algorithmic randomness*, as follows:[6]

- Eqs. (1.4) hold if $s \in 2^{\mathbb{N}}$ is $f^{\mathbb{N}}$*-random*; a property defined in Appendix A as a special case of what we call P-randomness, defined for so-called computable probability spaces (X,P).
- The set $\mathscr{R}$ of all $f^{\mathbb{N}}$-random sequences has $f^{\mathbb{N}}(\mathscr{R}) = 1$ (so one may take $\mathscr{N} = 2^{\mathbb{N}} \backslash \mathscr{R}$).

The following line of proof of these results is instructive and important, as it can be generalized.

1. The theory of *large deviations* provides bounds like the following:[7] for any $\varepsilon > 0$, one has

$$V_N(\varepsilon) := \{s \in 2^{\mathbb{N}} : |S_N(s) - 1/2| > \varepsilon\}; \qquad\qquad f^{\mathbb{N}}(V_N(\varepsilon)) \leq e^{-c(\varepsilon)N}, \tag{1.5}$$

   for some $c(\varepsilon) > 0$ and sufficiently large N.

2. The previous result yields

$$\sum_{N=1}^{\infty} f^{\mathbb{N}}(V_N(\varepsilon)) < \infty, \tag{1.6}$$

   so that the Borel–Cantelli lemma applies: $f^{\mathbb{N}}$-almost every $s \in 2^{\mathbb{N}}$ lies in only finitely many sets $V_N(\varepsilon)$, so that for all $\varepsilon > 0$ one has $|S_N(s) - 1/2| \leq \varepsilon$ for all but finitely many N. This implies the strong law of large numbers (1.4). See also footnote 45 for further details.

3. The theory of *algorithmic randomness* (see Appendix A) gives the following lemma, which trivially follows from the specific definition of P-randomness we use (see Definition A.1). It is therefore a consequence of any other equivalent definition of P-randomness one may encounter (such as the one named after P. Martin-Löf) and hence is also valid for all probability spaces and ensuing notions of P-randomness we encounter in this paper:

[6] Both bullets are due P. Martin-Löf (1966). See also Calude (2002), Theorem 6.57. Our proofs are different.

[7] First introductions to the theory of large deviations include Ellis (1995), Touchette (2009), and Landsman (2024). Ellis (1985), Dembo and Zeitouni (1998), and Rassoul-Agha and Seppäläinen (2015) are more advanced textbooks. For a coin toss (1.5) follows from Sanov's theorem or Cramér's theorem, which are equivalent in that case.

Lemma 1.1 *If (V_N) is a uniformly computable sequence of open subsets of $2^{\mathbb{N}}$ for which*

$$\sum_N f^{\mathbb{N}}(V_N) < \infty, \tag{1.7}$$

then each $f^{\mathbb{N}}$-random sequence $s \in 2^{\mathbb{N}}$ is contained only in finitely many V_N.

Taking $\varepsilon = 1/m$ with $m \in \mathbb{N}_*$ in (1.5), it follows from (1.6) that the $V_N(1/m)$ form such a sequence (whose uniform computability, as defined in Appendix A, is easy to prove). Hence by the same argument that led to the strong law of large numbers (1.4) we conclude that *each* $f^{\mathbb{N}}$-random sequence $s \in 2^{\mathbb{N}}$ satisfies this law. The "almost sure" clause has disappeared!

Adding dynamics and replacing the strong law of large numbers by some macroscopic equation (such as a toy Boltzmann equation), this example will be a model for our reasoning. Note that the property of $f^{\mathbb{N}}$-randomness is *sufficient* but not *necessary* for the strong law to hold (as obvious counterexamples show). In other words, such an assumption is stronger than necessary for the result and the question arises what this additional strength brings us. This will be addressed in §4.

The full Boltzmann equation is, at least for the moment, too complicated to be treated using algorithmic randomness. In particular, the necessity of the Boltzmann–Grad scaling limit seems to make pointwise results impossible, let alone algorithmically random ones. Nevertheless, we hope this hurdle will be overcome eventually, e.g. by modifying the dynamics. For now, we illustrate our approach in two well-known toy models of the Boltzmann equation, namely the (stochastic) Ehrenfest model and the (deterministic) Kac ring model, both of which do not necessarily involve a scaling limit. The former will be discussed in §3, whose results are almost entirely new.[8] The latter is the topic of §4, in which for completeness we rederive the pioneering results of Hiura and Sasa (2019), but also add what we consider a conceptually quite significant sharpening thereof.

Summarizing the lessons from our three cases, the ensuing picture of the irreversibility of macroscopic Boltzmann-like equations from reversible microscopic dynamics is as follows:[9]

1. In all rigorous derivations, functions $f(t,\cdot)$ supposed to solve a Boltzmann-like equation start their lives as sequences (f_N) of random variables on a space of $\sim N$ microscopic degrees of freedom (d.o.f.) equipped with a probability measure $\mathbb{P}_N$ (or on a space of infinitely many d.o.f. equipped with a single probability measure $\mathbb{P}$) under which each d.o.f. is independent at (say) $t = 0$. One then takes a suitable limit $N \to \infty$ of some version of f_N:

   (a) Starting with the work on the Boltzmann equation for dilute gases of hard spheres by Lanford (1975) for short times as completed by Gallagher, Saint-Raymond, and Texier (2014), and culminating in the recent work of Deng, Hani, and Ma (2024) for arbitrary times, this version of the random variable f_N is the *average* $\langle f_N \rangle_{\mathbb{P}_N}$, which coincides with the first (single-particle) correlation function in the BBGKY hierarchy.

   (b) For short times Cercignani, Illner, and Pulvirenti (2013), §4.6, and Villani (2013), based on Sznitman (1991), improve this to convergence *in probability* of $f_N(t,\cdot)$.

2. The choice of $\mathbb{P}_N$ under which all these authors derive the Boltzmann equation makes the positions and velocities of the N molecules comprising the gas i.i.d. random variables at some initial time $t = 0$. As explained in §2, "everything" including irreversibility follows from this assumption, combined with the convention of tracing motion *forward in time*. However, apart from its success in deriving Boltzmann-like equations, the i.i.d. assumption itself remains unjustified. This seems rarely mentioned, but is a major conceptual problem.

[8] Fun fact: these results were originally inspired by the work of A. Martin-Löf (1979), the brother of P. Martin-Löf.

[9] The (toy) Boltzmann equation is irreversible in the sense that if $f(t,\cdot)$ solves it for given initial conditions, then $f(-t,\cdot)$, combined with velocity reversal if applicable, does not. See also Appendix B for more information.

3. What we expect—and prove for the Ehrenfest and Kac ring models but defer judgement as to how or even whether it applies to the "real" Boltzmann equation—is that one may even take *pointwise* limits in the space of microscopic d.o.f. almost surely. Our approach is to refine these pointwise limits via algorithmic randomness of the microstates, and show that it is this kind of randomness that guarantees that $f(t,\cdot)$ satisfies a Boltzmann-like equation, and hence may be regarded as the physical and mathematical source of irreversibility. As just shown for the coin toss, any concept of algorithmic randomness of points in some space X is predicated on some probability measure $\mathbb{P}$ on X, which in all "Boltzmannian" cases studied here makes the degrees of freedom i.i.d. at $t = 0$. This choice of $\mathbb{P}$ is what is often seen as the assumption of (molecular) chaos,[10] but as such it doesn't tell us *which* sample configurations of the d.o.f. are chaotic. Our answer is that these are the $\mathbb{P}$-random ones.

4. However, since what is macroscopically irreversible is a *path* $t \mapsto f(t,\cdot)$, it seems more natural to talk about probability measures on microscopic paths rather than on their initial conditions. This suggests the use of Markov chains like the Ehrenfest model (or rather a slight variation thereof, which is better behaved), in which the above scenario can be realized. Here the microscopic state space is $2^{\mathbb{N}}$, whose elements can be coarse-grained into toy versions of Boltzmann's distribution function f. We will show that if a (sample) path in $2^{\mathbb{N}}$ is $\mathbb{P}$-random with respect to the Markovian probability measure $\mathbb{P}$ on the space of such paths, then the ensuing distribution function satisfies an irreversible toy Boltzmann equation. A similar result holds for the Kac ring model (Hiura and Sasa, 2019), now formulated either in terms of paths or, equivalently, initial conditions (since the model is deterministic).

5. In conclusion, what leads to irreversibility is the time asymmetry of the condition of $\mathbb{P}$-randomness relative to a suitable probability measure $\mathbb{P}$ on the space of sample paths for $t \geq 0$ that makes the particles i.i.d. at $t = 0$: the point is that $\mathbb{P}$-random paths satisfy the (toy) Boltzmann equation in question whereas their time reversals do not.

How explicit is the characterization of microstates as being "random"? There are two sides to this:

- On the one hand, the property of P-randomness of points in a computable probability space (X,B,P) is clearly and distinctly defined (see Appendix A), and some individual P-random objects can similarly be defined by some logical description (in standard set theory).[11]

- On the other hand, *Chaitin's second incompleteness theorem* states that if $s \in 2^{\mathbb{N}}$ is $f^{\mathbb{N}}$-random, then the value of all but finitely many digits is undecidable (à la Gödel) within any sufficiently comprehensive mathematical theory like ZF or even ZFC.[12]

Thus although "most" infinite binary sequences are random with respect to the probability measure $f^{\mathbb{N}}$, none of these can be explicitly *known* or *shown*; for doing so would blast their randomness.[13] This elusiveness of algorithmically random objects (which, based on undecidability, goes far beyond their obvious non-computability) seems a deep and unavoidable aspect of their existence.

[10] See for example Villani (2013). We will get back to the difference with Boltzmann's version in the next section.

[11] The most famous example is Chaitin's Ω, defined as the halting probability $\Omega = \sum_{p|p \text{ halts}} 2^{-|p|}$, where the sum is over all self-delimiting programs p that can run on a given Turing machine, cf. Downey and Hirschfeldt (2010), §3.13.

[12] See Calude (2002), Theorem 8.7. This is stated for Chaitin's Ω, but as pointed out by Landsman (2021) the proof holds for any $f^{\mathbb{N}}$-random sequence.

[13] This is not an artifact of the idealization inherent in the use of *infinite* binary sequences: for finite binary strings, *Chaitin's (first) incompleteness theorem* states that although countably many strings *are* random, this can be *proved* only for finitely many of these (Raatikainen, 1998).

2 Irreversibility of the original Boltzmann equation

We start with a quick review of Boltzmann's great papers from 1872 (in which he introduced the transport equation named after him as well as his negentropic H-function) and 1877 (in which he began his discretization and counting techniques.[14] Although the second looks quite different from the first and indeed for Boltzmann himself was a fresh start,[15] from a modern mathematical point of view the papers are closely related: both are based on a coarse-graining technique later recognized as a passage to what we now call an *empirical measure*,[16] see also (1.2).

- Denoting the space of probability measures on $\mathbb{R}^{2d}$ by $\mathrm{Prob}(\mathbb{R}^{2d})$, Boltzmann (1872) associated a one-particle distribution function $f_N \in \mathrm{Prob}(\mathbb{R}^{2d})$ to a microscopic N-particle configurations $(\mathbf{z}_1, \ldots \mathbf{z}_N) \in \mathbb{R}^{2dN}$, where $\mathbf{z} = (\mathbf{r}, \mathbf{v})$,[17] We (re)write his construction as

$$f_N(\mathbf{z}_1, \ldots, \mathbf{z}_N) := \frac{1}{N} \sum_{n=1}^{N} \delta_{\mathbf{z}_n}. \tag{2.1}$$

  If we equip $\mathbb{R}^{2dN}$ with a probability measure $\mathbb{P}_N$, as we will, then, again from a modern point of view, f_N is a random variable *on* $(\mathbb{R}^{2dN}, \mathbb{P}_N)$ taking values *in* $\mathrm{Prob}(\mathbb{R}^{2d})$. In physics one prefers probability *distributions* to *measures*: one would like to write some $\mu \in \mathrm{Prob}(\mathbb{R}^{2d})$ as $d\mu(\mathbf{z}) = \rho(\mathbf{z})d\mathbf{z}$, where ρ is a probability *distribution* on $\mathbb{R}^{2d}$. This formally works for $\mu = \delta_{\mathbf{z}_n}$ if we take $\rho(\mathbf{z}) = \delta(\mathbf{z} - \mathbf{z}_n)$, the Dirac delta-function, so that (2.1) becomes

$$df_N(\mathbf{z}_1, \ldots, \mathbf{z}_N)(\mathbf{z}) = \frac{1}{N} \sum_{n=1}^{N} \delta(\mathbf{z} - \mathbf{z}_n) d\mathbf{z}. \tag{2.2}$$

  In view of this, as long as we keep the official meaning (2.2) in mind, we may also write

$$f_N(\mathbf{z}; \mathbf{z}_1, \ldots, \mathbf{z}_N) = \frac{1}{N} \sum_{n=1}^{N} \delta(\mathbf{z} - \mathbf{z}_n). \tag{2.3}$$

  The point made after (2.1) may then be rephrased as follows: f_N has two very different kinds of arguments, namely $(\mathbf{z}_1, \ldots, \mathbf{z}_N)$, making f_N a random variable *defined on* $(\mathbb{R}^{2dN}, \mathbb{P}_N)$, and $\mathbf{z}$, making this random variable *take values in* the space of probability distributions on $\mathbb{R}^{2d}$.

- Boltzmann (1877) uses the same idea: in modern notation, let A^N be the set of N-particle configurations $\sigma : \{1, \ldots, N\} \to A$, where A is a finite set (whose elements Boltzmann took to be discrete energy levels),[18] and N is once again the number of particles. One now has a discrete version p_N of the empirical measure f_N in (2.1), which coarse-grains $\sigma \in A^N$ into

$$p_N(\sigma) := \frac{1}{N} \sum_{n=1}^{N} \delta_{\sigma(n)}. \tag{2.4}$$

[14] Boltzmann (1877) also initiated the study of their asymptotics, which was one of the sources of the modern probabilistic theory of large deviations. See references in footnote 7, as well Ellis (1995) in connection with Boltzmann.

[15] See the references in footnote 1, as well as the summary below.

[16] Our notation in both (2.1) and (2.4) is as follows. For any measure space X (equipped with a σ-algebra Σ we suppress), let $\mathrm{Prob}(X)$ be the space of probability measures on X. Any $x \in X$ defines a *point measure* $\delta_x \in \mathrm{Prob}(X)$ via $\delta_x(U) = 1$ if $x \in U$ and $\delta_x(U) = 0$ if $x \notin U$ (for any $U \in \Sigma$). In (2.1) we have $X = \mathbb{R}^{2d}$, and in (2.4) we have $X = A^N$.

[17] Boltzmann (1872) initially introduced f_N for homogeneous gases in which the spatial argument $\mathbf{r}$ is absent and the velocity $\mathbf{v}$ is replaced by the kinetic energy x, so that 'the number of molecules in unit volume whose kinetic energy at time t lies between x and $x + dx$ I will call $f(x,t)dx$' (p. 268 of the English translation). Later in the paper he introduces what we call $f_N(\mathbf{z})$ for inhomogeneous gases, with a similar interpretation.

[18] Putting his energy unit $\varepsilon = 1$ for simplicity, Boltzmann (1877) has $A = \{0, 1, \ldots, p\}$. Writing N for his d, his 'complexions' $(k_1, \ldots, k_N)$ therefore correspond to our microstates $\sigma \in A^N$. His 'distribution of states' (*Zustandsverteilung*) $(N_0, \ldots, N_p)$ corresponds to our Np_N, so that $N_k = p_N(k) = \sum_{n=1}^{N} \delta_{\sigma(n)k}$ is the number of molecules with energy k.

