Determinism and Indeterminism

Charlotte Werndl

Forthcoming in: The Oxford Handbook of Philosophy of Science.
Oxford: Oxford University Press, edited by Paul Humphreys

Abstract

This article focuses on three themes concerning determinism and indeterminism. The first theme is observational equivalence between deterministic and indeterministic models. Here I discuss several results about observational equivalence and present an argument on how to choose between deterministic and indeterministic models involving indirect evidence. The second theme is whether Newtonian physics is indeterministic. I argue that the answer depends on what one takes Newtonian mechanics to be, and I highlight how contemporary debates on this issue differ from those in the nineteenth century. The third major theme is how the method of arbitrary functions can be used to make sense of deterministic probabilities. I discuss various ways of interpreting the initial probability distributions and argue that they are best understood as physical, biological etc. quantities characterising the particular situation at hand. Also, I emphasise that the method of arbitrary functions deserves more attention than it has received so far.

Contents

1 Introduction

Determinism reigns when the state of the system at one time fixes the past and future evolution of the system. The question of determinism can be asked about real systems (i.e. whether the state of a real system at one time fixes the state of the system at all times) or about models of real systems (i.e. whether the state of a model at one time fixes the state of the model at all times). *Indeterminism* amounts simply to the negation of determinism. Of course, one usually uses models to arrive at claims about the deterministic characters of real systems, but, as we will see, the relationship between deterministic systems and models is not at all straightforward.

This article will focus on three major themes in the recent debate on determinism in the philosophy of science. Throughout the article, emphasis will not be just on summarising the debates, but also on presenting some novel criticism and arguments. The first major theme will be determinism, indeterminism and observational equivalence. Here I will critically discuss various notions of observational equivalence between deterministic and indeterministic systems, and whether there is underdetermination between deterministic and indeterministic models (Section 2). The second major theme will be whether Newtonian mechanics is indeterministic and how scientists' debate in the nineteenth century differs from the contemporary debate (Section 3). The third major theme will be how probabilities can arise in deterministic systems. Here I will stress the usefulness of the method of arbitrary functions for understanding deterministic probabilities (Section 4). The paper will end with a conclusion (Section 5).

2 Determinism, indeterminism and observational equivalence

2.1 Deterministic and indeterministic models

Consider the evolution of the temperature in London (which is assumed to take values between 0 and 30 degrees). A meteorologist measures the temperature over nine days and obtains a sequence of observations as shown in Figure 1. Meteorologists aim to find a model which reproduces these observations and correctly predicts the future temperature values. In this context, the question arises whether the temperature evolution is best described by a deterministic or an indeterministic model, and one might think that the observations only allow for one or the other. However, in several cases including the example of
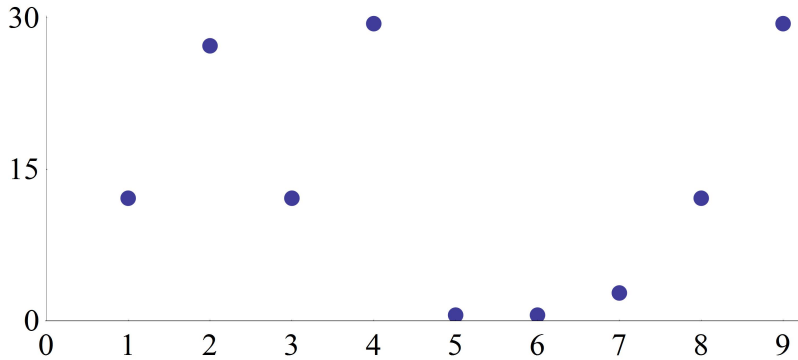
Figure 1: Observed temperature in London over nine days

the evolution of the temperature, both a deterministic and an indeterministic model is possible (i.e. there is observational equivalence between deterministic and indeterministic models). This raises the question which model is preferable and whether there is underdetermination. To tackle these questions, let me first introduce deterministic and indeterministic models.

The deterministic models we focus on are measure-theoretic deterministic models (M, T_t, p).[1] Here M is the set of possible states (the *phase space*), where $m \in M$ represents the state of the system. The functions $T_t : M \to M$ are the *evolution equations*, telling one that a state $m \in M$ evolves to $T_t(m)$ after t time steps ($t \in \mathbb{Z}$). p is a probability measure, assigning a probability to regions of M.[2] The *solution* through m represents a possible path of the deterministic system over time. Formally, it is the bi-infinite sequence $(...T_{-2}(m), T_{-1}(m), m, T_1(m), T_2(m)...)$. Clearly, because T_t are functions, the models (M, T_t, p) are deterministic: the initial state m fixes the past and future evolution of the system. These deterministic models are among the most important models in science (e.g., including all Newtonian models of energy-conserving systems).

In the observation of a deterministic system, a value is observed that is dependent on, but usually different from, the actual value (since observations cannot be done with infinite precision). Formally, an observation corresponds

---

[1]For technical details see Werndl (2009a, 2011). For simplicity, I focus on models with discrete time, but all that will be said carries over to models with continuous time (cf. Werndl 2011).

[2]There are various interpretations of this probability measure from being a physical quantity that describes the probability of finding a system in a certain region of phase space to the long-run average of the proportion of time a solution spends in a certain region (cf. Lavis 2011).
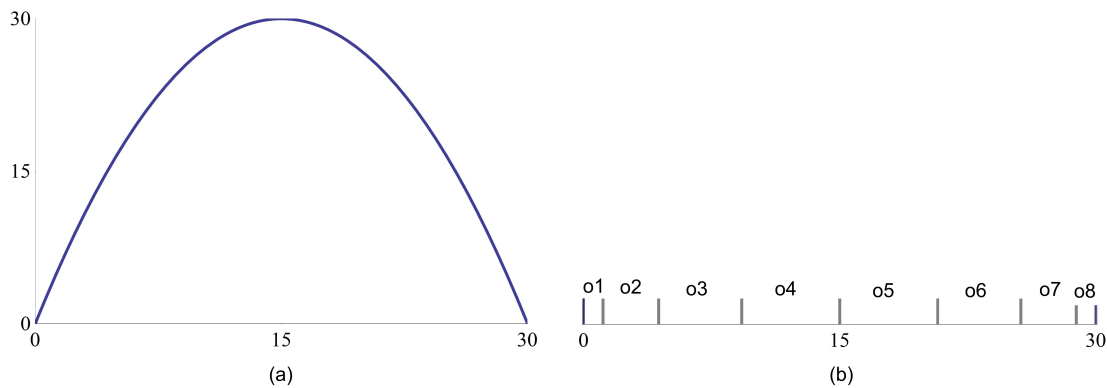
4

Figure 2: (a) the logistic map $T(m)$; (b) the observation function Φ_8

to an *observation function*, i.e. a function $\Phi : M \to M_O$, where $\Phi(m)$ represents the observed value (M_O is the set of all possible observation values). Since we can only make observations with finite precision, in what follows it is assumed that observation function take only finitely many values.

