5

ENTROPY

A GUIDE FOR THE PERPLEXED

Roman Frigg and Charlotte Werndl[1]

1 Introduction

Entropy is ubiquitous in physics, and it plays important roles in numerous other disciplines ranging from logic and statistics to biology and economics. However, a closer look reveals a complicated picture: entropy is defined differently in different contexts, and even within the same domain different notions of entropy are at work. Some of these are defined in terms of probabilities, others are not. The aim of this essay is to arrive at an understanding of some of the most important notions of entropy and to clarify the relations between them. In particular, we discuss the question what kind of probabilities are involved whenever entropy is defined in terms of probabilities: are the probabilities chances (i.e. physical probabilities) or credences (i.e. degrees of belief)?

After setting the stage by introducing the thermodynamic entropy (Sec. 2), we discuss notions of entropy in information theory (Sec. 3), statistical mechanics (Sec. 4), dynamical-systems theory (Sec. 5), and fractal geometry (Sec. 6). Omissions are inevitable; in particular, space constraints prevent us from discussing entropy in quantum mechanics and cosmology.[2]

2 Entropy in thermodynamics

Entropy made its first appearance in the middle of the nineteenth century in the context of thermodynamics (TD). TD describes processes like the exchange of heat between two bodies or the spreading of gases in terms of macroscopic variables like temperature, pressure, and volume. The centerpiece of TD is the Second Law of TD, which, roughly speaking, restricts the class of physically allowable processes in isolated systems to those that are not entropy-decreasing. In this section we introduce the TD entropy and the Second Law.[3] We keep

[1] The authors are listed alphabetically; the paper is fully collaborative.

[2] Hemmo & Shenker 2006 and Sorkin 2005 provide good introductions to the quantum and cosmological entropies, respectively.

[3] Our presentation follows Pippard 1966, pp. 19–23, 29–37. There are also many different (and nonequivalent) formulations of the Second Law (see Uffink 2001).

C. Beisbart & S. Hartmann (eds), *Probabilities in Physics*, Oxford University Press 2011, 115–42

this presentation short because the TD entropy is not a probabilistic notion and therefore falls, strictly speaking, outside the scope of this book.

The thermodynamic state of a system is characterized by the values of its thermodynamic variables; a state is an *equilibrium state* if, and only if (iff), all variables have well-defined and constant values. For instance, the state of a gas is specified by the values of temperature, pressure, and volume, and the gas is in equilibrium if these have well-defined values which do not change over time. Consider two equilibrium states A and B. A process that changes the state of the system from A to B is *quasistatic* iff it only passes through equilibrium states (i.e. if all intermediate states between A and B are also equilibrium states). A process is *reversible* iff it can be exactly reversed by an infinitesimal change in the external conditions. If we consider a cyclical process—a process in which the beginning and the end state are the same—a reversible process leaves the system *and its surroundings* unchanged.

The Second Law (in Kelvin's formulation) says that it is impossible to devise an engine which, working in a cycle, produces no effect other than the extraction of heat from a reservoir and the performance of an equal amount of mechanical work. It can be shown that this formulation implies that

$$\oint \frac{dQ}{T} \leq 0, \tag{1}$$

where dQ is the amount of heat put into the system and T is the system's temperature. This is known as *Clausius' Inequality*.

If the cycle is reversible, then the inequality becomes an equality. Trivially, this implies that for reversible cycles,

$$\int_A^B \frac{dQ}{T} = -\int_B^A \frac{dQ}{T} \tag{2}$$

for *any* paths from A to B and from B to A, and the value of the integrals only depends on the beginning and the end point.

We are now in a position to introduce the thermodynamic entropy S_{TD}. The leading idea is that the integral in Eqn (2) gives the entropy difference between A and B. We can then assign an absolute entropy value to every state of the system by choosing one particular state A (we can choose any state we please!) as the reference point, choosing a value for its entropy $S_{\text{TD}}(A)$, and then defining the entropy of all other points B by

$$S_{\text{TD}}(B) := S_{\text{TD}}(A) + \int_A^B \frac{dQ}{T}, \tag{3}$$

where the change of state from A to B is *reversible*.

What follows from these considerations about irreversible changes? Consider the following scenario: we first change the state of the system from A to B along a quasistatic irreversible path, and then go back from B to A along a quasistatic reversible path. It follows from Eqns (1) and (3) that

$$S_{\mathrm{TD}}(B) - S_{\mathrm{TD}}(A) \ \leq \ \int_A^B \frac{dQ}{T}. \tag{4}$$

If we now restrict attention to *adiathermal* processes (i.e. ones in which temperature is constant), the integral in Eqn (4) becomes zero and we have

$$S_{\mathrm{TD}}(B) \ \leq \ S_{\mathrm{TD}}(A).$$

This is often referred to as the Second Law, but it is important to point out that it is only a special version of it which holds for adiathermal processes.

S_{TD} has no intuitive interpretation as a measure of disorder, disorganization, or randomness (as is often claimed). In fact such considerations have no place in TD.

We now turn to a discussion of the information-theoretic entropy, which, unlike the S_{TD}, is a probabilistic concept. At first sight the information-theoretic and the thermodynamic entropy have nothing to do with each other. This impression will be dissolved in Sec. 4, when a connection is established via the Gibbs entropy.

3 Information theory

Consider the following situation (Shannon 1949). There is a source producing messages which are communicated to a receiver. The receiver registers them, for instance, on a paper tape.[4] The messages are discrete and sent by the source one after the other. Let $m = \{m_1, \ldots, m_n\}$ be the complete set of messages (in the sense that the source cannot send messages other than the m_i). The production of one message is referred to as a *step*.

When receiving a message, we gain information, and depending on the message, more or less information. According to Shannon's theory, information and uncertainty are two sides of the same coin: the more uncertainty there is, the more information we gain by removing the uncertainty. The literature's usage oscillates between 'information' and 'uncertainty,' and so will we.

Shannon's basic idea was to characterize the amount of information gained from the receipt of a message as a function which depends only on how likely the messages are. Formally, for $n \in \mathbb{N}$ let V_m be the set of all probability distributions $P = (p_1, \ldots, p_n) := \big(p(m_1), \ldots, p(m_n)\big)$ on $m_1, \ldots, m_n$ (i.e. $p_i \geq 0$ and $p_1 + \cdots + p_n = 1$). A reasonable measure of information is a function $S_{\mathrm{S,\,d}}(P)\colon V_m \to \mathbb{R}$ (where 'S' is for 'Shannon' and 'd' for 'discrete') which satisfies the following axioms (cf. Klir 2006, Sec. 3.2.2):

[4] We assume that the channel is noiseless and deterministic, meaning that there is a one-to-one correspondence between the input and the output.

1. *Continuity*. $S_{\mathrm{S,d}}(p_1, \ldots, p_n)$ is continuous in all its arguments $p_1, \ldots, p_n$.
2. *Additivity*. The information gained in two independent communications is the sum of the information of the communications, i.e. for $P = (p_1, \ldots, p_n)$ and $Q = (q_1, \ldots, q_k)$, we have: $S_{\mathrm{S,d}}(p_1 q_1, p_1 q_2, \ldots, p_n q_k) = S_{\mathrm{S,d}}(P) + S_{\mathrm{S,d}}(Q)$.
3. *Monotonicity*. For uniform distributions the uncertainty increases with n. That is, for any $P = (1/n, \ldots, 1/n)$ and $Q = (1/k, \ldots, 1/k)$, for arbitrary $k, n \in \mathbb{N}$, we have: if $k > n$, then $S_{\mathrm{S,d}}(Q) > S_{\mathrm{S,d}}(P)$.
4. *Branching*. The information is independent of how the process is divided into parts. That is, for $(p_1, \ldots, p_n)$ ($n \geq 3$), divide $m = \{m_1, \ldots, m_n\}$ into two blocks $A = (m_1, \ldots, m_s)$ and $B = (m_{s+1}, \ldots, m_n)$, and let $p_A = \sum_{i=1}^{s} p_i$ and $p_B = \sum_{i=s+1}^{n} p_i$. Then

$$\begin{aligned} S_{\mathrm{S,d}}(p_1, \ldots, p_n) \; &= \\ & S_{\mathrm{S,d}}(p_A, p_B) + p_A\, S_{\mathrm{S,d}}\left(\frac{p_1}{p_A}, \ldots, \frac{p_s}{p_A}\right) + p_B\, S_{\mathrm{S,d}}\left(\frac{p_{s+1}}{p_B}, \ldots, \frac{p_n}{p_B}\right).\end{aligned}$$ [5]

5. *Bit normalization*. By convention, the average information gained for two equally likely messages is one *bit* ('binary digit'): $S_{\mathrm{S,d}}(1/2, 1/2) = 1$.

There is exactly one function satisfying these axioms, the *discrete Shannon entropy*:[6]

$$S_{\mathrm{S,d}}(P) := -\sum_{i=1}^{n} p_i \log p_i,$$

where 'log' stands for the logarithm to the base of two.[7] Any change toward equalization of $p_1, \ldots, p_n$ leads to an increase of the uncertainty $S_{\mathrm{S,d}}$, which reaches its maximum, $\log n$, for $p_1 = \cdots = p_n = 1/n$. Furthermore, $S_{\mathrm{S,d}}(P) = 0$ iff all p_i but one equal zero.

What kind of probabilities are invoked in Shannon's scenario? Approaches to probability can be divided into two broad groups.[8] First, epistemic approaches take probabilities to be measures of degrees of belief. Those who subscribe to an objectivist epistemic theory take probabilities to be degrees of rational belief, whereby 'rational' is understood to imply that given the same evidence, all rational agents have the same degree of belief in any proposition. This is denied by those who hold a subjective epistemic theory, regarding probabilities as subjective degrees of belief that can differ between persons even if they are presented with the same body of evidence. Second, ontic approaches take probabilities

[5] For instance, for $\{m_1, m_2, m_3\}$, $P = (1/3, 1/3, 1/3)$, $A = \{m_1, m_2\}$, and $B = \{m_3\}$, branching means that $S_{\mathrm{S,d}}(1/3, 1/3, 1/3) = S_{\mathrm{S,d}}(2/3, 1/3) + 2/3\, S_{\mathrm{S,d}}(1/2, 1/2) + 1/3\, S_{\mathrm{S,d}}(1)$.

[6] There are other axioms that uniquely characterize the Shannon entropy (cf. Klir 2006, Sec. 3.2.2).

[7] We set $x \log x := 0$ for $x = 0$.

[8] For a discussion of the different interpretations of probability, see, for instance, Howson 1995, Gillies 2000, and Mellor 2005.

to be part of the 'furniture of the world.' The two most prominent ontic approaches are frequentism and the propensity view. On the frequentist approach, probabilities are long-run frequencies of certain events. On the propensity view, probabilities are tendencies or dispositions inherent in objects or situations.

The emphasis in information theory is on the receiver's amount of uncertainty about the next incoming message. This suggests that the $p(m_i)$ should be interpreted as epistemic probabilities (credences). While correct as a first stab, a more nuanced picture emerges once we ask the question of how the values of the $p(m_i)$ are set. Depending on how we understand the nature of the source, we obtain two very different answers. If the source itself is not probabilistic, then the $p(m_i)$ express the beliefs—and nothing but the beliefs—of receivers. For proponents of subjective probabilities these probabilities express the individual beliefs of an agent, and beliefs may vary between different receivers. The Objectivists insist that all rational agents must come to the same value assignment. This can be achieved, for instance, by requiring that $S_{\mathrm{S,d}}(P)$ be maximal, which singles out a unique distribution. This method, now known as Jaynes' Maximum-Entropy Principle, plays a role in statistical mechanics and will be discussed later.

Alternatively, the source itself can be probabilistic. The probabilities associated with the source have to be ontic probabilities of one kind or another (frequencies, propensities, etc.). In this case agents are advised to use the Principal Principle—roughly, the rule that a rational agent's credence for a certain event to occur should be equal to the ontic probability (chance) of that event to occur.[9] In Shannon's setting this means that the $p(m_i)$ have to be equal to the source's ontic probability of producing the message m_i. If this connection is established, the information of a *channel* is a measure of an objective property of a source.

It is worth emphasizing that $S_{\mathrm{S,d}}(P)$ is a technical conception of information, which should not be taken as an analysis of the various senses of 'information' in ordinary discourse. In ordinary discourse, information is often equated with knowledge, propositional content, or meaning. Hence 'information' is a property of a single message. But information, as understood in information theory, is *not* concerned with individual messages and their content; its focus is on *all* messages a source could possibly send. What makes a single message informative is not its meaning but the fact that it has been selected from a set of possible messages.

Given the probability distributions $P_m = (p_{m_1}, \ldots, p_{m_n})$ on $\{m_1, \ldots, m_n\}$, $P_s = (p_{s_1}, \ldots, p_{s_l})$ on $\{s_1, \ldots, s_l\}$, and the joint probability distribution $(p_{m_1,s_1},$

[9]The Principal Principle has been introduced by Lewis (1980); for a recent discussion see Frigg & Hoefer 2010.

$p_{m_1,s_2}, \ldots, p_{m_n,s_l})$[10] on $\{m_1s_1, m_1s_2, \ldots, m_ns_l\}$, the *conditional Shannon entropy* is defined as

$$S_{\mathrm{S,d}}(P_m \mid P_s) \;:=\; \sum_{j=1}^{l} p_{s_j} \sum_{k=1}^{n} \frac{p_{m_k,s_j}}{p_{s_j}} \log \frac{p_{m_k,s_j}}{p_{s_j}}. \tag{5}$$

It measures the average information received from a message m_k, given that a message s_j has been received before.

The Shannon entropy can be generalized to the continuous case. Let $p(x)$ be a probability density. The *continuous Shannon entropy* is (where 'S' is for 'Shannon' and 'c' for 'continuous')

$$S_{\mathrm{S,c}}(p) \;=\; -\int_{\mathbb{R}} p(x) \log p(x)\, dx, \tag{6}$$

if the integral exists. The generalization of (6) to densities of n variables $x_1, \ldots, x_n$ is straightforward. If $p(x)$ is positive, except for a set of Lebesgue measure zero, exactly on the interval $[a, b]$ (a, $b \in \mathbb{R}$), then $S_{\mathrm{S,c}}$ reaches its maximum, $\log(b-a)$, for $p(x) = 1/(b-a)$ in $[a, b]$, and is zero elsewhere. Intuitively, every change towards equalization of $p(x)$ leads to an increase in entropy. For probability densities which are, except for a set of measure zero, positive everywhere on $\mathbb{R}$, the question of the maximum is more involved. If the standard deviation is held fixed at value σ, then $S_{\mathrm{S,c}}$ reaches its maximum for a Gaussian $p(x) = \left(1/\sqrt{2\pi}\,\sigma\right) \exp(-x^2/2\sigma^2)$, and the maximum value of the entropy is $\log\left(\sqrt{2\pi e}\,\sigma\right)$ (Ihara 1993, Sec. 3.1; Shannon & Weaver 1949, pp. 88–9).

There is an important difference between the discrete and the continuous Shannon entropy. In the discrete case, the value of the Shannon entropy is uniquely determined by the probability measure over the messages. In the continuous case the value depends on the coordinates we choose to describe the messages. Hence the continuous Shannon entropy cannot be regarded as measuring information, since an information measure should not depend on the way in which we describe a situation. But usually we are interested in entropy differences rather than in absolute values, and it turns out that entropy *differences* are coordinate-independent and the continuous Shannon entropy can be used to measure differences in information (Ihara 1993, pp. 18–20; Shannon & Weaver 1949, pp. 90–1).[11]

We now turn to two other notions of information-theoretic entropy, namely Hartley's entropy and Rényi's entropy. The former preceded Shannon's entropy; the latter is a generalization of Shannon's entropy. One of the first accounts of

[10] The outcomes m_i and s_j are not assumed to be independent.

[11] This coordinate-dependence reflects a deeper problem: the uncertainty reduced by receiving a message of a continuous distribution is infinite and hence not measured by $S_{\mathrm{S,c}}$. Yet by approximating a continuous distribution by discrete distributions, one obtains that $S_{\mathrm{S,c}}$ measures differences in information (Ihara 1993, p. 17).

information was introduced by Hartley (1928). Assume that $m := \{m_1, \ldots, m_n\}$ ($n \in \mathbb{N}$) represents mutually exclusive possible alternatives and that one of the alternatives is true, but we do not know which one. How can we measure the amount of information gained when knowing which of these n alternatives is true; or, equivalently, how can we measure the uncertainty associated with these n possibilities? Hartley postulated that any function $S_{\mathrm{H}}\colon \mathbb{N} \to \mathbb{R}^+$ answering this question has to satisfy the following axioms:

1. *Monotonicity*. The uncertainty increases with n, that is, $S_{\mathrm{H}}(n) \leq S_{\mathrm{H}}(n+1)$ for all $n \in \mathbb{N}$.
2. *Branching*. The measure of information is independent of how the process is divided into parts: $S_{\mathrm{H}}(n.m) = S_{\mathrm{H}}(n) + S_{\mathrm{H}}(m)$, where '$n.m$' means that there are n times m alternatives.
3. *Normalization*. By convention, $S_{\mathrm{H}}(2) = 1$.

Again, there is exactly one function satisfying these axioms, namely $S_{\mathrm{H}}(n) = \log n$ (Klir 2006, p. 26), which is now referred to as the *Hartley entropy*.

On the face of it this entropy is based solely on the concept of mutually exclusive alternatives, and it does not invoke probabilistic assumptions. However, views diverge on whether this is the full story. Those who deny this argue that the Hartley entropy implicitly assumes that all alternatives have *equal* weight. This amounts to assuming that they have equal probability, and hence the Hartley entropy is a special case of the Shannon entropy, namely the Shannon entropy for the uniform distribution. Proponents of the former view argue that Hartley's notion of alternatives does not presuppose probabilistic concepts and is therefore independent of Shannon's (cf. Klir 2006, pp. 25–30).

The Rényi entropies generalize the Shannon entropy. As with the Shannon entropy, assume a probability distribution $P = (p_1, \ldots, p_n)$ over $m = \{m_1, \ldots, m_n\}$. Require of a measure of information that it satisfies all the axioms of the Shannon entropy except for branching. Rényi argues that, unlike in the cases of the other axioms, it is unclear whether a measure of information should satisfy branching and hence whether branching should be on the list of axioms (Rényi 1961). If the outcomes of two independent events with respective probabilities p and q are observed, we want the total received information to be the sum of the two partial informations. This implies that the amount of information received for message m_i is $-\log p_i$ (Jizba & Arimitsu 2004). If a weighted arithmetic mean is taken over the $-\log p_i$, we obtain the Shannon entropy. Now, is it possible to take another mean such that the remaining axioms about information are satisfied? If so, these quantities are also possible measures of the average information received. The general definition of a mean over $-\log p_i$ weighted by p_i is that it is of the form $f^{-1}\left(\sum_{i=1}^{n} p_i f(-\log p_i)\right)$ where f is a continuous, strictly monotonic, and invertible function. For $f(x) = x$ we obtain the Shannon entropy.

There is only one alternative mean satisfying the axioms, namely $f(x) = 2^{(1-q)x}$ ($q \in (0,\infty),\ q \neq 1$). It corresponds to the *Rényi entropy of order q*:

$$S_{\mathrm{R},q}(P) \ := \ \frac{1}{1-q} \log \sum_{k=1}^{n} p_k^q.$$

The limit of the Rényi entropy for $q \to 1$ gives the Shannon entropy, i.e. $\lim_{q\to 1} S_{\mathrm{R},q}(P) = \sum_{k=1}^{n} -p_k \log p_k$ (Jizba & Arimitsu 2004; Rényi 1961), and for this reason one sets $S_{\mathrm{R},1}(P) := \sum_{k=1}^{n} -p_k \log p_k$.

4 Statistical mechanics

Statistical mechanics (SM) aims to explain the behavior of macroscopic systems in terms of the dynamical laws governing their microscopic constituents.[12] One of the central concerns of SM is to provide a microdynamical explanation of the Second Law of TD. The strategy to achieve this goal is to first introduce a mechanical notion of entropy, then to argue that it is in some sense equivalent to the TD entropy, and finally to show that it tends to increase if its initial value is low. There are two nonequivalent frameworks in SM, one associated with Boltzmann and one with Gibbs. In this section we discuss the various notions of entropy introduced within these frameworks and briefly indicate how they have been used to justify the Second Law.