Thus $p_N(\sigma)$ is a probability distribution on A, and once again, if we equip A^N with some probability measure $\mathbb{P}_N$, then p_N is a random variable *on* $(A^N, \mathbb{P}_N)$ taking values *in* $\mathrm{Prob}(A)$. Since A is finite, the move from measures to distributions is trivial:[19] the analogue of (2.3) is now simply given by $p_N(a;\sigma) = \frac{1}{N}\sum_{n=1}^{N} \delta_{\sigma(n)a}$, where δ_{ba} is the usual Kronecker delta.

The above text was written with hindsight. Historically, apart from the jump from continuous to discrete variables there are two major differences between Boltzmann's 1872 and 1877 papers:

1. Although f_N is a probability distribution on the one-particle phase space $\mathbb{R}^{2d}$ and Boltzmann (1872) stressed this probabilistic interpretation right from the start, he did not see f_N as a random variable in its dependence on the N-particle coordinates $(\mathbf{z}_1, \ldots, \mathbf{z}_N)$, since he did not assume a probability distribution on the latter.[20] In contrast, Boltzmann (1877) explicitly assumes a uniform probability distribution on microstates, making p_N a random variable. This was the real and deep introduction of probability into the problem of irreversibility.

   Furthermore, the distinction between micro- and macrostates, which Uffink (2007), p. 977, even calls 'the decisive new element that allowed Boltzmann a complete reinterpretation of the notion and role of probability' is now explicit (whereas it was only implicit in 1872).

2. Nonetheless, in 1872 Boltzmann was closer (than in 1877) to what one really wants, namely to describe the evolution of an individual gas from some initial state, as opposed to describing generic or typical behaviour of an ensemble (helpful as that may be, too). Boltzmann's goal of introducing f_N in 1872 was to study its time evolution (with an eye on entropy and irreversibility), which aspect, curiously, was lacking for p_N in 1877. This time evolution was supposed to be governed by what we now call the *Boltzmann equation*

$$\partial_t f(t,\mathbf{z}) + \mathbf{v}\cdot\partial_{\mathbf{r}} f(t,\mathbf{z}) = C\left(f(t,\mathbf{z})\right), \tag{2.5}$$

   whose right-hand side is a quadratic expression in f incorporating two-body-collisions. As Boltzmann showed through his H-theorem, eq. (2.5) implies that f evolves irreversibly.

As we see it, the technique of algorithmic randomness gives the best of both worlds, combining Boltzmann's probabilistic philosophy of 1877 with the individual spirit of 1872. See below.

Boltzmann derives (2.5) by considering how the one-particle distribution function (2.1) changes under the effect of microscopic dynamics $\mathbf{z}_n \mapsto \mathbf{z}_n(t)$, which he took to be hard sphere collisions. Thus (without such formulae) he considered the time-dependent one-particle distribution function

$$f_N(t,\mathbf{z}_1,\ldots,\mathbf{z}_N) := \frac{1}{N}\sum_{n=1}^{N} \delta_{\mathbf{z}_n(t)} \qquad \text{or} \qquad f_N(t,\mathbf{z};\mathbf{z}_1,\ldots,\mathbf{z}_N) := \frac{1}{N}\sum_{n=1}^{N} \delta(\mathbf{z}-\mathbf{z}_n(t)), \tag{2.6}$$

[19] If X is finite (and Σ is just the power set of X), then any probability *measure* $\mathbb{P}$ on X is equivalently given by a probability *distribution* (or *mass*) P, where $\mathbb{P}(U) = \sum_{x\in U} P(x)$ and conversely $P(x) = \mathbb{P}(\{x\})$. In that case, the probability distribution corresponding to the point measure δ_x is the Kronecker delta $\delta_x(y) = \delta_{xy}$.

[20] Uffink (2007), p. 967, makes the following comment: 'It is clear that the conception of probability expounded here is thoroughly frequentist and that he takes 'the laws of probability' as empirical statements. Furthermore, probabilities can be fully expressed in mechanical terms: the probability distribution f is nothing but the relative number of particles whose molecular states lie within certain limits. (...) Indeed, it seems to me that Boltzmann's emphasis on the crucial role of probability in this paper is only intended to convey that probability theory provides a particularly useful and appropriate language for discussing mechanical problems in gas theory. There is no indication in this paper yet that probability theory could play a role by furnishing assumptions of a non-mechanical nature, i.e., independent of the equations of motion.' Likewise, Darrigol (2018), p. xviii, takes Boltzmann's 'probabilistic turn' to start only in 1876.

and tried to show that it has autonomous time evolution $f_N \mapsto f_N(t)$ (i.e., not depending on the underlying microstate). However, Boltzmann's reasoning, while groundbreaking, was neither conceptually nor mathematically rigorous.[21] In particular, his definition of the one-particle distribution function was only qualitative (cf. footnote 17) and he apparently failed to grasp that (2.5) can only hold exactly in what is nowadays called the Boltzmann–Grad limit (see below).

Based on earlier work by Grad (1958), Lanford (1975, 1976), Cercignani, Illner, and Pulvirenti (2013), Gallagher, Saint-Raymond, and Texier (2014), and others,[22] the program of rigorously deriving the Boltzmann equation from hard sphere dynamics was recently (about 150 years after Boltzmann!) completed by Deng, Hani, and Ma (2024), who extended earlier short-time derivations of the Boltzmann equation for the average value of f_N as defined in (2.6) with respect to i.i.d. initial conditions to arbitrary times (for which solutions exist). Their overall strategy remained Lanford's, who combined Boltzmann's ideas from 1872 and 1877 by interpreting his distribution function f_N from 1872 as a random variable in its dependence on the microscopic variables $\{\mathbf{z}_n\}$; as we noted, this was how Boltzmann saw his energy distribution function in 1877 (but not the one in 1872) in its dependence on a microstate.

For understanding our "randomness first" program, it is important to explain what has been achieved here mathematically,[23] and hence to recognize what remains unsatisfactory physically.[24]

1. One assumes Newton's laws (for finite N) with hard sphere dynamics, in which each molecule moves freely until it encounters another molecule, upon which they scatter elastically. Let ε be the diameter of the molecules, with non-overlapping domain $\mathscr{D}_N \subset \mathbb{R}^{2dN}$ on which $|\mathbf{r}_n - \mathbf{r}_m| \geq \varepsilon$ for all $1 \leq n \neq m$. There is a subset $\mathscr{Z}_N \subset \mathscr{D}_N$ of Lebesgue measure zero such that on $\mathscr{D}_N \backslash \mathscr{Z}_N$ the hard sphere dynamics exists for all times; in particular, collisions of more than two particles do not occur on $\mathscr{D}_N \backslash \mathscr{Z}_N$. To keep the notation simple we will always assume that $(\mathbf{z}_1, \ldots, \mathbf{z}_N) \in \mathscr{D}_N \backslash \mathscr{Z}_N$, which as $N \to \infty$ holds with probability $\to 1$.

2. At $t = 0$ the coordinates $\mathbf{z}_n$ of the N molecules are assumed independent and identically distributed (i.i.d.) via some one-particle distribution function $p(\mathbf{z})$, assumed Lipschitz continuous in both variables with Gaussian decay in $\mathbf{v}$, and normalized such that

$$\int_{\mathbb{R}^{2d}} d\mathbf{z}\, p(\mathbf{z}) = 1, \tag{2.7}$$

so that the microstates $(\mathbf{z}_1, \ldots, \mathbf{z}_N)$ become random variables with initial distribution

$$p^{(N)}(\mathbf{z}_1, \ldots, \mathbf{z}_N) = C_N \prod_{n=1}^{N} p(\mathbf{z}_n) 1_{\mathscr{D}_N \backslash \mathscr{Z}_N}(\mathbf{z}_1, \ldots, \mathbf{z}_N). \tag{2.8}$$

Here $C_N > 1$ is a normalization constant, which despite (2.7) is necessary because $\mathbb{R}^{2dN} \backslash \mathscr{D}_N$ has positive Lebesgue measure. However, it follows from point 3 below that the latter

[21] The need for mathematical rigour in deriving macroscopic laws from microscopic dynamics was formulated by Hilbert, who in the elaboration of his Sixth Problem stated that 'Boltzmann's work on the principles of mechanics suggests the problem of developing mathematically the limiting processes, there merely indicated, which lead from the atomistic view to the laws of motion of continua.' (Hilbert, 1900/2000, p. 419). In the physics literature Ehrenfest and Ehrenfest (1911) is an early reference to this end.

[22] See especially the work summarized by Bodineau et al. (2023).

[23] See also Bodineau et al. (2023) for a summary. The book-length survey by Villani (2002) remains worthwhile.

[24] For simplicity we ignore a complication: instead of taking the particle number N fixed and letting $N \to \infty$, Lanford (1975) and Deng, Hani, and Ma (2024) use the grand canonical ensemble and take N to be random variable whose expectation value is sent to infinity in the Boltzmann–Grad limit. This apparent complication in fact simplifies their proofs but could have been avoided (as they state themselves), and seems irrelevant for our purposes.

measure vanishes as $N \to \infty$, implying $C_N \to 1$, so to simplify our expressions we take $C_N = 1$ in what follows and also, as already explained, omit the restriction $(\mathbf{z}_1, \ldots, \mathbf{z}_N) \in \mathscr{D}_N \backslash \mathscr{Z}_N$. We denote the probability measure on $\mathbb{R}^{2dN}$ induced by the initial distribution (2.8) by $\mathbb{P}_N$.

3. The function $f(t, \mathbf{z})$ that will eventually satisfy the Boltzmann equation is a limit $f_N \to f$, where f_N is given by (2.6), whose initial value is the function p that occurs in (2.8), i.e.,

$$f(0, \mathbf{z}) = p(\mathbf{z}). \tag{2.9}$$

But what is meant by "$f = \lim_{N\to\infty} f_N$"? First, $\lim_{N\to\infty}$ is in fact a double limit called the *Boltzmann–Grad limit* (physically this is the *dilute gas limit*), in which one simultaneously takes $N \to \infty$ and $\varepsilon \to 0$ at constant $\varepsilon^{d-1}N$, where d is the spatial dimension.

Second, since $f_N(t, \mathbf{z}; \mathbf{z}_1, \ldots, \mathbf{z}_N)$ as defined in (2.6) is a function of both $\mathbf{z}$ and the random variable $(\mathbf{z}_1, \ldots, \mathbf{z}_N)$, assumed distributed by (2.8), whereas $f(t, \mathbf{z})$ is just a function of $\mathbf{z}$, one has to get rid of the dependence of f_N on $(\mathbf{z}_1, \ldots, \mathbf{z}_N)$. The most straightforward (but still highly technical!) approach to this is to take averages over $(\mathbf{z}_1, \ldots, \mathbf{z}_N)$ with respect to the probability measure $\mathbb{P}_N$ on $\mathbb{R}^{2dN}$, but we will argue for more powerful possibilities.

Furthermore, one has to get rid of the distributional character of $f_N(t, \mathbf{z}; \mathbf{z}_1, \ldots, \mathbf{z}_N)$ in $\mathbf{z}$ by smearing. Using (2.6), which contains (2.1) as a special case for $t = 0$, the pairing of f_N, seen as a distribution in $\mathbf{z}$ keeping $(\mathbf{z}_1, \ldots, \mathbf{z}_N)$ fixed, with a test function $h \in C_b(\mathbb{R}^{2d})$, is:[25]

$$\int_{\mathbb{R}^{2d}} f_N(t, \mathbf{z}; \mathbf{z}_1, \ldots, \mathbf{z}_N) h(\mathbf{z}) d\mathbf{z} \equiv \langle f_N(t, \cdot; \mathbf{z}_1, \ldots, \mathbf{z}_N), h \rangle = N^{-1} \sum_{n=1}^{N} h(\mathbf{z}_n(t)), \tag{2.10}$$

where the middle notation would be the more official one, but in what follows we prefer the more informal left-hand side (which avoids brackets). The limit $f_N \to f$ then reads

$$\left\langle \int_{\mathbb{R}^{2d}} f_N(t, \mathbf{z}; \cdot) h(\mathbf{z}) d\mathbf{z} \right\rangle_{\mathbb{P}_N} \to \int_{\mathbb{R}^{2d}} f(t, \mathbf{z}) h(\mathbf{z}) d\mathbf{z}, \tag{2.11}$$

for each $h \in C_b(\mathbb{R}^{2d})$. Weak convergence by smearing with h understood, eq. (2.11) can be written as

$$f(t, \mathbf{z}) = \lim_{N\to\infty} \langle f_N(t, \mathbf{z}) \rangle_{\mathbb{P}_N}. \tag{2.12}$$

4. The key result of Deng, Hani, and Ma (2024), then,[26] is that for all times t for which a solution of the Boltzmann equation (2.5) with initial value $p(\mathbf{z})$ exists, the limit (2.12) exists in the sense (2.11), upon which the limit function $f(t, \mathbf{z})$ satisfies the Boltzmann equation (2.5) with initial value $p(\mathbf{z})$, see (2.9), which had already defined $\mathbb{P}_N$ via (2.8).

They also prove *propagation of chaos*, in that the assumption (2.8) of exact factorization of $p^{(N)}$ at $t = 0$ implies that at all (admissible) later times it remains valid approximately: for any $k = 1, \ldots, N$ and given N, we define k-particle functions $f_k^{(N)}(t, \mathbf{z}_1, \ldots, \mathbf{z}_k)$ distributionally via the pairing

$$\int_{\mathbb{R}^{2dk}} f_k^{(N)}(t, \mathbf{z}_1, \ldots, \mathbf{z}_k) h(\mathbf{z}_1, \ldots, \mathbf{z}_k) d\mathbf{z}_1 \cdots d\mathbf{z}_k := \frac{1}{N^k} \left\langle \sum_{n_1, \ldots, n_k = 1}^{N} h(\mathbf{z}_{n_1}(t), \ldots, \mathbf{z}_{n_k}(t)) \right\rangle_{\mathbb{P}_N}, \tag{2.13}$$

[25] Since in our case f_N is a distribution of order zero, i.e., a measure, one does not need smooth test functions.

[26] What follows is not the way Deng, Hani, and Ma (2024) state their result. First, they do not invoke our (2.6) but work with the lowest distribution function in the BBGKY hierarchy, which they denote by $f_1(t, \mathbf{z})$ and which Uffink (2007) writes as $\rho_t^{(1)}(x)$. As Uffink shows in the argument leading to his eq. (164) on page 1037, his $\rho_t^{(1)}$ and hence their $f_1(t)$ coincides with our $\langle f_N(\mathbf{z}, t) \rangle_{\mathbb{P}_N}$. See also Uffink and Valente (2015). This implies what we call their key result above. Second, they prove convergence $f_N \to f$ in $L^1(\mathbb{R}^{2d})$, which implies weak convergence in Prob$(\mathbb{R}^{2d})$.

where the sum is only over k-tuples $n_1,\dots,n_k$ with distinct entries. For example, eq. (2.8) gives

$$f_k^{(N)}(0,\mathbf{z}_1,\dots,\mathbf{z}_k) = p(\mathbf{z}_1)\cdots p(\mathbf{z}_k). \tag{2.14}$$

Propagation of chaos then means that these k-particle functions factorize at all times, i.e.,[27]

$$\lim_{N\to\infty} \left\| f_k^{(N)}(t,\mathbf{z}_1,\dots,\mathbf{z}_k) - \prod_{n=1}^{k} f(t,\mathbf{z}_n) \right\|_{L^1(\mathbb{R}^{2dk})} = 0. \tag{2.15}$$

We now discuss the conceptual structure of the derivation. In a physically very careful (but mathematically incomplete) discussion, Grad (1958), §11, already made a crucial point:

> It is possible to specify the exceptional set on which $f_2^{(N)}(\mathbf{z}_1,\mathbf{z}_2,t)$ does *not* converge to $f(\mathbf{z}_1,t)f(\mathbf{z}_2,t)$ [as $N\to\infty$] rather precisely and verify that this set can in fact be ignored.[28] This exceptional set, on which chaos is not satisfied, has its own special interest. It consists, roughly speaking, of pairs of points $(\mathbf{z}_1,\mathbf{z}_2)$ that have recently collided. It is the existence of this exceptional set which provides the connection between the reversibility of LIOUVILLE'S equation and the irreversibility of BOLTZMANN'S equation. (Grad, 1958, p. 224)

And here is the main conceptual comment by Gallagher, Saint-Raymond, and Texier (2014):

> [The irreversibility of the Boltzmann equation] will actually appear in the limiting process as an arbitrary choice of the time direction (encoded in the distinction between pre-collisional and post-collisional particles), and more precisely as an arbitrary choice of the initial time, which is the only time for which one has a complete information on the correlations. The point is that the joint probability of having particles of velocity (v_1,v_2) [...] before the collision is assumed to be equal to $f(t,x,v_1)\cdot f(t,x,v_2)$, [...] meaning that particles should be independent before collision. (Gallagher, Saint-Raymond, and Texier, 2014, Remark 2.2.1)

The physical picture behind the mathematical derivation of the Boltzmann equation may therefore be summarized as follows. The one-particle distribution function $f_N(t,\cdot)$ develops and coarse-grains these initial microscopic conditions forward in time. The key point, which Boltzmann (1877) already made but did not emphasize or even played down,[29] is that two-body collisions contribute to a distribution function $f_N(t,\cdot)$ that (approximately) solves the Boltzmann equation (and exactly does so in the limit $N\to\infty$) *provided the molecules in question are uncorrelated before they collide* (after which they *are* correlated). This assumption is Boltzmann's *Stosszahlansatz*.