Lorenz (1964) used the logistic map to model the evolution of the temperature. Since this is a very simple model, I will use it for illustration purposes. More specifically, the phase space M of the logistic map is [0,30] (representing the temperature values between 0 and 30). The temperature at day $t + 1$ is obtained from the temperature m at day t by the equation (cf. Figure 2):

$$T(m) = 4m(1 - \frac{m}{30}). \tag{1}$$

Hence the evolution equations are given by $T_t(m) := T^t(m)$ (where T^t is the t-th iterate of m).[3] The probability measure assigns to a region A of $[0, 30]$ the value:

$$p(A) = \int_A \frac{30}{\pi \sqrt{\frac{m}{30}(1 - \frac{m}{30})}} dm. \tag{2}$$

Consider the observation function Φ_8 of the logistic map with eight values $\Phi_8(m) = o1 = 0.570905$ for $0 \leq m < 1.14181$, $\Phi_8(m) = o2 = 2.767605$ for $1.14181 \leq m < 4.3934$, $\Phi_8(m) = o3 = 6.82605$ for $4.3934 \leq m < 9.2587$, $\Phi_8(m) = o4 = 12.12935$ for $9.2587 \leq m < 15$, $\Phi_8(m) = o5 = 17.87015$ for $15 \leq m < 20.7403$, $\Phi_8(m) = o6 = 23.17345$ for $20.7403 \leq m < 25.6066$,

---

[3]The logistic map is only forward deterministic. That is, the state of the model at one time determines all the future states but not the past states. Nothing hinges on this and all the results presented in this section carry over to systems that are only forward deterministic.

$\Phi_8(m) = o7 = 27.2324$ for $25.6066 \leq m < 28.8582$ and $\Phi_8(m) = o8 = 29.4291$ for $28.8582 \leq m \leq 30$. Figure 2(b) shows this observation function. Suppose that the initial temperature is 0.3355. Then the first nine iterates coarsegrained by the observation function Φ_8 are: $o4$, $o7$, $o4$, $o8$, $o1$, $o1$, $o2$, $o4$, $o8$. This is the sequence shown in Figure 1. Consequently, the time series of Figure 1 can be derived from observations of the logistic system.

The indeterministic models we will focus on are *stochastic models* $\{Z_t\}$ with a finite number of outcomes E (representing systems that evolve according to probabilistic laws). Here Z_t represents the outcome of the system at time t. Probability distributions characterise the probabilistic behaviour of the stochastic system: e.g., the probability distribution $P(Z_t = e)$ gives one the probability that the outcome of the system is e at time t; joint probability distributions $P(Z_s = e$ and $Z_t = f)$ tell one the probabilities of outcomes at different times; conditional probability distributions $P(Z_s = e$ given that $Z_s = f)$ tell one the probability that the outcome is e at s given that it was f at t (for any s and t). A *realisation* represents a possible evolution of the stochastic system over time. Formally, it is a bi-infinite sequence $(...Z_{-2}, Z_{-1}, Z_0, Z_1, Z_2...)$. A stochastic model is indeterministic in the sense that given the initial outcome, there are several outcomes that might follow (and these possibilities are measured by probabilities). Most indeterministic models in science are stochastic models. For stochastic systems observations are also modelled by *observation functions*, i.e., functions $\Gamma : E \rightarrow E_O$, where $\Gamma(e)$ represents the observed value (E_O is the set of all possible observed values).

Probably the best-known stochastic models are *Bernoulli models* (representing a sequence of identically distributed random experiments, where the outcomes are independent of each other, like for a sequence of coin tosses or a sequence of throwing dice). Widely used in science are also Markov models (representing a sequence of identically distributed random experiments, where the next outcome only depends on the previous outcome). Consider the following specific Markov model $\{V_t\}$: there are eight possible states $o1 = 0.570905$ $o2 = 2.767605$, $o3 = 6.82605$, $o4 = 12.12935$, $o5 = 17.87015$, $o6 = 23.17345$, $o7 = 27.2324$, $o8 = 29.4291$, which each have probability $1/8$. Each state can be followed by two other states, and the probability that a state is followed by any the two other states is $1/2$. More specifically, $o1$ can be followed by $o1$ or $o2$, $o2$ by $o3$ or $o4$, $o3$ by $o5$ or $o6$, $o4$ by $o7$ or $o8$, $o5$ by $o7$ or $o8$, $o6$ by $o5$ or $o6$, $o7$ by $o3$ or $o4$, $o8$ by $o1$ or $o2$. For one of the realisations of $\{V_t\}$ the entries from time 0 to 9 are: $o4$, $o7$, $o4$, $o8$, $o1$, $o1$, $o2$, $o4$, $o8$. This is the sequence shown in Figure 1. Therefore, the time series shown in Figure 1 can also derive from a Markov model. Recall that the very same time series can

also arise from the logistic system. This raises the question of observational equivalence, to which we now turn.

2.2 Observational equivalence

There is observational equivalence when *the deterministic model, relative to an observation function* Φ, *and the stochastic model, relative to an observation function* Γ *give the same predictions* (formally, this kind of observational equivalence is called manifest isomorphism – cf. Werndl 2009a). What it means to give the same predictions needs further elaboration. The predictions obtained from a stochastic model are the probability distributions over its realisations coarse-grained by Γ. Recall that a probability measure p is defined on the phase space of a deterministic model. Therefore, the predictions derived relative to an observation function Φ are the probability distributions over the solutions coarse-grained by Φ. Therefore 'give the same predictions' means that (i) the possible values of the observation function Γ of the stochastic model and of the observation function Φ of the deterministic model are the same, and (ii) the probability distributions over the realisations of the stochastic model coarse-grained by Γ and the probability distributions over the solutions of the deterministic model coarse-grained by Φ are the same.

Given a deterministic model (M, T_t, p) and an observation function $\Phi :$ $M \to M_O$, can an observationally equivalent stochastic model be found? Yes: $\{Z_t\} := \{\Phi(T_t)\}$ is a stochastic model, constructed by applying the observation function to the deterministic model. The possible values of $\{\Phi(T_t)\}$ are the same as the possible observed values of (M, T_t, p). Further, the realisations of $\{\Phi(T_t)\}$ and the solutions of (M, T_t, p) coarse-grained by Φ have the same probability distributions. Hence *there is observational equivalence* between the stochastic model $\{\Phi(T_t)\}$ (assuming that all values can be observed, i.e. that Γ is the identity function) and the deterministic model (M, T_t, p), relative to Φ. One might wonder whether $\{\Phi(T_t)\}$ has only trivial probabilities (0 and 1) because it derives from applying an observation function to a deterministic model. However, importantly, this is *not* so. It can be shown that in several cases the stochastic model $\{\Phi(T_t)\}$ is *nontrivial* (i.e., there are probabilities assigned to outcomes that are strictly between 0 and 1) (Werndl, 2009a, 2011).