SM deals with systems consisting of a large number of microconstituents. A typical example of such a system is a gas, which is made up of a large number n of particles of mass m confined to a vessel of volume V. And in this essay we restrict attention to gases. Furthermore we assume that the system is isolated from its environment and hence that its total energy E is conserved. The behavior of such systems is usually modeled by continuous measure-preserving dynamical systems. We discuss such systems in detail in the next section; for the time being it suffices to say that the phase space of the system is $6n$-dimensional, having three position and three momentum coordinates for every particle. This space is called the system's *γ-space* X_γ. Then x_γ denotes a vector in X_γ, and the x_γ are called *microstates*. The set X_γ is the Cartesian product of n copies of the 6-dimensional phase space of one particle, called the particle's *μ-space* X_μ.[13] In what follows, $x_\mu = (x, y, z, p_x, p_y, p_z)$ denotes a vector in X_μ; moreover, we use $\vec{r} = (x, y, z)$ and $\vec{p} = (p_x, p_y, p_z)$.[14]

In a seminal paper published in 1872 Boltzmann set out to show that the Second Law of TD is a consequence of the collisions between the particles of

[12]For an extended discussion of SM, see Frigg 2008, Sklar 1993, and Uffink 2007.

[13]This terminology has been introduced by Ehrenfest & Ehrenfest-Afanassjewa (1911) and has been used since then. The subscript 'μ' here stands for 'molecule' and has nothing to do with a measure.

[14]We use momentum rather than velocity since this facilitates the discussion of the connection of Boltzmann entropies with other entropies. One could also use the velocity $\vec{v} = \vec{p}/m$.

a gas. The distribution $f(x_\mu, t)$ specifies the fraction of particles in the gas whose position and momentum lies in the infinitesimal interval $(x_\mu,\, x_\mu + dx_\mu)$ at time t. In 1860 Maxwell had shown that for a gas of *identical* and *noninteracting* particles in equilibrium the distribution had to be what is now called the *Maxwell–Boltzmann distribution*:

$$f(x_\mu, t) \;=\; \frac{\chi_V(\vec{r})\,(2\pi mkT)^{-3/2}}{\|V\|}\;\exp\Big(-\frac{\vec{p}^{\,2}}{2mkT}\Big),$$

where $\vec{p}^{\,2} := p_x^2 + p_y^2 + p_z^2$, the factor k is Boltzmann's constant, T the temperature of the gas, $\|V\|$ the volume of the vessel, and $\chi_V(\vec{r})$ the characteristic function of the set V (it is 1 if $\vec{r} \in V$, and 0 otherwise).

The state of a gas at time t is described by a distribution $f(x_\mu, t)$, and the dynamics of the gas can be studied by considering how this distribution evolves over time. To this end, Boltzmann introduced the quantity

$$\mathrm{H_B}(f) \;:=\; \int_{X_\mu} f(x_\mu, t)\log f(x_\mu, t)\, dx_\mu$$

(where 'B' is for 'Boltzmann'), and set out to prove on the basis of mechanical assumptions about the collisions of gas molecules that $\mathrm{H_B}(f)$ must decrease monotonically over the course of time and that it reaches its minimum at equilibrium, where $f(x_\mu, t)$ becomes the Maxwell–Boltzmann distribution. This result, which is derived using the *Boltzmann Equation*, is known as the H-*Theorem*, and it is generally regarded as problematic.[15]

The problems of the H-Theorem are not our concern. What matters is that the *fine-grained Boltzmann entropy* $S_{\mathrm{B,f}}$ (also *continuous Boltzmann entropy*) is proportional to $\mathrm{H_B}(f)$:

$$S_{\mathrm{B,f}}(f) \;:=\; -kn\,\mathrm{H_B}(f). \tag{7}$$

Therefore, if the H-Theorem were true, it would establish that the Boltzmann entropy increased monotonically and reached a maximum once the system's distribution becomes the Maxwell–Boltzmann distribution. Thus, if we associated the Boltzmann entropy with the thermodynamic entropy, this would amount to a justification of the Second Law.

How are we to interpret the distribution $f(x_\mu, t)$? As introduced, $f(x_\mu, t)$ reflects the distribution of the particles: it says what fraction of the particles in the gas are located in a certain region of the phase space. So it can be interpreted as an (approximate) actual distribution, involving no probabilistic notions. But $f(x_\mu, t)$ can also be interpreted probabilistically, as specifying the probability that a particle drawn at random (with replacement) from the gas is located in a particular part of the phase space. This probability is most naturally interpreted in a frequentist way: if we keep drawing molecules at random from the gas, then

[15] See Emch & Liu 2002, pp. 92–105, and Uffink 2007, pp. 962–74.

$f(x_\mu, t)$ gives us the relative frequency of molecules drawn from a certain region of phase space.

In response to criticism of his 1872 derivation, Boltzmann presented an altogether different approach to justifying the Second Law in 1877.[16] Since energy is conserved and the system is confined to volume V, each state of a particle lies within a finite subregion $X_{\mu,\mathrm{a}}$ of X_μ, the accessible region of X_μ. Now we *coarse-grain* $X_{\mu,\mathrm{a}}$, i.e. we choose a partition $\omega = \{\, \omega_i \mid i = 1, \ldots, l \,\}$ of $X_{\mu,\mathrm{a}}$.[17] The cells ω_i are taken to be rectangular with respect to the position and momentum coordinates and of equal volume $\delta\omega$, i.e. $\mu(\omega_i) = \delta\omega$, for all $i = 1, \ldots, l$, where μ is the Lebesgue measure on the 6-dimensional phase space of one particle. The *coarse-grained microstate*, also called *arrangement*, is a specification of which particle's state lies in which cell of ω.

The macroscopic properties of a gas (e.g. temperature, pressure) do not depend on which specific molecule is in which cell of the partition but are determined solely by the number of particles in each cell. A specification of how many particles are in each cell is called a *distribution* $D = (n_1, \ldots, n_l)$, meaning that n_1 particles are in cell ω_1, etc. Clearly, $\sum_{j=1}^{l} n_j = n$. We label the different distributions with a discrete index i and denote the i^{th} distribution by D_i. The ratio D_i/n can be interpreted in the same way as $f(x_\mu, t)$ above.

Several arrangements correspond to the same distribution. More precisely, elementary combinatorial considerations show that

$$G(D) := \frac{n!}{n_1! \cdots n_l!} \tag{8}$$

arrangements are compatible with a given distribution D. The so-called *coarse-grained Boltzmann entropy* (also *combinatorial entropy*) is defined as (where 'B' is for 'Boltzmann' and 'ω' denotes the partition)

$$S_{\mathrm{B},\omega}(D) := k \log G(D). \tag{9}$$

Since $G(D)$ is the number of arrangements compatible with a given distribution and the logarithm is an increasing function, $S_{\mathrm{B},\omega}(D)$ is a natural measure for the number of arrangements that are compatible with a given distribution in the sense that the greater $S_{\mathrm{B},\omega}(D)$, the more arrangements are compatible with a given distribution D. Hence $S_{\mathrm{B},\omega}(D)$ is a measure of how much we can infer about the arrangement of a system on the basis of its distribution. The higher $S_{\mathrm{B},\omega}(D)$, the less information a distribution conveys about the arrangement of the system.

[16] See Uffink 2007, pp. 974–83, and Frigg 2008, pp. 107–13. Frigg (2009, 2010b) provides a discussion of Boltzmann's use of probabilities.

[17] We give a rigorous definition of a partition in the next section.

Boltzmann then postulated that the distribution with the highest entropy was the equilibrium distribution, and that systems had a natural tendency to evolve from states of low to states of high entropy. However, as later commentators, most notably Ehrenfest & Ehrenfest-Afanassjewa (1911), pointed out, for the latter to happen, further dynamical assumptions (e.g. ergodicity) are needed. If such assumptions are in place, the n_i evolve so that $S_{\mathrm{B},\omega}(D)$ increases and then stays close to its maximum value most of the time. This has engendered a large literature, covering many aspects. Two recent reviews are Lavis 2004 and 2008.

There is a third notion of entropy in the Boltzmannian framework, and this notion is preferred by contemporary Boltzmannians.[18] We now consider X_γ rather than X_μ. Since there are constraints on the system, its state will lie within a finite subregion $X_{\gamma,\mathrm{a}}$ of X_γ, the accessible region of X_γ.[19]

If the gas is regarded as a macroscopic object rather than as a collection of molecules, its state can be characterized by a small number of macroscopic variables such as temperature, pressure, and density. These values are then usually coarse-grained so that all values falling into a certain range are regarded as belonging to the same macrostate. Hence the system can be described as being in one of a finite number of macrostates M_i ($i = 1, \ldots, m$). The set of the M_i is complete in that at any given time t the system must be in exactly one M_i. It is a basic posit of the Boltzmann approach that a system's macrostate supervenes on its fine-grained microstate, so that a change in the macrostate must be accompanied by a change in the fine-grained microstate. Therefore, to every given microstate x_γ there corresponds *exactly one* macrostate $M(x_\gamma)$. But many different microstates can correspond to the same macrostate. We therefore define

$$X_{M_i} \;:=\; \Big\{\, x_\gamma \in X_{\gamma,\mathrm{a}} \;\Big|\; M_i = M(x_\gamma) \,\Big\} \qquad (i = 1, \ldots, m),$$

which is the subset of $X_{\gamma,\mathrm{a}}$ consisting of all microstates that correspond to macrostate M_i. The X_{M_i} are called *macroregions*. Clearly, they form a partition of $X_{\gamma,\mathrm{a}}$.

The Boltzmann entropy of a macrostate M is (where 'B' is for 'Boltzmann' and 'm' is for 'macrostate')[20]

$$S_{\mathrm{B,m}}(M) \;:=\; k \log \mu(X_M). \tag{10}$$

[18] See, for instance, Goldstein 2001 and Lebowitz 1999b.

[19] These constraints include conservation of energy. Therefore, $X_{\gamma,\mathrm{a}}$ is $(6n-1)$-dimensional. This causes complications because the measure μ needs to be restricted to the $(6n-1)$-dimensional energy hypersurface and the definitions of macroregions become more complicated. In order to keep things simple, we assume that $X_{\gamma,\mathrm{a}}$ is $6n$-dimensional. For the $(6n-1)$-dimensional case, see Frigg 2008, pp. 107–14.

[20] See e.g. Goldstein 2001, p. 43, and Lebowitz 1999b, p. 348.

Hence $S_{B,m}(M)$ measures the portion of the system's γ-space that is taken up by microstates that correspond to M. Consequently, $S_{B,m}(M)$ measures how much we can infer about where in γ-space the system's microstate lies: the higher $S_{B,m}(M)$, the larger the portion of the γ-space in which the system's microstate could be.

Given this notion of entropy, the leading idea is to argue that the dynamics is such that $S_{B,m}$ increases. That is, the evolution of $x_\gamma \in X_{\gamma,a}$ is such that the sequence of macrostates $M(x_\gamma)$ gives increasing $S_{B,m}(M_\gamma)$.

Most contemporary Boltzmannians aim to achieve this by arguing that entropy-increasing behavior is *typical*; see, for instance, Goldstein 2001. These arguments are the subject of ongoing controversy (see Frigg 2009, 2010b).

We now turn to a discussion of the interrelationships between the various entropy notions introduced so far. Let us begin with $S_{B,\omega}$ and $S_{B,m}$. The former is a function of a distribution over a partition of $X_{\mu,a}$, while $S_{B,m}$ takes cells of a partition of $X_{\gamma,a}$ as arguments. The crucial point to realize is that each distribution corresponds to a well-defined region of $X_{\gamma,a}$: for the choice of a partition of $X_{\mu,a}$ induces a partition of $X_{\gamma,a}$ (because X_γ is the Cartesian product of n copies of X_μ). Hence any D_i determines a unique region $X_{D_i} \subseteq X_{\gamma,a}$ so that all states $x_\gamma \in X_{D_i}$ have distribution D_i:

$$X_{D_i} \;:=\; \Big\{ x_\gamma \in X_\gamma \;\Big|\; D(x_\gamma) = D_i \Big\}, \tag{11}$$

where $D(x_\gamma)$ is the distribution determined by the state x_γ (via the arrangement that x_γ determines—cf. the discussion of Eqn 8). Because all cells have measure $\delta\omega$, Eqns (8) and (11) imply:

$$\mu(X_{D_i}) \;=\; G(D_i)\,(\delta\omega)^n. \tag{12}$$

Given this, the question of the relation between $S_{B,\omega}$ and $S_{B,m}$ comes down to the question of how the X_{D_i} and the X_{M_i} relate. Since there are no canonical procedures to define what we mean by 'macrostate,' and hence to construct the X_{M_i}, one can use the above considerations about how distributions determine regions to construct the X_{M_i}, making $X_{D_i} = X_{M_i}$ true by definition. So one can say that, conceptually speaking, $S_{B,\omega}$ is a special case of $S_{B,m}$ (or that it is a concrete realization of the more abstract notion of $S_{B,m}$). If $X_{D_i} = X_{M_i}$, Eqns (10) and (12) imply:

$$S_{B,m}(M_i) \;=\; k \log G(D_i) + kn\log(\delta\omega). \tag{13}$$

Hence $S_{B,m}(M_i)$ equals $S_{B,\omega}$ up to an additive constant.

How are $S_{B,m}$ and $S_{B,f}$ related? Assume that $X_{D_j} = X_{M_j}$, that the system is large, and that there are many particles in each cell ($n_j \gg 1$ for all j), which allows us to use Stirling's Formula: $n! \approx \sqrt{2\pi n}\,(n/e)^n$. Plugging Eqn (8) into Eqn (13) yields (Tolman 1938, Ch. 4):

$$\log \mu(X_{M_j}) \;\approx\; n\log n - \sum_{i=1}^{l} n_i \log n_i + n\log(\delta\omega). \tag{14}$$

Clearly, for the n_i used in the definition of $S_{\mathrm{B},\omega}$ we have

$$n_i \;\approx\; \tilde{n}_i(t) \;:=\; n\int_{\omega_i} f(x_\mu, t)\,dx_\mu.$$

Unlike the n_i, the $\tilde{n}_i$ need not be integers. If $f(x_\mu, t)$ does not vary much in each cell ω_i, we find:

$$\sum_{i=1}^{l} n_i \log n_i \;\approx\; n\mathrm{H}_\mathrm{B} + n\log n + n\log(\delta\omega). \tag{15}$$

Comparing (14) and (15) yields $-nk\mathrm{H}_\mathrm{B} \approx k\log\mu(X_{M_j})$, i.e. $S_{\mathrm{B,m}} \approx S_{\mathrm{B,f}}$. Hence, for large numbers of particles, $S_{\mathrm{B,m}}$ and $S_{\mathrm{B,f}}$ are approximately equal.

How are $S_{\mathrm{B,m}}$ and the Shannon entropy related? According to Eqn (14),

$$S_{\mathrm{B,m}}(M_j) \;\approx\; -k\sum_{i=1}^{l} n_i \log n_i + C(n, \delta\omega),$$

where $C(n, \delta\omega)$ is a constant depending on n and $\delta\omega$. Introducing the quotients $p_j := n_j/n$, we find

$$S_{\mathrm{B,m}}(M_j) \;\approx\; -nk\sum_{i=1}^{l} p_i \log p_i + \tilde{C}(n, \delta\omega), \tag{16}$$

where $\tilde{C}(n, \delta\omega)$ is a constant depending on n and $\delta\omega$. The quotients p_i are finite relative frequencies for a particle being in ω_i. The p_i can be interpreted as the probability of finding a randomly chosen particle in cell ω_i. Then, if we regard the ω_i as messages, $S_{\mathrm{B,m}}(M_i)$ is equivalent to the Shannon entropy up to the multiplicative constant nk and the additive constant $\tilde{C}$.

Finally, how does $S_{\mathrm{B,f}}$ relate to the TD entropy? The TD entropy of an ideal gas is given by the Sackur–Tetrode Formula

$$S_{\mathrm{TD}} \;=\; nk\log\left(\left(\frac{T}{T_0}\right)^{3/2}\frac{V}{V_0}\right), \tag{17}$$

where T_0 and V_0 are the temperature and the volume of the gas at reference point E (see Reiss 1965, pp. 89–90). One can show that $S_{\mathrm{B,f}}$ for the Maxwell–Boltzmann distribution is equal to Eqn (17) up to an additive constant (Emch & Liu 2002, p. 98; Uffink 2007, p. 967). This is an important result. However, it is an open question whether this equivalence holds for systems with interacting particles, that is, for systems different from ideal gases.

We now turn our attention to Gibbsian SM. The object of study in the Gibbs approach is not an individual system (as in the Boltzmann approach) but an ensemble—an uncountably infinite collection of independent systems that are all governed by the same equations, but whose states at a time t differ. The ensemble is specified by an everywhere positive density function $\rho(x_\gamma, t)$ on the system's γ-space: $\rho(x_\gamma, t)\, dx_\gamma$ is the infinitesimal fraction of systems in the ensemble whose state lies in the $6n$-dimensional interval $(x_\gamma, x_\gamma + dx_\gamma)$. The time-evolution of the ensemble is then associated with changes in the density function in time.

Thus $\rho(x_\gamma, t)$ is a probability density, so that the probability at time t of finding the state of a system in region $R \subseteq X_\gamma$ is

$$p_t(R) = \int_R \rho(x_\gamma, t)\, dx_\gamma.$$

The *fine-grained Gibbs entropy* (also known as *ensemble entropy*) is defined as (where 'G' is for 'Gibbs' and 'f' is for 'fine-grained')

$$S_{\mathrm{G,f}}(\rho) := -k\int_{X_\gamma} \rho(x_\gamma, t) \log \rho(x_\gamma, t)\, dx_\gamma.$$

How to interpret $\rho(x_\gamma, t)$ (and hence $p_t(R)$) is far from clear. Edwin Jaynes proposed to interpret $\rho(x_\gamma, t)$ epistemically; we turn to his approach to SM below. There are (at least) two possible ontic interpretations: a frequency interpretation and a time-average interpretation. On the frequency interpretation one thinks about an ensemble as analogous to an urn, but rather than containing balls of different colors the ensemble contains systems in different microstates (Gibbs 1981, p. 163). The density $\rho(x_\gamma, t)$ specifies the frequency with which we draw systems in a certain microstate. On the time-average interpretation, $\rho(x_\gamma, t)$ encodes the fraction of time that the system would spend, in the long run, in each of the various regions of the phase space if it was left to its own. Although plausible at first blush, both interpretations face serious difficulties, and it is unclear whether these can be met (see Frigg 2008, pp. 153–5).

If we regard $S_{\mathrm{G,f}}(\rho)$ as equivalent to the TD entropy, then $S_{\mathrm{G,f}}(\rho)$ is expected to increase over time (during an irreversible adiathermal process) and to assume a maximum in equilibrium. However, systems in SM are Hamiltonian, and it is a consequence of an important theorem of Hamiltonian mechanics, *Liouville's Theorem*, that $S_{\mathrm{G,f}}$ is a constant of motion: $dS_{\mathrm{G,f}}/dt = 0$. So $S_{\mathrm{G,f}}$ remains constant, and hence the approach to equilibrium cannot be described in terms of an increase in $S_{\mathrm{G,f}}$.

The standard way to solve this problem is to instead consider the coarse-grained Gibbs entropy. This solution was suggested by Gibbs (1902, Ch. 12) himself and has since been endorsed by many (e.g. Penrose 1970). Consider a partition ω of X_γ where the cells ω_i are of equal volume $\delta\omega$. The *coarse-grained* *density* $\bar{\rho}(x_\gamma, t)$ is defined as the density that is uniform within each cell, taking as its value the average value in this cell:

$$\bar{\rho}_\omega(x_\gamma, t) \;:=\; \frac{1}{\delta\omega}\int_{\omega(x_\gamma)} \rho(x'_\gamma, t)\, dx'_\gamma,$$

where $\omega(x_\gamma)$ is the cell in which x_γ lies. We can now define the *coarse-grained Gibbs entropy* (where 'G' stands for 'Gibbs' and 'ω' for the partition):

$$S_{\mathrm{G},\,\omega}(\rho) \;:=\; S_{\mathrm{G,\,f}}(\bar{\rho}_\omega) \;=\; -k\int_{X_\gamma} \bar{\rho}_\omega \log \bar{\rho}_\omega \, dx_\gamma.$$

One can prove that $S_{\mathrm{G},\,\omega} \geq S_{\mathrm{G,\,f}}$; the equality holds iff the fine-grained distribution is uniform over the cells of the coarse-graining (see Lavis 2004, p. 229; Wehrl 1978, p. 672). The coarse-grained density $\bar{\rho}_\omega$ is not subject to Liouville's Theorem and is not a constant of motion. So $\bar{\rho}_\omega$ could, in principle, increase over time.[21]

How do the two Gibbs entropies relate to the other notions of entropy introduced so far? The most straightforward connection is between the Gibbs entropy and the continuous Shannon entropy, which differ only by the multiplicative constant k. This realization provides a starting point for Jaynes's (1983) information-based interpretation of SM, at the heart of which lies a radical reconceptualization of SM. On his view, SM is about *our knowledge* of the world, not about the world. The probability distribution represents our state of knowledge about the system and not some matter of fact about the system: $\rho(x_\gamma, t)$ represents our lack of knowledge about a microstate of a system given its macrocondition, and entropy is a measure of how much knowledge we lack.