The initial distribution $\mathbb{P}_N$ in the modern derivation trivially enforces this assumption at $t=0$, since it erases past correlations by the very choice of the (i.i.d.) initial conditions. The idea, then, is that what is impossible by assumption at $t=0$, namely the existence of correlations between particles about to collide, remains unlikely at later times, indeed so unlikely in the Boltzmann–Grad limit that the possibility can be ignored. This is achieved by what is called 'propagation of chaos': although after a first round of collisions correlations have arisen, the probability of recollision converges to zero,[30] and hence collisions that do take place (with nonzero probability) are still collisions between uncorrelated particles, so that Boltzmann equation continues to hold.

[27] Eq. (2.15) follows from eq. (1.18) in Deng, Hani, and Ma (2024).

[28] What Grad means is that the measure of these sets converges to 0 as $N\to\infty$. His singular 'set' may be confusing.

[29] See Grad (1958), Uffink (2007, 2022), Brown, Myrvold, and Uffink (2009), Uffink and Valente (2015), and Darrigol (2018) for the relevant history and philosophy, and Gallagher, Saint-Raymond, and Texier (2014) and related work summarized by Bodineau et al. (2023) for full mathematical rigour concerning this point.

[30] This is very hard to prove rigorously, but it seems natural in the said limit, which makes the gas infinitely dilute.

Boltzmann (1872), on the other hand, had to (and did) assume the *Stosszahlansatz* for all times. But, apart from technicalities, this is not the only difference between Boltzmann and the modern setting: Boltzmann had the configuration of an individual gas in mind and so his distribution functions were defined by *counting* the number of particles with specific properties (regarding their positions and velocities). His idea of 'correlation' was therefore the lack of factorization of the *actual* number of molecules in a certain two-body configuration as a product two numbers defined by the corresponding one-body configurations.[31] Again, then, his reasoning was not probabilistic.

Compared with Boltzmann (1872) the result of Deng, Hani, and Ma (2024), being a statement about averages (see footnote 26), is undoubtedly highly significant, but it does not explain which *individual* microstates give rise to macroscopic evolutions (approximately) satisfying the Boltzmann equation.[32] As illustrated in §1 for the fair coin toss, this deficiency can be partially solved by appealing to algorithmic randomness. For the Boltzmann equation, we would ideally have that

$$\lim_{N\to\infty} f_N(t,\mathbf{z};\mathbf{z}_1,\ldots,\mathbf{z}_N) = f(t,\mathbf{z}) \tag{2.16}$$

exists and satisfies the Boltzmann equation in some sense for all algorithmically random configurations $(\mathbf{z}_1,\mathbf{z}_2,\ldots)$ of infinitely many particles, with respect to some probability measure $\mathbb{P}$ on these microscopic d.o.f. The limiting distribution function $f(t,\mathbf{z})$ can only solve the Boltzmann equation in the Boltzmann–Grad limit, in which the diameter ε varies with N, vanishing in the limit $N\to\infty$. Consequently, the probability measure $\mathbb{P}$ should necessarily also vary with N, which means we cannot have a fixed notion of randomness.[33] Another way to view the problem is that there exists no hard sphere dynamics with $N=\infty$, since the Boltzmann–Grad limit would demand $\varepsilon = 0$, upon which hard sphere collisions do not make sense anymore. This problem may be circumvented by switching from hard spheres and dynamics based on collisions to point particles interaction via pair potentials.[34] If such a dynamics can be directly defined for $N=\infty$, then one could simply fix a probability measure $\mathbb{P}$ and study the evolution of the one-particle distribution function induced by a $\mathbb{P}$-random microstate $(\mathbf{z}_1,\mathbf{z}_2,\ldots)$. In that case, we hope that algorithmic randomness will be relevant to the derivation of the full Boltzmann equation, along the lines of the two toy models studied in the remainder of this paper.[35]

Furthermore, in the modern approach the decisive i.i.d. (or indeed "chaotic") initial distribution (2.8) remains to be clarified. So far, its success is *a posteriori*, in being a sufficient assumption for the derivation of the Boltzmann equation. *A priori*, it cannot be justified by probabilistic considerations, since it defines what we mean by probability (both in the usual and in the algorithmic randomness approach). We leave this as an open problem, bewildered this has been ignored so far.

We finally comment on time reversal. Running the movie backward in time would show the correlated outgoing particles as correlated incoming particles giving rise to uncorrelated outgoing particles, which violates the (probabilistic) *Stosszahlansatz* and hence the Boltzmann equation (albeit in favour of the *anti-Boltzmann equation*, in which the sign of the collision term changes).[36]

[31] See also Brown, Myrvold, and Uffink (2009) and Uffink and Valente (2015).

[32] Seen as a statement about the BBGKY hierarchy, Cercignani, Illner, and Pulvirenti (2013), §4.6, show how one might improve this to a statement about convergence in probability, i.e., a weak law of large numbers. But this still does not give information on individual microstates.

[33] To make sense of (2.16), a random $(\mathbf{z}_1,\mathbf{z}_2,\ldots)$ should have the property that the spheres with centres $(\mathbf{r}_1,\ldots,\mathbf{r}_N)$ and diameter ε should be non-overlapping. Clearly, this means that the probability measure should change as $\varepsilon\to 0$.

[34] See Gallagher, Saint-Raymond, and Texier (2014), who carry out the short-time derivation for that case also.

[35] Another approach would be to keep hard sphere collisions and learn how to apply algorithmic randomness in the presence of a scaling limit, which could first be explored in a toy model.

[36] See Uffink and Valente (2015) and Ardourel (2017) for a discussion about the source of irreversibility in Lanford's proof. These authors–indeed practically all authors on this topic–apparently disagree, or at best say different things. Lanford's proof indeed lacks transparency and by current mathematical standards even completeness and rigour.

This process is dynamically possible, but it is blocked by the choice of i.i.d. initial conditions for the Boltzmann equation. Hence we also side with Grad on the arrow of time. For his previous quotation continues as follows:[37]

> Suppose we suddenly "turn off" LIOUVILLE'S equation, reverse all velocities, and proceed again with the solution of LIOUVILLE'S equation. We retrace our steps and find, for example, that [minus the entropy] H_γ is increasing instead of decreasing. The reason is that, even though chaos is satisfied almost everywhere, the small exceptional set on which it is not satisfied has now become exactly the set on which the value of $f_2^{(N)}$ determines the course of f. (...)
>
> As a final description of this situation we remark that *there is no time's arrow*. If we prepare a state we will observe that H_γ, decreases; the same will be true if we follow the prepared state backwards in time. Or, if we observe an isolated system until it achieves a certain preassigned, unlikely state, we will observe that H_γ immediately begins to decrease; or if on observing this state, we reverse all velocities, we will again observe that H_γ decreases. No matter which direction in time we look, H_γ can only decrease. We can obtain an increase of H_γ if we reverse the time only if we have specifically selected the given state not on its own merits but as the successor to another unlikely state. (Grad, 1958, pp. 224, 228–229).

3 Irreversibility in the Ehrenfest model

Predating their famous conceptual analysis of statistical mechanics (Ehrenfest and Ehrenfest, 1911), the oldest toy model for the Boltzmann equation is what we now call the *Ehrenfest model* (Ehrenfest and Ehrenfest, 1907), which is subject of a large literature.[38] We distinguish two versions, in each of which $N \in \mathbb{N}$ is the total number of balls that are distributed over two different urns, labeled U_0 and U_1. Both are time-homogeneous Markov chains on finite state spaces.

- *Microscopic version*: this has state space $A_N = 2^N$, so that a microstate $a \in A_N$ is a function

$$a : \{1,\ldots,N\} \to \{0,1\}, \tag{3.1}$$

  which specifies the location of each ball: $a_n \in \{0,1\}$ for each $n = 1,\ldots,N$ says that ball number n is in urn a_n. The transition probabilities are given by

$$P_{ab} = 1/N \text{ iff } b_n = a_n \text{ for all } n \text{ except one;} \tag{3.2}$$

$$P_{ab} = 0 \text{ otherwise.} \tag{3.3}$$

  At each time step a single ball is randomly and uniformly chosen from the set of N balls, and is then moved to the other urn. This immediately gives (3.2) and (3.3). The stationary probability measure $\pi \in \text{Prob}(A_N)$ is unique (by irreducibility) and is given by the flat distribution

$$\pi_a = 2^{-N} \text{ for each } a \in A_N. \tag{3.4}$$

[37] *Mutatis mutandis*, the discussion by Roberts (2022), Chapter 5, partly based on Price (1996), is relevant here, e.g.:

'In our universe, an oscillating charge and a coordinated ring of oscillating charges are equally likely: if the first occurs, the second occurs as well, for example when the waves produced by the single oscillating charge are absorbed into the environment. More generally, the time reversal invariance of electromagnetism guarantees that if one solution is possible, then the time reversed solution is possible as well. One might like to add that an oscillating charge is more likely to occur 'earlier than' a coordinated ring of charges in time. But, in which temporal direction are we looking when we say this? If we respond, "towards (what we normally call) the future", then we have just assumed what we were trying to prove, that there is a preferred direction of time for formulating such statements.' (pp. 121–122).

The context is the temporal asymmetry between the usual outgoing shells of electromagnetic waves produced by an oscillating charge, and incoming waves produced by a coordinated ring of oscillating charges, which are never observed.

[38] See e.g., quite differently, Johnson and Kotz (1977), Baldovin, Caprini, and Vulpiani (2019), and Bricmont (2022).

- *Macroscopic version*: here the state space is $M_N = \{0, 1, \ldots, N\}$, where $m \in M_N$ ($0 \leq m \leq N$) is the number of balls in U_1 (so that $N - m$ is the number of balls in U_0). The transition probabilities come from the following picture, derived from the microscopic model: if U_1 contains m balls the probability that the jumping ball is in U_1 is m/N, in which case the number of balls in U_1 after the jump equals $m-1$.The probability that the jumping ball is in U_0 equals $1-(m/N)$; after its jump to U_1 the number of balls in U_1 is $m+1$. This gives

$$P_{m,m-1} := \frac{m}{N} \ (m \geq 1); \qquad P_{m,m+1} := \frac{N-m}{N} \ (m < N); \qquad P_{mn} = 0 \ (n \neq m \pm 1). \tag{3.5}$$

The stationary probability measure $\pi \in \mathrm{Prob}(M_N)$ is again unique, given by the binomial distribution

$$\pi_m = 2^{-N} \begin{pmatrix} N \\ m \end{pmatrix}. \tag{3.6}$$

Though both versions may be defined separately, the macroscopic version can also be derived from the microscopic version. The connection is given by the coarse-graining map

$$m : A_N \to M_N; \qquad\qquad m(a) = \sum_{n=1}^{N} a_n. \tag{3.7}$$

Denoting the microscopic Markov chain by $(X(t))_{t\in\mathbb{N}}$, governed by the transition probabilities (3.2) - (3.3), and defining

$$Y(t) = \sum_{n=1}^{N} X_n(t), \tag{3.8}$$

the image of the microscopic process under (3.7), the process $(Y(t))_{t\in\mathbb{N}}$ is a Markov chain with transition probabilities (3.5)–note that an image of a Markov chain is not in general again Markov.

The one-particle distribution "function" of the Ehrenfest model, analogous to (2.1), is simply the coarse-graining map (3.7) divided by N, giving the proportion of balls in U_1. This is just a single number rather than a function, because a probability distribution on $\{0,1\}$ is determined by its value at 1. Adding time evolution, cf. (2.6), we write

$$f_N(t) = \frac{1}{N} \sum_{n=1}^{N} X_n(\lfloor Nt \rfloor), \tag{3.9}$$

for $t \in [0, \infty)$, where, following A. Martin-Löf (1979), we have rescaled the time variable, which is necessary to obtain a sensible limit as $N \to \infty$.[39] Note that the random variable (3.9) depends on the *path* of microstates (though we suppress this dependence in the notation for the time being, as is common in probability theory), unlike (2.6) which depends on a *single* microstate. As before, we would like an autonomous evolution equation for $f_N(t)$, and the most straightforward way to achieve this is to consider the average

$$\overline{f}_N(t) := \langle f_N(t) \rangle_{\mathbb{P}_N} = \sum_{m=0}^{N} \frac{m}{N} \mathbb{P}_N \left(f_N(t) = \frac{m}{N} \right). \tag{3.10}$$

Here $\mathbb{P}_N$ is the law of the microscopic process $(X(t))_{t\in\mathbb{N}}$ with N particles (not scaled in time), which is determined by the transition probabilities (3.2) - (3.3) together with an initial distribution

[39] In deriving Brownian motion as a limit of random walks one rescales $t \mapsto \lfloor Nt \rfloor$ as in (3.9), but instead of $1/N$ in (3.9) one has $1/\sqrt{N}$ and furthermore $\{0,1\}$ is replaced by $\{-1,1\}$. The limit process then remains stochastic.

on A_N at $t = 0$. Using (3.5), we can derive the time evolution of $\overline{f}_N$. If $Nt \in \mathbb{N}$, then

$$\begin{aligned}
\overline{f}_N(t+1/N) &= \sum_{m=0}^{N} \frac{m}{N} \mathbb{P}_N \left(f_N(t+1/N) = \frac{m}{N} \right) \\
&= \sum_{\ell,m=0}^{N} \frac{m}{N} \mathbb{P}_N \left(f_N(t+1/N) = \frac{m}{N} \mid f_N(t) = \frac{\ell}{N} \right) \mathbb{P}_N \left(f_N(t) = \frac{\ell}{N} \right) \\
&= \sum_{\ell,m=0}^{N} \frac{m}{N} \cdot P_{\ell m} \mathbb{P}_N \left(f_N(t) = \frac{\ell}{N} \right) \\
&= \sum_{\ell=0}^{N} \left(\frac{(\ell-1)}{N} \cdot \frac{\ell}{N} + \frac{(\ell+1)}{N} \cdot \frac{N-\ell}{N} \right) \mathbb{P}_N \left(f_N(t) = \frac{\ell}{N} \right) \\
&= \sum_{l=0}^{N} \left(\frac{\ell}{N} + \frac{1}{N} - \frac{2\ell}{N^2} \right) \mathbb{P}_N \left(f_N(t) = \frac{\ell}{N} \right) \\
&= \overline{f}_N(t) + \frac{1}{N}(1 - 2\overline{f}_N(t)).
\end{aligned} \tag{3.11}$$

Solving this equation for t such that $Nt \in \mathbb{N}$ and using $f_N(t) = f_N(\lfloor Nt \rfloor / N)$, we obtain

$$\overline{f}_N(t) - \frac{1}{2} = \left(\overline{f}_N(0) - \frac{1}{2} \right) \left(1 - \frac{2}{N} \right)^{\lfloor Nt \rfloor}. \tag{3.12}$$

Provided $\lim_{N\to\infty} \overline{f}_N(0)$ exists, eq. (3.12) implies that $\overline{f}(t) := \lim_{N\to\infty} \overline{f}_N(t)$ also exists pointwise for any $t \in [0, \infty)$, and is given by

$$\overline{f}(t) - \frac{1}{2} = \left(\overline{f}(0) - \frac{1}{2} \right) e^{-2t}, \tag{3.13}$$

which exhibits an exponentially fast approach to the equilibrium value $\overline{f} = \frac{1}{2}$ as $t \to \infty$. One may also take $\lim_{N\to\infty}$ before solving (3.11) by rewriting it as

$$\frac{\overline{f}_N(t+1/N) - \overline{f}_N(t)}{1/N} = 1 - 2\overline{f}_N(t). \tag{3.14}$$

As $N \to \infty$, this yields, at least heuristically, A. Martin-Löf's (1979) toy Boltzmann equation

$$\frac{d\overline{f}(t)}{dt} = 1 - 2\overline{f}(t). \tag{3.15}$$

The above derivation is an embryonic version of the main result of Deng, Hani, and Ma (2024), at least in the sense that both are statements about averages. However, while their result crucially depends on the choice of i.i.d. initial conditions, here we only have the mild demand that $\lim_{N\to\infty} \overline{f}_N(0)$ exists. The initial condition seems to be washed out by the stochasticity of the dynamics, since the same does not hold for deterministic dynamics (as we will see in §4).