This result can be illustrated with the example of the evolution of the temperature. We know that we can describe the evolution of the temperature by the logistic map (equation 1). The set of possible values of the observation function Φ_8 is the same as the set of all possible outcomes of the stochastic model $\{V_t\} = \{\Phi_8(T_t)\}$. The probability distributions of this stochas-

tic model are determined by applying Φ_8 to the logistic map and hence are identical to those of $\{\Phi_8(T_t)\}$. Instances of these identical probability distributions are: $p(\Phi_8(T_t) = o1) = P(Z_t = o1)$ or $p(\Phi_8(T_{t+1}) = o2$ given that $\Phi_8(T_t) = o1) = P(Z_{t+1} = o2$ given that $Z_t = o2)$ for all $t \in \mathbb{N}$. Hence the conclusion is that the logistic map, relative to Φ_8, and the stochastic model $\{\Phi_8(T_t)\}$ (when all values are observed) are observationally equivalent. Indeed, $\{\Phi_8(T_t)\}$ is the Markov model $\{V_t\}$. Thus the logistic map, relative to Φ_8, and the Markov model $\{V_t\}$ are observationally equivalent. So we have found an explanation of why the time series of Figure 1 can arise from both the logistic map and the $\{V_t\}$.

In an insightful paper Berlanger (2013) investigates whether manifest isomorphism can serve as a *purely mathematical* notion of observational equivalence. He concludes that the answer is negative because the specific choice of the observation function will depend on the context and the physical situation at hand. While I agree with his conclusion, I do not think that this shows that there is anything wrong about manifest isomorphism. Observational equivalence is about observations. Hence it is only desirable that manifest isomorphism is not a purely mathematical notion and that the physical situation at hand will influence the choice of the observation function. What is important is that there are investigations of physical phenomena that can be regarded as instances of manifest isomorphism. This is certainly the case. For instance, chaos theory is about deterministic systems that are nevertheless unpredictable and show irregular and random behaviour (cf. Werndl, 2009b). And in the context of chaos theory many scientists reported that they first described a physical phenomenon with a stochastic model only to find later that the data can also be regarded as deriving from a deterministic system (e.g. Shaw, 1984).

The results I have presented so far only show that there can be observational equivalence between deterministic and stochastic models. Yet one might still doubt that *stochastic models arising in scientific theorising* (in short: stochastic models in science) *can be observationally equivalent to deterministic models arising in scientific theorising* (in short: deterministic models in science). If such doubts were justified, then one could divide the probability distributions found in science into two groups: the ones deriving from observations of deterministic systems in science, and the ones deriving from observations of stochastic systems in science. Then one might argue that if the observed probability distributions are of the type of stochastic models in science, this amounts to evidence for a stochastic model, and if they are of the type of deterministic models, this provides evidence for an underlying deterministic system. Clearly, such an argument only works if stochastic models in

science cannot be observationally equivalent to deterministic models in science.

Indeed, this is what Kolmogorov believed. More specifically, Kolmogorov introduced the Kolmogorov-Sinai entropy to measure the amount of information produced by a stochastic model and a deterministic model, and he expected that deterministic models in science have positive entropy and that stochastic models in science have zero entropy. But when Kolomogorov tried to prove this conjecture, he failed. A few years later it was found that many deterministic systems in science including Newtonian systems have positive Kolmogorov-Sinai entropy (cf. Sinai 1989, 835–837; Werndl 2011). In conclusion, Kolmogorov's attempt of separating deterministic models in science from stochastic models in science failed. Indeed, *many deterministic models in science including Newtonian models are observationally equivalent (i.e. manifestly isomorphic) to stochastic models in science.* To come back to our example of the evolution of the temperature, the logistic map (a deterministic model in science) relative to the observation function Φ_8 is observationally equivalent to the Markov model $\{V_t\}$ (Markov models are widely used in science). Relative to the coarser observation function $\Phi_4(m)$, where $\Phi_4(m) = 2.1967$ for $0 \leq m < 4.3934$, $\Phi_4(m) = 9.6967$ for $4.3934 \leq m < 15$, $\Phi_4(m) = 20.3033$ for $15 \leq m < 25.6066$, $\Phi_4(m) = 27.8033$ for $25.6066 \leq m \leq 30$, the logistic map is even observationally equivalent to a Bernoulli model with two outcomes (i.e. to a series of coin tosses) (Werndl, 2009a, 2011).

One might think that *only if certain coarse* observation functions are applied to deterministic models in science, one obtains stochastic models in science and that fine-enough observations of deterministic systems in science should yield probability distributions that do not derive from stochastic models in science. In other words, one might doubt that deterministic models in science are observationally equivalent to stochastic models in science *at every observation level.*

The idea of observational equivalence at every observation level can be spelled out in various ways (cf. Werndl, 2011). Here I will only discuss results about the most commonly used notion referring to $(\varepsilon_1, \varepsilon_2)$-*congruence*, which was introduced by the mathematician Ornstein (for technical details, see Ornstein and Weiss, 1991; Werndl, 2009a, 2011). For a sufficiently small $\varepsilon_1 \geq 0$, one will not be able to distinguish states of the deterministic system which are less than the distance ε_1 apart. Further, suppose that for sufficiently small $\varepsilon_2 \geq 0$ one will not be able to distinguish differences in probabilities of less than ε_2. Then a deterministic model and a stochastic model are $(\varepsilon_1, \varepsilon_2)$-congruent (i.e. give the same predictions at level $(\varepsilon_1, \varepsilon_2)$) iff there is a one to

one correspondence between the solutions of the deterministic model and the realisations of the stochastic model such that the state of the deterministic model and the outcome of the stochastic model are at all time points within distance ε_1 except for a set of probability smaller than ε_2.

The following can be shown for many deterministic models in science, including our example of the logistic map (the evolution of the temperature) and several Newtonian models: for any arbitrary $\epsilon_1 > 0$ and $\varepsilon_2 > 0$ there is a Markov model which is $(\varepsilon_1, \varepsilon_2)$-congruent to the deterministic model (Ornstein and Weiss, 1991; Werndl, 2009a). From this the conclusion is drawn that the doubts raised above cannot be substantiated: deterministic models in science can indeed be observationally equivalent at every observation level to stochastic models in science.

Berlanger (2013) has argued that $(\varepsilon_1, \varepsilon_2)$-congruence is not sufficient for observational equivalence because the set of points where the models differ by ε_2 (henceforth called the ε_2-set) is only restricted in its probability measure and *not in its distribution*. He goes on to construct examples of $(\varepsilon_1, \varepsilon_2)$-congruent models which differ at regular time intervals and argues that they are not observationally equivalent because they differ systematically and detectably.

Berlanger's argument is spot on. It shows that in order to arrive at a valuable notion of observational equivalence, $(\varepsilon_1, \varepsilon_2)$-congruence needs to be *strengthened* by adding the condition that the ε_2-set is *distributed randomly* (to match our expectations of random experimental noise). Indeed, for the examples discussed by Ornstein and Weiss (1991) and Werndl (2009, 2011) it is easy to see that they cannot differ at regular time intervals as in Berlanger's counterexample.[4] Still, there remains the question whether for the examples of $(\varepsilon_1, \varepsilon_2)$-congruence discussed in the literature, the ε_2-set is really distributed in a way that matches our expectations of random noise.

In our context it is important that it can be shown that several deterministic models in science (including the logistic map) are $(\varepsilon_1, 0)$-congruent to Markov models for every $\varepsilon_1 \geq 0$. Here Berlanger's concerns do not arise because there is no exceptional set of positive measure ε_2 where the models differ. To conclude, *there are indeed deterministic models in science that are observationally equivalent at every observation level to stochastic models in science.*

---

[4]These examples are chaotic (strongly mixing), which implies that regular differences as in Berlanger's counterexample are impossible.