Jaynes then postulated that to make predictions we should always use the distribution that maximizes uncertainty under the given macroscopic constraints. This means that we are asked to find the distribution for which the Gibbs entropy is maximal, and then use this distribution to calculate expectation values of the variables of interest. This prescription is now known as *Jaynes' Maximum-Entropy Principle*. Jaynes could show that this principle recovers the standard SM distributions (e. g. the microcanonical distribution for isolated systems).

The idea behind this principle is that we should always choose the distribution that is maximally noncommittal with respect to the missing information, because by not doing so we would make assertions for which we have no evidence. Although intuitive at first blush, the Maximum-Entropy Principle is fraught with controversy (see, for instance, Howson & Urbach 2006, pp. 276–88).[22]

[21] There is a thorny issue under which conditions the coarse-grained entropy actually increases (see Lavis 2004).

[22] For a discussion of Jaynes' take on nonequilibrium SM, see Sklar 1993, pp. 255–7. Furthermore, Tsallis (1988) proposed a way of deriving the main distributions of SM which is very similar to Jaynes',

A relation between $S_{\mathrm{G,f}}(\rho)$ and the TD entropy can be established only case by case. $S_{\mathrm{G,f}}(\rho)$ coincides with S_{TD} in relevant cases arising in practice. For instance, the calculation of the entropy of an ideal gas from the microcanonical ensemble yields equation (17)—up to an additive constant (Kittel 1958, p. 39).

Finally, how do the Gibbs and Boltzmann entropies relate? Let us start with the fine-grained entropies $S_{\mathrm{B,f}}$ and $S_{\mathrm{G,f}}$. Assume that the particles are identical and noninteracting. Then $\rho(x_\gamma, t) = \prod_{i=1}^{n} \rho_i(x_\mu, t)$, where ρ_i is the density pertaining to particle i. Then

$$S_{\mathrm{G,f}}(\rho) \; := \; -kn\int_{X_\mu} \rho_1(x_\mu, t) \log \rho_1(x_\mu, t)\, dx_\mu, \tag{18}$$

which is *formally* equivalent to $S_{\mathrm{B,f}}$ (7). The question is how ρ_1 and f relate, since they are different distributions. Our f is the distribution of n particles over the phase space; ρ_1 is a one-particle function. Because the particles are identical and noninteracting, we can apply the Law of Large Numbers to conclude that it is very likely that the probability of finding a given particle in a particular region of phase space is approximately equal to the proportion of particles in that region. Hence $\rho_1 \approx f$ and $S_{\mathrm{G,f}} \approx S_{\mathrm{B,f}}$.

A similar connection exists between the coarse-grained entropies $S_{\mathrm{G,m}}$ and $S_{\mathrm{B},\omega}$. If the particles are identical and noninteracting, one finds

$$S_{\mathrm{G},\omega} \; = \; -kn\sum_{i=1}^{l}\int_{\omega_i} \frac{\Omega_i}{\delta\omega} \log \frac{\Omega_i}{\delta\omega}\, dx_\mu \; = \; -kn\sum_{i=1}^{l} \Omega_i \log \Omega_i + C(n, \delta\omega),$$

where $\Omega_i = \int_{\omega_i} \rho_1\, dx_\mu$. This is *formally* equivalent to $S_{\mathrm{B,m}}$ (16), which in turn is equivalent (up to an additive constant) to $S_{\mathrm{B},\omega}$ (9). Again for large n we can apply the Law of Large Numbers to conclude that it is very likely that $\Omega_i \approx p_i$ and $S_{\mathrm{G,m}} = S_{\mathrm{B},\omega}$.

It is crucial for the connections between the Gibbs and the Boltzmann entropy that the particles are identical and noninteracting. It is unclear whether the conclusions hold if these assumptions are relaxed.[23]

5 Dynamical-systems theory

In this section we focus on the main notions of entropy in dynamical-systems theory, namely the Kolmogorov–Sinai entropy (KS entropy) and the topological

based on establishing a connection between what is now called the Tsallis entropy and the Rényi entropy. A similar attempt using only the Rényi entropy has been undertaken by Bashkirov (2006).

[23] Jaynes (1965) argues that the Boltzmann entropy differs from the Gibbs entropy except for noninteracting and identical particles. However, he *defines* the Boltzmann entropy as (18). As argued, (18) is equivalent to the Boltzmann entropy if the particles are identical and noninteracting, but this does not appear to be generally the case. So Jaynes' (1965) result seems useless.

entropy.[24] They occupy center stage in *chaos theory*—a mathematical theory of deterministic yet irregular and unpredictable, or even random, behavior.[25]

We begin by briefly recapitulating the main tenets of dynamical-systems theory.[26] The two main elements of every dynamical system are a set X of all possible states x, the *phase space* of the system, and a family of transformations $T_t\colon X \to X$ mapping the phase space to itself. The parameter t is time, and the transformations $T_t(x)$ describe the *time-evolution* of the system's instantaneous state $x \in X$. For the systems we have discussed in the last section, X consists of the positions and momenta of all particles in the system and T_t is the time-evolution of the system under the dynamical laws. If t ranges over the positive real numbers and zero (i.e. $t \in \mathbb{R}_0^+$), the system's dynamics is *continuous*. If t ranges over the natural numbers including zero (i.e. $t \in \mathbb{N}_0$), the dynamics is *discrete*.[27] The family T_t defining the dynamics must have the structure of a semigroup where $T_{t_1+t_2}(x) = T_{t_2}\big(T_{t_1}(x)\big)$ for all t_1, t_2 either in $\mathbb{R}_0^+$ (continuous time) or in $\mathbb{N}_0$ (discrete time).[28] The continuous trajectory through x is the set $\{\, T_t(x) \mid t \in \mathbb{R}_0^+ \,\}$; the discrete trajectory through x is the set $\{\, T_t(x) \mid t \in \mathbb{N}_0 \,\}$.

Continuous time-evolutions often arise as solutions to differential equations of motion (such as Newton's or Hamilton's). In dynamical-systems theory the class of allowable equations of motion is usually restricted to ones for which solutions exist and are unique for all times $t \in \mathbb{R}$. Then $\{\, T_t \mid t \in \mathbb{R} \,\}$ is a group, where $T_{t_1+t_2}(x) = T_{t_2}\big(T_{t_1}(x)\big)$ for all t_1, $t_2 \in \mathbb{R}$, and is often called a *flow*. In what follows we only consider continuous systems that are flows.

For discrete systems the maps defining the time-evolution neither have to be injective nor surjective, and so $\{\, T_t \mid t \in \mathbb{N}_0 \,\}$ is only a semigroup. All T_t are generated as iterative applications of the single map $T_1 := T\colon X \to X$ because $T_t := T^t$, and we refer to the $T_t(x)$ as *iterates of* x. Iff T is invertible, T_t is defined both for positive and negative times and $\{\, T_t \mid t \in \mathbb{Z} \,\}$ is a group.

It follows that all dynamical systems are *forward-deterministic*: any two trajectories that agree at one instant of time agree at all *future* times. If the dynamics of the system is invertible, the system is deterministic *tout court*: any two trajectories that agree at one instant of time agree at all times (Earman 1971).

[24]There are also a few other, less important entropies in dynamical-systems theory, e.g. the Brin–Katok local entropy (see Mañé 1987).

[25]For a discussion of the kinds of randomness in chaotic systems, see Berkovitz, Frigg & Kronz 2006 and Werndl 2009a, 2009b, 2009d.

[26]For more details, see Cornfeld, Fomin & Sinai 1982 and Petersen 1983.

[27]The reason not to choose $t \in \mathbb{Z}$ is that some maps, e.g. the logistic map, are not invertible.

[28]$S = \{a, b, c, \ldots\}$ is a *semigroup* iff there is a multiplication operation '$\cdot$' on S so that (i) $a \cdot b \in S$ for all a, $b \in S$; (ii) $a \cdot (b \cdot c) = (a \cdot b) \cdot c$ for all a, b, $c \in S$; (iii) there is an $e \in S$ such that $e \cdot a = a \cdot e = a$ for all $a \in S$. A semigroup as defined here is also called a *monoid*. If for every $a \in S$ there is an $a^{-1} \in S$ so that $a^{-1} \cdot a = a \cdot a^{-1} = e$, then S is a *group*.

Two kinds of dynamical systems are relevant for our discussion: measure-theoretical and topological dynamical ones. A *topological dynamical system* has a metric defined on X.[29] More specifically, a *discrete topological dynamical system* is a triple (X, d, T) where d is a metric on X and $T\colon X \to X$ is a mapping. *Continuous topological dynamical systems* (X, d, T_t) ($t \in \mathbb{R}$) are defined accordingly, where T_t is the above semigroup. Topological systems allow for a wide class of dynamical laws since the T_t do not have to be either injective or surjective.

A *measure-theoretical dynamical system* is one whose phase space is endowed with a measure.[30] More specifically, a *discrete measure-theoretical dynamical system* (X, Σ, μ, T) consists of a phase space X, a σ-algebra Σ on X, a measure μ, and a measurable transformation $T\colon X \to X$. If T is *measure-preserving*, i.e. $\mu(T^{-1}(A)) = \mu(A)$ for all $A \in \Sigma$, where $T^{-1}(A) := \{\, x \in X \mid T(x) \in A \,\}$, we have a *discrete measure-preserving dynamical system*. It only makes sense to speak of measure-preservation if T is surjective. Therefore, we suppose that the T in measure-preserving systems is surjective. However, we do not presuppose that it is injective, because some important maps are not injective, e.g. the logistic map.

A *continuous measure-theoretical dynamical system* is a quadruple (X, Σ, μ, T_t), where $\{\, T_t \mid t \in \mathbb{R}_0^+ \,\}$ is the above semigroup of transformations which are measurable on $X \times \mathbb{R}_0^+$, and the other elements are as above. Such a system is a *continuous measure-preserving dynamical system* if T_t is measure-preserving for all t (again, we presuppose that all T_t are surjective).

We make the (common) assumption that the measure of measure-preserving systems is normalized: $\mu(X) = 1$. The motivation for this is that normalized measures are probability measures, making it possible to use probability calculus. This raises the question of how to interpret these probabilities. This issue is particularly thorny because it is widely held that there cannot be ontic probabilities in deterministic systems: either the dynamics of a system is deterministic or chancy, but not both. This dilemma can be avoided if one interprets probabilities epistemically, i.e. as reflecting lack of knowledge. As we saw in the previous section, this is what Jaynes did in SM. Although sensible in some situations, this interpretation is clearly unsatisfactory in others. Roulette wheels and dice are paradigmatic examples of chance setups, and it is widely held that there are ontic chances for certain events to occur: the chance of getting a '3' when throwing a die is 1/6, and this is so because of how the world *is* and it has nothing to do with what we happen to *know* about it. Yet, from a mechanical point of view, these are deterministic systems. Consequently, there must be ontic interpretations of probabilities in deterministic systems. There are at least three options available.

[29] For an introduction to metric spaces, see Sutherland 2002.

[30] See Halmos 1950 for an introduction to measures.

The first is the time-average interpretation already mentioned above: the probability of an event E is the fraction of time that the system spends (in the long run) in the region of X associated with E (Falconer 1990, p. 254; Werndl 2009d). The ensemble interpretation defines the measure of a set A at time t as the fraction of solutions starting from some set of initial conditions that are in A at t. A third option is the so-called Humean Best-System Analysis originally proposed by Lewis (1980). Roughly speaking, this interpretation is an elaboration of (finite) frequentism. Lewis' own assertions notwithstanding, this interpretation works in the context of deterministic systems (Frigg & Hoefer 2010).

Let us now discuss the notions of volume-preservation and measure-preservation. If the preserved measure is the Lebesgue measure, the system is *volume-preserving*. If the system fails to be volume-preserving, then it is *dissipative*. Being dissipative is not the failure of measure-preservation with respect to *any* measure (as a common misconception has it); it is nonpreservation of the *Lebesgue* measure. In fact many dissipative systems preserve measures. More precisely, if (X, Σ, λ, T) (or $(X, \Sigma, \lambda, T_t)$) is dissipative ($\lambda$ is the Lebesgue measure), often, although not always, there exists a measure $\mu \neq \lambda$ such that (X, Σ, μ, T) (resp. (X, Σ, μ, T_t)) is measure-preserving. The Lorenz system and the logistic maps are cases in point.

A *partition* $\alpha = \{ \alpha_i \mid i = 1, \ldots, n \}$ of (X, Σ, μ) is a collection of nonempty, nonintersecting measurable sets that cover X, that is: $\alpha_i \neq \emptyset$ for all $i \in \{1, \ldots, n\}$, $\alpha_i \cap \alpha_j = \emptyset$ for all $i \neq j$, and $X = \bigcup_{i=1}^{n} \alpha_i$. The α_i are called *atoms*. If α is a partition, $T_t^{-1}\alpha := \{ T_t^{-1}\alpha_i \mid i = 1, \ldots, n \}$ is a partition too. The set $T_t\alpha := \{ T_t\alpha_i \mid i = 1, \ldots, n \}$ is a partition iff T_t is invertible. Given two partitions $\alpha = \{ \alpha_i \mid i = 1, \ldots, n \}$ and $\beta = \{ \beta_j \mid j = 1, \ldots, m \}$, the *join* $\alpha \vee \beta$ is defined as $\{ \alpha_i \cap \beta_j \mid i = 1, \ldots, n;\ j = 1, \ldots, m \}$.

This concludes our brief recapitulation of dynamical-systems theory. The rest of this section concentrates on measure-preserving systems. This is not very restrictive because many systems, including all deterministic Newtonian systems, many dissipative systems, and all chaotic systems (Werndl 2009d), fall into this class.

Let us first discuss the KS entropy. Given a partition $\alpha = \{\alpha_1, \ldots, \alpha_k\}$, let $\mathrm{H}(\alpha) := -\sum_{i=1}^{k} \mu(\alpha_i) \log \mu(\alpha_i)$. For a discrete system (X, Σ, μ, T) consider

$$\mathrm{H}_n(\alpha, T) \ := \ {}^{1}/_{n}\, \mathrm{H}\big(\alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-n+1}\alpha\big).$$

The limit $\mathrm{H}(\alpha, T) := \lim_{n\to\infty} \mathrm{H}_n(\alpha, T)$ exists, and the KS entropy is defined as the supremum over all partitions α (Petersen 1983, p. 240):

$$S_{\mathrm{KS}}(X, \Sigma, \mu, T) \ := \ \sup_{\alpha} \mathrm{H}(\alpha, T). \tag{19}$$

For a continuous system (X, Σ, μ, T_t) it can be shown that for any $t_0 \neq 0$ with $-\infty < t_0 < \infty$,

$$S_{\mathrm{KS}}(X, \Sigma, \mu, T_{t_0}) \;=\; |t_0|\, S_{\mathrm{KS}}(X, \Sigma, \mu, T_1),$$

where $S_{\mathrm{KS}}(X, \Sigma, \mu, T_{t_0})$ is the KS entropy of the discrete system $(X, \Sigma, \mu, T_{t_0})$ and $S_{\mathrm{KS}}(X, \Sigma, \mu, T_1)$ is the KS entropy of the discrete system (X, Σ, μ, T_1) (Cornfeld, Fomin & Sinai 1982). Consequently, the KS entropy of a continuous system (X, Σ, μ, T_t) is defined as $S_{\mathrm{KS}}(X, \Sigma, \mu, T_1)$, and when discussing the meaning of the KS entropy it suffices to focus on (19).[31]

How can the KS entropy be interpreted? There is a fundamental connection between dynamical-systems theory and information theory, as follows. For a dynamical system (X, Σ, μ, T) each $x \in X$ produces, relative to a partition α, an infinite string of messages $m_0 m_1 m_2 \ldots$ in an alphabet of k letters via the coding $m_j = \alpha_i$ iff $T^j(x) \in \alpha_i$ $(j \geq 0)$. Assume that (X, Σ, μ, T) is interpreted as the source. Then the output of the source are the strings $m_0 m_1 m_2 \ldots$ If the measure is interpreted as a probability density, we have a probability distribution over these strings. Hence the whole apparatus of information theory can be applied to these strings.[32] In particular, notice that $\mathrm{H}(\alpha)$ is the Shannon entropy of $P = \big(\mu(\alpha_1), \ldots, \mu(\alpha_k)\big)$ and so measures the average information of the message α_i.

In order to motivate the KS entropy, consider for $\alpha := \{\alpha_1, \ldots, \alpha_k\}$ and $\beta := \{\beta_1, \ldots, \beta_l\}$:

$$\mathrm{H}(\alpha \mid \beta) \quad := \quad \sum_{j=1}^{l} \mu(\beta_j) \sum_{i=1}^{k} \frac{\mu(\alpha_i \cap \beta_j)}{\mu(\beta_j)} \log \frac{\mu(\alpha_i \cap \beta_j)}{\mu(\beta_j)}.$$

Recalling the definition of the conditional Shannon entropy (5), we see that $\mathrm{H}\big(\alpha \mid \bigvee_{k=1}^{n} T^{-k}\alpha\big)$ measures the average information received about the present state of the system whatever n past states have already been recorded. It is proven (Petersen 1983, pp. 241–2) that

$$S_{\mathrm{KS}}(X, \Sigma, \mu, T) \;=\; \sup_{\alpha} \lim_{n \to \infty} \mathrm{H}\big(\alpha \mid \textstyle\bigvee_{k=1}^{n} T^{-k}\alpha\big). \tag{20}$$

Hence the KS entropy is linked to the Shannon entropy; namely it measures the highest average information received about the present state of the system relative to a coding α given the past states that have been received.

Clearly, Eqn (20) implies that

$$S_{\mathrm{KS}}(X, \Sigma, \mu, T) \;=\; \sup_{\alpha} \lim_{n \to \infty} {}^{1}\!/_{n} \sum_{k=1}^{n} \mathrm{H}\big(\alpha \mid \textstyle\bigvee_{i=1}^{k} T^{-i}\alpha\big).$$

[31] For experimental data the KS entropy, and also the topological entropy (discussed later), is rather hard to determine. For details, see Eckmann & Ruelle 1985 and Ott 2002; see also Shaw 1985, where the question is discussed how to define a quantity similar to the KS entropy for dynamical systems with added noise.

[32] For details, see Frigg 2004 and Petersen 1983, pp. 227–34.

Hence the KS entropy can be also interpreted as the highest average of the average information gained about the present state of the system relative to a coding α whatever past states have been received (Frigg 2004; 2006a).

This is not the only connection to the Shannon entropy: Let us regard *strings of length* n ($n \in \mathbb{N}$) produced by the dynamical system relative to a coding α as messages. The set of all possible n-strings relative to α is $\beta = \{\beta_1, \ldots, \beta_h\} := \alpha \vee T^{-1}\alpha \vee \cdots \vee T^{-n+1}\alpha$ (where $h \in \mathbb{N}$), and the probability distribution of these possible strings of length n is $\mu(\beta_i)$ $(1 \leq i \leq h)$. Hence $\mathrm{H}_n(\alpha, T)$ measures the average amount of information which the system produces *per step* over the first n steps relative to the coding α, and $\lim_{n\to\infty} \mathrm{H}_n(\alpha, T)$ measures the average amount of information that the system can produce per step relative to a coding (cf. Petersen 1983, pp. 227–34).

A positive KS entropy is often linked to chaos. The interpretations just discussed provide a rationale for this: The Shannon information measures uncertainty, and this uncertainty is a form of unpredictability (Frigg 2004). Hence a positive KS entropy means that relative to some codings the behavior of the system is unpredictable.

Kolmogorov (1958) was the first to connect dynamical-systems theory with information theory. Based on Kolmogorov's work, Sinai (1959) introduced the KS entropy. One of Kolmogorov's main motivations was the following.[33] Kolmogorov conjectured that while the deterministic systems used in science produce no information, the stochastic processes used in science do produce information, and the KS entropy was introduced to capture the property of producing positive information. It was a big surprise when it was found that also several deterministic systems used in science, e.g. some Newtonian systems etc., have positive KS entropy. Hence this attempt of separating deterministic systems from stochastic processes failed (Werndl 2009a).

Due to lack of space we cannot discuss another, quite different, interpretation of the Kolmogorov–Sinai entropy, where $\sup_\alpha \mathrm{H}(\alpha, T)$ is a measure of the highest average rate of exponential divergence of solutions relative to a partition as time goes to infinity (Berger 2001, pp. 117–18). This implies that if $S_{\mathrm{KS}}(X, \Sigma, \mu, T) > 0$, there is exponential divergence and thus unstable behavior on some regions of phase space, explaining the link to chaos. This interpretation does *not* require that the measure is interpreted as probability.