Both the microscopic and the macroscopic Ehrenfest models are time symmetric in the specific way defined in Appendix B, since both satisfy (B.8) and hence (B.7); in both cases the instantaneous time-reversal map T is the identity. However, the fact that our definition (B.7) of time-reversal invariance in time-homogeneous Markov chains forces the initial condition μ to be stationary (and hence maintains the system in equilibrium) may still leave room for irreversible behaviour if one picks a different initial condition. This does not happen in the microscopic

Ehrenfest model, but in the macroscopic version, take some $\tau > 0$ and for each $t \in [0, \tau]$ define the time-reversed distribution function by

$$\overline{f}_N^R(t) := \overline{f}_N(\tau - t), \tag{3.16}$$

cf. (3.10) and (B.2). Using (3.13), $\overline{f}^R(t) = \lim_{N\to\infty} \overline{f}_N^R(t)$ exists for any $t \in [0, \tau]$ and is given by

$$\overline{f}^R(t) - \frac{1}{2} = \left(\overline{f}^R(0) - \frac{1}{2}\right) e^{2t}, \tag{3.17}$$

which instead of (3.15) solves the toy anti-Boltzmann equation

$$\frac{d\overline{f}^R(t)}{dt} = 1 + 2\overline{f}^R(t). \tag{3.18}$$

In particular, this macroscopic theory fails to be time-reversal invariant in the sense that $\overline{f}$ and $\overline{f}^R$ satisfy different equations (3.15) and (3.18), respectively. As already noted in connection with the real Boltzmann equation, this does not provide an "arrow of time": whatever properties $\overline{f}$ has towards the future, $\overline{f}^R$ has towards the past. The most telling such property is entropy increase: a natural (and successful) choice here is the Shannon entropy

$$S(t) = -\overline{f}(t)\log\overline{f}(t) - (1-\overline{f}(t))\log(1-\overline{f}(t)), \tag{3.19}$$

where $\overline{f}$ is seen a probability distribution p on $\{0,1\}$ by identifying $\overline{f}$ with $p(1)$; this interpretation follows from the definition of f_N as the one-particle distribution function, i.e., $1/N$ times the number of balls in urn U_1, and the convergence $f_N \to \overline{f}$, which holds on average. Then

$$\frac{dS}{dt} = (2\overline{f}(t) - 1)\log\left(\frac{\overline{f}(t)}{1-\overline{f}(t)}\right), \tag{3.20}$$

by the toy Boltzmann equation (3.15). Hence $dS/dt \geq 0$ with S reaching its maximum value at the equilibrium value $\overline{f} = 1/2$, where $dS/dt = 0$. All this is analogous to the real Boltzmann equation.

Nonetheless, despite the fame of the Ehrenfest model and their success in modeling certain aspects of the real Boltzmann equation, the underlying assumption that any at time step *exactly one ball jumps* is physically quite curious, for how are the other balls supposed to "know" that they must stay where they are? Randomness of the dynamics is no excuse, since this "knowledge" is supposed to exist for any sample path. In addition, the derivation of macroscopic evolution equations in the Ehrenfest model requires time scaling, analogous to the necessity of the Boltzmann–Grad limit in deriving the Boltzmann equation. As indicated at the end of §2, this forms a problem for our program of deriving macroscopic evolution equations from microscopic algorithmic randomness: there seems to be no version of the Ehrenfest model with $N = \infty$.[40]

For these reasons we switch to a variant of the Ehrenfest model model introduced by Hauert, Nagler, and Schuster (2004), which we call the *modified Ehrenfest model*. Here the state space remains $A_N = 2^N$, again with the interpretation of balls jumping between two urns U_0 and U_1, but

[40] A possible solution to this problem is to restrict to a finite number of particles N and to show that random paths of microstates satisfy (3.15) approximately. Since our main notion of algorithmic randomness, as given in Definition A.1, is trivial for finite spaces, 'randomness' should here be taken to mean *Kolmogorov randomness* (see Appendix A). Results in this direction have already been obtained by Calude (2002), Chapter 5 for the law of large numbers for a finite number of fair coin tosses. Since macroscopic evolution equations are essentially consequences of the law of large numbers, we expect these results to generalize to the non-equilibrium setting.

now each ball is independent in that for some fixed $p \in [0,1]$ it has its own transition probabilities, given by the 2×2 matrix

$$(\pi_{ij}) = \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix}, \qquad i,j = 0,1, \tag{3.21}$$

and by definition the overall transition probability matrix between $a, b \in 2^N$ is given by

$$P_{ab}^{(N)} = \prod_{n=1}^{N} \pi_{a_n b_n}. \tag{3.22}$$

This defines a Markov chain in the usual way upon supplying an initial distribution. The original Ehrenfest model is strictly speaking not a special case of this modified model, but the latter gives similar results upon setting $p = 1/N$. This emerges most clearly in the macroscopic behaviour of the modified model. For example, if U_1 contains m balls, then the probability of a single ball jumping from U_1 to U_0 is approximately $(m/N)/e$ and the probability of a single ball jumping from U_0 to U_1 is approximately $(1-m/N)/e$. This differs from (3.5) by a factor $1/e$, stemming from the fact that transitions in which multiple balls jump are now also possible.

Unlike in the original Ehrenfest model, in the modified model rescaling of time is not necessary to derive a sensible macroscopic evolution equation in the limit $N \to \infty$, since the transition probabilities (3.21) do not scale with N, and we indeed do not rescale time so as to be able to take pointwise limits and refine these through algorithmic randomness. Consequently, the time evolution of the "one-particle distribution"

$$f_N(t) = \frac{1}{N} \sum_{n=0}^{N-1} X_n(t), \tag{3.23}$$

where $t \in \mathbb{N}$, will be a difference equation rather than a differential equation. As before, we first derive the time evolution of the average

$$\overline{f}_N(t) = \langle f_N(t) \rangle_{\mathbb{P}_N}, \tag{3.24}$$

where $\mathbb{P}_N$ is the law of $(X(t))_{t \in \mathbb{N}}$ with an arbitrary initial condition. Since each $X_n(t)$ takes values in $\{0,1\}$, we may write

$$\overline{f}_N(t) = \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{P}_N(X_n(t) = 1). \tag{3.25}$$

Since $(X_n(t))_{t \in \mathbb{N}}$ is a Markov chain with transition probabilities (3.21) for each n, we have

$$\begin{aligned} \mathbb{P}_N(X_n(t+1) = 1) &= \mathbb{P}_N(X_n(t) = 0)\pi_{01} + \mathbb{P}_N(X_n(t) = 1)\pi_{11} \\ &= \mathbb{P}_N(X_n(t) = 1) + p(1 - 2\mathbb{P}_N(X_n(t) = 1)), \end{aligned} \tag{3.26}$$

which is a master equation.[41] By linearity, this implies that (3.25) satisfies the difference equation

$$\overline{f}_N(t+1) = \overline{f}_N(t) + p(1 - 2\overline{f}_N(t)), \tag{3.27}$$

which is solved by

$$\overline{f}_N(t) - \frac{1}{2} = \left(\overline{f}_N(0) - \frac{1}{2} \right) (1-2p)^t. \tag{3.28}$$

[41] For $p = 1/2$ the master equation degenerates into $\mathbb{P}_N(X_n(t+1) = 1) = 1/2$, so that any initial distribution of $X_n(0)$ immediately collapses into equilibrium after a single time step.

If $p = 1/2$, this formula still holds for $t > 0$ if the convention $0^0 = 1$ is used. From now on we assume $p \in (0, 1/2)$ to avoid pathological situations. With this assumption, (3.28) represents a monotonic and exponentially fast approach to equilibrium. Again,

$$\overline{f}(t) := \lim_{N\to\infty} \overline{f}_N(t) \tag{3.29}$$

exists if and only if $\overline{f}(0) = \lim_{N\to\infty} \overline{f}_N(0)$ exists, and in that case $\overline{f}$ satisfies

$$\overline{f}(t+1) = \overline{f}(t) + p(1 - 2\overline{f}(t)), \tag{3.30}$$

which is the same as the toy Boltzmann equation (3.27) but now without the subscript.

As in the original Ehrenfest model and similarly to the real Boltzmann equation, the toy Boltzmann equation (3.30) is an equation for the *average* of the one-particle distribution function. But unlike the previous two cases, the modified Ehrenfest model allows us to formulate the dynamics directly with an infinite number of particles. The above strategy of starting with a Markov chain with state space $\{0, 1\}$ and transition probabilities (3.21), and then taking a product of copies of this process also works to this end. The strong law of large numbers can then be formulated and proved, and may subsequently be refined by algorithmic randomness, as we now show.

A stochastic process $(X(t))_{t\in\mathbb{T}}$ with state space E may be represented using the product space $E^{\mathbb{T}}$ as underlying probability space, with a probability measure uniquely determined by the process.[42] This is called the *Kolmogorov representation*, which allows us to represent stochastic processes by probability measures on product spaces. In particular, with $E = \{0, 1\}$ and $\mathbb{T} = \mathbb{N}$ or an initial segment of $\mathbb{N}$ (so that $0 \in \mathbb{T}$), we let $(2^{\mathbb{T}}, \mathbb{P}_1)$ be the Kolmogorov representation of the two-state Markov chain with transition probabilities (3.21) with initial distribution p given by $p(1) = \alpha$ for some fixed $\alpha \in [0, 1]$, though we suppress the dependence of $\mathbb{P}_1$ (and all ensuing constructions) on α; we keep the notation $\mathbb{T}$ in order to avoid confusion with another $\mathbb{N}$ below.

The modified Ehrenfest model with $N = \infty$ is then simply defined as the product of countably infinitely many copies of $(2^{\mathbb{T}}, \mathbb{P}_1)$, so that the ensuing probability space $(\Omega, \mathbb{P})$ has

$$\Omega = \prod_{n=0}^{\infty} 2^{\mathbb{T}} = (2^{\mathbb{T}})^{\mathbb{N}} \cong (2^{\mathbb{N}})^{\mathbb{T}}; \qquad\qquad \mathbb{P} = \prod_{n=0}^{\infty} \mathbb{P}_1. \tag{3.31}$$

The isomorphism means that we regard $2^{\mathbb{N}}$ as the state space of a stochastic process $(X(t))_{t\in\mathbb{T}}$ characterized by $\mathbb{P}$, with $(\Omega, \mathbb{P})$ its Kolmogorov representation; in other words, whereas elements of $(2^{\mathbb{T}})^{\mathbb{N}}$ are by definition functions $\omega : \mathbb{N} \to 2^{\mathbb{T}}$, we see these as functions $\omega : \mathbb{T} \to 2^{\mathbb{N}}$ (so we identify $\omega(n)(t)$ with $\omega_n(t)$). By construction, the process $(X(t))_{t\in\mathbb{T}}$ has the following properties:

- At each fixed $n \in \mathbb{N}$, the $(X_n(t))_{t\in\mathbb{T}}$ form a two-state Markov chain with transition probabilities and initial distribution p given by $p(1) = \alpha$;
- For each fixed $t \in \mathbb{T}$, the $(X_n(t))_{n\in\mathbb{N}}$ are i.i.d.

Because we have infinite number of particles, the one-particle distribution function is now a limit

$$f(t, \omega) = \lim_{N\to\infty} f_N(t, \omega); \qquad\qquad f_N(t, \omega) = \frac{1}{N} \sum_{n=0}^{N-1} X_n(t, \omega), \tag{3.32}$$

[42] This construction, due to Kolmogorov, is completely standard in probability theory. See for example Dudley (1989), §8.2, or Klenke (2020), Chapter 14. Here $E^{\mathbb{T}}$ is equipped with the smallest σ-algebra $\mathscr{F}$ that makes each evaluation map $X_t : E^{\mathbb{T}} \to E$, $x \mapsto x(t)$ measurable, given some σ-algebra on the state space E that is part of its definition as a measurable space.

where we have reinstated the dependence on the microscopic path ω in the notation, cf. (2.6),[43] and the notation $f(t,\omega)$ is reserved for those ω for which the limit exists. This is not the case for every path $\omega \in \Omega$, but $f(t,\omega)$ does exist for every $\mathbb{P}$-random path, as we will now show.

Using the theory of large deviations, we can prove that for $\mathbb{P}$-almost all $\omega \in \Omega$ this limit converges to a function $f(t)$ satisfying the toy Boltzmann equation

$$f(t+1) = f(t) + p(1-2f(t)). \tag{3.33}$$

More precisely, Cramér's theorem, which we already used to derive (1.5), in which $q = f$ with $f(0) = f(1) = 1/2$, also works for possibly biased priors q and gives almost identical results as the fair case: for any $m \in \mathbb{N}_*$, one has

$$V_N(m) := \{s \in 2^{\mathbb{N}} : |S_N(s) - q| > 1/m\}; \qquad q^{\mathbb{N}}(V_N(m)) \leq e^{-d(m)N}, \tag{3.34}$$

for some $d(m) > 0$ and sufficiently large N, using the same notation as in §1. Now define

$$f(t) = \langle f_N(t,\cdot)\rangle_{\mathbb{P}}. \tag{3.35}$$

By the same reasoning leading to (3.30), $f(t)$ satisfies (3.33). By our choice of the initial condition of $\mathbb{P}_1$, $f(0) = \alpha$. Let $\mathbb{T} \subset \mathbb{N}$ be an initial segment. Applying Cramér's theorem for biased priors to the sequence $(X_n(t))_{n\in\mathbb{N}}$, which is i.i.d. by construction, uniformly in $t \in \mathbb{T}$ (which is a trivial extension, since $\mathbb{T}$ is finite), we have

$$W_N(m) := \left\{\omega \in (2^{\mathbb{N}})^{\mathbb{T}} : \|f_N(\cdot,\omega) - f\|_\infty > 1/m\right\}; \qquad \mathbb{P}(W_N(m)) \leq e^{-d'(m)N}, \tag{3.36}$$

for $d'(m) > 0$ and sufficiently large N, where

$$\|f_N(\cdot,\omega) - f\|_\infty := \sup_{t\in\mathbb{T}} |f_N(t,\omega) - f(t)|; \tag{3.37}$$

one could also state a similar result for each single $t \in \mathbb{T}$, and take the conjunction of these. It then follows from the Borel–Cantelli lemma that $\mathbb{P}$-almost surely on Ω,

$$\lim_{N\to\infty} f_N(\cdot,\omega) = f. \tag{3.38}$$

To strengthen the clause "for $\mathbb{P}$-almost every ω" in this result to our desirable clause "for all $\mathbb{P}$-random ω" we may again use the reasoning based on Lemma 1.1, which gave us a similar strengthening of the strong law of large numbers for a fair coin toss. As we noted, this lemma holds for any computable probability space. Thus we obtain the main result in this section:[44]

Theorem 3.1 *Let the initial condition $\alpha \in [0,1]$, which parametrizes $\mathbb{P}_1$ and hence $\mathbb{P}$, be a computable real. For each initial segment $\mathbb{T} \subset \mathbb{N}$ and each $\mathbb{P}$-random path $\omega \in (2^{\mathbb{N}})^{\mathbb{T}}$, we have*

$$\lim_{N\to\infty} f_N(\cdot,\omega) = f, \tag{3.39}$$

in norm, cf. (3.36)*, where $f(t)$ is a solution of the toy Boltzmann equation*

$$f(t+1) = f(t) + p(1-2f(t)), \tag{3.40}$$

for $t \in \mathbb{T}$ with initial condition $f(0) = \alpha$.