2.3   Choice and underdetermination

As shown above, there are cases where deterministic and stochastic models are observationally equivalent. Let one of the cases be given where the deterministic model (M, T_t, p) relative to Φ and the stochastic model $\{\Psi(T_t)\}$ relative to Γ are observationally equivalent. Here there is a choice between different models and the question arises: *is the deterministic model or the stochastic model preferable relative to evidence?* If the data equally supports the deterministic and the stochastic model, there is *underdetermination*. To illustrate this with our example: here the question arises whether to choose the logistic map or the Markov model $\{\Phi_8(T_t)\}$ to describe the evolution of the temperature in London.

Suppes (1993) and Suppes and de Barros (1996) argue that there is underdetermination in these cases. Yet more care is needed. In particular, in order to answer the question which model is preferable, one needs to specify the class of observations under consideration. The two main cases are (i) the *currently possible* observations given the available technology (which is the kind of choice arising in practice), and the (ii) the observations which are *possible in principle* (where it is assumed that there are no limits, in principle, on observational accuracy) (cf. Werndl, 2013a).

Let me first consider case (ii), i.e. the observations which are possible in principle. Here one quickly sees that *there is no underdetermination*. If always finer observations can be made, then the deterministic model is preferable (since only the deterministic model allows that always finer observations can be made). On the other hand, suppose the possible observations show that there are no other states apart from those corresponding to the values of a certain observation function Ψ. Then the stochastic model $\{Z_t\} = \{\Psi(T_t)\}$ is preferable because only this model does not have more states. Hence there is no underdetermination. Winnie (1998) and Wüthrich (2011) also present an argument along these lines to argue against Supper's (1993) underdetermination thesis.

Let me turn to case (i), i.e. the currently possible observations. To avoid a trivial answer, assume that Φ is at least as fine as the currently possible observations and that hence it is not possible to find out whether there are more states than the ones corresponding to the values given by Ψ. In other words, the predictions of the deterministic model and stochastic model $\{Z_t\} = \{\Psi(T_t)\}$ agree at all currently possible observation levels. To provide an example, if $\Psi := \Phi_8$ corresponds to an observation at least as fine as the

currently possible observations, then the logistic map and $\{\Psi_8(T_t)\}$ will give the same predictions at all currently possible observation levels.

Werndl (2013a) argues that underdetermination can still be avoided in the case most commonly discussed in the literature, i.e. the choice between deterministic models derived from Newtonian theory and stochastic models obtained by applying observation functions to these deterministic models. Her argument involves the idea of *indirect evidence*, which is best explained with an example. Galileo's theory is only about the motion of bodies on Earth and Kepler's theory is only about the motion of planets. So data about planets cannot be derived from Galileo's theory (with help of standard auxiliary assumptions). Still, data about planets support Kepler's theory and, with Newtonian mechanics as a bridge, they provide indirect evidence for Galileo's theory. As emphasised by Laudan and Leplin (1991), indirect evidence can block underdetermination. For instance, suppose there is a hypothesis H from which the same predictions are derivable than from Galileo's theory but which does not follow from Newtonian mechanics (or another general theory). In this case there is no underdetermination because only Galileo's theory (but not H) is supported by indirect evidence. Along these lines Werndl (2013) argues that for deterministic models derived from Newtonian mechanics there is indirect evidence from other similar Newtonian models. Yet, for the stochastic models there is no indirect evidence. Hence the deterministic models are preferable, and there is no underdetermination. To illustrate this argument with our simple example of the evolution of the temperature: suppose that the logistic map were derivable from the generally well-confirmed theory of fluid dynamics but the stochastic model not. Then the deterministic model would receive indirect evidence from other similar models of fluid mechanics but the stochastic model would not. Hence in this case the deterministic model would be preferable.

Let us now turn to the second main theme of this article, where the concern is an altogether different version of indeterminism that does *not* involve any probability distributions.

3   Indeterminism in Newtonian physics

3.1   Examples of indeterminism

In the past decades the question whether the equations of Newtonian physics are deterministic or not has received much interest in the philosophy of science
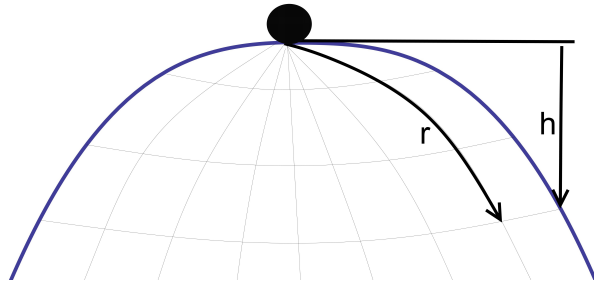
Figure 3: The mass sitting at the highest point on Norton's dome

community. Contra to popular belief that Newtonian physics is deterministic, the answer to this question is not clear cut.

Two kinds of examples have been discussed in the literature that are taken to show that Newtonian mechanics is indeterministic. The first class of examples are systems *where the initial conditions of the bodies do not uniquely fix their solutions*. For example, Norton (2003, 2008) discusses a system where a point particle of unit mass is moving on a dome of the shape shown in Figure 3. The dome is rotationally symmetric about the origin $r = 0$, which corresponds to the highest point of the dome. The shape of the dome is specified by

$$h(r) = \frac{2}{3g}r^{\frac{3}{2}}, \tag{3}$$

describing how far the surface of the dome lies below the highest point, where r is the radial distance coordinate at the surface of the dome. The mass is accelerated by the gravitational force along the surface. At any point of the surface the gravitational force tangential to the surface is directed radially outward and is assumed to be given by

$$F = \frac{d(gh)}{dr} = r^{\frac{1}{2}} \tag{4}$$

(there is no tangential force at $r = 0$). Recall Newton's second law of motion: $F = m.a$ (i.e. the force equals the mass times the acceleration). When this law is applied to the radial acceleration d^2r/dt, one obtains

$$\frac{d^2r}{dt} = r^{\frac{1}{2}}. \tag{5}$$

If the mass sits initially (at $t = 0$) at the highest point of the dome, there is an obvious solution given by $r(t) = 0$ for all times t. Yet, unexpectedly, there

is another class of solutions given by:

$$r(t) \;=\; \frac{1}{144}(t-T)^4 \;\text{ for } t \geq T \tag{6}$$

$$r(t) \;=\; 0 \;\text{ for } t \leq T, \tag{7}$$

where $T \geq 0$ is an arbitrary constant.[5] Hence the evolution of the mass sitting initially at the highest point of the dome is not determined: it can stay on the highest point forever or start moving down the surface at an arbitrary point of time.