There is also another connection of the KS entropy to exponential divergence of solutions. The Lyapunov exponents of x measure the mean exponential divergence of solutions originating near x, where the existence of positive Lyapunov exponents indicates that, in some directions, solutions diverge exponentially on average. *Pesin's Theorem* states that under certain assumptions

[33] Another main motivation was to make progress on the problem of which systems are probabilistically equivalent (Werndl 2009c).

$S_{\mathrm{KS}}(X, \Sigma, \mu, T) = \int_X S(x)\, d\mu$, where $S(x)$ is the sum of the positive Lyapunov exponents of x. Another important theorem which should be mentioned is *Brudno's Theorem*, which states that if the system is ergodic and certain other conditions hold, $S_{\mathrm{KS}}(X, \Sigma, \mu, T)$ equals the algorithmic complexity (a measure of randomness) of almost all solutions (Batterman & White 1996).

To sum up, the interpretations of the KS entropy as measuring exponential divergence are not connected to any other notion of entropy or to what these notions are often believed to capture, such as information (Grad 1961, pp. 323–34; Wehrl 1978, pp. 221–4). To conclude, the only link of the KS entropy to entropy notions is with the Shannon entropy.

Let us now discuss the topological entropy, which is always defined only for discrete systems. It was first introduced by Adler, Konheim & McAndrew (1965); later Bowen (1971) introduced two other, equivalent definitions.

We first turn to Adler, Konheim & McAndrew's definition. Let (X, d, T) be a topological dynamical system where X is compact and $T\colon X \to X$ is a continuous function which is surjective.[34] Let U be an *open cover of* X, i.e. a set $U := \{U_1, \ldots, U_k\}$ ($k \in \mathbb{N}$) of open sets such that $\bigcup_{i=1}^{k} U_i \supseteq X$.[35] An open cover $V = \{V_1, \ldots, V_l\}$ is said to be an *open subcover of* an open cover U iff $V_j \in U$ for all j $(1 \le j \le l)$. For open covers $U = \{U_1, \ldots, U_k\}$ and $V = \{V_1, \ldots, V_l\}$ let $U \vee V$ be the open cover $\{\, U_i \cap V_j \mid 1 \le i \le k;\, 1 \le j \le l \,\}$. Now for an open cover U let $N(U)$ be the minimum of the cardinality of an open subcover of U and let $h(U) := \log N(U)$. The following limit exists (Petersen 1983, pp. 264–5):

$$h(U, T) \;:=\; \lim_{n \to \infty} \frac{h\big(U \vee T^{-1}(U) \vee \cdots \vee T^{-n+1}(U)\big)}{n},$$

and the topological entropy is

$$S_{\mathrm{top,A}}(X, d, T) \;:=\; \sup_{U} h(U, T). \tag{21}$$

$h(U, T)$ measures how the open cover U spreads out under the dynamics of the system. Hence $S_{\mathrm{top,A}}(X, d, T)$ is a measure for the highest possible spreading of an open cover under the dynamics of the system. In other words, the topological entropy measures how the map T scatters states in phase space (Petersen 1983, p. 266). Note that this interpretation does not involve any probabilistic notions.

Having positive topological entropy is often linked to chaotic behavior. For a compact phase space a positive topological entropy indicates that relative

[34] T is required to be surjective because only then it holds that for any open cover U also $T^{-t}(U)$ $(t \in \mathbb{N})$ is an open cover.

[35] Every open cover of a compact set has a finite subcover; hence we can assume that U is finite.

to some open covers the system scatters states in phase space. If scattering is regarded as indicating chaos, a positive entropy indicates that there is chaotic motion on some regions of the phase space. But there are many dynamical systems whose phase space is not compact; then $S_{\text{top, A}}(X, d, T)$ cannot be applied to distinguish chaotic from nonchaotic behavior.

How does the topological entropy relate to the Kolmogorov–Sinai entropy? Let (X, d, T) be a topological dynamical system where X is compact and T is continuous, and denote by $M_{(X,d)}$ the set of all measure-preserving dynamical systems (X, Σ, μ, T) where Σ is the Borel σ-algebra of (X, d).[36] Then (Goodwyn 1972)

$$S_{\text{top, A}}(X, d, T) = \sup_{(X,\Sigma,\mu,T) \in M_{(X,d)}} S_{\text{KS}}(X, \Sigma, \mu, T).$$

Furthermore, it is often said that the topological entropy is an *analogue* of the KS entropy (e.g. Bowen 1970, p. 23; Petersen 1983, p. 264), but without providing an elaboration of the notion of analogy at work. An analogy is more than a similarity. Hesse (1963) distinguishes two kinds of analogy, material and formal. Two objects stand in material analogy if they share certain intrinsic properties; they stand in formal analogy if they are described by the same mathematical expressions but do not share any other intrinsic properties (see also Polyá 1954). This leaves the question of what it means for *definitions* to be analogous. We say that definitions are *materially/formally analogous* iff there is a material/formal analogy between the objects appealed to in the definition.

The question then is whether $S_{\text{top, A}}(X, d, T)$ is analogous to the KS entropy. Clearly, they are formally analogous: Relate open covers U to partitions α, $U \vee V$ to $\alpha \vee \beta$, and $h(U)$ to $\mathrm{H}(\alpha)$. Then, $h(U, T) = \lim_{n\to\infty}\big(U \vee T^{-1}(U) \vee \cdots \vee T^{-n+1}(U)\big)/n$ corresponds to $\mathrm{H}(\alpha, T) = \lim_{n\to\infty} \mathrm{H}\big(\alpha \vee T^{-1}(\alpha) \vee \cdots \vee T^{-n+1}(\alpha)\big)/n$, and $S_{\text{top, A}}(X, d, T) = \sup_U h(U, T)$ corresponds to $S_{\text{KS}}(X, \Sigma, \mu, T) = \sup_\alpha h(\alpha, T)$. However, these definitions are not materially analogous. First, $\mathrm{H}(\alpha)$ can be interpreted as corresponding to the Shannon entropy, but $h(U)$ cannot because of the absence of probabilistic notions in its definition. This seems to link it more to the Hartley entropy, which also does not explicitly appeal to probabilities: we could regard $h(U)$ as the Hartley entropy of a subcover V of U with the least elements (cf. Sec. 3). However, this does not work because, except for the trivial open cover X, no open cover represents a set of *mutually exclusive* possibilities. Second, $h(U)$ measures the logarithm of the minimum number of elements of U needed to cover X, but $\mathrm{H}(\alpha)$ has no similar interpretation, e.g. it is not the logarithm of the number of elements of the partition α. Thus $S_{\text{top, A}}(X, d, T)$ and the KS entropy are not materially analogous.

[36] The Borel σ-algebra of a metric space (X, d) is the σ-algebra generated by all open sets of (X, d).

Bowen (1971) introduced two definitions which are equivalent to Adler, Konheim & McAndrew's definition. Because of lack of space, we cannot discuss them here (see Petersen 1983, pp. 264–7). What matters is that there is neither a formal nor a material analogy between the Bowen entropies and the KS entropy. Consequently, all we have is a formal analogy between the KS entropy and the topological entropy (21), and the claims in the literature that the KS entropy and the topological entropy are analogous are to some extent misleading. Moreover, we conclude that the topological entropy does not capture what notions of entropy are often believed to capture, such as information, and that none of the interpretations of the topological entropy is similar in interpretation to another notion of entropy.

6 Fractal geometry

It was not until the late 1960s that mathematicians and physicists started to systematically investigate irregular sets—sets that were traditionally considered as pathological. Mandelbrot coined the term *fractal* to denote these irregular sets. Fractals have been praised for providing a better representation of many natural phenomena than figures of classical geometry, but whether this is true remains controversial (cf. Falconer 1990, p. xiii; Mandelbrot 1983; Shenker 1994).

Fractal dimensions measure the irregularity of a set. We will discuss those fractal dimensions which are called *entropy dimensions*. The basic idea underlying fractal dimensions is that a set is a fractal *if* its fractal dimension is greater than its usual topological dimension (which is an integer). Yet the converse is not true: there are fractals where the relevant fractal dimensions do equal the topological dimension (Falconer 1990, pp. xx–xxi and Ch. 3; Mandelbrot 1983, Sec. 39).

Fractals arise in many different contexts. In particular, in dynamical-systems theory, scientists frequently focus on invariant sets, i.e. sets A for which $T_t(A) = A$ for all t, where T_t is the time-evolution. And *invariant sets are often fractals*. For instance, many dynamical systems have attractors, i.e. invariant sets which are asymptotically approached by neighboring states in the course of dynamic evolution. Attractors are sometimes fractals, e.g. the Lorenz and the Hénon attractor.

The following idea underlies the various definitions of a dimension of a set F. For each $\varepsilon > 0$ we take some sort of measurement of the set F at the level of resolution ε, yielding a real number $M_\varepsilon(F)$, and then we ask how $M_\varepsilon(F)$ behaves as ε goes to zero. If $M_\varepsilon(F)$ obeys the power law

$$M_\varepsilon(F) \approx c\varepsilon^{-s}, \tag{22}$$

for some constants c and s as ε goes to zero, then s is called the *dimension* of F. From (22) it follows that as ε goes to zero,

$$\log M_\varepsilon(F) \approx \log c - s \log \varepsilon.$$

Consequently,

$$s = \lim_{\varepsilon \to 0} \frac{\log M_\varepsilon(F)}{-\log \varepsilon}. \tag{23}$$

If $M_\varepsilon(F)$ does not obey a power law (22), one can consider instead of the limit in (23) the limit superior and the limit inferior (cf. Falconer 1990, p. 36).

Some fractal dimensions are called entropy dimensions, namely the box-counting dimension and the Rényi entropy dimensions. Let us start with the former. Assume that $\mathbb{R}^n$ is endowed with the usual Euclidean metric d. Given a nonempty and bounded subset $F \subseteq \mathbb{R}^n$, let $B_\varepsilon(F)$ be the smallest number of balls of diameter ε that cover F. The following limit, if it exists, is called the *box-counting dimension* but is also referred to as the *entropy dimension* (Edgar 2008, p. 112; Falconer 1990, p. 38; Hawkes 1974, p. 704; Mandelbrot 1983, p. 359):

$$\mathrm{Dim}_B(F) := \lim_{\varepsilon \to 0} \frac{\log B_\varepsilon(F)}{-\log \varepsilon}. \tag{24}$$

There are several equivalent formulations of the box-counting dimension. For instance, for $\mathbb{R}^n$ consider the boxes defined by the ε-coordinate mesh with elements:

$$\big[m_1\varepsilon, (m_1+1)\varepsilon\big) \times \cdots \times \big[m_n\varepsilon, (m_n+1)\varepsilon\big), \tag{25}$$

where $m_1, \ldots, m_n \in \mathbb{Z}$. Then if we define $B_\varepsilon(F)$ as the number of boxes in the ε-coordinate mesh that intersect F and again take the limit as $\varepsilon \to 0$, then the dimension obtained is equal to that in (24) (Falconer 1990, pp. 38–9). As we would expect, typically, for sets of classical geometry the box-counting dimension is integer-valued and for fractals it is non-integer-valued.[37]

For instance, how many squares of side-length $\varepsilon = 1/2^n$ are needed to cover the unit square $U = [0,1] \times [0,1]$? The answer is $B_{1/2^n}(U) = 2^{2n}$. Hence the box-counting dimension is $\lim_{n\to\infty}\big(\log 2^{2n} / -\log 1/2^n\big) = 2$. As another example we consider the *Cantor dust*, a well-known fractal. Starting with the unit interval $C_0 = [0,1]$, the set C_1 is obtained by removing the middle third from $[0,1]$, then C_2 is obtained by removing from C_1 the middle third of each of the intervals of C_1, and so on (see Fig. 1). The Cantor dust C is defined as $\bigcap_{k=0}^{\infty} C_k$. By setting $\varepsilon = 1/3^n$ and by considering the ε-coordinate mesh, we see that $B_{1/3^n}(C) = 2^n$. Hence

$$\mathrm{Dim}_B(C) := \lim_{n\to\infty} \frac{\log 2^n}{-\log 1/3^n} = \frac{\log 2}{\log 3} \approx 0.6309.$$

The box-counting dimension can readily be interpreted as the value of the coefficient s such that $B_\varepsilon(F)$ obeys the power law $B_\varepsilon(F) \approx c\varepsilon^{-s}$ as ε goes to

[37] The box-counting dimension has the shortcoming that even compact countable sets can have positive dimension. Therefore, the definition is often modified, but we will not go into details (cf. Edgar 2008, p. 213; Falconer 1990, pp. 37 and 44–6).

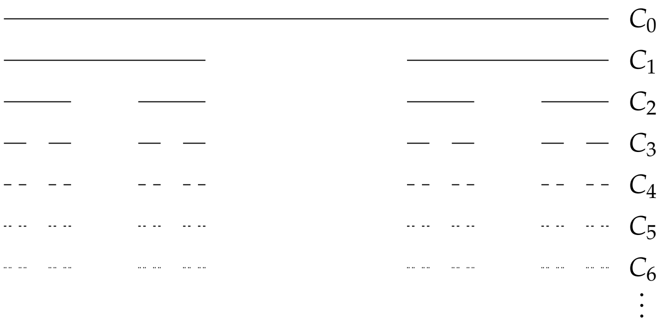

FIG. 1. The Cantor Dust.

zero. That is, it measures how 'spread out' the set is when examined at an infinitesimally small scale. However, this interpretation does not link to any entropy notions. So is there such a link?

Indeed there is (surprisingly, we have been unable to identify this argument in print).[38] Consider the box-counting dimension, where $B_\varepsilon(F)$ is the number of boxes in the ε-coordinate mesh that intersect F. Assume that each of these boxes represents a possible outcome and that we want to know what the actual outcome is. This assumption is sometimes natural. For instance, when we are interested in the dynamics on an invariant set F of a dynamical system we might ask: in which of the boxes of the ε-coordinate mesh that intersect F is the state of the system? Then the information gained when we learn which box the system occupies is quantified by the Hartley entropy $\log B_\varepsilon(F)$, as discussed in Sec. 3. Hence the box-counting dimension measures how the Hartley information grows as ε goes to zero. Thus there is a link between the box-counting dimension and the Hartley entropy.

Let us now turn to the Rényi entropy dimensions. Assume that $\mathbb{R}^n$ $(n \geq 1)$ is endowed with the usual Euclidean metric. Let $(\mathbb{R}^n, \Sigma, \mu)$ be a measure space where Σ contains all open sets of $\mathbb{R}^n$ and where $\mu(\mathbb{R}^n) = 1$. First, we need to introduce the notion of the support of the measure μ, which is the set on which the measure is concentrated. Formally, the *support* of μ is the smallest closed set X such that $\mu(\mathbb{R}^n \backslash X) = 0$. For instance, when measuring the dimension of a set F, the support of the measure is typically F. We assume that the support of μ is contained in a bounded region of $\mathbb{R}^n$.

Consider the ε-coordinate mesh of $\mathbb{R}^n$ (25). Let B_ε^i $(1 \leq i \leq m,\ m \in \mathbb{N})$ be the boxes that intersect the support of μ, and let $Z_{q,\varepsilon} := \sum_{i=1}^{m} \mu(B_\varepsilon^i)^q$. The Rényi

[38]Moreover, Hawkes (1974, p. 703) refers to $\log B_\varepsilon(F)$ as 'ε-entropy.' This is backed up by Kolmogorov & Tihomirov 1961; Kolmogorov & Tihomirov justify calling $\log B_\varepsilon(F)$ 'entropy' by an appeal to Shannon's Source-Coding Theorem. However, as they themselves observe, this justification relies on assumptions that have no relevance for scientific problems.

entropy dimension of order q $(-\infty < q < \infty,\ q \neq 1)$ is defined to be

$$\mathrm{Dim}_q \ := \ \lim_{\varepsilon\to 0}\left(\frac{1}{q-1}\frac{\log Z_{q,\varepsilon}}{\log \varepsilon}\right),$$

and the Rényi entropy dimension of order 1 is

$$\mathrm{Dim}_1 \ := \ \lim_{\varepsilon\to 0}\lim_{q\to 1}\left(\frac{1}{q-1}\frac{\log Z_{q,\varepsilon}}{\log \varepsilon}\right),$$

if the limit exists.

It is not hard to see that if $q < q'$, then $\mathrm{Dim}_{q'} \leq \mathrm{Dim}_q$ (cf. Beck & Schlögl 1995, p. 117). The cases $q = 0$ and $q = 1$ are of particular interest. Because $\mathrm{Dim}_0 = \mathrm{Dim}_B(\text{support}\,\mu)$, the Rényi entropy dimensions are a generalization of the box-counting dimension. And for $q = 1$ it can be shown (Rényi 1961) that

$$\mathrm{Dim}_1 \ = \ \lim_{\varepsilon\to 0}\frac{\sum_{i=1}^{m} -\mu(B_\varepsilon^i)\log\mu(B_\varepsilon^i)}{-\log\varepsilon}.$$

Since $\sum_{i=1}^{m} -\mu(B_\varepsilon^i)\log\mu(B_\varepsilon^i)$ is the *Shannon entropy* (cf. Sec. 3), Dim_1 is called the *information dimension* (Falconer 1990, p. 260; Ott 2002, p. 81).

The Rényi entropy dimensions are often referred to as 'entropy dimensions' *simpliciter* (e.g. Beck & Schlögl 1995, pp. 115–16). Before turning to a rationale for this name, let us state the motivation of the Rényi entropy dimensions that is usually given. The number q determines how much weight we assign to μ: the higher q, the greater the influence of boxes with larger measure. So the Rényi entropy dimensions measure the coefficient s such that $Z_{q,\varepsilon}$ obeys the power law $Z_{q,\varepsilon} \approx c\varepsilon^{-(1-q)s}$ as ε goes to zero. That is, Dim_q measures how 'spread out' the support of μ is when it is examined at an infinitesimally small scale and when the weight of the measure is q (Beck & Schlögl 1995, p. 116; Ott 2002, pp. 80–5). Consequently, when the Rényi entropy dimensions differ for different q, this is a sign of a *multifractal*, i.e. a set with different scaling behavior for different q (see Falconer 1990, pp. 254–64). This motivation does not refer to entropy notions.

Yet there is an obvious connection of the Rényi entropy dimensions for $q > 0$ to the Rényi entropies (cf. Sec. 3).[39] We proceed analogously to the case of the box-counting dimension. Namely, assume that each of the boxes of the ε-coordinate mesh which intersect the support of μ represent a possible outcome. Further, assume that the probability that the outcome is in the box B_i is $\mu(B_i)$. Then the information gained when we learn which box the system occupies can be quantified by the Rényi entropies H_q. Consequently, each Rényi entropy dimension for $q \in (0,\infty)$ measures how the information grows as ε goes to zero. For $q = 1$ we get a measure of how the Shannon information grows as ε goes to zero.

[39]Surprisingly, we have not found this motivation in print.

7 Conclusion

This essay has been concerned with some of the most important notions of entropy. The interpretations of these entropies have been discussed and their connections have been clarified. Two points deserve attention. First, all notions of entropy discussed in this essay, except the thermodynamic and the topological entropy, can be understood as variants of some information-theoretic notion of entropy. However, this should not distract us from the fact that different notions of entropy have different meanings and play different roles. Second, there is no preferred interpretation of the probabilities that figure in the different notions of entropy. The probabilities occurring in information-theoretic entropies are naturally interpreted as epistemic probabilities, but ontic probabilities are not ruled out. The probabilities in other entropies, for instance the different Boltzmann entropies, are most naturally understood ontically. So when considering the relation between entropy and probability there are no simple and general answers, and a careful case-by-case analysis is the only way forward.

Acknowledgments

We would like to thank David Lavis, Robert Batterman, Jeremy Butterfield, and two anonymous referees for comments on earlier drafts of this essay. We are grateful to Claus Beisbart and Stephan Hartmann for editing this book.

REFERENCES

Abrams, M. (2000). Short-run mechanistic probability. Talk given at Philosophy of Science Association conference, November 2000. ⟨http://members.logical.net/~marshall⟩.

Accardi, L. & Cecchini, C. (1982). Conditional expectations in von Neumann algebras and a theorem of Takesaki. *Journal of Functional Analysis* **45**, 245–73.

Adler, R., Konheim, A. & McAndrew, A. (1965). Topological entropy. *Transactions of the American Mathematical Society* **114**, 309–19.

Aharonov, Y., Anandan, J. & Vaidman, L. (1993). Meaning of the wave function. *Physical Review A* **47**, 4616–26.