[43] Due to the construction (3.31), the random variable $X_n(t)$ is simply the evaluation function $X_n(t,\omega) = \omega_n(t)$; the real content of the microscopic process lies in the probability measure $\mathbb{P}$.

[44] The computability requirements enabling Lemma 1.1 are easily checked. First, the topological space $(2^{\mathbb{T}})^{\mathbb{N}}$ becomes effective by numbering the cylinder sets. Second, if $f(0)$ is computable, then also $f(t)$ is computable, which makes the family $(W_N(m))_N$ uniformly computable by saturating the inequality in its definition by computable opens.

In other words, $\mathbb{P}$-randomness of a *microscopic* path ω is sufficient for the convergence of the corresponding *macroscopic* truncated distribution function $f_N(\cdot,\omega)$ to a solution of the toy Boltzmann equation, where $\mathbb{P}$ is a probability measure on paths developing from i.i.d. initial conditions via the (Markovian) stochastic dynamics of the modified Ehrenfest model.

Proof. This just summarizes our reasoning so far. Large deviation theory gives (3.34) applied to the sequence $(X_n(t))_{n\in\mathbb{N}}$ at fixed $t\in\mathbb{T}$ with probabililty distribution $q=f(t)$ on $\{0,1\}$, and hence (3.36), since only finitely many times are involved. The appearance of f is explained by (3.35) and the argument leading to (3.30) showed that f satisfies (3.40). The inequality in (3.36) enables Lemma 1.1, which gives the implication

$$\omega\in(2^{\mathbb{N}})^{\mathbb{T}} \text{ is } \mathbb{P}\text{-random} \Rightarrow \omega\in W_N(m) \text{ for finitely many } N. \tag{3.41}$$

Since m is arbitrary,[45] this then implies (3.39). □

What are the implications for time reversal? Recall that our goal is to show that some (algorithmic) microscopic randomness assumption (which incorporates the initial conditions) lies behind irreversible macroscopic evolution equations, like the Boltzmann equation. Hence it would be natural that such an assumption is also the source of irreversibility. In particular, the appropriate concept of $\mathbb{P}$-randomness of (microscopic) paths should not be invariant under time reversal, and indeed, whereas the set of $\mathbb{P}$-random paths has probability one, the time reversals of these paths should have very low (or even zero) probability, as they are never observed. In our modified Ehrenfest model, this indeed follows from the contrapositive of the implication (3.41) on which our proof was based. Writing $\mathbb{T}=\{0,1,2,\ldots,\tau\}$, the time reversal ω^R of a path $\omega\in\Omega$ is

$$\omega^R(t)=\omega(\tau-t), \tag{3.42}$$

cf. (B.2) with $T=\mathrm{id}$. Since $f_N(t,\omega^R)=f_N(\tau-t,\omega)$, if ω is $\mathbb{P}$-random, then

$$\lim_{N\to\infty} f_N(t,\omega^R)=f(\tau-t). \tag{3.43}$$

by Theorem 3.1, with $f(t)$ a solution of the toy Boltzmann equation (3.40). If ω^R were $\mathbb{P}$-random too, then $f(\tau-t)$ would have to satisfy (3.40), which it clearly does not (unless already $f(0)=1/2$). Hence ω^R is not $\mathbb{P}$-random. Moreover, taking also the proof into account yields some $\varepsilon>0$ and a subsequence (N_k) such that $\omega^R\in W_{N_k}(\varepsilon)$ for all k, so that (3.36) implies in addition that such time-reversed paths are exponentially unlikely in N as $N\to\infty$.

As in the original Ehrenfest model, the result (3.30) on convergence of averages does not dependent on the initial condition, expect for the mild demand of convergence at $t=0$. Moreover, Theorem 3.1 even holds without the assumption that $(X_n(t))_{t\in\mathbb{T}}$ has the same distribution for each $n\in\mathbb{N}$, which is baked into the construction (3.31).[46] Making the distribution of $X_n(0)$ deterministic for each $n\in\mathbb{N}$, this implies that *any* initial microstate may lead to irreversible behaviour, which is a big difference with deterministic dynamics. This raises the question what the role of the *Stosszahlansatz* in the modified Ehrenfest model could be, either in Boltzmann's original form in or the modern probabilistic form (cf. §2). Let us discuss various perspectives on this.

1. While every initial microstate $s\in 2^{\mathbb{N}}$ may lead to irreversible behaviour, it does not follow that each path starting from s behaves irreversibly, as obvious counterexamples show; instead, the analogue of Boltzmann's original *Stosszahlansatz* in the modified Ehrenfest model

[45] Since the finite subset $S\subset\mathbb{N}$ for which $\omega\in W_N(m)$ iff $N\in S$ depends on m, let us be careful: the implication (3.41) is $\forall_m\exists_{S\subset\mathbb{N}_f}\forall_{N\notin S}\forall_{t\in\mathbb{T}}|f_N(t,\omega)-f(t)|\le 1/m$, where $\mathbb{T}$ is finite. For all m and t this gives $\limsup_{N\to\infty} f_N(t,\omega)\le f(t)+1/m$ and $\liminf_{N\to\infty} f_N(t,\omega)\ge f(t)-1/m$, so that $\lim_{N\to\infty} f_N(t,\omega)=f(t)$ for all $t\in\mathbb{T}$, which is (3.39).

[46] But this time one cannot use the (standard version of) Cramér's theorem anymore to prove this.

exists on the level of paths. Recalling the notation (3.32), the *Stosszahlansatz* states that the proportion of balls in U_1 which jump to U_0 in the next time step is $pf(t,\omega)$, and similarly equals $p(1-f(t,\omega))$ for balls jumping from U_0 to U_1. From this assumption, one obtains

$$\begin{aligned} f(t+1,\omega)-f(t,\omega) &= p(1-f(t,\omega))-pf(t,\omega) \\ &= p(1-2f(t,\omega)), \end{aligned} \tag{3.44}$$

which is a toy Boltzmann equation. The *Stosszahlansatz* of the modified Ehrenfest model can actually be rigorously formulated and proven to hold for all $\mathbb{P}$-random $\omega\in\Omega$, once again not necessarily assuming that the balls have identically distributed Markov chains. This rigorous formulation of the *Stosszahlansatz* for balls jumping from U_1 to U_0 is

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}X_n(t,\omega)(1-X_n(t+1,\omega)) = p\cdot\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}X_n(t,\omega), \tag{3.45}$$

which can be rewritten as

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}X_n(t,\omega)(Y_n(t,\omega)-p)=0, \tag{3.46}$$

where $Y_n(t,\cdot)$ is the random variable with value $Y_n(t,\omega)=1$ if the n'th ball jumps during the time step $t\to t+1$, and $Y_n(t,\omega)=0$ otherwise. Clearly, $X_n(t,\cdot)$ and $Y_n(t,\cdot)$ are independent and $\langle Y_n(t,\cdot)\rangle_{\mathbb{P}}=p$. Hence one recognizes the structure of the law of large numbers in (3.46). Indeed, just as in Theorem 3.1 this equality can be proved for all $\mathbb{P}$-random $\omega\in\Omega$.

2. While the assumption of i.i.d. initial conditions can be relaxed as indicated, independence does remain a crucial assumption. Even in the above-mentioned scenario with a fixed initial microstate, in which each $X_n(0)$ is deterministic, independence is satisfied (since deterministic random variables are trivially independent). In fact, under the assumption of exchangeability at $t=0$, which is a natural assumption in view of our ignorance of microscopic conditions, i.i.d. initial conditions are *necessary* for laws of large numbers such as (3.39). This gives at least a mathematical, but not yet a conceptual answer to the question of why i.i.d. initial conditions are used in deriving macroscopic evolution equations.

   In a more general setting, let $(X_n)_{n\in N}$ be an exchangeable sequence of random variables taking values in $\{0,1\}$ with common mean μ. In particular, exchangeability already implies that the sequence is identically distributed. Setting $S_N=N^{-1}\sum_{n=0}^{N-1}X_n$, the weak law of large numbers

$$\lim_{N\to\infty}\mathbb{P}(|S_N-\mu|\geq\varepsilon)=0 \tag{3.47}$$

   is then equivalent to $\lim_{N\to\infty}\mathrm{Var}(S_N)=0$; one implication makes use Chebyshev's inequality, the other follows from bounding $\mathrm{Var}(S_N)$ by $\mathbb{P}(|S_N-\mu|\geq\varepsilon)$. Using exchangeability,

$$\mathrm{Var}(S_N)=\frac{\sigma^2}{N}+\left(1-\frac{1}{N}\right)\rho, \tag{3.48}$$

   where $\sigma^2=\mathrm{V}(X_1)$ and $\rho=\mathrm{Cov}(X_1,X_2)$. Hence $\lim_{N\to\infty}\mathrm{Var}(S_N)=0$ iff $\rho=0$, which for random variables taking values in $\{0,1\}$ is equivalent to pairwise independence. Using de Finetti's theorem,[47] this can be upgraded to mutual independence, from which we conclude

[47] See Cifarelli and Muliere (2002) for an introduction to de Finetti's theorem. See also Klenke (2020), §§12.3, 13.4.

that $(X_n)_{n\in\mathbb{N}}$ are i.i.d. To see this, use de Finetti's theorem to write

$$\mathbb{P}(X_0 = x_0, \ldots, X_{N-1} = x_{N-1}) = \int_0^1 \prod_{n=0}^{N-1} \theta^{x_i}(1-\theta)^{1-x_i}\, d\nu(\theta), \tag{3.49}$$

for some probability measure ν on $[0,1]$. Using this, as well as pairwise independence $\mathbb{P}(X_0 = x_0, X_1 = x_1) = \mathbb{P}(X_0 = x_0)\mathbb{P}(X_1 = x_1)$, we obtain

$$\int_0^1 \theta^2\, d\nu(\theta) = \left(\int_0^1 \theta\, d\nu(\theta)\right)^2. \tag{3.50}$$

This can only be the case if ν is deterministic, which implies that (3.49) factorizes, i.e. the random variables $(X_n)_{n\in\mathbb{N}}$ are mutually independent. Applying this reasoning to the modified Ehrenfest model, we see that (3.39) (which implies the weak law of large numbers) implies that $(X_n(0))_{n\in\mathbb{N}}$ are i.i.d., under the assumption of exchangeability. Note that this has nothing to do with the dynamics, since we only need to look at the initial time $t = 0$.

4 Irreversibility in the Kac ring model

A similar scenario may be realized in the deterministic Kac ring model.[48] Since this model has already been studied from a similar point of view in the pinoneering paper by Hiura and Sasa (2019), we will be brief. We rederive their main result and add some relevant insights. We again start with a microscopic description. For finite N, the microscopic state space of the Kac ring is

$$A_N := 2^{2N+1} \times 2^{2N+1}, \tag{4.1}$$

whose elements we write as pairs (x,y) consisting of two kinds of microscopic degrees of freedom. The component x_n represents the colour of a ball at the n'th site, which is white if $x_n = 1$ and black if $x_n = 0$. For the other kind of microscopic variables, we say that site n is *marked* if $y_n = 1$ and *unmarked* if $y_n = 0$. The microscopic dynamics is defined by the following map:

$$\varphi : A_N \to A_N; \qquad \varphi(x,y)(n) = (x_{n-1} + y_{n-1}(1 - 2x_{n-1}), y_n), \tag{4.2}$$

whose t-fold iteration we denote by φ_t. The subscripts here are always taken modulo N, in effect enforcing periodic boundary conditions, whence the name 'ring'. The interpretation of (4.2) is as follows. At each time step every ball moves one site in a fixed direction of the ring. If it leaves a marked site its colour changes; otherwise its colour remains unchanged. The set of marked sites does not change. The dynamics of the Kac ring is reversible, cf. (B.1), with instantaneous time-reversal map $T : A_N \to A_N$ given by

$$T(x,y)(n) = (x_{-n}, y_{-n-1}). \tag{4.3}$$

The Kac ring model may also conveniently be described using the coordinates (η, ε) defined by

$$(\eta,\varepsilon) \in \{-1,1\}^{2N+1} \times \{-1,1\}^{2N+1}; \qquad \eta_n = 2x_n - 1; \qquad \varepsilon_n = -2y_n + 1. \tag{4.4}$$

In this representation, η is often interpreted as a configuration of spins and ε as a configuration of scatterers. Using these variables, the dynamics becomes much simpler:

$$\varphi(\eta,\varepsilon)(n) = (\eta_{n-1}\varepsilon_{n-1}, \varepsilon_n); \qquad \varphi_t(\eta,\varepsilon)(n) = (\eta_{n-t}\varepsilon_{n-t}\cdots\varepsilon_{n-1}, \varepsilon_n). \tag{4.5}$$

[48] The original reference for the Kac ring model is Kac (1959). Useful literature includes Maes, Netočný, and Shergelashvili (2007), De Bièvre and Parris (2017), and Jebeile (2020). Our take is quite different.

Because the Kac ring has two kinds of microscopic degrees of freedom, there are two one-particle distribution functions acting as macroscopic variables, one for balls and the other for markers:

$$f_N(x,y) := \frac{1}{2N+1}\sum_{n=-N}^{N} x_n; \qquad\qquad s_N(x,y) = \frac{1}{2N+1}\sum_{n=-N}^{N} y_n. \tag{4.6}$$

The first gives the proportion of white balls, the second the proportion of marked sites. Note that the latter is non-dynamical. Writing $\varphi_t(x,y)(n) = (x_n(t), y_n(t))$, with of course $y_n(t) = y_n$, the time evolution of f_N reads

$$f_N(t,x,y) := f_N(\varphi_t(x,y)) = \frac{1}{2N+1}\sum_{n=-N}^{N} x_n(t). \tag{4.7}$$

Scaling in time, though an option, is not necessary to obtain limits of these macroscopic quantities as $N \to \infty$. As in all previous cases, we introduce probability by assuming an i.i.d. initial distribution over microstates (x,y), in which pairs (x_n, y_n) are i.i.d. as n runs through $\mathbb{Z}$. Thus we equip A_N with the product Bernoulli measure

$$\mathbb{P}_N = \alpha^{2N+1} \times \beta^{2N+1}, \tag{4.8}$$

where we label probability measures on $\{0,1\}$ by their means. Getting rid of the dependence on the microstate by averaging, i.e., defining

$$\overline{f}_N(t) = \langle f_N(t,\cdot)\rangle_{\mathbb{P}_N}; \qquad\qquad \overline{s}_N = \langle s_N\rangle_{\mathbb{P}_N}, \tag{4.9}$$

one readily calculates, most conveniently using the representation (4.5), that

$$\overline{f}_N(t) - \frac{1}{2} = \left(\alpha - \frac{1}{2}\right)(1-2\beta)^t; \qquad\qquad \overline{s}_N = \beta. \tag{4.10}$$

Hence $\overline{f}_N(t)$ satisfies the toy Boltzmann equation

$$\overline{f}_N(t+1) = \overline{f}_N(t) + \beta(1 - 2\overline{f}_N(t)), \tag{4.11}$$

with initial condition $\overline{f}(0) = \alpha$ and so does $\overline{f}(t) := \lim_{N\to\infty} \overline{f}_N(t)$.[49] This equation is clearly irreversible, in the sense that, trivially, if $\overline{f}_N(t)$ satisfies (4.11), then its time reversal $\overline{f}_N(-t)$ does not (and no choice of the instantaneous time-reversal T can save this in the macroscopic model).