Mathematicians have of course investigated under what conditions unique solutions exist for differential equations. In this context, a crucial condition is *local Lipschitz continuity*. Intuitively speaking, a local Lipschitz continuous function is limited in how fast it can change. More formally: a function $F(r)$ is *locally Lipschitz* continuous when, for every initial state x in the domain, there is a neighbourhood U of x such that for the restriction of F to U it holds that for all r, s in U:

$$|F(r) - F(s)| \leq K|r - s|. \tag{8}$$

The Picard-Lindelöf theorem from the theory of ordinary differential equations is a key theorem about the existence and uniqueness of differential equations. It states that if the force F is locally Lipschitz continuous on its domain, then there is a unique maximally extended solution to the equation of motion (cf. Arnol'd, 1992). Most differential equations used in science fulfill the conditions of the Picard-Lindelöf theorem. Since for Norton's dome there is a failure of uniqueness, Norton's dome is not locally Lipschitz continuous. More specifically, for Norton's dome $F(r) = r^{\frac{1}{2}}$ (cf. equations 4 and 5), and for $F(r) = r^{\frac{1}{2}}$ condition (8) fails for $x = 0$ when the particle is at rest at the highest point of the dome.

The second kind of examples that are taken to show that Newtonian mechanics is indeterministic concern *space invaders*. There is no upper bound to the speed of particles in Newtonian mechanics. Hence there is the possibility of space invaders, i.e. particles zooming in from spatial infinity in a finite amount of time. For instance, Xia (1992) proved that there can be space invaders for $n \geq 5$ particles where each of the n particles is subject to the gravitational influence of the other $n - 1$ particles and there are no other forces present. Hence for an universe that it empty at t_0 there are the following two possibilities. First, obviously, it is possible that the universe remains empty forever.

---

[5](6) and (7) solve Newton's second law (5) because $\frac{d^2 r}{dt} = \frac{d^2(\frac{1}{144}(t-T)^4)}{dt} = \frac{4*3}{144}(t-T)^2 = \frac{1}{12}(t-T)^2 = (\frac{1}{144}(t-T)^4)^{\frac{1}{2}} = r^{\frac{1}{2}}$ for $t \geq T$, and $\frac{d^2 r}{dt} = \frac{d^2 0}{dt} = 0 = 0^{\frac{1}{2}} = r^{\frac{1}{2}}$ for $t \leq T$.

However, second, it is also possible that after t_0 there are, say, five particles in the universe with interactions as discussed in Xia (1992). Hence Newtonian particle mechanics is an indeterministic theory: the fate of the universe after t_0 is not determined: it can stay empty forever or be occupied by five particles.

3.2 Is Newtonian physics indeterministic?

Based on these examples, various authors such as Earman (1986) and Norton (2003, 2008) have argued that Newtonian physics is indeterministic. Others objected to this conclusion and defended the claim that Newtonian physics is deterministic. For instance, Korolev (2007) argues for local Lipschitz continuity as an implicit assumption of Newtonian physics. In Korolev (2007) one also finds the idea that it is improper idealizations that lead to indeterminism and that once these improper idealizations are abandoned, the resulting systems will be deterministic. For instance, if the dome were not completely rigid (and in reality it certainly is not completely rigid), then it would deform in a way that guarantees local Lipschitz continuity and hence would prevent indeterminism. Zinkernagel (2010) claims that Norton's dome arises from an incorrect application of the first law. In essence, what Zinkernagel requires is that the first law is understood in a way that every force has a first "cause" and that forces are not turned on smoothly from zero to non-zero magnitude. Applied to Norton's dome this yields the conclusion that the particle at rest at the top of the dome must stay at rest because for the other solutions the forces are turned on smoothly from zero to non-zero magnitude.

While there is much that has been learned from this debate, I would agree with Fletcher (2012) that the question "Is Newtonian mechanics indeterministic" is too simplistic. The answer to this question depends on what one takes Newtonian mechanics to be, and there is no unequivocal answer to this question (cf. Malament, 2008; Wilson, 2009). There are various different conceptions of Newtonian physics, which are all useful in their own way and none of them is a priori privileged over the others. For instance, while an applied mathematician may only consider forces with certain continuity properties, a physicist may focus on a class of properties of models that can be investigated through experiments. There are still interesting scientific and philosophical questions about determinism to be answered, for instance, whether a certain configuration of matter allows for space invaders or what role Lipschitz indeterminism plays in fluid dynamics. Yet these questions will be about precisely specified versions of Newtonian physics rather than about "the Newtonian physics".

This does not mean that there is little that has been learnt from this

debate. On the contrary, the analysis of the various indeterministic systems and the detailed investigations into the reasons why determinism obtains or fails has been highly insightful. To provide a few examples, we have learned that the violation of Lipschitz continuity is benign in the sense that the force on a ball rolling off the table is not Lipschitz at the point where it looses contact (Fletcher, 2012). Zinkernagel's (2010) analysis has shown us that are two different understandings of Newton's first law (one where it is required that every force has a first "cause" and one where this is not the case) and that some cases of indeterminism can be avoided by requiring that every force has a first "cause". Furthermore, we have learned that trying to exclude the indeterministic examples by forbidding certain kinds of idealizations is unlikely to succeed as virtually all of the idealizations are used elsewhere without any complaint (Fletcher 2012; Wilson 2009).

3.3   Determinism and indeterminism: past and present

Although examples of indeterminism in Newtonian physics have attracted much attention in the philosophy community in the past decades, these examples are nothing new. They have already been discussed in the nineteenth century by scientists such as Poisson, Duhamel and Bertrand. In an interesting paper van Strien (2014) compares the current debates to those in the nineteenth century and argues:

> [...] nineteenth century conceptions of determinism were essentially different from the contemporary conception of determinism in classical physics. Contemporary philosophers of physics largely regard determinism as a property of the equations of physics, specically as the statement that for each system there are equations of motion with unique solutions for given initial conditions. However, I show that in the nineteenth century, this claim was not strongly established, and that the authors that I discuss from this period treated determinism in an essentially different way. Specifically, from their arguments it appears that they thought that determinism could hold even in cases in which the equations of physics did not have a unique solution for given initial conditions. [...] Apparently, for these nineteenth century authors, whether or not there was determinism in physical reality did not necessarily depend on whether the equations of physics had unique solutions. This indicates that for them, determinism was not an idea based on the properties of the equations of physics, but rather an a priori principle that was possibly based on metaphysical considerations about causality or

16

the principle of sufficient reason; rather than a result derived from science, determinism was a presupposition of science, that had to be upheld even if it was not reflected in the equations (van Strien 2014, 168).

There are two main claims made by van Strien. The first claim is that nineteenth century conceptions of determinism were *essentially different* from those in use today. The second claim is that determinism was not a result derived from science but an *a priori principle* possibly based on metaphysical considerations about causality or the principle of sufficient reasons. Let me discuss these claims in turn.

I agree with van Strien that the nineteenth century discussion has a very different focus, but I would not say that their conceptions of determinism were different. As discussed in the introduction, determinism is the idea that the state of the system at one time determines the future and past evolution of the system, and this idea can be applied either to models and equations or to physical systems. So I would rather say that while all had the same idea of determinism in mind, the focus is different: the debate in the nineteenth century centres on determinism in real physical systems, but the current debate focuses on determinism in equations or models (as opposed to real systems).