Albert, D. Z. (1992). *Quantum Mechanics and Experience*. Cambridge, Mass.: Harvard University Press.

——— (2000). *Time and Chance*. Cambridge, Mass.: Harvard University Press.

Albert, M. (1992). Die Falsifikation statistischer Hypothesen. *Journal for General Philosophy of Science* **23**, 1–32.

——— (2002). Resolving Neyman's Paradox. *British Journal for the Philosophy of Science* **53**, 69–76.

Araki, H. (1964). Type of von Neumann algebra associated with free field. *Progress in Theoretical Physics* **32**, 956–65.

Bacciagaluppi, G. (2005). A conceptual introduction to Nelson's mechanics. In *Endophysics, Time, Quantum and the Subjective: Proceedings of the ZiF Interdisciplinary Research Workshop* (eds R. Buccheri, A. C. Elitzur & M. Saniga), pp. 367–88. Singapore: World Scientific.

——— (2009). Is logic empirical? In *Handbook of Quantum Logic and Quantum Structures: Quantum Logic* (eds K. Engesser, D. M. Gabbay & D. Lehmann), pp. 49–78. Amsterdam: Elsevier.

——— & Dickson, M. (1999). Dynamics for modal interpretations. *Foundations of Physics* **29**, 1165–1201.

——— & Valentini, A. (2009). *Quantum Theory at the Crossroads: Reconsidering the 1927 Solvay Conference*. Cambridge: Cambridge University Press.

Bailer-Jones, D. M. (2002). Models, metaphors, and analogies. In *The Blackwell Companion to the Philosophy of Science* (eds P. Machamer & M. Silberstein), pp. 108–27. Oxford: Blackwell.

——— (2003). When scientific models represent. *International Studies in the Philosophy of Science* **17**, 59–75.

Ballentine, L. (1998). *Quantum Mechanics: A Modern Development*. Singapore: World Scientific.

Barnum, H., Caves, C., Finkelstein, J., Fuchs, C. & Schack, R. (2000). Quantum probability from decision theory? *Proceedings of the Royal Society of London A* **456**, 1175–82.

Barrett, J. (1999). *The Quantum Mechanics of Minds and Worlds*. Oxford: Oxford University Press.

——— (2007). Information processing in generalized probabilistic theories. *Physical Review A* **75**, 032304.

———, Linden, N., Massar, S., Pironio, S., Popescu, S. & Roberts, D. (2005). Non-local correlations as an information-theoretic resource. *Physical Review A* **71**, 022101.

——— & Pironio, S. (2005). Popescu–Rohrlich correlations as a unit of nonlocality. *Physical Review Letters* **95**, 140401.

Bashkirov, A. G. (2006). Rényi entropy as a statistical entropy for complex systems. *Theoretical and Mathematical Physics* **149**, 1559–73.

Bassi, A. & Ghirardi, G.C. (2003). Dynamical reduction models. *Physics Reports* **379**, 257–426.

Batterman, R. & White, H. (1996). Chaos and algorithmic complexity. *Foundations of Physics* **26**, 307–36.

Bayes, T. (1763). Essay towards solving a problem in the doctrine of chances. *Philosophical Transactions of the Royal Society of London* **53**, 370–418. Repr. in *Biometrika* **45** (1958), 293–315.

Beck, C. & Schlögl, F. (1995). *Thermodynamics of Chaotic Systems*. Cambridge: Cambridge University Press.

Bell, J. L. & Machover, M. (1977). *A Course in Mathematical Logic*. Amsterdam: North-Holland.

Bell, J. S. (1964). On the Einstein–Podolsky–Rosen Paradox. *Physics* **1**, 195–200. Repr. in Bell 1987c, pp. 14–21.

——— (1966). On the problem of hidden variables in quantum mechanics. *Reviews of Modern Physics* **38**, 447–52. Repr. in Bell 1987c, pp. 1–13.

——— (1977). Free variables and local causality. *Lettres épistémologiques* **15**. Repr. in *Dialectica* **39** (1985), 103–6, and in Bell 1987c, pp. 100–4. Page references are to the second reprint.

——— (1980). Atomic-cascade photons and quantum-mechanical nonlocality. *Comments on Atomic and Molecular Physics* **9**, 121–6. Repr. in Bell 1987c, pp. 105–10. Page references are to the reprint.

——— (1987a). Are there quantum jumps? In *Schrödinger: Centenary Celebration of a Polymath* (ed. W. M. Kilmister), pp. 41–52. Cambridge: Cambridge University Press. Repr. in Bell 1987c, pp. 201–12.

——— (1987b). On wave packet reduction in the Coleman–Hepp model. In Bell 1987c, pp. 45–51.

——— (1987c). *Speakable and Unspeakable in Quantum Mechanics*. Cambridge:

Cambridge University Press. 2nd edn 2004.

Beller, M. (1990). Born's probabilistic interpretation: A case study of 'concepts in flux.' *Studies in History and Philosophy of Science* **21**, 563–88.

Beltrametti, E. G. & Cassinelli, G. (1981). *The Logic of Quantum Mechanics*. Reading, Mass.: Addison–Wesley.

Bennett, J. (2003). *A Philosophical Guide to Conditionals*. Oxford: Oxford University Press.

Berger, A. (2001). *Chaos and Chance: An Introduction to Stochastic Aspects of Dynamics*. New York: de Gruyter.

Berkovitz, J., Frigg, R. & Kronz, F. (2006). The ergodic hierarchy, randomness and Hamiltonian chaos. *Studies in History and Philosophy of Modern Physics* **37**, 661–91.

Berndl, K. (1996). Global existence and uniqueness of Bohmian mechanics. In *Bohmian Mechanics and Quantum Theory: An Appraisal* (eds J. T. Cushing, A. Fine & S. Goldstein), pp. 77–86. Dordrecht: Kluwer.

———, Dürr, D., Goldstein, S., Peruzzi, P. & Zanghì, N. (1995). On the global existence of Bohmian mechanics. *Communications in Mathematical Physics* **173**, 647–73.

Bernoulli, D. (1738). *Hydrodynamica*. Basel: J. R. Dulsecker. Excerpt transl. into English by J. P. Berryman in *The Kinetic Theory of Gases: An Anthology of Classic Papers with Historical Commentary* (ed. S. G. Brush), pp. 57–66 (London: Imperial College Press, 2003).

Bernoulli, J. (1713). *Ars Conjectandi*. Basel: Thurnisius. Repr. in *Die Werke von Jakob Bernoulli*, Vol. 3 (ed. B. L. van der Waerden), edited by Naturforschende Gesellschaft in Basel (Basel: Birkhäuser, 1975).

Bigelow, J. C. (1976). Possible worlds foundations for probability. *Journal of Philosophical Logic* **5**, 299–320.

Bitbol, M. (1996). *Schrödinger's Philosophy of Quantum Mechanics*. Dordrecht: Kluwer.

Blackwell, D. & Dubins, L. (1962). Merging of opinions with increasing information. *Annals of Statistical Mathematics* **33**, 882–6.

Bohm, D. (1952). A suggested interpretation of the quantum theory in terms of 'hidden' variables, I and II. *Physical Review* **85**, 166–79, 180–93.

——— & Hiley, B. (1993). *The Undivided Universe: An Ontological Interpretation of Quantum Theory*. London: Routledge.

——— & Vigier, J.-P. (1954). Model of the causal interpretation in terms of a fluid with irregular fluctuations. *Physical Review Letters* **96**, 208–16.

Bohr, N. (1913). On the constitution of atoms and molecules, Part I. *Philosophical Magazine* **26**, 1–24.

Boltzmann, L. (1868). Studien über das Gleichgewicht der lebendigen Kraft zwischen bewegten materiellen Punkten. *Wiener Berichte* **58**, 517–60. Repr. in

Boltzmann 1909, Vol. I, pp. 49–96.

——— (1871). Einige allgemeine Sätze über Wärmegleichgewicht. *Wiener Berichte* **63**, 679–711. Repr. in Boltzmann 1909, Vol. I, pp. 259–87.

——— (1872). Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen. *Wiener Berichte* **66**, 275–370. Repr. in Boltzmann 1909, Vol. I, pp. 316–402.

——— (1877). Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung resp. den Sätzen über das Wärmegleichgewicht. *Wiener Berichte* **76**, 373–435. Repr. in Boltzmann 1909, Vol. II, pp. 164–223.

——— (1894). On the application of the determinantal relation to the kinetic theory of gases. Repr. in Boltzmann 1909, Vol. III, pp. 520–5.

——— (1909). *Wissenschaftliche Abhandlungen*, Vols I–III. Leipzig: Barth.

——— (1964). *Lectures on Gas Theory.* Berkeley, Calif.: University of California Press.

——— (1974). *Theoretical Physics and Philosophical Problems: Selected Writings*, Vol. 5. Dordrecht & Boston, Mass.: Reidel.

——— & Nabl, J. (1905). Kinetische Theorie der Materie. In *Encyklopädie der Mathematischen Wissenschaften mit Einschluß ihrer Anwendungen*, Vol. V-1 (ed. F. Klein), pp. 493–557. Leipzig: Teubner.

Borek, R. (1985). Representations of the current algebra of a charged massless Dirac field. *Journal of Mathematical Physics* **26**, 339–44.

Born, M. (1926a). Zur Quantenmechanik der Stoßvorgänge. *Zeitschrift für Physik* **37**, 863–7.

——— (1926b). Quantenmechanik der Stoßvorgänge. *Zeitschrift für Physik* **38**, 803–27.

——— (1964). The statistical interpretations of quantum mechanics. In *Nobel Lectures: Physics (1942–1962)* (ed. Nobelstiftelsen), pp. 256–67. Amsterdam: Elsevier.

Bowen, R. (1970). Topological entropy and Axiom A. In *Global Analysis: Proceedings of the Symposium of Pure Mathematics* **14**, 23–41. Providence, R.I.: American Mathematical Society.

——— (1971). Periodic points and measures for Axiom A diffeomorphisms. *Transactions of the American Mathematical Society* **154**, 377–97.

Bratteli, O. & Robinson, D.W. (1987). *Operator Algebras and Quantum Statistical Mechanics*, Vol. 1, 2nd edn. Berlin, Heidelberg, New York: Springer.

——— & ——— (1997). *Operator Algebras and Quantum Statistical Mechanics*, Vol. 2, 2nd edn. Berlin, Heidelberg, New York: Springer.

Bricmont, J. (1995). Science of chaos or chaos in science? *Physicalia* **17**, 159–208.

——— (2001). Bayes, Boltzmann and Bohm: Probabilities in Physics. In *Chance in Physics: Foundations and Perspectives* (eds J. Bricmont, D. Dürr, M. C.

Galavotti, G.C. Ghirardi, F. Petruccione & N. Zanghì), pp. 3–21. Berlin & New York: Springer.

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 Modern Physics* **40**, 174–91.

——— & Timpson, C. G. (2006). Why special relativity should not be a template for a fundamental reformulation of quantum mechanics. In Demopoulos & Pitowsky 2006, pp. 29–41.

Brush, S. G. (1976). *The Kind of Motion We Call Heat: A History of the Kinetic Theory of Gases in the 19th Century*, Vol. 6. Amsterdam & New York: North-Holland.

——— & Hall, N. S. (2003). *The Kinetic Theory of Gases: An Anthology of Classic Papers with Historical Commentary*, Vol. 1. London: Imperial College Press.

Bub, J. (1977). Von Neumann's Projection Postulate as a probability conditionalization rule in quantum mechanics. *Journal of Philosophical Logic* **6**, 381–90.

——— (1997). *Interpreting the Quantum World*. Cambridge: Cambridge University Press.

——— (2007a). Quantum information and computation. In Butterfield & Earman 2007, pp. 555–660.

——— (2007b). Quantum probabilities as degrees of belief. *Studies in History and Philosophy of Modern Physics* **38**, 232–54.

——— & Pitowsky, I. (2010). Two dogmas of quantum mechanics. In *Many Worlds? Everett, Quantum Theory & Reality* (eds S. Saunders, J. Barrett, A. Kent & D. Wallace), pp. 433–59. Oxford: Oxford University Press. *arXiv e-print quant-ph/0712.4258*.

——— & Stairs, A. (2009). Contextuality and nonlocality in 'no signaling' theories. *Foundations of Physics* **39**, 690–711. *arXiv e-print quant-ph/0903.1462*.

Buchholz, D., D'Antoni, C. & Fredenhagen, K. (1987). The universal structure of local algebras. *Communications in Mathematical Physics* **111**, 123–35.

——— & Doplicher, S. (1984). Exotic infrared representations of interacting systems. *Annales de l'Institut Henri Poincaré : Physique théorique* **32**, 175–84.

———, ——— & Longo, R. (1986). On Noether's Theorem in quantum field theory. *Annals of Physics* **170**, 1–17.

Busch, P. (2003). Quantum states and generalized observables: A simple proof of Gleason's Theorem. *Physical Review Letters* **91** (12), 120403.

———, Grabowski, M. & Lahti, P. (1995). *Operational Quantum Physics*. Berlin: Springer.

Butterfield, J. & Earman, J. (2007). *Philosophy of Physics*. Handbook of the Philosophy of Science. Amsterdam & Oxford: North-Holland.

Cabello, A. (2003). Kochen–Specker Theorem for a single qubit using positive-operator-valued measures. *Physical Review Letters* **90**, 190401.

Callender, C. (1997). What is 'the problem of the direction of time'? *Philosophy of Science* **64** (4), Supplement, S 223–34.

——— (1999). Reducing thermodynamics to statistical mechanics: The case of entropy. *Journal of Philosophy* **96** (7), 348–73.

——— (2004). Measures, explanations, and the past: Should 'special' initial conditions be explained? *British Journal for the Philosophy of Science* **55** (2), 195–217.

——— (2007). The emergence and interpretation of probability in Bohmian mechanics. *Studies in History and Philosophy of Modern Physics* **38**, 351–70.

——— (2010). The Past Hypothesis meets gravity. In Ernst & Hüttemann 2010, pp. 34–58.

——— & Cohen, J. (2006). There is no special problem about scientific representation. *Theoria* **55**, 7–25.

——— & ——— (2010). Special sciences, conspiracy and the better Best System Account of lawhood. *Erkenntnis* **73**, 427–47.

Campbell, L. & Garnett, W. (1884). *The Life of James Clerk Maxwell*. London: Macmillan.

Cartwright, N. (1983). *How the Laws of Physics Lie*. Oxford: Oxford University Press.

——— (1999). *The Dappled World: A Study of the Boundaries of Science*. Cambridge: Cambridge University Press.

Caticha, A. & Giffin, A. (2006). Updating probabilities. In *Bayesian Inference and Maximum Entropy Methods in Science and Engineering* (ed. A. Mohammad-Djafari), AIP Conference Proceedings, Vol. 872, pp. 31–42. *arXiv e-print physics/0608185v1*.

Caves, C. M., Fuchs, C. A. & Schack, R. (2002). Quantum probabilities as Bayesian probabilities. *Physical Review A* **65**, 022305.

———, ——— & ——— (2007). Subjective probability and quantum certainty. *Studies in History and Philosophy of Modern Physics* **38**, 255–74.

———, ———, Manne, K. K. & Renes, J. M. (2004). Gleason-type derivations of the quantum probability rule for generalized measurements. *Foundations of Physics* **34** (2), 193–209.

Clauser, J. F., Horne, M. A., Shimony, A. & Holt, R. A. (1969). Proposed experiment to test local hidden-variable theories. *Physical Review Letters* **23**, 880–4.

Clifton, R. (1993). Getting contextual and nonlocal elements-of-reality the easy way. *American Journal of Physics* **61**, 443–7.

——— (1995). Independently motivating the Kochen–Dieks modal interpretation of quantum mechanics. *British Journal for the Philosophy of Science* **46**, 33–57.

——— (2000). The modal interpretation of algebraic quantum field theory.

Physics Letters A **271**, 167–77.

——— & Halvorson, H. (2001). Are Rindler quanta real? Inequivalent particle concepts in quantum field theory. *British Journal for the Philosophy of Science* **52**, 417–70.

Cohen, J. & Callender, C. (2009). A better Best System Account of lawhood. *Philosophical Studies* **145**, 1–34.

Conway, J. H. & Kochen, S. (2006). The Free Will Theorem. *Foundations of Physics* **36**, 1441–73.

——— & ——— (2009). The Strong Free Will Theorem. *Notices of the American Mathematical Society* **56**, 226–32.

Cooke, R., Keane, M. & Moran, W. (1984). An elementary proof of Gleason's Theorem. *Mathematical Proceedings of the Cambridge Philosophical Society* **98**, 117–28.

Cornfeld, I., Fomin, S. & Sinai, Y. (1982). *Ergodic Theory*. Berlin: Springer.

da Costa, N. C. A. & French, S. (1990). The model-theoretic approach in philosophy of science. *Philosophy of Science* **57**, 248–65.

Daley, D. J. & Vere-Jones, D. (1988). *An Introduction to the Theory of Point Processes*, Vol. 2. Berlin: Springer. 2nd edn 2008.

Davey, K. (2008). The justification of probability measures in statistical mechanics. *Philosophy of Science* **75** (1), 28–44.

Davies, E. B. (1976). *Quantum Theory of Open Systems*. New York: Academic Press.

Davies, P. C. W. (1974). *The Physics of Time Asymmetry.* Berkeley, Calif.: University of California Press.

de Broglie, L. (1928). La nouvelle dynamique des quanta. In *Electrons et photons : Rapports et discussions du cenquième Conseil de Physique*, pp. 105–41. Paris: Gauthier-Villars.

——— (2009 [1928]). The new dynamics of quanta. Transl. in *Quantum Mechanics at the Crossroads: Reconsidering the 1927 Solvay Conference* (eds G. Bacciagaluppi & A. Valentini), pp. 341–71. Cambridge: Cambridge University Press.

de Finetti, B. (1931a). Probabilismo. *Logos* **14**, 163–219. Translated as 'Probabilism: A critical essay on the theory of probability and on the value of science,' in *Erkenntnis* **31** (1989), pp. 169–223.

——— (1931b). Sul significato soggettivo della probabilità. *Fundamenta Mathematica* **17**, 298–329.

——— (1964). Foresight: Its logical laws, its subjective sources. In *Studies in Subjective Probability* (eds H. E. Kyburg & H. E. Smokler), pp. 93–158. New York: John Wiley & Sons.

——— (1972). *Probability, Induction and Statistics*. New York: John Wiley & Sons.

——— (1974). Bayesianism: Its unifying role for both the foundations and applications of statistics. *International Statistical Review* **42**, 117–30.

Demopoulos, W. & Pitowsky, I. (eds) (2006). *Physical Theory and its Interpretation: Essays in Honor of Jeffrey Bub*. Western Ontario Series in Philosophy of Science. Dordrecht: Springer.

de Muynck, W. (2007). POVMs: A small but important step beyond standard quantum mechanics. In *Beyond the Quantum* (eds T. Nieuwenhuizen, B. Mehmani, V. Špička, M. Aghdami & A. Khrennikov), pp. 69–79. Singapore: World Scientific.

Denbigh, K. G. & Denbigh, J. (1985). *Entropy in Relation to Incomplete Knowledge*. Cambridge: Cambridge University Press.

de Oliveira, C. R. & Werlang, T. (2007). Ergodic hypothesis in classical statistical mechanics. *Revista Brasileira de Ensino de Física* **29**, 189–201.

Deutsch, D. (1999). Quantum theory of probability and decisions. *Proceedings of the Royal Society of London A* **455**, 3129–37. *arXiv e-print quant-ph/0990.6015*.

Dickson, M. (1995). An empirical reply to empiricism: Protective measurement opens the door for quantum realism. *Philosophy of Science* **62**, 122–40.

——— (1998). *Quantum Chance and Nonlocality*. Cambridge: Cambridge University Press.

——— (2001). Quantum logic is alive $\wedge$ (it is true $\vee$ it is false). *Philosophy of Science* **68**, Supplement, S 274–87.

——— (2007). Non-relativistic quantum mechanics. In Butterfield & Earman 2007, pp. 275–415.

——— & Dieks, D. (2009). Modal interpretations of quantum mechanics. In *The Stanford Encyclopedia of Philosophy* (ed. E. N. Zalta). ⟨http://plato.stanford.edu/archives/spr2009/entries/qm-modal⟩.

Dieks, D. (2000). Consistent histories and relativistic invariance in the modal interpretation of quantum physics. *Physics Letters A* **265**, 317–25.

——— (2007). Probability in modal interpretations of quantum mechanics. *Studies in History and Philosophy of Modern Physics* **38**, 292–310.