As in the modified Ehrenfest model, it is possible to go beyond averages and prove almost sure pointwise results, which can subsequently be converted into randomness statements. This was achieved by Hiura and Sasa (2019), Theorem 3.5. For completeness we now state and (re)prove their result, using a different argument that connects well with techniques used elsewhere in our paper. Instead of (4.1) the state space $A_{\mathbb{Z}} = 2^{\mathbb{Z}} \times 2^{\mathbb{Z}}$ is used, and on it the infinite analogue of (4.8),

$$\mathbb{P} = \alpha^{\mathbb{Z}} \times \beta^{\mathbb{Z}}. \tag{4.12}$$

The dynamics (4.2) remains the same, but without the periodic boundary conditions. The interpretation is now one of coloured balls moving on an infinite $1d$ lattice $\mathbb{Z}$, which Hiura and Sasa (2019) call the 'Kac chain'. As in the modified Ehrenfest model, the relevant macroscopic quantities are now the limits of (4.6) as $N \to \infty$.

[49] This toy Boltzmann equation is the same as in the modified Ehrenfest model, cf. (3.27). The latter may indeed be seen as a stochastic version of the Kac ring model, in which the markers are continually randomly added and erased.

Theorem 4.1 *Let $\alpha, \beta \in [0,1]$ be computable reals. For each $t \in \mathbb{N}$ and each $\mathbb{P}$-random microstate $(x,y) \in A_{\mathbb{Z}}$, we have*

$$\lim_{N\to\infty} f_N(t,x,y) = f(t); \qquad \lim_{N\to\infty} s_N(x,y) = \beta, \tag{4.13}$$

where $f(t)$ is the solution of the toy Boltzmann equation

$$f(t+1) = f(t) + \beta(1-2f(t)), \tag{4.14}$$

with initial condition $f(0) = \alpha$.

Proof. The arguments after (1.4) suffice to prove the second limit in (4.13) and also the first at $t = 0$, since $(x_n(0))_{n\in\mathbb{Z}}$ and $(y_n)_{n\in\mathbb{Z}}$ are both i.i.d. sequences by the choice of $\mathbb{P}$. For $t > 0$, this is no longer the case: correlations develop between the colours of the balls. But these correlations are weak enough to derive a large deviation principle.[50] More precisely, the sequence $(x_n(t))_{n\in\mathbb{Z}}$ has the property that it only depends on $(t-1)$, which means that the two sequences

$$(\ldots, x_{n-1}(t), x_n(t)); \qquad (x_k(t), x_{k+1}(t), \ldots) \tag{4.15}$$

are independent if $k - n > t - 1$; correlations only reach so far, as is most easily seen in the representation (4.5). Combined with stationarity, i.e., invariance of the distribution of $(x_n(t))_{n\in\mathbb{Z}}$ under shifts in n, this property suffices to prove a large deviation principle at $t > 0$. Setting

$$f(t) = \langle f_N(t,\cdot)\rangle_{\mathbb{P}}, \tag{4.16}$$

this implies

$$U_N(m) = \{(x,y) \in A_{\mathbb{Z}} : |f_N(t,x,y) - f(t)| > 1/m\}; \qquad \mathbb{P}(U_N(m)) \leq e^{-C(m)N}. \tag{4.17}$$

The remainder of the proof is then the same as in the modified Ehrenfest model: from (4.17) one derives pointwise $\mathbb{P}$-almost sure convergence via the Borel–Cantelli lemma, and then pointwise convergence for each $\mathbb{P}$-random (x,y) via Lemma 1.1. □

If one extends $\mathbb{P}$ to a probability measure on (discrete) paths in $A_{\mathbb{Z}}$, simply by taking the probability of some path to be the same as the probability of its initial condition, the situation is analogous to the modified Ehrenfest model: each $\mathbb{P}$-random path induces a one-particle distribution function $f(t)$ that satisfies the toy Boltzmann equation (4.14), whereas the time-reversed path does not (unless it is constant in time) and has exponentially small probability (and zero probability at $N = \infty$). The status of the *Stosszahlansatz* in the Kac model is also analogous to the case of the modified Ehrenfest model, as discussed at the end of §3:

1. The analogue of the *Stosszahlansatz* of Boltzmann (1872) is the following assumption: for each $t \in \mathbb{N}$, the proportion of white balls about to leave a marked site is equal to $\beta f(t,x,y)$ and similarly equals $\beta(1 - f(t,x,y))$ for black balls. Here we have defined, cf. (4.13),

$$f(t,x,y) = \lim_{N\to\infty} f_N(t,x,y), \tag{4.18}$$

for all t and (x,y) for which this limit exists; by Theorem 4.1, this is the case for all $\mathbb{P}$-random (x,y) at all t. As before, if this assumption holds for (x,y), then $f(t,x,y)$ satisfies

[50] See Dembo and Zeitouni (1998), Section 6.4, which studies large deviation principles of stationary processes which satisfy mixing conditions (roughly speaking, a process is mixing if its future values eventually become independent of past values). The eventual independence property in the main text is a strong form of this. In particular, it makes Assumption 6.4.1 in Dembo and Zeitouni (1998) true, enabling their Theorem 6.4.4.

the toy Boltzmann equation (4.14). The precise formulation of the *Stosszahlansatz* for white balls, then, is

$$\lim_{N\to\infty}\frac{1}{2N+1}\sum_{n=-N}^{N}x_n(t)y_n=\beta\cdot\lim_{N\to\infty}\frac{1}{2N+1}\sum_{n=-N}^{N}x_n(t), \tag{4.19}$$

which is equivalent to

$$\lim_{N\to\infty}\frac{1}{2N+1}\sum_{n=-N}^{N}x_n(t)(y_n-\beta)=0. \tag{4.20}$$

This again has the form of a law of large numbers, which, using the same line of proof as (4.13), can be shown to hold for all $\mathbb{P}$-random $(x,y)\in A_{\mathbb{Z}}$.[51] This shows how also in this model the *Stosszahlansatz* of Boltzmann (1872) is a consequence of the deeper assumption of microscopic randomness relative to an i.i.d. probability distribution on microstates.

2. The necessity of i.i.d. initial conditions under the assumption of exchangeability follows from the general reasoning at the end of §2, since the Kac chain model can be formulated in terms of random variables taking values in $\{0,1\}$.

Finally, we analyze the strength of our randomness assumption. The previous arguments show that $\mathbb{P}$-randomness for suitable $\mathbb{P}$ is sufficient for Boltzmann-type equations; but so are $\mathbb{P}$-almost sure arguments. These are valid under weaker assumptions, because the set of all $\mathbb{P}$-random points in a computable probability space $(X,\mathbb{P})$ is just one example of a probability-one subset.[52] We now clarify this issue at least in the Kac chain model; but our conclusion seems more wide-ranging.

Our argument is based on ergodic theory. As a warm-up, let us recall how the strong law of large numbers for a fair coin toss and the much stronger fact that almost every sequence is Borel normal follow from Birkhoff's pointwise ergodic theorem. This is a relevant exercise, because the former is the simplest version of what we would call (correct) macroscopic behaviour, whilst the latter provides additional pointwise almost sure "macroscopic" laws. The setting is a dynamical system (X,Σ,P,T) where (X,Σ,P) is a probability space and $T:X\to X$ is a (not necessarily invertible) measurable and measure-preserving map, i.e., $P(T^{-1}B)=P(B)$ for each $B\in\Sigma$) that in addition is ergodic, in the sense that $T^{-1}(B)=B$ implies $P(B)=0$ or $P(B)=1$. Then

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\delta_{T^nx}(B)=P(B), \tag{4.21}$$

for P-almost every $x\in X$ and every $B\in\Sigma$, where we use the same notation as in (1.2), i.e., $\delta_{T^nx}(B)=1$ if $T^nx\in B$ and $\delta_{T^nx}(B)=0$ if $T^nx\notin B$. Apply this to the case in which

$$X=2^{\mathbb{N}};\qquad\qquad \Sigma=\mathscr{F};\qquad\qquad P=f^{\mathbb{N}};\qquad\qquad T=S, \tag{4.22}$$

where S is the shift map, given by $(Ss)_n=s_{n+1}$. This system is well known to be ergodic.[53] Take

$$B_1=\{s\in 2^{\mathbb{N}}\mid s_0=1\}. \tag{4.23}$$

[51] Since the proportion of markers s_N is also a random variable, a more proper formulation of the *Stosszahlansatz* should have $\lim_{N\to\infty}s_N(x,y)\cdot\lim_{N\to\infty}f_N(t,x,y)$ instead of $\beta\cdot\lim_{N\to\infty}f_N(t,x,y)$. Even in this form, it still holds for all $\mathbb{P}$-random $(x,y)\in A_{\mathbb{Z}}$, but is a bit more tedious to prove.

[52] For example, take the uniform probability f on $2^{\mathbb{N}}$, with induced Bernoulli probability measure $f^{\mathbb{N}}$ on $X=2^{\mathbb{N}}$. Then the set $\mathscr{R}_f$ of all $f^{\mathbb{N}}$-random sequences has $f^{\mathbb{N}}(\mathscr{R}_f)=1$ and its elements s have the "right" macroscopic behaviour (1.4). But so does a sequence like $010101\cdots$, and one can easily construct a countable number of such computable sequences with the "right" macroscopic behaviour. These are of course far from $f^{\mathbb{N}}$-random.

[53] A good reference for all of the above is for example Viana and Oliveira (2016), Chapter 4. Mackey (1974) is a beautiful historical survey of the interaction between ergodic theory, statistical mechanics, and probability theory.

Then $\delta_{S^n s}(B_1) = s_n$ and $f^{\mathbb{N}}(B_1) = 1/2$, so that (4.21) is the strong law of large numbers (1.4). Moreover, take any $\sigma \in 2^N$ and put $B = B_\sigma = [\sigma]_N$, cf. (A.4), which includes the case (4.23) by choosing $N = 1$ and $\sigma = 1$. Then $\delta_{S^n s}([\sigma]_N)$ equals 1 iff $s_n \cdots s_{n+N-1} = \sigma$, so that the left-hand side of (4.21) with $B = B_\sigma$ equals the asymptotic relative frequency of the string σ in s. Furthermore,

$$f^{\mathbb{N}}(B_\sigma) = f^{\mathbb{N}}(\sigma) = 2^{-N}, \tag{4.24}$$

so that (4.21) is the claim that $f^{\mathbb{N}}$-almost every $s \in 2^{\mathbb{N}}$ is Borel normal.[54]

Of course, this is the kind of result we wish to sharpen according to our familiar pattern of replacing the "for $f^{\mathbb{N}}$-almost every $s \in 2^{\mathbb{N}}$" clause by "for all $f^{\mathbb{N}}$-random $s \in 2^{\mathbb{N}}$". This can be done in several ways,[55] but in the spirit of the above story we invoke the *effective ergodic theorem.*[56] Let (X, B, P) be a computable probability space (see Appendix A) and $T : X \to X$ a computable (and hence continuous) P-preserving map. Let (X, B, P, T) be ergodic. Under these assumptions:[57]

1. For each computable open set $V \subset X$ and all P-random $x \in X$ one has

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \delta_{T^n x}(V) = P(V). \tag{4.25}$$

2. A point $x \in X$ is P-random iff (4.25) holds for every computable open $V \subset X$.

Since $(2^{\mathbb{N}}, B, f^{\mathbb{N}}, S)$, where B is defined in Appendix A, satisfies the assumptions, it follows from part 2 and the choice $V = [\sigma]_N$ in (4.25) that each $f^{\mathbb{N}}$-random sequence $s \in 2^{\mathbb{N}}$ is Borel normal.

After this preparation, we return to the Kac chain and apply the effective ergodic theorem (4.25). Our key observation is that the first equation in (4.13) may be rewritten as a two-sided version of the ergodic limit (4.25). To this end, take

$$X = A_{\mathbb{Z}}; \quad P = \mathbb{P} = \alpha^{\mathbb{Z}} \times \beta^{\mathbb{Z}}; \quad V = \varphi_{-t} U; \quad U = \{(x,y) \in A_{\mathbb{Z}} \mid x_0 = 1\}; \quad T = S, \tag{4.26}$$

where S is the (double) bilateral shift

$$S(x,y)(n) = (x_{n+1}, y_{n+1}), \tag{4.27}$$

$\varphi_{-t} U$ denotes the preimage of U under φ_t, and $x \in X$ in (4.25) is now of course $(x,y) \in A_{\mathbb{Z}}$. This system is well known to be ergodic,[58] and since S is invertible (unlike its counterpart in the unilateral case), we can extend (4.25) to a sum from $n = -N$ to $n = N$ with normalization $1/(2N+1)$. The point is that we may now identify

$$\frac{1}{2N+1} \sum_{n=-N}^{N} \delta_{T^n x}(V) = f_N(t,x,y); \qquad P(V) = \frac{1}{2} + \left(\alpha - \frac{1}{2}\right)(1-2\beta)^t, \tag{4.28}$$

[54] See e.g. Calude (2002), Definition 6.53. Borel normality of $x \in [0,1]$ was the first serious attempt to define randomness. It was flawed because *Champernowne's number* $0.12345678910111213\cdots$ is Borel nomal (in base 10) but hardly random, in view of the simple rule behind it. Likewise, in base 2 the binary analogue of Champernowne's number, i.e., $0.1101110010111011111000\cdots \in 2^{\mathbb{N}}$, is Borel normal but computable and hence it cannot be $f^{\mathbb{N}}$-random.

[55] See for example Calude (2002), Theorem 6.57, for a direct argument.

[56] See especially Franklin et al. (2012), Theorem 6. Similar results appear in Galatolo, Hoyrup, and Rojas (2010), Theorem 3.2.2, and Bienvenu et al. (2012), Theorem 8. The first author to prove such results was V'yugin (1997). See also the reviews by Towsner (2020) and V'yugin (2022). At the cost of additional computability assumptions, this theorem sharpens Birkhoff's pointwise ergodic theorem, which yields the first claim for P-almost every $x \in X$. As such, this is one of the key results in algorithmic randomness that replaces "for P-almost all x" by "for all P-random x".

[57] Note that since $P(X) = 1$ and $\delta_x(X \backslash V) = 1 - \delta_x(V)$, all assumptions as well as the statement of the theorem may equivalently be stated for computable closed sets, i.e., complements of computable opens (our sets will be clopen).