About the second claim: van Strien is right to emphasise that metaphysical considerations about causality or the principle of sufficient reason possibly (even likely) influenced the thinking of scientists in the nineteenth century. Still, I would not go so far to claim that determinism was an a priori principle. It seems likely that part of scientists' reason in the nineteenth century to upheld determinism was *the empirical success of deterministic equations*. In other words, there was (and still is no) evidence that the indeterminism showing up in examples such as Norton's dome or the space invaders is a real feature of physical systems (cf. Fletcher 2012; Wilson 2009). Instead, there was a vast amount of evidence confirming deterministic equations, and it is likely that this contributed to the general belief that physical systems are governed by deterministic laws.[6]

Let me finally turn to the third main topic of this article: the question of deterministic probabilities.

---

[6]As Wilson (2009) describes in detail, physicists coming across indeterministic equations often find that there are gaps in the mathematical description of the physical system, and that once these gaps are closed, determinism is regained.
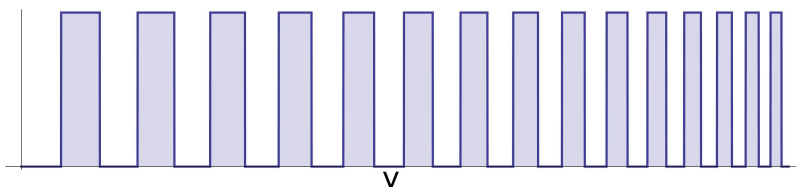
Figure 4: Conversion of initial velocities v into grey and white outcomes

4 Deterministic probabilities

4.1 The method of arbitrary functions

Philosophers often puzzle how stable *ontic probabilities*, i.e. probabilities that are real features of the world, can arise out of *deterministic equations*. The *method of arbitrary functions* is philosophically important because it shows how this is possible. It has an illustrious history and has been advocated, amongst others, by Hopf, Poincare, Reichenbach and von Kries (cf. von Plato, 1983).

Let me give a simple example introducing the method of arbitrary functions (cf. Strevens, 2011). Consider a very simple wheel of fortune, i.e. a wheel painted in equal numbers of small equal-sized white and grey sections. The wheel is spun with a certain initial velocity and when it comes to rest a fixed pointer indicates the outcome (white or grey). Our immediate judgement is that the probability of the outcome 'grey' is 1/2, even though the wheel is governed by deterministic equations.[7]

This judgement can be substantiated by analysing the wheel of fortune in more detail. The first component we have to look at is how the *dynamics* of the wheel converts initial velocities into outcomes. Figure 4 shows the conversion of initial velocities into grey and white outcomes. The crucial feature that emerges here is what Strevens (2003) calls *microconstancy*, i.e. that for small ranges of initial velocities the proportion of initial velocities leading to the outcome 'grey' is 1/2 and, likewise, the proportion of initial velocities leading to the outcome 'white' is 1/2. The second component we have to look at is the *preparation of the wheel of fortune in a certain initial velocity*. This preparation is modelled by a probability distribution p over initial velocities. Of course, we usually know very little about this initial distribution, and different

---

[7]If quantum effects crop up, the wheel will be governed by equations which are approximately deterministic (everything that will be said in this paper carries over to this case).
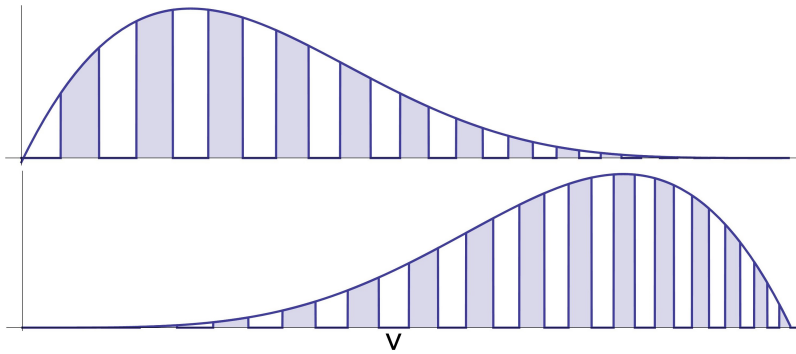
Figure 5: Probabilities of 'grey' and 'white' for two different initial probability densities

ways of spinning the wheel (by different persons or the same person in different contexts) will correspond to different initial probability distributions. Yet this does not matter. All we need is the plausible assumption that the class of probability densities p we might possibly employ are what Strevens (2003) calls *macroperiodic*, i.e. do not fluctuate drastically on a very small region. Then, as illustrated by Figure 5 for two different initial probability densities, the probabilities for the outcomes 'grey' and 'white' will approximately be $1/2$.[8] To sum up, there are stable probabilities for the wheel of fortune even though the wheel is governed by deterministic equations. What explains these stable probabilities is that (i) the dynamics is microconstant, and that (ii) the class of possible probability distributions describing the preparation of the wheel are macro-periodic.

The method of arbitrary functions is particularly relevant when there is a *class of possible initial densities* (e.g., because the method of preparation differs from scientist to scientist[9]). It is not meant to apply to all situations where there are ontic probabilities but only to certain cases. The prime examples to which the method of arbitrary functions has been successfully applied are games of chance (cf. Strevens, 2003; von Plato, 1983). It has also been suggested to apply this method to probabilities in statistical mechanics, ecology and the social sciences (the method has not been proven to apply to realistic systems of these disciplines since the relevant mathematics is extremely diffi-

---

[8]The name 'arbitrary functions' is fitting in the sense that a large class of initial probability densities assigns probability $1/2$ to the outcomes 'grey' and 'white'. Yet it is also misleading in the sense that only certain initial densities lead to these probabilities. Hence at least some plausibility arguments need to be given for the assumption that the initial densities are macroperiodic.

[9]Of course, the class of possible densities can also arise from purely natural processes.

cult) (Abrams, 2012; Strevens, 2003; Werndl, 2010). In particular, although occasionally discussed (e.g. Myrvold 2012, forthcoming), the method of arbitrary functions deserves more attention in *statistical mechanics*. A common complaint in statistical mechanics has been that different ways of preparing, say, a gas, result in different initial probability distributions, and that thus identifying the microcanonical measure (restricted to a macroregion) with the initial probability density is besides the point (Leeds, 1989; Maudlin, 2007; Werndl, 2013). Clearly, the method of arbitrary functions is ideally suited to make sense of the idea that there is a class of possible initial probability distributions which all lead to (approximately) the same macroscopic behaviour.

Let us now turn to the question of how to interpret the probabilities that arise by applying the method of arbitrary functions and, related to this, how to interpret the initial probability distributions.

4.2   Interpretational issues

Myrvold (2011, 2014) suggests applying the method of arbitrary functions to statistical mechanics. He interprets the initial probability distributions as representing agents' possible rational credences about finding the system in a certain initial state. He argues that the resulting probabilities of outcomes that arise out of the dynamics are approximately equal to the measure assigned to these outcomes by the microcanonical measure. Hence the probabilities obtained by the method of arbitrary functions combine epistemic (credences) and physical considerations (the deterministic dynamics). To reflect this, Myrvold (2011, 2014) calls these probabilities '*epistemic chances*'. Myrvold's discussion is insightful. In particular, his point that all that is needed is that the standard probability measures in statistical mechanics assign probabilities to outcomes that are effectively the same as the correct probabilities (but need not be exactly the same) cannot be stressed enough.