Doplicher, S., Figliolini, F. & Guido, D. (1984). Infrared representations of free Bose fields. *Annales de l'Institut Henri Poincaré : Physique Théorique* **41**, 49–62.

Dorato, M. & Esfeld, M. (2010). GRW as an ontology of dispositions. *Studies in History and Philosophy of Modern Physics* **41**, 41–9.

Drory, A. (2008). Is there a reversibility paradox? Recentering the debate on the thermodynamic time arrow. *Studies in History and Philosophy of Modern Physics* **39**, 889–913.

Dunn, M. (1993). Star and perp: Two treatments of negation. *Philosophical Perspectives* **7** (*Language and Logic*, ed. J. E. Tomberlin), pp. 331–57. Atascadero, Calif.: Ridgeview.

Dürr, D., Goldstein, S. & Zanghì, N. (1992a). Quantum chaos, classical randomness, and Bohmian mechanics. *Journal of Statistical Physics* **68**, 259–70.

———, ——— & ——— (1992b). Quantum equilibrium and the origin of absolute uncertainty. *Journal of Statistical Physics* **67**, 843–907.

———, ——— & ——— (1996). Bohmian mechanics as the foundation of quantum mechanics. In *Bohmian Mechanics and Quantum Theory: An Appraisal* (eds J. Cushing, A. Fine & S. Goldstein), pp. 21–44. Dordrecht: Kluwer.

Duwell, A. (2007). Reconceiving quantum mechanics in terms of information-theoretic constraints. *Studies in History and Philosophy of Modern Physics* **38**, 181–201.

Eagle, A. (2004). Twenty-one arguments against propensity analyses of probability. *Erkenntnis* **60**, 371–416.

———, ed. (2010). *Philosophy of Probability: Contemporary Readings*. London: Routledge.

Earman, J. (1971). Laplacian determinism, or Is this any way to run a universe? *Journal of Philosophy* **68**, 729–44.

——— (1986). *A Primer on Determinism*. Dordrecht: Reidel.

——— (1987). The problem of irreversibility. In *PSA 1986: Proceedings of the 1986 Biennial Meeting of the Philosophy of Science Association*, Vol. II: *Symposia and Invited Papers* (eds A. Fine & P. Machamer), pp. 226–33. East Lansing, Mich.: Philosophy of Science Association.

——— (1992). *Bayes or Bust? A Critical Examination of Bayesian Confirmation Theory.* Cambridge, Mass.: MIT Press.

——— (2004). Determinism: What we have learned and what we still don't know. In *Freedom and Determinism* (eds J. K. Campbell *et al.*), pp. 21–46. Cambridge, Mass.: MIT Press.

——— (2006). The 'Past Hypothesis': Not even false. *Studies in History and Philosophy of Modern Physics* **37** (3), 399–430.

——— (2007). Aspects of determinism in modern physics. In Butterfield & Earman 2007, pp. 1369–1434.

——— (2008a). How determinism can fail in classical physics and how quantum physics can (sometimes) provide a cure. *Philosophy of Science* **75**, 817–29.

——— (2008b). Superselection rules for philosophers. *Erkenntnis* **69**, 377–414.

——— (2009). Essential self-adjointness: Implications for determinism and the classical–quantum correspondence. *Synthese* **169**, 27–50.

——— & Rédei, M. (1996). Why ergodic theory does not explain the success of equilibrium statistical mechanics. *British Journal for the Philosophy of Science* **47**, 63–78.

——— & Ruetsche, L. (2005). Relativistic invariance and modal interpretations. *Philosophy of Science* **72**, 557–83.

Eckmann, J.-P. & Ruelle, D. (1985). Ergodic theory of chaos and strange attractors. *Reviews of Modern Physics* **57**, 617–54.

Edgar, G. (2008). *Measure, Topology, and Fractal Geometry*. New York: Springer.

Ehrenfest, P. & Ehrenfest-Afanassjewa, T. (1911). Begriffliche Grundlagen der statistischen Auffassung in der Mechanik. In *Encyklopädie der Mathematischen Wissenschaften mit Einschluß ihrer Anwendungen*, Vol. IV-4.II (eds F. Klein & C. Müller). English transl.: *The Conceptual Foundations of the Statistical Approach in Mechanics*, Ithaca, N.Y.: Cornell University Press, 1959.

Einstein, A. (1905). Über die von der molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen. *Annalen der Physik* **17**, 549–60.

Elby, A. & Bub, J. (1994). Triorthogonal uniqueness theorem and its relevance to the interpretation of quantum mechanics. *Physical Review A* **49**, 4213–16.

Elga, A. (2004). Infinitesimal chances and the laws of nature. *Australasian Journal of Philosophy* **82**, 67–76.

Emch, G. G. (1972). *Algebraic Methods in Statistical Mechanics and Quantum Field Theory*. New York: John Wiley & Sons.

——— (2007a). Models and the dynamics of theory-building in physics I: Modeling strategies. *Studies in History and Philosophy of Modern Physics* **38** (3), 558–85.

——— (2007b). Models and the dynamics of theory-building in physics II: Case studies. *Studies in History and Philosophy of Modern Physics* **38** (4), 683–723.

——— & Liu, C. (2002). *The Logic of Thermostatistical Physics*. Heidelberg & Berlin: Springer.

Engesser, K., Gabbay, D. M. & Lehmann, D. (2007). *Handbook of Quantum Logic and Quantum Structures: Quantum Structures*. Amsterdam: Elsevier.

Ernst, G. & Hüttemann, A. (eds) (2010). *Time, Chance, and Reduction: Philosophical Aspects of Statistical Mechanics*. Cambridge: Cambridge University Press.

Everett, H., III (1957). 'Relative state' formulation of quantum mechanics. *Review of Modern Physics* **29**, 454–62.

Falconer, K. (1990). *Fractal Geometry: Mathematical Foundations and Applications*. New York: John Wiley & Sons.

Falkenburg, B. & Muschik, W. (1998). *Models, Theories and Disunity in Physics*. Frankfurt am Main: Klostermann. *Philosophia Naturalis* **35** (Special Issue).

Feller, W. (1968). *An Introduction to Probability Theory and its Applications*, Vols 1 & 2, 3rd edn. New York: John Wiley & Sons.

Fetzer, J. (1971). Dispositional probabilities. *Boston Studies in the Philosophy of Science* **8**, 473–82.

——— (1974). A single case propensity theory of explanation. *Synthese* **28**, 171–98.

——— (1981). Probability and explanation. *Synthese* **48**, 371–408.

——— (1983a). Probabilistic explanations. In *PSA 1982: Proceedings of the 1982 Biennial Meeting of the Philosophy of Science Association*, Vol. 2: *Symposia and Invited Papers* (eds P. D. Asquith & T. Nickles), pp. 194–207. East Lansing, Mich.: Philosophy of Science Association.

——— (1983b). Probability and objectivity in deterministic and indeterministic situations. *Synthese* **57**, 367–86.

Feynman, R. (1967). *The Character of Physical Law*. Cambridge, Mass.: MIT Press.

Fine, A. (1982a). Hidden variables, joint probability, and the Bell Inequalities. *Physical Review Letters* **48**, 291–5.

——— (1982b). Joint distributions, quantum correlations, and commuting observables. *Journal of Mathematical Physics* **23**, 1306–9.

Fine, T. L. (1973). *Theories of Probability: An Examination of Foundations*. New York & London: Academic Press.

Foulis, D. J. & Bennett, M. K. (1994). Effect algebras and unsharp quantum logic. *Foundations of Physics* **24**, 1325–46.

——— & Greechie, R. J. (2007). Quantum logic and partially ordered abelian groups. In Engesser *et al.* 2007, pp. 215–84.

Friedman, K. & Shimony, A. (1971). Jaynes's maximum entropy prescription and probability theory. *Journal of Statistical Physics* **3**, 381–4.

Frigg, R. (2004). In what sense is the Kolmogorov–Sinai entropy a measure for chaotic behaviour?—Bridging the gap between dynamical systems theory and communication theory. *British Journal for the Philosophy of Science* **55**, 411–34.

——— (2006a). Chaos and randomness: An equivalence proof of a generalised version of the Shannon entropy and the Kolmogorov–Sinai entropy for Hamiltonian dynamical systems. *Chaos, Solitons and Fractals* **28**, 26–31.

——— (2006b). Scientific representation and the semantic view of theories. *Theoria* **55**, 49–65.

——— (2008). A field guide to recent work on the foundations of statistical mechanics. In *The Ashgate Companion to Contemporary Philosophy of Physics* (ed. D. Rickles), pp. 99–196. Aldershot & Burlington, Vt.: Ashgate.

——— (2009). Typicality and the approach to equilibrium in Boltzmannian statistical mechanics. *Philosophy of Science* **76**, Supplement, S 997–1008.

——— (2010a). Probability in Boltzmannian statistical mechanics. In Ernst & Hüttemann 2010, pp. 92–118.

——— (2010b). Why typicality does not explain the approach to equilibrium. In *Probabilities, Causes and Propensities in Physics* (ed. M. Suárez). Synthese Library, Vol. 347. Berlin: Springer, *to appear*.

——— & Hartmann, S. (eds) (2007). *Probabilities in Quantum Mechanics*. Special issue of *Studies in History and Philosophy of Modern Physics* **38**, 231–456.

——— & ——— (2009). Models in science. In *The Stanford Encyclopedia of Philosophy* (ed. E. N. Zalta). ⟨http://plato.stanford.edu/archives/sum2009/entries/models-science⟩.

——— & Hoefer, C. (2007). Probability in GRW theory. *Studies in History and Philosophy of Modern Physics* **38**, 371–89.

——— & ——— (2010). Determinism and chance from a Humean perspective. In *The Present Situation in the Philosophy of Science* (eds D. Dieks, W. González, S. Hartmann, M. Weber, F. Stadler & T. Uebel). Berlin & New York: Springer.

Frisch, M. (2007). Causation, counterfactuals and entropy. In *Causation, Physics, and the Constitution of Reality: Russell's Republic Revisited* (eds H. Price & R. Corry), pp. 351–95. Oxford: Oxford University Press.

Fuchs, C. A. (2001). Quantum foundations in the light of quantum information. In *Proceedings of the* NATO *Advanced Research Workshop on Decoherence and its Implications in Quantum Computation and Information Transfer* (eds A. Gonis & P. Turchi), pp. 38–82. Amsterdam: IOS Press. *arXiv e-print quant-ph/0106166*.

——— (2002a). Quantum mechanics as quantum information (and only a little more). *arXiv e-print quant-ph/0205039*.

——— (2002b). The anti-Växjö interpretation of quantum mechanics. *arXiv e-print quant-ph/0204146*.

——— (2003). *Notes on a Paulian Idea: Foundational, Historical, Anecdotal and Forward-looking Thoughts on the Quantum*. Växjö, Sweden: Växjö University Press. *arXiv e-print quant-ph/0105039*.

Gaifman, H. & Snir, M. (1980). Probabilities over rich languages, testing and randomness. *Journal of Symbolic Logic* **47**, 495–548.

Galavotti, M. C. (1991). The notion of subjective probability in the work of Ramsey and de Finetti. *Theoria* **57** (3), 239–59.

——— (2005). *Philosophical Introduction to Probability*. Stanford, Calif.: CSLI.

Garber, E. (1973). Aspects of the introduction of probability into physics. *Centaurus* **17**, 11–40.

———, Brush, S. G. & Everitt, C. W. F. (eds) (1986). *Maxwell on Molecules and Gases*. Cambridge, Mass.: MIT Press.

———, ——— & ——— (eds) (1995). *Maxwell on Heat and Statistical Mechanics: On 'Avoiding All Personal Enquiries' of Molecules.* Bethlehem, Pa. & London: Lehigh University Press.

Gardiner, C.W. (2004). *Handbook of Stochastic Methods for Physics, Chemistry and the Natural Sciences*, 3rd edn. Berlin etc.: Springer.

Ghirardi, G.C. (2009). Collapse theories. In *The Stanford Encyclopedia of Philosophy* (ed. E. N. Zalta). ⟨http://plato.stanford.edu/entries/qm-collapse⟩.

———, Rimini, A. & Weber, T. (1986). Unified dynamics for microscopic and macroscopic systems. *Physical Review D* **34**, 470–91.

Gibbs, J.W. (1902). *Elementary Principles in Statistical Mechanics: Developed with Especial Reference to the Rational Foundation of Thermodynamics*. New Haven, Conn.: Yale University Press. Repr. Mineola, N.Y.: Dover, 1960, and Woodbridge, Conn.: Ox Bow Press, 1981.

Giere, R. N. (1973). Objective single case probabilities and the foundation of statistics. In *Logic, Methodology and Philosophy of Science IV: Proceedings of the*

Fourth International Congress for Logic, Methodology and Philosophy of Science, Bucharest, 1971 (eds P. Suppes, L. Henkin, G. C. Moisil & A. Joja), pp. 467–83. Amsterdam: North-Holland.

——— (1988). *Explaining Science: A Cognitive Approach*. Chicago, Ill.: University of Chicago Press.

——— (2004). How models are used to represent. *Philosophy of Science* **71**, 742–52.

Gillespie, C. C. (1963). Intellectual factors in the background of analysis by probabilities. In *Scientific Change* (ed. A. C. Crombie), pp. 431–53, 499–502. London: Heinemann.

Gillies, D. A. (1971). A falsifying rule for probability statements. *British Journal for the Philosophy of Science* **22**, 231–61.

——— (1973). *An Objective Theory of Probability*. London: Methuen.

——— (2000a). *Philosophical Theories of Probability*. London: Routledge.

——— (2000b). Varieties of propensity. *British Journal for the Philosophy of Science* **51**, 807–35.

Gleason, A. M. (1957). Measures on the closed subspaces of a Hilbert space. *Journal of Mathematics and Mechanics* **6**, 885–93.

Goldstein, S. (2001). Boltzmann's approach to statistical mechanics. In *Chance in Physics: Foundations and Perspectives* (eds J. Bricmont, D. Dürr, M. C. Galavotti, G.C. Ghirardi, F. Petruccione & N. Zanghì), pp. 39–54. Berlin & New York: Springer.

——— (2006). Bohmian mechanics. In *Stanford Encyclopedia of Philosophy* (ed. E. N. Zalta). ⟨http://plato.stanford.edu/entries/qm-bohm⟩.

——— & Lebowitz, J. L. (2004). On the (Boltzmann) entropy of non-equilibrium systems. *Physica D: Nonlinear Phenomena* **193**, 53–66.

——— & Struyve, W. (2007). On the uniqueness of quantum equilibrium in Bohmian mechanics. *Journal of Statistical Physics* **128**, 1197–1209.

———, Tausk, D. V., Tumulka, R. & Zanghì, N. (2010). What does the Free Will Theorem actually prove? *Notices of the American Mathematical Society* **57**, 1451–3. *arXiv e-print quant-ph/0905.4641v1*.

Goodwyn, L. (1972). Comparing topological entropy with measure-theoretic entropy. *American Journal of Mathematics* **94**, 366–88.

Grad, H. (1961). The many faces of entropy. *Communications in Pure and Applied Mathematics* **14**, 323–54.

Graham, N. (1973). The measurement of relative frequency. In *The Many-Worlds Interpretation of Quantum Mechanics* (eds B. S. DeWitt & N. Graham), pp. 229–53. Princeton, N.J.: Princeton University Press.

Greaves, H. (2004). Understanding Deutsch's probability in a deterministic multiverse. *Studies in History and Philosophy of Modern Physics* **35**, 423–56.

——— (2007). The Everettian epistemic problem. *Studies in History and Philosophy of Modern Physics* **38** (1), 120–52.

——— & Myrvold, W. (2010). Everett and evidence. In Saunders *et al.* 2010, pp. 264–304.

Gregory, O. (1825). *Mathematics for Practical Men*. London: Baldwin, Cradock, and Joy. 3rd edn, revised and enlarged by H. Law (London: J. Weale, 1848).

Greiner, W., Neise, L. & Stücker, H. (1993). *Thermodynamik und Statistische Mechanik*. Leipzig: Harri Deutsch.

Grünbaum, A. (1963). *Philosophical Problems of Space and Time*. New York: Alfred A. Knopf.

Gudder, S. (2007). Quantum probability. In Engesser *et al.* 2007, pp. 121–46.

——— & Greechie, R. (2002). Sequential products on effect algebras. *Reports on Mathematical Physics* **49**, 87–111.

——— & Latrémolière, F. (2008). Characterization of the sequential product on quantum effects. *Journal of Mathematical Physics* **49**, 052106.

Guttman, Y. M. (1999). *The Concept of Probability in Statistical Physics*. Cambridge: Cambridge University Press.

Haag, R. (1996). *Local Quantum Physics*, 2nd edn. New York: Springer.

Hacking, I. (1975). *The Emergence of Probability*. Cambridge: Cambridge University Press.

——— (1990). *The Taming of Chance*. Cambridge: Cambridge University Press.

——— (2001). *An Introduction to Probability and Inductive Logic*. Cambridge: Cambridge University Press.

Hájek, A. (1996). 'Mises *Redux*'—*Redux*: Fifteen arguments against finite frequentism. *Erkenntnis* **45**, 209–27.

——— (2003). Conditional probability is the very guide of life. In *Probability Is the Very Guide of Life: The Philosophical Uses of Chance* (eds H. Kyburg, jr. & M. Thalos), pp. 183–203. La Salle, Ill.: Open Court.

——— (2007). The reference class problem is your problem too. *Synthese* **156**, 563–85.

——— (2009). Fifteen arguments against hypothetical frequentism. *Erkenntnis* **70**, 211–35.

——— (2010). Interpretations of probability. In *The Stanford Encyclopedia of Philosophy* (ed. E. N. Zalta), Spring 2010 edition. ⟨http://plato.stanford.edu/entries/probability-interpret⟩.

Halmos, P. (1950). *Measure Theory*. New York & London: Van Nostrand.

Halvorson, H. (2001). On the nature of continuous physical quantities in classical and quantum mechanics. *Journal of Philosophical Logic* **30**, 27–50.

——— (2004). Complementarity of representations in quantum mechanics. *Studies in History and Philosophy of Modern Physics* **35**, 45–56.

——— & Clifton, R. (1999). Maximal beable subalgebras of quantum-mechanical observables. *International Journal of Theoretical Physics* **38**, 2441–84.

——— & ——— (2000). Generic Bell Correlation between arbitrary local algebras in quantum field theory. *Journal of Mathematical Physics* **41**, 1711–17.

Hamhalter, J. (2003). *Quantum Measure Theory*. Dordrecht: Kluwer.

Hardy, L. (2001). Quantum theory from five reasonable axioms. *arXiv e-print quant-ph/0101012*.

——— (2002). Why quantum theory? In *Non-locality and Modality* (eds T. Placek & J. Butterfield), NATO Science Series, pp. 61–73. Dordrecht: Kluwer.

Harman, P. M. (ed.) (1990). *The Scientific Letters and Papers of James Clerk Maxwell*, Vol. I: *1846–1862*. Cambridge: Cambridge University Press.

——— (1998). *The Natural Philosophy of James Clerk Maxwell.* Cambridge: Cambridge University Press.

Hartley, R. (1928). Transmission of information. *Bell System Technical Journal* **7**, 535–63.

Hartmann, S. & Suppes, P. (2010). Entanglement, upper probabilities and decoherence in quantum mechanics. In *EPSA Philosophical Issues in the Sciences, Launch of the European Philosophy of Science Association*, Vol. 2 (eds M. Suárez, M. Dorato & M. Rédei), pp. 93–103. Dordrecht: Springer.

Hawkes, J. (1974). Hausdorff measure, entropy, and the independence of small sets. *Proceedings of the London Mathematical Society* **28**, 700–23.

Heisenberg, W. (1958). *Physics and Philosophy*. London: Penguin.

Held, C. (2006). The Kochen–Specker Theorem. In *Stanford Encyclopedia of Philosophy* (ed. E. N. Zalta). ⟨http://plato.stanford.edu/entries/kochen-specker⟩.

Hellman, G. (2008). Interpretations of probability in quantum mechanics: A case of 'experimental metaphysics.' In *Quantum Reality, Relativistic Causality, and Closing the Epistemic Circle: Essays in Honour of Abner Shimony* (eds W. Myrvold & J. Christian), pp. 211–27. The Western Ontario Series in Philosophy of Science, Vol. 73. Amsterdam: Springer.

Hemmo, M. & Pitowsky, I. (2007). Quantum probability and many worlds. *Studies in History and Philosophy of Modern Physics* **38**, 333–50.

——— & Shenker, O. (2006). Von Neumann's entropy does not correspond to thermodynamic entropy. *Philosophy of Science* **73**, 153–74.

Henderson, L. (2010). Bayesian updating and information gain in quantum measurements. In *Philosophy of Quantum Information and Entanglement* (eds A. Bokulich & G. Jaeger), pp. 151–67. Cambridge: Cambridge University Press.