[58] And, as we have seen, effective provided α and β are computable reals, with S easily shown to be computable.

upon which (4.25) indeed turns into the first equation in (4.13). The first part of (4.28) follows from (4.7) and the equalities

$$S^{-n}U = \{(x,y) \in A_{\mathbb{Z}} \mid x_n = 1\}; \qquad \delta_{S^n(x,y)}(\varphi_{-t}U) = \delta_{\varphi_t(x,y)}(S^{-n}U) = x_n(t), \tag{4.29}$$

of which the second holds since S and φ commute. The second part of (4.28) comes down to the fact that

$$\mathbb{P}(\varphi_{-t}U) = \mathbb{P}(x_0(t) = 1) = \langle x_0(t) \rangle_{\mathbb{P}}, \tag{4.30}$$

the average being computed using (4.5) in precisely the same way as in (4.9) - (4.10). Generalizing (4.26) and the previous reasoning to arbitrary computable opens $U \subset A_{\mathbb{Z}}$, eq. (4.25) becomes

$$\lim_{N\to\infty} \frac{1}{2N+1} \sum_{n=-N}^{N} \delta_{\varphi_t(x,y)}(S^{-n}U) = \mathbb{P}(\varphi_{-t}U). \tag{4.31}$$

It is tempting, then, to regard (4.31) as U ranges over all computable opens as a set of "macroscopic laws" of the Kac chain, sufficiently large that (x,y) is $\mathbb{P}$-random iff all these macroscopic laws are satisfied.[59] The left-hand side of (4.31) is a macroscopic quantity, while the right-hand side is its expected time evolution. The existence of non-$\mathbb{P}$-random microstates whose induced one-particle distribution functions nevertheless satisfy the toy Boltzmann equation, which possibility was noted by Hiura and Sasa (2019), is then explained by the fact that these microstates fail to satisfy other macroscopic laws. Even the class of macroscopic laws just defined via the effective ergodic theorem (4.25) does not exhaust the set of all macroscopic laws of the Kac chain model.[60]

A Algorithmic randomness

Triggered by ideas of Kolmogorov, the concept of P-randomness (where P is a probability measure) was introduced by Martin-Löf (1966) for the case $X = 2^{\mathbb{N}}$ of infinite binary sequences with the uniform Bernoulli measure $P = f^{\mathbb{N}}$. A more general definition of P-randomness is as follows:[61]

Definition A.1 *1. A* computable *(or* effective*) topological space X is a topological space with a countable base $\mathscr{B} \subset \mathscr{O}(X)$, numbered by a bijection $B : \mathbb{N} \xrightarrow{\cong} \mathscr{B}$. In such a space:*

(a) An open set $V \in \mathscr{O}(X)$ in X is computable *(effective) if*

$$V = \bigcup_{n\in\mathbb{N}} B(f(n)), \tag{A.1}$$

for some (total) computable function $f : \mathbb{N} \to \mathbb{N}$.

[59] This characterization of $\mathbb{P}$-randomness holds for any $t \in \mathbb{N}$, in particular for $t = 0$. Hence if (x,y) is such that its macroscopic values are all correct at $t = 0$ in the sense of (4.31), then (x,y) is $\mathbb{P}$-random, which implies that its macroscopic values will also be correct (in the same sense) at all future times $t > 0$.

[60] Bienvenu et al. (2012) prove two different effective ergodic theorems. Apart from (4.25), which is their Theorem 8, they also show that $\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(T^n x) = \int_X f dP$, for arbitrary lower semicomputable functions $f : X \to [0,\infty)$ and arbitrary P-random points $x \in X$ (which is their Theorem 10). These theorems are equivalent in the classical case, where f is required to be in $L^1(X,P)$, but *prima facie* the left-hand side adds further macroscopic laws.

[61] We mainly follow Hertling and Weihrauch (2003), §3, which also defines computability of P (whose definition we omit). There is an alternative formalism of computable metric spaces due to Hoyrup and Rojas (2009), which seems equivalent at least in the cases like $X = 2^{\mathbb{N}}$ etc. studied in this paper. We also replace the original definition of P-randomness based on Martin-Löf (1966) by an easy reformulation due to Solovay. See for example Theorem 6.37 in Calude (2002) or Theorem 6.2.8 in Downey and Hirschfeldt (2010).

(b) A sequence (V_n) *of opens* $V_n \in \mathscr{O}(X)$ *is* uniformly computable *if*

$$V_n = \bigcup_{m\in\mathbb{N}} B(g(n,m)), \tag{A.2}$$

for some (total) computable function $g : \mathbb{N}\times\mathbb{N} \to \mathbb{N}$.

2. *A* computable probability space (X,B,P) *is a computable topological space* X *with a computable probability measure* P.

3. *A (Solovay)* test *in such a space is a uniformly computable sequence of opens* (V_n) *for which*

$$\sum_n P(V_n) < \infty. \tag{A.3}$$

4. *A point* $x \in X$ *is* P-random *if for each test* (V_n) *one has* $x \in V_n$ *for only finitely many* N.

A key example is $X = 2^{\mathbb{N}}$ with $2 = \{0,1\}$. Its elements are maps $s : \mathbb{N} \to \{0,1\}$, and the "right" topology $\mathscr{O}(2^{\mathbb{N}})$ on this set (for our purposes) is generated by the so-called *cylinder sets*

$$[\sigma]_N = \{s \in 2^{\mathbb{N}} \mid s_{|N} = \sigma\}, \tag{A.4}$$

where $N \in \mathbb{N}$ and $\sigma \in 2^N$, and $s_{|N} \in 2^N$ is the restriction of s to $\{1,\ldots,N\}$. These cylinder sets, which by definition form a base $\mathscr{B}$ of the topology $\mathscr{O}(2^{\mathbb{N}})$, form a countable subset of the power set of $2^{\mathbb{N}}$ which may be numbered for example via the lexicographical order $L : \mathbb{N} \xrightarrow{\cong} 2^*$, where

$$2^* = \cup_{N\in\mathbb{N}} 2^N \tag{A.5}$$

is the set of all finite binary strings. This numbering, in turn, gives a bijection

$$B : \mathbb{N} \xrightarrow{\cong} \mathscr{B}; \qquad\qquad n \mapsto L(n)2^{\mathbb{N}}, \tag{A.6}$$

which may be used to turn $2^{\mathbb{N}}$ into a computable space. In the usual approach to probability theory on $2^{\mathbb{N}}$ the cylinder sets likewise generate the standard σ-algebra $\mathscr{F}$ on which probability measures may be defined; $\mathscr{F}$ is also the smallest σ-algebra that makes all evaluation maps

$$X_n : 2^{\mathbb{N}} \to \{0,1\}; \qquad\qquad X_n(s) = s_n, \tag{A.7}$$

measurable, where $s \in 2^{\mathbb{N}}$ and $n \in \mathbb{N}$. By the Carathéodory extension theorem, any probability measure on $\mathscr{F}$ is determined by its values on $\mathscr{B}$. For example, some prior q on $\{0,1\}$ (specified by $q(1) \in [0,1]$ so that $q(0) = 1-q(1)$), such as the uniform (unbiased) probability $q = f$, i.e.,

$$f(0) = f(1) = 1/2, \tag{A.8}$$

induces the corresponding Bernoulli probability measure $q^{\mathbb{N}}$ on $2^{\mathbb{N}}$ via

$$q^{\mathbb{N}}([\sigma]_N) = q^N(\sigma) = \prod_{n=1}^{N} q(\sigma_n). \tag{A.9}$$

In particular, $f^{\mathbb{N}}([\sigma]_N) = 2^{-N}$ for any $\sigma \in 2^N$. Then $(2^{\mathbb{N}},B,q^{\mathbb{N}})$ is computable iff q is computable (as a real). This is all we need (including at times the slight variation of using $\mathbb{Z}$ instead of $\mathbb{N}$).

Our unusual definition of P-randomness (chosen because it greatly facilitates our arguments) perhaps needs some clarification. Kolmogorov's original concept of randomness of *finite* binary

strings σ was quite intuitive,[62] namely that σ is random iff it is incompressible, i.e., σ is random if its shortest description is σ itself.[63] Here the vague concept of a "description" is made precise using computability theory; roughly speaking, σ is random if the shortest computer program that outputs σ is at least as long as σ itself. More technically, the *prefix Kolmogorov complexity* $K_T(\sigma)$ of $\sigma \in A^*$ is defined as the length of the shortest program running on some universal prefix Turing machine T that outputs σ and then halts:[64] fix some universal prefix Turing machine T, and define

$$K_T(\sigma) := \min\{|x| : x \in 2^*, T(x) = \sigma\}, \tag{A.10}$$

i.e., the length of the shortest program running on T that outputs σ. Different choices of T give different numbers $K_T(\sigma)$ only up to a σ-independent constant; reflecting this arbitrariness we call $\sigma \in A^*$ *c-Kolmogorov random* (relative to T), where $c \in \mathbb{N}$ is some σ-independent constant, iff

$$K_T(\sigma) \geq |\sigma| - c. \tag{A.11}$$

On the other hand, being defined as a minimum, for any string σ the number $K_T(\sigma)$ can't be much longer than $|\sigma|$, so that roughly speaking, for long strings (such that c is insignificant relative to $|\sigma|$) σ is (Kolmogorov) random iff $K(\sigma) \approx |\sigma|$, omitting the reference to T. One justification of Definition A.1.4, then, is a theorem due to Schnorr stating that an infinite binary sequence $s \in 2^{\mathbb{N}}$ is $f^{\mathbb{N}}$-random iff each finite initial segment $s_{|N} \in 2^N$ is uniformly Kolmogorov random,[65] in the sense that there is a constant $c \in \mathbb{N}$ (which may depend on s but not on N) such that for all N,

$$K(s_{|N}) \geq N - c. \tag{A.12}$$

There is also a version of this result for for arbitrary computable measures P on $2^{\mathbb{N}}$, viz.

Theorem A.2 (Levin, Gács) *Let P be a computable probability measure on $2^{\mathbb{N}}$. Then $s \in 2^{\mathbb{N}}$ is P-random iff there is a constant $c \in \mathbb{N}$ such that for all N,*

$$K(s_{|N}) \geq -\log_2(P([s_{|N}])) - c. \tag{A.13}$$

If $P = f^{\mathbb{N}}$, then $P([s_{|N}]) = 2^{-N}$, and we recover Schnorr's theorem just quoted. See Gács (1979).

On top of this justification via randomness of finite sequences, Definition A.1.4 may also be motivated on its own terms, namely via examples of the kind (1.5), where $\varepsilon = 1/m$ with $m \in \mathbb{N}$ to make things computable. The idea, originally due to P. Martin-Löf (1966), is that each member V_N of such a test (V_N) detects some pattern in some finite part of a sequence $s \in 2^{\mathbb{N}}$, which occurs just in case that $s \in V_N$. For example, in case of (1.5) this pattern gives $s_{|N}$ an unusually large number of either 0's or 1's (relative to the given probability f^N). If this pattern perseveres, that is, if $s \in V_N$ for an infinite number of V_N, then s cannot be random. Hence also this motivation relies on the analysis of finite initial segments of s (this aspect may not immediately be obvious from Definition A.1.4, but unfolding the computability requirement would make it clear).[66]

[62] The original references are Kolmogorov (1965, 1968). Li and Vitányi (2008) is the standard textbook in the field.

[63] See van Lambalgen (1987), Porter (2012), Eagle (2019), and Landsman (2020) for further history and analysis.

[64] That is, the domain $D(T)$ of T consists of a prefix subset of 2^*, so if $x \in D(T)$ then $y \notin D(T)$ whenever $x \prec y$.

[65] See Downey and Hirschfeldt, Theorems 6.2.3 and 6.2.8.

[66] See Li and Vitányi (2008), §2.5.2, for further information. Calude (2002), §6.2, shows that for $X = 2^{\mathbb{N}}$ (and more generally for $A^{\mathbb{N}}$ with finite A) one may assume that V_N is a cylinder set, i.e., it is essentially defined in 2^*.

B Time reversal

In deterministic theories defined over $\mathbb{R}$ as the time axis, with state space $\mathscr{S}$, time evolution is given by an $\mathbb{R}$-action on $\mathscr{S}$, providing trajectories $x \mapsto x(t) \equiv \varphi_t(x)$. Following Roberts (2022), we say that this dynamics is *time symmetric* (or *time-reversal invariant*) if the following is true:

- There is an invertible *instantaneous time-reversal map* $T : \mathscr{S} \to \mathscr{S}$ with $T^{-1} = T$, such that

$$\varphi_t \circ T = T \circ \varphi_{-t}. \qquad (t \in \mathbb{R}). \tag{B.1}$$

In that case, define the time reversed path $x^R(t)$ of $x(t)$, running from $x_0 = x(0)$ to $x_f = x(\tau)$, by

$$x^R(t) := T\varphi_{-t}(x_f), \tag{B.2}$$

which runs from $x^R(0) \equiv x_0^R = Tx_f$ to $x^R(\tau) = Tx_0$. The reason for time symmetry, then, is that because of (B.1) the path $x^R(t)$ is also given by $x^R(t) = \varphi_t(x_0^R)$, so that, being given by φ_t, it satisfies the equations of motion, too (which are supposed to be implicit in specifying the dynamics φ_t).

For example, in Newtonian physics one has $x = (q, v)$ and T is the velocity inversion map

$$T(q, v) = (q, -v). \tag{B.3}$$

In quantum mechanics on $\mathbb{R}^d$ one has $\psi \in L^2(\mathbb{R}^d)$ and T is famously complex conjugation, i.e.,

$$T\psi = \overline{\psi}. \tag{B.4}$$

This also works if $t \geq 0$, in which case we constrain t in (B.2) to $t \in [0, \tau]$, and even if $t \in \mathbb{N}$.

In the non-deterministic case we restrict ourselves to time-homogenous Markov chains $(X(t))_{t\in\mathbb{N}}$ with state space $\mathscr{S}$ (and time $t \in \mathbb{N}$). First, note that the concept of a Markov chain (or process) is by itself independent of the direction of time:[67] this is because the defining Markov condition

$$\mathbb{P}(X(t) = x \mid X(t_1) = x_1, \ldots, X(t_n) = x_n) = \mathbb{P}(X(t) = x \mid X(t_1) = x_1), \tag{B.5}$$

for all n, $i = 1, \ldots, n$, all x_i, and t_i such that $t > t_1 > \cdots > t_n$, is equivalent to the same condition for $t < t_1 < \cdots < t_n$. Trajectories, then, are sample paths $(X(0) = x_0, \ldots, X(\tau) = x_f)$, whose time reversal according to (B.2) is $(X(0) = Tx_f, \ldots, X(\tau) = Tx_0)$. The standard definition of time symmetry of a Markov chain is to ask that, for all times $(t_0, \ldots, \tau)$ and states $(x_0, \ldots, x_f)$,

$$\mathbb{P}(X(0) = Tx_f, \ldots, X(\tau) = Tx_0) = \mathbb{P}(X(0) = x_0, \ldots, X(\tau) = x_f). \tag{B.6}$$

This is easily shown to be equivalent to the chain satisfying a twisted detailed balance condition

$$\mu_{Ty} P_{TyTx} = \mu_x P_{xy}, \tag{B.7}$$

where $P_{xy} = \mathbb{P}(X(1) = y \mid X(0) = x)$ and $\mu_x = \mathbb{P}(X(0) = x)$ as usual. If $T = \mathrm{id}$, this simply reads

$$\mu_y P_{yx} = \mu_x P_{xy}. \tag{B.8}$$

Summing over x gives $\mu = \mu P$ (since P is a stochastic matrix), so that μ is a stationary distribution. Hence (B.6) forces the distribution $\mu(t)$ of $X(t)$ to be stationary and hence equal to the distribution $\mu \equiv \mu(0)$ of $X(0)$. For example, eq. (B.8) holds for both the microscopic and the macroscopic Ehrenfest models, where $T = \mathrm{id}$, and indeed each has a unique stationary distribution given by (3.4) and (3.6), respectively.[68] Hence condition (B.6) is satisfied in both versions of the model.

[67] See e.g. Brémaud (2020), Remark 2.1.5.

[68] In both cases the stationary distribution is unique because the models are irreducible as Markov chains.

References

[1] Ardourel, V. (2017). Irreversibility in the derivation of the Boltzmann equation. *Foundations of Physics* 47, 471–489. https://doi.org/10.1007/s10701-017-0072-9.

[2] Badino, M. (2011). Mechanistic slumber vs. statistical insomnia: the early history of Boltzmann's H-theorem (1868–1877). *The European Physical Journal H* 36, 353–378. https://doi.org/10.1140/epjh/e2011-10048-5.

[3] Baldovin, M., Caprini, L., Vulpiani, A. (2019). Irreversibility and typicality: A simple analytical result for the Ehrenfest model. *Physica A* 524, 422–429. https://doi.org/10.1016/j.physa.2019.04.188.