The only concern I have is that credences cannot be invoked to explain how systems behave as they do. Myrvold (2014, 33) anticipates this concern when he writes:

> There is a connection, however, between the epistemic considerations we have invoked, and what would be required of an explanation of relaxation to equilibrium. The processes that are responsible for relaxation to equilibrium are also the processes that are responsible for knowledge about the system's past condition of non-equilibrium becoming useless to the agent.

There certainly is this link. Despite this, in my view, the initial probability distributions of the method of arbitrary functions should be primarily understood not as credences but as ontic probabilities describing the preparation of the system (since there is a fact of the matter in which states a system is prepared) (cf. Rosenthal, 2012). The probabilities of the method of arbitrary functions arise from these ontic initial probability distributions, while the epistemic chances arising from credences as propounded by Myrvold are *derivative*.

Rosenthal (2010, 2012) explores the intriguing idea that for the method of arbitrary functions probability is defined in terms of the *phase space structure* (various small contiguous regions, on each of which the proportion of the outcomes is the same). The idea is that for such a phase space structure the probability of an outcome simply amounts to the proportion of initial conditions of phase space that lead to this outcome. To illustrate this idea with the wheel of fortune: the probability of 'grey' is 1/2 because half of the initial conditions lead to the outcome 'grey' (cf. Figure 2). This idea is doomed to failure because the phase space structure *plus* certain initial probability distributions are needed to arrive at stable probabilities (Rosenthal also seems to acknowledge this). For instance, it is possible to construct a machine that prepares the wheel of fortune with an initial speed that always leads to the outcome 'grey'. Hence in this case the probability of 'grey' will not be 1/2.

Instead, what seems promising is the idea that the initial probability distributions are negligible in the following sense: whenever there are stable probabilities at the macroscopic level arising *for various different ways of preparing the system*, then one can expect that the probabilities derive from the *phase space structure*. The underlying reasoning is that if the same probabilities arise for various ways of preparing a system, then one can expect that the initial probability distributions are macroperiodic. If the initial probability distributions were not macroperiodic, slight changes in the way the system is prepared would lead to different macroscopic probabilities, contradicting the existence of stable probabilities.[10]

Let me make the side remark that whenever the structure of phase space is discussed, the assumption is that the phase space is composed of *various* small contiguous regions, on each of which the proportion of the outcomes is the same (e.g. Abrams, 2012; Rosenthal, 2010, 2012; Strevens, 2003). How-

---

[10]As Rosenthal (2010, 2012) remarks, this argument assumes that our usual phase spaces and their metric properties are privileged (in the sense that probabilities over distorted representations of standard physical quantities cannot be expected to be macroperiodic).

ever, for stable probabilities to arise, various small contiguous regions are not necessary. Indeed, there are cases where there is just one such region. For instance, consider a funnel located precisely above a central nail (i.e. a Galton board with just one level). When the ball is poured into the funnel, the ball either bounces to the left or to the right of the nail. Assuming that there is only a vertical velocity component when pouring the ball into the funnel, the phase space consists of just one contiguous region with initial conditions corresponding to two outcomes (landing 'right' and landing 'left' of the nail). Still, there are stable probabilities for the outcomes 'right' and 'left' (both 1/2) because the class of initial probability distributions describing the various ways of preparing the system contains only (approximately) symmetric distributions.

Strevens (2011) takes a different route. He claims that for nearly all long series of trials of the system the initial conditions form macroperiodically distributed sets. As Strevens stresses, this means that the initial distributions just represent actual occurrences of initial states and have nothing to do with *probabilities*. He regards this as desirable because then under his interpretation of probability based on the method of arbitrary functions probabilities arise from *non-probabilistic facts*. At the same time, by appealing to a robustness condition (that nearly all long series of trials produce macroperiodically distributed sets), he avoids the major pitfalls of finite frequentism.

Streven's proposal is worthwhile. Still, there are some problems. First, Strevens makes the 'nearly all'-condition more precise by stating that the actual distributions are macro-periodic in nearly all relevantly close possible worlds, and he appeals to the Lebesgue measure to quantify this claim. There is the worry that it is nontrivial to justify the Lebesgue measure as the correct measure over close possible worlds (cf. Rosenthal, 2010, 2012). Further, it is not clear to me whether it is formally possible to assign a measure to 'ways of altering the actual world' (the exact formal details are not spelled out by Strevens). Also, consider again the initial velocities of the wheel of fortune. When the wheel is spun by the same person in a specific context again and again, the frequency distribution of the initial velocities approximates the shape of a certain density. For these reasons scientists usually postulate a probability density describing the probability of preparing the wheel with a certain initial velocity. This probability density is predictively useful: it is (usually) found to give accurate predictions about the future frequencies of initial velocities produced by the person. However, because under Strevens' account the initial velocities are nothing more than actual occurrences of initial states, there cannot be any such predictive power.

In my view, there is a way to avoid these problems. Namely, I suggest to interpret the initial distributions simply as *probability distributions that are physical, biological etc. quantities characterising the particular situation at hand* (as suggested by Szabò, 2007; see also Sober, 2010). That is, the concept of probability reduces to ordinary physical, biological etc. quantities, and the word "probability" is a collective term whose precise meaning depends on the context of its application. Accordingly, the probabilities arising from the method of arbitrary functions also correspond to physical, biological etc. quantities that characterise particular situations.These probabilities are ontic and, as desired, they support counterfactuals about future predictions.[11]

A final remark: Werndl (2013b) put forward an account of how to understand typicality measures in statistical mechanics. The main idea is that there is a class of initial probability distributions of interest, which corresponds to the possible ways of preparing a system (elements of this class are assumed to be translation-continuous or translation-close). The typicality measure can then be used to make claims about all initial probability distributions of interest. E.g. if the claim is that typical initial conditions show thermodynamic-like behaviour, this implies that for all initial probability distributions of interest the probability of thermodynamic-like behaviour is close to one. Werndl (2013b) does not mention the method of arbitrary functions. Still, formally, the justification of the typicality measures involves (next to other ingredients) the application of the method of arbitrary functions to outcomes that have probability close to one. The slogan 'typicality is not probability' is correct here in the sense that the typicality measure is a measure at the microlevel and does not amount to a probability over the initial states (such as the probability of finding the system in a certain state).

4.3   Puzzles about deterministic probabilities resolved

Philosophers have often doubted that there can be ontic probabilities in a deterministic world (see, e.g., Schaffer, 2007, for a recent paper expressing such doubts). As our discussion has shown, the method of arbitrary functions provides an explanation of how probabilities can arise out of determinism.

---

[11]If the initial distributions are understood as summarising actual frequencies, the method of arbitrary functions is also compatible with an account of probabilities as Humean chances (Frigg and Hoefer, 2010, 2014). As for all Humeans, it is not trivial to say what it means to strike the best balance between simplicity, strength and fit. The account I propose does not have to deal with such difficulties.