Herapath, J. (1821). On the causes, laws and phenomena of heat, gases, gravitation. *Annals of Philosophy* Ser. 2, **1**, 273–93.

Herschel, J. F. W. (1850). Quételet on probabilities. *Edinburgh Review* **92**, 1–

57. Also in J. F. W. Herschel, *Essays from the Edinburgh and Quarterly Reviews*, London: Longman, Brown, Green, Longmans, and Roberts, 1857, pp. 365–465.

Hesse, M. (1953). Models in physics. *British Journal for the Philosophy of Science* **4**, 198–214.

——— (1963). *Models and Analogies in Science*. London: Sheed and Ward.

——— (2001). Models and analogies. In *A Companion to the Philosophy of Science* (ed. W. H. Newton-Smith), pp. 299–307. Oxford: Blackwell.

Hoefer, C. (2003a). Causal determinism. In *Stanford Encyclopedia of Philosophy* (ed. E. N. Zalta). ⟨http://plato.stanford.edu/entries/determinism-causal⟩.

——— (2003b). For fundamentalism. *Philosophy of Science* (*PSA Supplement 2002*) **70**, 1401–12.

——— (2007). The third way on objective probability: A sceptic's guide to objective chance. *Mind* **116** (463), 549–96.

——— (2010). *Chance in the World*. Draft book manuscript.

Holland, P. (1993). *The Quantum Theory of Motion: An Account of the de Broglie–Bohm Causal Interpretation of Quantum Mechanics*. Cambridge: Cambridge University Press.

Honerkamp, J. (1994). *Stochastic Dynamical Systems: Concepts, Numerical Methods, Data Analysis*. Weinheim: VCH Verlagsgesellschaft.

Hopf, E. (1934). On causality, statistics and probability. *Journal of Mathematics and Physics* **13**, 51–102.

Horwich, P. (1987). *Asymmetries in Time: Problems in the Philosophy of Science*. Cambridge, Mass.: MIT Press.

Howson, C. (1995). Theories of probability. *British Journal for the Philosophy of Science* **46**, 1–32.

——— & Urbach, P. (1989). *Scientific Reasoning: The Bayesian Approach*. La Salle, Ill.: Open Court.

——— & ——— (2006). *Scientific Reasoning: The Bayesian Approach*, 2nd edn. La Salle, Ill.: Open Court.

Huang, K. (1963). *Statistical Mechanics*. New York: John Wiley & Sons.

Hughes, R. I. G. (1989). *The Structure and Interpretation of Quantum Mechanics*. Cambridge, Mass.: Harvard University Press.

——— (1997). Models and representation. *Philosophy of Science* (Proceedings) **64**, S 325–36.

Hughston, L. P., Jozsa, R. & Wootters, W. K. (1993). A complete classification of quantum ensembles having a given density matrix. *Physics Letters A* **183**, 14–18.

Humphreys, P. (2004). *Extending Ourselves: Computational Science, Empiricism, and Scientific Method*. New York: Oxford University Press.

Ihara, S. (1993). *Information Theory for Continuous Systems*. London: World

Scientific.

Janssen, M. (2009). Drawing the line between kinematics and dynamics in special relativity. *Studies in History and Philosophy of Modern Physics* **40**, 26–52.

Jauch, J. M. (1960). Systems of observables in quantum mechanics. *Helvetica Physica Acta* **33**, 711–26.

——— & Misra, B. (1961). Supersymmetries and essential observables. *Helvetica Physica Acta* **34**, 699–709.

Jaynes, E. T. (1957). Information theory and statistical mechanics. *Physical Review* **106**, 620–30.

——— (1965). Gibbs vs. Boltzmann entropies. *American Journal of Physics* **33**, 391–8. Also in Jaynes 1983, pp. 77–86.

——— (1968). Prior probabilities. *IEEE Transactions on Systems Science and Cybernetics* **4**, 227–41.

——— (1979). Where do we stand on maximum entropy? In *The Maximum Entropy Formalism* (eds R. D. Levine & M. Tribus), pp. 15–118. Cambridge, Mass.: MIT Press.

——— (1983). *Papers on Probability, Statistics and Statistical Physics* (ed. R. D. Rosenkrantz). Dordrecht: Reidel.

Jeffrey, R. C. (1967). *The Logic of Decision*, 2nd edn. New York: McGraw–Hill.

——— (1977). Mises redux. In *Basic Problems in Methodology and Linguistics* (eds R. E. Butts & J. Hintikka), pp. 213–22. Dordrecht: D. Reidel. Repr. in Jeffrey, R. C., *Probability and the Art of Judgment*, Cambridge: Cambridge University Press, 1992, pp. 192–202.

——— (2004). *Subjective Probability: The Real Thing*. Cambridge: Cambridge University Press.

Jizba, P. & Arimitsu, T. (2004). The world according to Rényi: Thermodynamics of multifractal systems. *Annals of Physics* **312**, 17–59.

Jones, N. S. & Masanes, L. (2005). Interconversion of nonlocal correlations. *Physical Review A* **72**, 052312.

Jordan, P. (1927). Philosophical foundations of quantum theory. *Nature* **119**, 566–9.

Joyce, J. M. (2005). How probabilities reflect evidence. *Philosophical Perspectives* **19**, 153–78.

——— (2009). Accuracy and coherence: Prospects for an alethic epistemology of partial belief. In *Degrees of Belief* (eds F. Huber & C. Schmidt-Petri), pp. 263–97. Dordrecht: Kluwer.

Kac, M. (1959). *Probability and Related Topics in the Physical Sciences*. London: Interscience.

Kadison, R. V. & Ringrose, J. R. (1997a). *Fundamentals of the Theory of Operator Algebras*, Vol. 1: *Elementary Theory*. Providence, R.I.: American Mathematical Society.

——— & ——— (1997b). *Fundamentals of the Theory of Operator Algebras*, Vol. 2: *Advanced Theory*. Providence, R.I.: American Mathematical Society.

Kant, I. (1781/87 [1999]). *Critique of Pure Reason*, transl. P. Guyer & A. Wood. Cambridge: Cambridge University Press.

Kendall, M. G. & Stuart, A. (1979). *The Advanced Theory of Statistics*, 4th edn. London: Griffin.

Kerscher, M., Mecke, K., Schmalzing, J., Beisbart, C., Buchert, T. & Wagner, H. (2001). Morphological fluctuations of large-scale structure: The PSCz survey. *Astronomy and Astrophysics* **373**, 1–11.

Keynes, J. M. (1921). *A Treatise on Probability*. London: Macmillan & Co.

Khinchin, A. I. (1949). *Mathematical Foundations of Statistical Mechanics*. Mineola, N.Y.: Dover.

Kittel, C. (1958). *Elementary Statistical Mechanics*. Mineola, N.Y.: Dover.

Klir, G. (2006). *Uncertainty and Information: Foundations of Generalized Information Theory*. Hoboken, N.J.: John Wiley & Sons.

Kochen, S. & Specker, E. (1967). The problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics* **17**, 59–87.

Kolmogorov, A. N. (1956). *Foundations of the Theory of Probability*, 2nd English edn. New York: Chelsea.

——— (1958). A new metric invariant of transitive dynamical systems and automorphisms of Lebesgue spaces. *Doklady Academii Nauk SSSR* **119**, 861–4.

——— & Tihomirov, V. (1961). ε-entropy and ε-capacity of sets in functional spaces. *American Mathematical Society Translations* **17**, 277–364.

Kopersky, G. (2010). Models. In *Internet Encyclopedia of Philosophy* (eds J. Fieser & B. Dowden). ⟨http://www.iep.utm.edu/models⟩.

Kroes, P. (1989). Structural analogies between physical systems. *British Journal for the Philosophy of Science* **40**, 145–54.

Krüger, L., Daston, L. J., Heidelberger, M., Gigerenzer, G. & Morgan, M. S. (eds) (1990). *The Probabilistic Revolution*, Vols 1 & 2. Cambridge, Mass.: MIT Press.

Kullback, S. (1959). *Information Theory and Statistics*. New York: John Wiley & Sons.

Landau, L. & Lifshitz, E. (1976). *Mechanics*, 3rd edn. New York: Butterworth-Heineman.

Langevin, P. (1908). Sur la théorie du mouvement brownien. *Comptes rendus de l'Académie des Sciences* **146**, 530–3. English transl. in: D. S. Lemons & A. Gythiel, Paul Langevin's 1908 paper 'On the Theory of Brownian Motion,' *American Journal of Physics* **65** (1997), 1079–81.

Laplace, P. S. (1814). *Essai philosophique sur les probabilités*. Paris: Courcier. Transl. from the 5th French edn by A. I. Dale as *Philosophical Essay on Probability*, Berlin: Springer, 1995.

Lavis, D. A. (2004). The spin-echo system reconsidered. *Foundations of Physics*

34, 669–88.

——— (2005). Boltzmann and Gibbs: An attempted reconciliation. *Studies in History and Philosophy of Modern Physics* **36**, 245–73.

——— (2008). Boltzmann, Gibbs, and the concept of equilibrium. *Philosophy of Science* **75**, 682–96.

——— & Bell, G. M. (1999). *Statistical Mechanics of Lattice Systems*, Vol. 1: *Closed-Form and Exact Solutions*. Berlin: Springer.

——— & Milligan, P. J. (1985). The work of E. T. Jaynes on probability, statistics and statistical physics. *British Journal for the Philosophy of Science* **36**, 193–210.

Lebowitz, J. L. (1993). Boltzmann's entropy and time's arrow. *Physics Today* **46**, 32–8.

——— (1994). Time's arrow and Boltzmann's entropy. In *Physical Origins of Time Asymmetry* (eds J. J. Halliwell, J. Pérez-Mercarder & W. H. Zurek), pp. 131–46. Cambridge: Cambridge University Press.

——— (1999a). Microscopic origins of irreversible macroscopic behaviour. *Physica A* **263**, 516–27.

——— (1999b). Statistical mechanics: A selective review of two central issues. *Review of Modern Physics* **71**, S 346–57.

Leeds, S. (2003). Foundations of statistical mechanics—Two approaches. *Philosophy of Science* **70**, 126–44.

Leitgeb, H. & Pettigrew, R. (2010a). An objective justification of Bayesianism I: Measuring inaccuracy. *Philosophy of Science* **77**, 201–35.

——— & ——— (2010b). An objective justification of Bayesianism II: The consequences of minimizing inaccuracy. *Philosophy of Science* **77**, 236–72.

Lemons, D. S. (2002). *An Introduction to Stochastic Processes in Physics*. Baltimore, Md. & London: Johns Hopkins University Press.

Lenhard, J. (2006). Models and statistical inference: The controversy between Fisher and Neyman–Pearson. *British Journal for the Philosophy of Science* **57**, 69–91.

Lewis, D. (1980). A subjectivist's guide to objective chance. In *Studies in Inductive Logic and Probability*, Vol. II (ed. R. C. Jeffrey), pp. 263–93. Berkeley, Calif.: University of California Press. Repr. in Lewis 1986, pp. 83–131.

——— (1986). *Philosophical Papers*, Vol. II. Oxford: Oxford University Press.

——— (1994). Humean supervenience debugged. *Mind* **103** (412), 473–90.

——— (1999). Why conditionalize? In D. Lewis, *Papers in Metaphysics and Epistemology*, pp. 403–7.

Lewis, P. J. (2005). Probability in Everettian quantum mechanics. University of Miami Preprint, available at the Pitt Phil Sci Archive. ⟨http://philsci-archive.pitt.edu/archive/00002716⟩.

——— (2007). Uncertainty and probability for branching selves. *Studies in History and Philosophy of Modern Physics* **38**, 1–14. Available at the Pitt Phil Sci

Archive. ⟨http://philsci-archive.pitt.edu/archive/00002636⟩.
Loève, M. (1963). *Probability Theory*, 3rd edn. New York: Van Nostrand.
Loewer, B. (2001). Determinism and chance. *Studies in History and Philosophy of Modern Physics* **32**, 609–20.
——— (2004). David Lewis's Humean theory of objective chance. *Philosophy of Science* **71** (5), 1115–25.
Lucas, L. & Unterweger, M. (2000). Comprehensive review and critical evaluation of the half-life of tritium. *Journal of Research of the National Institute of Standards and Technology* **105**, 541–9.
Lüders, G. (1951). Über die Zustandsänderung durch den Meßprozeß. *Annalen der Physik* **8**, 322–8.
Maeda, S. (1989). Probability measures on projections in von Neumann algebras. *Reviews in Mathematical Physics* **1**, 235–90.
Magnani, L., Nersessian, N. J. & Thagard, P. (eds) (1999). *Model-Based Reasoning in Scientific Discovery*. Dordrecht: Kluwer.
——— & ——— (eds) (2002). *Model-Based Reasoning: Science, Technology, Values*. Dordrecht: Kluwer.
Mahnke, R., Kaupužs, J. & Lubashevsky, I. (2009). *Physics of Stochastic Processes: How Randomness Acts in Time*. Weinheim: Wiley–VCH.
Malament, D. B. & Zabell, S. L. (1980). Why Gibbs phase averages work—The role of ergodic theory. *Philosophy of Science* **47** (3), 339–49.
Mandelbrot, B. B. (1983). *The Fractal Geometry of Nature*. New York: Freeman.
Mañé, R. (1987). *Ergodic Theory and Differentiable Dynamics*. Berlin: Springer.
Margenau, H. (1950). *The Nature of Physical Reality*. New York: McGraw–Hill.
Masanes, L., Acin, A. & Gisin, N. (2006). General properties of nonsignaling theories. *Physical Review A* **73**, 012112.
Maudlin, T. (1994). *Quantum Nonlocality and Relativity: Metaphysical Intimations of Modern Physics*. Oxford: Blackwell.
——— (1995). Three measurement problems. *Topoi* **14**, 7–15.
——— (2007). What could be objective about probabilities? *Studies in History and Philosophy of Modern Physics* **38**, 275–91.
Maxwell, J. C. (1860). Illustrations of the dynamical theory of gases. *Philosophical Magazine* **19**, 19–32; **20**, 21–37. Also in Garber, Brush & Everitt 1986, pp. 285–318.
——— (1867). On the dynamical theory of gases. *Philosophical Transactions of the Royal Society of London* **157**, 49–88. Repr. in *The Kinetic Theory of Gases: An Anthology of Classic Papers with Historical Commentary*, Part II: *Irreversible Processes* (ed. S. G. Brush), pp. 197–261, Oxford: Pergamon Press, 1966, and in Garber, Brush & Everitt 1986, pp. 419–72.
——— (1879). On Boltzmann's theorem on the average distribution of energy in a system of material points. *Transactions of the Cambridge Philosophical Society*

12, 547–70. Also in Garber, Brush & Everitt 1995, pp. 357–86.

Maynard Smith, J. & Szathmáry, E. (1999). *The Origins of Life: From the Birth of Life to the Origin of Language.* Oxford & New York: Oxford University Press.

Mayo, D. G. (1996). *Error and the Growth of Experimental Knowledge*. Chicago, Ill.: University of Chicago Press.

McClintock, P. V. E. & Moss, F. (1989). Analogue techniques for the study of problems in stochastic nonlinear dynamics. In *Noise in Nonlinear Dynamical Systems*, Vol. 3: *Experiments and Simulations* (eds F. Moss & P. V. E. McClintock), pp. 243–74. Cambridge: Cambridge University Press.

Mellor, D. H. (1969). Chance. *The Aristotelian Society, Supplementary Volume* **43**, 11–34.

——— (1971). *The Matter of Chance*. Cambridge: Cambridge University Press.

——— (2005). *Probability: A Philosophical Introduction*. London: Routledge.

Menon, T. (2010). The Conway–Kochen Free Will Theorem. Manuscript.

Miller, D. W. (1994). *Critical Rationalism: A Restatement and Defence*. Chicago, Ill. & La Salle, Ill.: Open Court.

Mohrhoff, U. (2004). Probabilities from envariance. *International Journal of Quantum Information* **2**, 221–30.

Morgan, M. S. & Morrison, M. (1999a). Models as mediating instruments. In Morgan & Morrison 1999b, pp. 10–37.

——— & ——— (eds) (1999b). *Models as Mediators: Perspectives on Natural and Social Sciences*. Cambridge: Cambridge University Press.

Nelson, E. (1966). Derivation of the Schrödinger Equation from Newtonian mechanics. *Physical Review* **150**, 1079–85.

——— (1985). *Quantum Fluctuations*. Princeton, N.J.: Princeton University Press.

Newman, J. R. (1956). *The World of Mathematics*. New York: Simon & Schuster. Reissued Mineola, N.Y.: Dover, 2000.

Nielsen, M. A. & Chuang, I. (2000). *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press.

North, J. (forthcoming). Time in thermodynamics. In *The Oxford Handbook of Time* (ed. C. Callender). Oxford: Oxford University Press.

Norton, J. D. (1999). A quantum-mechanical supertask. *Foundations of Physics* **29**, 1265–1302.

——— (2008). The dome: An unexpectedly simple failure of determinism. *Philosophy of Science* **75**, 786–98.

Ott, E. (2002). *Chaos in Dynamical Systems*. Cambridge: Cambridge University Press.

Papoulis, A. (1984). *Probability, Random Variables, and Stochastic Processes*. New York: McGraw–Hill.

Parker, D. N. (2006). *Thermodynamics, Reversibility and Jaynes' Approach to Statistical Mechanics*. Ph.D. Thesis, University of Maryland.

Pauli, W. (1927). Über Gasentartung und Paramagnetismus. *Zeitschrift für Physik* **43**, 81–102.

Pearle, P. (1989). Combining stochastic dynamical state-vector reduction with spontaneous localization. *Physical Review A* **39**, 2277–89.

Peebles, P. J. E. (1980). *The Large-Scale Structure of the Universe*. Princeton, N.J.: Princeton University Press.

Penrose, R. (1970). *Foundations of Statistical Mechanics*. Oxford: Oxford University Press.

Petersen, K. (1983). *Ergodic Theory*. Cambridge: Cambridge University Press.

Pippard, A. B. (1966). *The Elements of Classical Thermodynamics.* Cambridge: Cambridge University Press.

Pitowsky, I. (1989). *Quantum Probability—Quantum Logic*. Lecture Notes in Physics, Vol. 321. Berlin: Springer.

——— (2003). Betting on the outcomes of measurements: A Bayesian theory of quantum probability. *Studies in History and Philosophy of Modern Physics* **34**, 395–414.

——— (2006). Quantum mechanics as a theory of probability. In Demopoulos & Pitowsky 2006, pp. 213–39.

Polyá, G. (1954). *Mathematics and Plausible Reasoning*, Vol. II: *Patterns of Plausible Inference*. Princeton, N.J.: Princeton University Press.

Popescu, S. & Rohrlich, D. (1994). Causality and non-locality as axioms for quantum mechanics. *Foundations of Physics* **24**, 379.

Popper, K. R. (1955). Two autonomous axiom systems for the calculus of probabilities. *British Journal for the Philosophy of Science* **21**, 51–7.

——— (1957). The propensity interpretation of the calculus of probability, and the quantum theory. In *Observation and Interpretation: A Symposium of Philosophers and Physicists* (ed. S. Körner), pp. 65–70, 88–9. London: Butterworths.

——— (1959). The propensity interpretation of probability. *British Journal for the Philosophy of Science* **10**, 25–42.

——— (1967). Quantum mechanics without 'the observer.' In *Quantum Theory and Reality* (ed. M. Bunge), pp. 1–12. New York: Springer.

——— (1982). *Quantum Theory and the Schism in Physics*. Totowa, N.J.: Rowan & Littlefield.

——— (1990). *A World of Propensities*. Bristol: Thoemmes.

Price, H. (2006). Probability in the Everett World: Comments on Wallace and Greaves. University of Sydney Preprint. Available at the Pitt Phil Sci Archive. ⟨http://philsci-archive.pitt.edu/archive/00002719⟩.

Prugovečki, E. (1981). *Quantum Mechanics in Hilbert Space*, 2nd edn. New York: Academic Press.

Quételet, A. (1846). *Lettres á S.A.R. le duc régnant du Saxe-Coburg et Gotha sur la théorie des probabilités*. Brussels: Hayez.

Rae, A. I. M. (2009). Everett and the Born Rule. *Studies in History and Philosophy of Modern Physics* **40** (3), 243–50.

Ramsey, F. P. (1926). Truth and probability. In *Studies in Subjective Probability* (eds H. Kyburg & H. Smokler), pp. 63–92. New York: John Wiley & Sons.

Rédei, M. (1992). When can non-commutative statistical inference be Bayesian? *International Studies in Philosophy of Science* **6**, 129–32.