[4] Bienvenu, L., Day, A.R., Hoyrup, M., Mezhirov, I., Shen, A. (2012). A constructive version of Birkhoff's ergodic theorem for Martin-Löf random points. *Information and Computation* 210, 21–30. https://doi.org/10.1016/j.ic.2011.10.006.

[5] Bodineau, T., Gallagher, I., Saint-Raymond, L., Simonella, S. (2023). On the dynamics of dilute gases. *European Mathematical Society Magazine* 128, 13–22. https://doi.org/10.4171/MAG/124.

[6] Boltzmann, L. (1872). Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen. *Wiener Berichte* 66, 275–370. Reprinted in Boltzmann, L., *Wissenschaftliche Abhandlungen, Vol. I*, paper 23, ed. F. Hasenöhrl (Chelsea, 1969). English translation: S. Brush, *The Kinetic Theory of Gases: An Anthology of Classic Papers with Historical Commentary*, pp. 262–349 (Imperial College Press, 2003).

[7] Boltzmann, L. (1877). Über die Beziehung dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respektive den Sätzen über das Wärmegleichgewicht. *Wiener Berichte* 76, 373–435. Reprinted in Boltzmann, L., *Wissenschaftliche Abhandlungen, Vol. II*, paper 39, ed. F. Hasenöhrl (Chelsea, 1969). English translation: *Entropy* 17, 1971–2009 (2015). https://www.mdpi.com/1099-4300/17/4/1971.

[8] Brémaud, P. (2020). *Markov Chains: Gibbs Fields, Monte Carlo Simulation and Queues* (Springer).

[9] Bricmont, J. (2020). Probabilistic explanations and the derivation of macroscopic laws. *Statistical Mechanics and Scientific Explanation: Determinism, Indeterminism, and Laws of Nature*, ed. Allori, V., pp. 31–64 (World Scientific).

[10] Bricmont, J. (2022). *Making Sense of Statistical Mechanics* (Springer).

[11] Brown, H. R., Myrvold, W., Uffink, J. (2009). Boltzmann's H-theorem, its discontents, and the birth of statistical mechanics. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 40, 174–191. https://doi.org/10.1016/j.shpsb.2009.03.003.

[12] Brush, S.G. (1974). The development of the kinetic theory of gases: VIII. Randomness and irreversibility. *Archive for History of Exact Sciences* 12, 1–88. https://www.jstor.org/stable/41133384.

[13] Calude, C.S. (2002). *Information and Randomness: An Algorithmic Perspective*, 2nd Edition (Springer).

[14] Cercignani, C., Illner, R., Pulvirenti, M. (2013). *The Mathematical Theory of Dilute Gases* (Springer).

[15] Cifarelli, D.M., Muliere, P. (2002). Exchangeable sequences of random variables. https://www.muliereconference.unibocconi.eu/docs/sezione6/6_5_Exchangeable.pdf.

[16] Darrigol, O. (2018). Atoms, Mechanics, and Probability. Ludwig Boltzmann's Statistico-Mechanical Writings: An Exegesis (Oxford University Press).

[17] De Bièvre, S., Parris, P.E. (2017). A rigourous demonstration of the validity of Boltzmann's scenario for the spatial homogenization of a freely expanding gas and the equilibration of the Kac ring. *Journal of Statistical Physics* 168, 772–793. https://doi.org/10.1007/s10955-017-1834-7.

[18] Dembo, A., Zeitouni, A. (1998). *Large Deviations: Techniques and Applications. Second Edition* (Springer).

[19] Deng, Y., Hani, Z., Ma, X. (2024). Long time derivation of the Boltzmann equation from hard sphere dynamics. https://arxiv.org/pdf/2408.07818.

[20] Deng, Y., Hani, Z., Ma, X. (2025). Hilbert's sixth problem: Derivation of fluid equations via Boltzmann's kinetic theory. https://arxiv.org/pdf/2503.01800.

[21] Downey, R., Hirschfeldt, D.R. (2010). *Algorithmic Randomness and Complexity* (Springer).

[22] Dudley, R.M. (1989). *Real Analysis and Probability* (Wadsworth and Brooks/Cole).

[23] Eagle, A. (2019). Chance versus randomness. *The Stanford Encyclopedia of Philosophy (Spring 2019 Edition)*, ed. Zalta, E.N. https://plato.stanford.edu/archives/spr2019/entries/chance-randomness/.

[24] Ehrenfest, P., Ehrenfest, T. (1907). Uber zwei bekannte Einwände gegen das Boltzmannsche H-theorem. *Physikalische Zeitschrift* 8, 311–314. https://www.lorentz.leidenuniv.nl/IL-publications/sources/Ehrenfest_07b.pdf.

[25] Ehrenfest, P., Ehrenfest, T. (1911). Begriffliche Grundlagen der statistischen Auffassung in der Mechanik. *Encyklopädie der mathematischen Wissenschaften*, eds. Klein, F., Müller, C., Vol. IV-32, pp. 1–90 (Vieweg–Teubner Verlag). https://www.lorentz.leidenuniv.nl/IL-publications/sources/Ehrenfest_1911b.pdf. Translation: *The Conceptual Foundations of the Statistical Approach in Mechanics* (Cornell University Press, 1959 and Dover, 1990).

[26] Ellis, R.S. (1985). *Entropy, Large Deviations, and Statistical Mechanics* (Springer).

[27] Ellis, R.S. (1995). An overview of the theory of large deviations and applications to statistical mechanics. *Scandinavian Actuarial Journal* 1, 97–142. https://doi.org/10.1080/03461238.1995.10413952.

[28] Ellis, R.S. (1999). The theory of large deviations: From Boltzmann's 1877 calculation to equilibrium macrostates in 2D turbulence. *Physica D* 133, 106–136. https://doi.org/10.1016/S0167-2789(99)00101-3.

[29] Franklin, J., Greenberg, N., Miller, J., Ng, K. M. (2012). Martin-Löf random points satisfy Birkhoff's ergodic theorem for effectively closed sets. *Proceedings of the American Mathematical Society* 140, 3623–3628. https://doi.org/10.1090/S0002-9939-2012-11179-7.

[30] Franklin, J.Y. , Porter, C.P., eds. (2020). *Algorithmic Randomness: Progress and Prospects* (Cambridge University Press).

[31] Gács, P. (1979). Exact expressions for some randomness tests. *Theoretical Computer Science 4th GI Conference*, ed. Weihrauch, K., *Lecture Notes in Computer Science* 67, 124–131 (Springer).

[32] Galatolo, S., Hoyrup, M., Rojas, C. (2010). Effective symbolic dynamics, random points, statistical behavior, complexity and entropy. *Information and Computation* 208, 23–41. https://doi.org/10.1016/j.ic.2009.05.001.

[33] Gallagher, I., Saint-Raymond, L., Texier, B. (2014). *From Newton to Boltzmann: Hard Spheres and Short-range Potentials* (European Mathematical Society). https://arxiv.org/abs/1208.5753.

[34] Grad, H. (1958). Principles of the kinetic theory of gases. *Thermodynamik der Gase/Thermodynamics of Gases*, ed. Flügge, S., pp. 205–294 (Springer).

[35] Hauert, C., Nagler, J., Schuster, H. G. (2004). Of dogs and fleas: the dynamics of N uncoupled two-state systems. *Journal of Statistical Physics* 116, 1453–1469. https://doi.org/10.1023/B:JOSS.0000041725.70622.c4.

[36] Hertling, P., Weihrauch, K. (2003). Random elements in effective topological spaces with measure. *Information and Computation* 181, 32–56. https://doi.org/10.1016/S0890-5401(02)00034-2.

[37] Hilbert, D. (1900/2000). Mathematische Probleme. *Nachrichten von der Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-Physikalische Klasse* 1900, 253–297. http://eudml.org/doc/58479. Translation: Mathematical problems, *Bulletin of the American Mathematical Society* 8, 437–479 (1902). Reprinted in *Bulletin of the American Mathematical Society (New Series)* 37, 407–436 (2000). https://doi.org/10.1090/S0002-9904-1902-00923-3.

[38] Hiura, K., Sasa, S. (2019). Microscopic reversibility and macroscopic irreversibility: from the viewpoint of algorithmic randomness. *Journal of Statistical Physics* 177, 727–751. https://doi.org/10.1007/s10955-019-02387-0.

[39] Hoyrup, M., Rojas, C. (2009). Computability of probability measures and Martin-Löf randomness over metric spaces. *Information and Computation* 207, 830–847. https://doi.org/10.1016/j.ic.2008.12.009.

[40] Jebeile, J. (2020). The Kac Ring or the art of making idealisations. *Foundations of Physics* 50, 1152–1170. https://doi.org/10.1007/s10701-020-00373-1.

[41] Johnson, N.L., Kotz, S. (1977). *Urn Models and their Applications: An Approach to Modern Discrete Probability Theory* (Wiley).

[42] Kac, M. (1959). *Probability and Related Topics in Physical Sciences* (Wiley).

[43] Klenke, A. (2020). *Probability Theory: A Comprehensive Course* (Springer).

[44] Kolmogorov, A.N. (1965). Three approaches to the quantitative definition of information. *Problems of Information Transmission* 1, 3–11. http://alexander.shen.free.fr/library/Kolmogorov65_Three-Approaches-to-Information.pdf.

[45] Kolmogorov, A.N. (1968). Logical Basis for information theory and probability theory. *IEEE Transactions on Information Theory* 14, 662–664. http://dx.doi.org/10.1109/TIT.1968.1054210.

[46] Landsman, K. (2020). Randomness? What randomness?, *Foundations of Physics* 50, 61–104. https://doi.org/10.1007/s10701-020-00318-8.

[47] Landsman, K. (2021). Indeterminism and undecidability. *Undecidability, Uncomputability, and Unpredictability*, eds. Aguirre, A., Merali, Z., Sloan, D., pp. 17–46 (Springer). https://arxiv.org/pdf/2003.03554.pdf.

[48] Landsman, K. (2024). *Entropy and Large Deviations. Lecture Notes.* https://www.math.ru.nl/~landsman/Entropy-2024-final-na.pdf.

[49] Lanford, O.E. (1975). Time evolution of large classical systems. *Dynamical Systems, Theory and Applications. Lecture Notes in Theoretical Physics, vol. 38*, ed. Moser, J., pp. 1–111 (Springer).

[50] Lanford, O.E. (1976). On the derivation of the Boltzmann equation. *Astérisque* 40, 117–137. https://www.numdam.org/article/AST_1976__40__117_0.pdf.

[51] Lazarovici, D., Reichert, P. (2015). Typicality, irreversibility and the status of macroscopic laws. *Erkenntnis* 80, 689–716. https://doi.org/10.1007/s10670-014-9668-z.

[52] Lebowitz, J.L. (1999). Statistical mechanics: A selective review of two central issues. *Reviews of Modern Physics* 71, S346–S357. https://doi.org/10.1103/RevModPhys.71.S346.

[53] Li, M., Vitányi, P.M.B. (2008). *An Introduction to Kolmogorov Complexity and Its Applications. Third Edition* (Springer).

[54] Mackey, G. W. (1974). Ergodic theory and its significance for statistical mechanics and probability theory. *Advances in Mathematics* 12, 178–268. https://doi.org/10.1016/S0001-8708(74)80003-4.

[55] Mackey, M.C. (1992). *Time's Arrow: The Origins of Thermodynamic Behavior* (Springer).

[56] Maes, C., Netočný, K., Shergelashvili, B. (2007). A selection of nonequilibrium Issues. *Lecture Notes in Mathematics* 1970, 247–306. https://doi.org/10.1007/978-3-540-92796-9_6.

[57] Martin-Löf, A. (1979). Statistical mechanics and the foundations of thermodynamics. *Lecture Notes in Physics* 101, 1–120 (Springer). https://doi.org/10.1007/BFb0119604.

[58] Martin-Löf, P. (1966). The definition of random sequences. *Information and Control* 9, 602–619. https://doi.org/10.1016/S0019-9958(66)80018-9.

[59] Porter, C.P. (2012). *Mathematical and Philosophical Perspectives on Algorithmic Randomness*. PhD Thesis, University of Notre Dame. https://www.cpporter.com/wp-content/uploads/2013/08/PorterDissertation.pdf.

[60] Price, H. (1996). *Time's Arrow and Archimedes' Point: New Directions for the Physics of Time* (Oxford University Press).

[61] Raatikainen, P. (1998). On interpreting Chaitin's incompleteness theorem, *Journal of Philosophical Logic* 27, 569–586. https://doi.org/10.1023/A:1004305315546.

[62] Rassoul-Agha, F., Seppäläinen, T. (2015). *A Course on Large Deviations with an Introduction to Gibbs Measures* (AMS).

[63] Roberts, B.W. (2022). *Reversing the Arrow of Time* (Cambridge University Press).

[64] Sklar, L. (1993). *Physics and Chance: Philosophical Issues in the Foundations of Statistical Mechanics* (Cambridge University Press).

[65] Sznitman, A. S. (1991). Topics in propagation of chaos. *Lecture Notes in Mathematics* 1464, 165–251. Ecole d'été de probabilités de Saint-Flour XIX—1989 (pp. 165-251). Berlin, Heidelberg: Springer Berlin Heidelberg.

[66] Spohn, H. (2012). *Large Scale Dynamics of Interacting Particles* (Springer).

[67] Touchette, H. (2009). The large deviation approach to statistical mechanics. *Physics Reports* 478, 1–69. https://doi.org/10.1016/j.physrep.2009.05.002.

[68] Towsner, H. (2020). Algorithmic randomness in ergodic theory. Franklin and Porter, pp. 40–57.

[69] Uffink, J. (2001). Bluff your way in the second law of thermodynamics. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics* 32, 305–394. https://doi.org/10.1016/S1355-2198(01)00016-8.

[70] Uffink, J. (2007). Compendium of the foundations of classical statistical physics. *Handbook of the Philosophy of Science. Vol. 2: Philosophy of Physics, Part B*, eds. Butterfield, J., Earman, J., pp. 923–1074 (North-Holland).

[71] Uffink, J. (2022). Boltzmann's work in statistical physics. *The Stanford Encyclopedia of Philosophy (Summer 2022 Edition)*, ed. Zalta, E.N. https://plato.stanford.edu/archives/sum2022/entries/statphys-Boltzmann/.

[72] Uffink, J., Valente, G. (2015). Lanford's Theorem and the emergence of irreversibility. *Foundations of Physics* 45, 404–438. https://doi.org/10.1007/s10701-015-9871-z.

[73] van Lambalgen, M. (1987). *Random Sequences*. PhD Thesis, University of Amsterdam. https://www.academia.edu/23899015/RANDOM_SEQUENCES.

[74] van Strien, M. (2013). The nineteenth century conflict between mechanism and irreversibility. *Studies in History and Philosophy of Modern Physics* 44, 191–205. https://doi.org/10.1016/j.shpsb.2013.05.004.

[75] Viana, M., Oliveira, K. (2016). *Foundations of Ergodic Theory* (Cambridge University Press).

[76] Villani, C. (2002). A review of mathematical topics in collisional kinetic theory. *Handbook of Mathematical Fluid Dynamics, Volume 1*, eds. Friedlander, S., Serre, D., pp. 71–306 (Elsevier).

[77] Villani, C. (2013). (Ir)reversibility and Entropy. *Time. Progress in Mathematical Physics, Vol. 63*, ed. Duplantier, B., pp. 19–79 (Birkhäuser).

[78] V'yugin, V. (1997). Effective convergence in probability and an ergodic theorem for individual random sequences. *SIAM Theory of Probability and Its Applications* 42, 39–50. https://doi.org/10.1137/S0040585X97975915.

[79] V'yugin, V. (2022). Ergodic theorems for algorithmically random points. https://arxiv.org/pdf/2202.13465.pdf.

[80] Zuchowski, L. (2024). *From Randomness and Entropy to the Arrow of Time* (Cambridge University Press).