——— & Summers, S. (2007). Quantum probability theory. *Studies in History and Philosophy of Modern Physics* **38**, 390–417. *arXiv e-print quant-ph/0601158*.

Redhead, M. (1974). On Neyman's Paradox and the theory of statistical tests. *British Journal for the Philosophy of Science* **25**, 265–71.

——— (1980). Models in physics. *British Journal for the Philosophy of Science* **31**, 145–63.

——— (1987). *Incompleteness, Nonlocality, and Realism*. Oxford: Clarendon Press.

Reichenbach, H. (1935). *Wahrscheinlichkeitslehre*. Leiden: A. W. Sijthoff.

——— (1948). The Principle of Anomaly in quantum mechanics. *Dialectica* **2**, 337–50.

——— (1949). *The Theory of Probability*. Berkeley, Calif.: University of California Press.

——— (1971). *The Direction of Time*. Berkeley, Calif.: University of California Press. Repr. (ed. M. Reichenbach) Mineola, N.Y.: Dover, 1999.

Reiss, H. (1965). *Methods of Thermodynamics*. Mineola, N.Y.: Dover.

Rényi, A. (1961). On measures of entropy and information. In *Proceedings of the Fourth Berkeley Symposium of Mathematical Statistics and Probability* (ed. J. Neyman), pp. 547–61. Berkeley, Calif.: University of California Press.

Ridderbos, K. (2002). The coarse-graining approach to statistical mechanics: How blissful is our ignorance? *Studies in History and Philosophy of Modern Physics* **33**, 65–77.

Robert, C. P. (1994). *The Bayesian Choice*. New York etc.: Springer.

Rosenthal, J. (2010). The natural-range conception of probability. In Ernst & Hüttemann 2010, pp. 71–91.

Ruetsche, L. (2003). Modal semantics, modal dynamics, and the problem of state preparation. *International Studies in the Philosophy of Science* **17**, 25–41.

Ryder, J. M. (1981). Consequences of a simple extension of the Dutch book argument. *British Journal for the Philosophy of Science* **32**, 164–7.

Salmon, W. C. (1967). *The Foundations of Scientific Inference*. Pittsburgh, Pa.: University of Pittsburgh Press.

——— (1979). Propensities: A discussion review of D. H. Mellor, *The Matter of Chance*. *Erkenntnis* **14**, 183–216.

Saunders, S. (1995). Time, quantum mechanics, and decoherence. *Synthese* **102**, 235–66.

——— (1996a). Relativism. In *Perspectives on Quantum Reality* (ed. R. Clifton), pp. 125–42. Dordrecht: Kluwer.

——— (1996b). Time, quantum mechanics, and tense. *Synthese* **107**, 19–53.

——— (1998). Time, quantum mechanics, and probability. *Synthese* **114**, 373–404.

——— (2004). Derivation of the Born Rule from operational assumptions. *Proceedings of the Royal Society of London A* **460**, 1771–88.

——— (2005). What is probability? In *Quo Vadis Quantum Mechanics?* (eds A. Elitzur, S. Dolev & N. Kolenda), pp. 209–38. Berlin: Springer.

———, Barrett, J., Kent, A. & Wallace, D. (eds) (2010). *Many Worlds? Everett, Quantum Theory, and Reality*. Oxford: Oxford University Press.

——— & Wallace, D. (2008). Branching and uncertainty. *British Journal for the Philosophy of Science* **59**, 293–305.

Savage, L. J. (1954). *The Foundations of Statistics*. New York: John Wiley & Sons.

——— (1972). *The Foundations of Statistics*, 2nd edn. Mineola, N.Y.: Dover.

Schack, R., Brun, T. A. & Caves, C. M. (2001). Quantum Bayes Rule. *Physical Review A* **64**, 014305.

Schaffer, J. (2007). Deterministic chance? *British Journal for the Philosophy of Science* **58**, 113–40.

Schlosshauer, M. & Fine, A. (2005). On Zurek's derivation of the Born Rule. *Foundations of Physics* **35**, 197–213.

Schrödinger, E. (1926a). Quantisierung als Eigenwertproblem (erste Mitteilung). *Annalen der Physik* **79**, 361–76.

——— (1926b). Quantisierung als Eigenwertproblem (zweite Mitteilung). *Annalen der Physik* **79**, 489–527.

——— (1935a). Discussion of probability relations between separated systems. *Proceedings of the Cambridge Philosophical Society* **31**, 555–63.

——— (1935b). The present situation in quantum mechanics. *Naturwissenschaften* **23**, 807–12, 823–8, 844–9. Repr. in Wheeler & Zurek 1983, pp. 152–67.

——— (1950). Irreversibility. *Proceedings of the Royal Irish Academy* **53 A**, 189–95.

Segal, I. (1959). The mathematical meaning of operationalism in quantum mechanics. In *Studies in Logic and the Foundations of Mathematics* (eds L. Henkin, P. Suppes & A. Tarski), pp. 341–52. Amsterdam: North-Holland.

Seidenfeld, T. (1986). Entropy and uncertainty. *Philosophy of Science* **53**, 467–91.

———, Schervish, M. & Kadane, J. (1995). A representation of partially ordered preferences. *Annals of Statistics* **23**, 2168–2217.

Sewell, G. (1986). *Quantum Theory of Collective Phenomena*. Oxford: Oxford University Press.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal* **27**, 379–423, 623–56.

——— & Weaver, W. (1949). *The Mathematical Theory of Communication*. Urbana, Ill., Chicago, Ill. & London: University of Illinois Press.

Shaw, R. (1985). *The Dripping Faucet as a Model Chaotic System*. Santa Cruz, Calif.: Aerial Press.

Shen, J. & Wu, J. (2009). Sequential product on standard effect algebra $\mathcal{E}(H)$. *Journal of Physics A* **42**, 345203.

Shenker, O. (1994). Fractal geometry is not the geometry of nature. *Studies in History and Philosophy of Modern Physics* **25**, 967–81.

Shimony, A. (1985). The status of the Principle of Maximum Entropy. *Synthese* **63**, 55–74.

——— (2009a). Bell's Theorem. In *Stanford Encyclopedia of Philosophy* (ed. E. N. Zalta). ⟨http://plato.stanford.edu/entries/bell-theorem⟩.

——— (2009b). Probability in quantum mechanics. In *Compendium of Quantum Physics* (eds D. Greenberger, K. Hentschel & F. Weinert), pp. 492–7. Berlin: Springer.

———, Horne, M. A. & Clauser, J. F. (1976). Comment on 'The theory of local beables.' *Lettres épistémologiques* **13**, 1–8. Repr. in *Dialectica* **39** (1985), pp. 97–102.

Sinai, Y. (1959). On the concept of entropy for dynamical systems. *Doklady Akademii Nauk SSSR* **124**, 768–71.

Sklar, L. (1993). *Physics and Chance: Philosophical Issues in the Foundations of Statistical Mechanics*. Cambridge & New York: Cambridge University Press.

——— (2006). Why does the standard measure work in statistical mechanics? In *Interactions: Mathematics, Physics and Philosophy, 1860–1930* (eds V. F. Hendricks, K. F. Jørgensen, J. Lützen & S. A. Pedersen), pp. 307–20. Boston Studies in the Philosophy of Science, Vol. 251. Dordrecht: Springer.

Skyrms, B. (1999). *Choice and Chance: An Introduction to Inductive Logic*, 4th edn. Belmont, Calif.: Wadsworth.

Sober, E. (2010). Evolutionary theory and the reality of macro-probabilities. In *The Place of Probability in Science: In Honor of Ellery Eells (1953–2006)* (eds E. Eells & J. H. Fetzer), pp. 133–61. Boston Studies in the Philosophy of Science, Vol. 284. Heidelberg: Springer.

Sorkin, R. (2005). Ten theses on black hole entropy. *Studies in History and Philosophy of Modern Physics* **36**, 291–301.

Spekkens, R. (2005). Contextuality for preparations, transformations, and unsharp measurements. *Physical Review A* **71**, 052108.

Spiegelhalter, D. & Rice, K. (2009). Bayesian statistics. *Scholarpedia* **4** (8), 5230. ⟨http://www.scholarpedia.org/article/Bayesian_statistics⟩.

Spohn, H. (1991). *Large Scale Dynamics of Interfacing Particles.* Berlin & Heidelberg: Springer.

Sprenger, J. (2009). Statistics between inductive logic and empirical science.

Journal of Applied Logic **7**, 239–50.

——— (2010). Statistical inference without frequentist justifications. In *EPSA Epistemology and Methodology of Science: Launch of the European Philosophy of Science Association*, Vol. I (eds M. Suárez, M. Dorato & M. Rédei), pp. 289–97. Berlin: Springer.

Stigler, S. M. (1982). Thomas Bayes's Bayesian inference. *Journal of the Royal Statistical Society Series A* **145**, 250–8.

——— (1999). *Statistics on the Table: The History of Statistical Concepts and Methods*. Cambridge, Mass.: Harvard University Press.

Stoyan, D. & Stoyan, H. (1994). *Fractals, Random Shapes and Point Fields: Methods of Geometrical Statistics*. Chichester: John Wiley & Sons.

Streater, R. F. (2000). Classical and quantum probability. *Journal of Mathematical Physics* **41**, 3556–3603.

Strevens, M. (2003). *Bigger than Chaos: Understanding Complexity through Probability*. Cambridge, Mass.: Harvard University Press.

——— (2006). Probability and chance. In *The Encyclopedia of Philosophy*, 2nd edn (ed. D. M. Borchert), Vol. 8, pp. 24–40.
Detroit, Mich.: Macmillan Reference USA.

——— (2009). *Depth: An Account of Scientific Explanation*. Cambridge, Mass.: Harvard University Press.

Suárez, M. (2004). An inferential conception of scientific representation. *Philosophy of Science* **71**, 767–79.

——— (2009). Propensities in quantum mechanics. In *Compendium of Quantum Physics* (eds D. Greenberger, K. Hentschel & F. Weinert), pp. 502–5. Berlin: Springer.

Sunder, V. (1986). *An Invitation to von Neumann Algebras.* Berlin: Springer.

Suppes, P. (1993). The transcendental character of determinism. *Midwest Studies in Philosophy* **18**, 242–57.

——— & Zanotti, M. (1981). When are probabilistic explanations possible? *Synthese* **48**, 191–9.

Sutherland, W. (2002). *Introduction to Metric and Topological Spaces*. Oxford: Oxford University Press.

Swoyer, C. (1991). Structural representation and surrogative reasoning. *Synthese* **81**, 449–508.

Takesaki, M. (1972). Conditional expectations in von Neumann algebras. *Journal of Functional Analysis* **9**, 306–21.

——— (2003). *Theory of Operator Algebras*, Vols 2 & 3. Berlin: Springer.

Teller, P. (1973). Conditionalization and observation. *Synthese* **26**, 218–58.

Timpson, C. (2008a). Philosophical aspects of quantum information theory. In *The Ashgate Companion to Contemporary Philosophy of Physics* (ed. D. Rickles), pp. 197–261. Aldershot & Burlington, Vt.: Ashgate. *arXiv e-print quant-ph/*

0611187.

——— (2008b). Quantum Bayesianism: A study. *Studies in History and Philosophy of Modern Physics* **39**, 579–609. *arXiv e-print quant-ph/0804.2047.*

——— (2010). *Quantum Information Theory and the Foundations of Quantum Mechanics.* Oxford: Oxford University Press.

Tolman, R. C. (1938). *The Principles of Statistical Mechanics.* Oxford: Oxford University Press. Reissued Mineola, N.Y.: Dover, 1979.

Torretti, R. (2007). The problem of time's arrow historico-critically reexamined. *Studies in History and Philosophy of Modern Physics* **38** (4), 732–56.

Tsallis, C. (1988). Possible generalization of Boltzmann–Gibbs statistics. *Journal of Statistical Physics* **52**, 479–87.

Tsirelson, B. S. (1980). Quantum generalizations of Bell's Inequality. *Letters in Mathematical Physics* **4**, 93–100.

Tumulka, R. (2006). A relativistic version of the Ghirardi–Rimini–Weber model. *Journal of Statistical Physics* **125**, 821–40.

——— (2007). Comment on 'The Free Will Theorem.' *Foundations of Physics* **37**, 186–97.

Uffink, J. (1995). Can the Maximum Entropy Principle be explained as a consistency requirement? *Studies in History and Philosophy of Modern Physics* **26**, 223–61.

——— (1996). The constraint rule of the Maximum Entropy Principle. *Studies in History and Philosophy of Modern Physics* **27**, 47–79.

——— (1999). How to protect the interpretation of the wave function against protective measurements. *Physical Review A* **60**, 3474–81.

——— (2001). Bluff your way in the Second Law of Thermodynamics. *Studies in History and Philosophy of Modern Physics* **32**, 305–94.

——— (2004). Boltzmann's work in statistical physics. In *The Stanford Encyclopedia of Philosophy* (ed. E. N. Zalta). ⟨http://plato.stanford.edu/entries/statphys-Boltzmann⟩.

——— (2007). Compendium of the foundations of classical statistical physics. In Butterfield & Earman 2007, pp. 923–1074.

Uhlhorn, U. (1963). Representation of symmetry transformations in quantum mechanics. *Arkiv Fysik* **23**, 307–40.

Vaidman, L. (1998). On schizophrenic experiences of the neutron or Why we should believe in the many-worlds interpretation of quantum mechanics. *International Studies in the Philosophy of Science* **12**, 245–61.

——— (2002). Many-worlds interpretation of quantum mechanics. In *The Stanford Encyclopedia of Philosophy* (ed. E. N. Zalta). ⟨http://plato.stanford.edu/archives/fall2008/entries/qm-manyworlds⟩.

Valente, G. (2007). Is there a stability problem for Bayesian noncommutative probabilities? *Studies in History and Philosophy of Modern Physics* **38**, 832–43.

Valentini, A. (1991a). Signal-locality, uncertainty, and the Sub-quantum H-Theorem I. *Physics Letters A* **156** (1,2), 5–11.

——— (1991b). Signal-locality, uncertainty, and the Sub-quantum H-Theorem II. *Physics Letters A* **158** (1,2), 1–8.

——— & Westman, H. (2005). Dynamical origin of quantum probabilities. *Proceedings of the Royal Society of London A* **461**, 253–72.

van Fraassen, B. C. (1980). *The Scientific Image*. Oxford: Oxford University Press.

——— (1991). *Quantum Mechanics: An Empiricist View*. Oxford: Clarendon Press.

van Kampen, N. G. (1981). *Stochastic Processes in Physics and Chemistry*. Amsterdam: North-Holland.

van Lith, J. H. (2001a). Ergodic theory, interpretations of probability and the foundations of statistical mechanics. *Studies in History and Philosophy of Modern Physics* **32**, 581–94.

——— (2001b). Stir in stillness: A study in the foundations of equilibrium statistical mechanics. Ph.D. Thesis, Utrecht University. ⟨http://igitur-archive.library.uu.nl/dissertations/1957294/title.pdf⟩.

von Mises, R. (1928). *Probability, Statistics and Truth*. London: George Allen and Unwin. Page references are to the 2nd, revised English edn, prepared by H. Geiringer, New York: Macmillan, 1957.

von Neumann, J. (1955). *Mathematical Foundations of Quantum Mechanics*. Princeton, N.J.: Princeton University Press.

von Plato, J. (1982). The significance of the ergodic decomposition of stationary measures for the interpretation of probability. *Synthese* **53**, 419–32.

——— (1983). The method of arbitrary functions. *British Journal for the Philosophy of Science* **34**, 37–47.

——— (1989a). De Finetti's earliest works on the foundations of probability. *Erkenntnis* **31**, 263–82.

——— (1989b). Probability in dynamical systems. In *Logic, Methodology and Philosophy of Science VIII: Proceedings of the Eighth International Congress of Logic, Methodology and Philosophy of Science, Moscow, 1987* (eds J. E. Fenstad, I. T. Frolov & R. Hilpinen), pp. 427–43. Studies in Logic and the Foundations of Mathematics, Vol. 126. Amsterdam etc.: North-Holland.

——— (1994). *Creating Modern Probability*. Cambridge: Cambridge University Press.

Wallace, D. (2002). Worlds in the Everett interpretation. *Studies in History and Philosophy of Modern Physics* **33**, 637–61.

——— (2003a). Everett and structure. *Studies in History and Philosophy of Modern Physics* **34**, 87–105.

——— (2003b). Everettian rationality: Defending Deutsch's approach to probability in the Everett interpretation. *Studies in History and Philosophy of*

Modern Physics **34** (3), 415–40.
——— (2006). Epistemology quantised: Circumstances in which we should come to believe in the Everett interpretation. *British Journal for the Philosophy of Science* **57** (4), 655–89.
——— (2007). Quantum probability from subjective likelihood: Improving on Deutsch's proof of the Probability Rule. *Studies in History and Philosophy of Modern Physics* **38**, 311–32.
——— (2010a). Gravity, entropy, and cosmology: In search of clarity. *British Journal for the Philosophy of Science* **61**, 513–40.
——— (2010b). How to prove the Born Rule. In Saunders *et al.* 2010, pp. 237–63.
——— (forthcoming). *The Emergent Multiverse: Quantum Mechanics according to the Everett Interpretation*. Oxford: Oxford University Press.
——— & Timpson, C. G. (2010). Quantum mechanics on spacetime I: Spacetime state realism. *British Journal for the Philosophy of Science* **61**, 697–727.
Wehrl, A. (1978). General properties of entropy. *Reviews of Modern Physics* **50**, 221–59.
Weisberg, M. (2007). Who is a modeler? *British Journal for the Philosophy of Science* **58**, 207–33.
Werndl, C. (2009a). Are deterministic descriptions and indeterministic descriptions observationally equivalent? *Studies in History and Philosophy of Modern Physics* **40**, 232–42.
——— (2009b). Deterministic versus indeterministic descriptions: Not that different after all? In *Reduction, Abstraction, Analysis: Proceedings of the 31th International Ludwig Wittgenstein-Symposium in Kirchberg, 2008* (eds A. Hieke & H. Leitgeb), pp. 63–78. Frankfurt: Ontos.
——— (2009c). Justifying definitions in matemathics—going beyond Lakatos. *Philosophia Mathematica* **17**, 313–40.
——— (2009d). What are the new implications of chaos for unpredictability? *British Journal for the Philosophy of Science* **60**, 195–220.
Wessels, L. (1981). What was Born's statistical interpretation?, In *PSA 1980: Proceedings of the 1980 Biennial Meeting of the Philosophy of Science Association*, Vol. 2: *Symposia and Invited Papers* (eds P. D. Asquith & R. N. Giere), pp. 187–200. East Lansing, Mich.: Philosophy of Science Association.
Wheeler, J. A. & Zurek, W. H. (eds) (1983). *Quantum Theory and Measurement*. Princeton, N.J.: Princeton University Press.
Wigner, E. (1959). *Group Theory and its Applications to Quantum Mechanics of Atomic Spectra*. New York: Academic Press.
Williamson, J. (2009). Philosophies of probability. In *Handbook of the Philosophy of Mathematics* (ed. A. Irvine), pp. 493–533. Amsterdam: North Holland.
——— (2010). *In Defence of Objective Bayesianism*. Oxford: Oxford University Press.

Winnie, J. A. (1997). Deterministic chaos and the nature of chance. In *The Cosmos of Science: Essays of Exploration* (eds J. Earman & J. D. Norton), pp. 299–324. Pittsburgh, Pa.: University of Pittsburgh Press.

Winsberg, E. (2004a). Can conditionalizing on the 'Past Hypothesis' militate against the reversibility objections? *Philosophy of Science* **71** (4), 489–504.

——— (2004b). Laws and statistical mechanics. *Philosophy of Science* **71** (5), 707–18.

Wüthrich, C. (2006). *Approaching the Planck Scale from a Generally Relativistic Point of View: A Philosophical Appraisal of Loop Quantum Gravity*. Ph.D. dissertation, University of Pittsburgh.

Yngvason, J. (2005). The role of type III factors in quantum field theory. *Reports on Mathematical Physics* **55**, 135–47.

Zabell, S. (2005). *Symmetry and its Discontents*. Cambridge: Cambridge University Press.

Zurek, W. H. (1993). Preferred states, predictability, classicality, and the environment-induced decoherence. *Progress in Theoretical Physics* **89**, 281–312.

——— (2003a). Decoherence, einselection, and the quantum origins of the classical. *Reviews of Modern Physics* **75**, 715–75.

——— (2003b). Environment-assisted invariance, entanglement, and probabilities in quantum physics. *Physical Review Letters* **90**, 120404.

——— (2005). Probabilities from entanglement, Born's Rule $p_k = |\psi_k|^2$ from envariance. *Physical Review A* **71**, 052105.