A prominent worry about deterministic probabilities is that they lead to a violation of the *Principal Principle*, which establishes a connection between credences and chances (e.g. Schaffer, 2007). According to this principle, a rational agent's credence in the occurrence of an event E should equal the chance of E as long as the agent has no inadmissible knowledge about the truth of E. Formally: for all events E, all P and all K

$$cr_t(E\,|P\&K) = p \qquad (9)$$

where 'cr_t' stands for the credence of the agent at time t, P is the proposition that the chance of E is p and K is any admissible proposition. The crucial question here is what an 'admissible proposition' amounts to. Lewis (1986) suggested that historical information about the exact state of a system up to time t as well as information about the laws of nature are always admissible. But given deterministic laws, this implies that the exact future state of the system can be predicted. Hence the credences in equation (9) can only be 0 and 1 and for nontrivial deterministic probabilities p this leads to a violation of the Principal Principle. Commonly, the conclusion drawn from this is that there exist no nontrivial deterministic probabilities (since a violation of the Principal Principle is deemed to be unacceptable).

This argument is too quick and there are better ways to characterise an admissible proposition. Following Frigg and Hoefer (2014, 4) let a proposition K be *admissible* with respect to event E and chance setup S iff "K contains only the sort of information whose impact on reasonable credence about E, if any, comes entirely by way of impact on credence about the chances of those outcomes". With this modified notion of admissability, as desired, the Prinicipal Principle (equation 9) comes out as true. As Glynn (2010) stresses, when making rational decisions, it would be a real *loss* if we could not rely on macroscopic probabilities such as those of the outcome 'Tail' for a coin toss, the outcome 'Black' for a wheel of fortune etc. Glynn argues that chances are level-relative and that only the initial history and laws at the specific *level of reality* should count as admissible (this amounts to a special case of Frigg and Hoefer's general definition of 'admissibility').

Probabilities arising from deterministic equations for games of chance, in statistical mechanics etc., are often said to be *epistemic* in the sense that if we had precise knowledge about the initial conditions and the deterministic laws, then we would not need them. Our discussion puts these claims into the right perspective. The probabilities of the method of arbitrary functions are *not* epistemic in the sense that they are credences. The only sense in which they are epistemic is that they are relative to a certain *level of reality*. If the

system were described at the level where determinism reigns, we would indeed not need any probabilities (cf. Frigg and Hoefer, 2010; Glynn, 2010; Rosenthal, 2011).

5 Conclusion

This article centred on three major themes in the recent discussion on determinism in philosophy of science. The first major theme was determinism, indeterminism and observational equivalence. Here I first critically discussed various notions of observational equivalence and then presented results about the observational equivalence between deterministic and indeterministic models. I also put forward an argument on how to choose between deterministic and indeterministic models involving the idea of indirect evidence. The second major theme concerned the question whether Newtonian physics is indeterministic. I argued that the answer to this question depends on what one takes Newtonian mechanics to be. Further, I discussed how contemporary debates on this issue differ from those of scientists in the nineteenth century. In particular, I pointed out that the focus of scientists in the nineteenth century was on investigating whether determinism holds for real systems rather than just for certain equations or models. The third major theme was how the method of arbitrary functions can be used to make sense of deterministic probabilities. Here I discussed various ways of interpreting the initial probability distributions and I argued that they are best understood as physical, biological etc. quantities characterising the particular situation at hand. I also emphasised that the method of arbitrary functions deserves more attention than it has received so far and that it is among the most promising contenders for understanding probabilities in certain fields, e.g. classical statistical mechanics.

The topic of determinism is a very old one, but it is still very much alive: There is a lot that has to be learnt in the recent decades and there is a lot that still remains to be discovered.

References

Abrams, M. (2012). "Mechanistic probability." *Synthese* 187 (2): 343-375.

Arnol'd, V.I. (1992). *Ordinary differential equations* (Berlin: Springer).

Berlanger, C. (2013). "On two mathematical definitions of observational equiv-

alence: manifest isomorphism and epsilon-congruence reconsidered." *Studies in History and Philosophy of Science Part B* 44 (2): 69-76.

Earman, J. (1986). *A primer on determinism* (Berlin and New York: Springer).

Frigg, R., and Hoefer, C. "Determinism and chance from a Humean perspective." In *The Present Situation in the Philosophy of Science*, edited by F. Stadler, pp. 351-372. Dordrecht: Springer, 2010.

Frigg, R. and Hoefer, C. (2014). "The Best Humean System for Statistical Mechanics." Forthcoming in: *Erkenntnis*.

Fletcher, S. (2012). "What counts as Newtonian system – the view from Norton's dome." *European Journal for the Philosophy of Science* 2 (3): 275-297.

Glynn, L. (2010). "Deterministic chance." *The British Journal for the Philosophy of Science* 61 (1): 51-80.

Korolev, A.V. (2007). "Indeterminism, asymptotic reasoning, and time irreversibility in classical physics." *Philosophy of Science* 74 (5): 943-956.

Laudan, L. and Leplin, J. (1991) Empirical equivalence and underdetermination". *The Journal of Philosophy* 88, 449-472.

Leeds, S. (1989). "Malament and Zabell on Gibbs phase space averaging." *Philosophy of Science* 56: 325-340.

Lewis, D. (1986). "A subjectivit's guide to objective chance". In *Studies in Inductive Logic and Probability, Volume 2*, edited by R.C. Jeffrey, pp. 83-132. Berkeley: University of California Press.

Lorenz, E. (1964). "The problem of deducing the climate from the governing equations." *Tellus* 16 (1): 1-11.

Malament, D. (2008). "Norton's slippery slope." *Philosophy of Science* 75 (5): 799-816.

Maudlin, T. (2007). "What could be objective about probabilities?" *Studies in History and Philosophy of Modern Physics* 38: 275-291.

Myrvold, W. "Deterministic laws and epistemic chances." In *Probability in*

Physics, edited by Y. Ben-Menahem and M. Hemmo, pp. 73-85. New York: Springer, 2012.

Myrvold, W. "Probabilities in statistical mechanics" Forthcoming in *Oxford Handbook of Probability and Philosophy*, edited by C. Hitchcock and A. Hajek. Oxford: Oxford University Press.

Norton, J. (2003). "Causation as folk science." *Philosopher's Imprint* 3(4): 1-22.

Norton, J. (2008). "The dome: An unexpectedly simple failure of determinism." *Philosophy of Science* 75 (5): 786-798.

Ornstein, D. and Weiss, B. (1991). "Statistical Properties of Chaotic Systems." *Bulletin of the American Mathematical Society* 24: 11-116.

Rosenthal, J. "The natural-range conception of probability." In *Time, Chance and Reduction, Philosophical Aspects of Statistical Mechanics*, edited by G. Ernst and A. Httemann, pp. 71-91. Cambridge: Cambridge University Press, 2010.

Rosenthal, J. (2012). "Probabilities as ratios of ranges in initial state spaces." *Journal of Logic, Language and Information* 21: 217-236.

Schaffer, J. (2007): "Deterministic Chance?" *The British Journal for the Philosophy of Science* 58: 113-40.

Shaw, R. (1984). *The Dripping Faucet* (Santa Cruz: Aerial Press).

Sinai, Y.G. (1989). "Kolmogorov's Work on Ergodic Theory." *The Annals of Probability* 17: 833–839.

Sober, E. "Evolutionary theory and the reality of macro probabilities." In *Probability in Science*, edited by E. Eells and J. Fetzner, pp. 133-162. Heidelberg: Springer, 2010.

Strevens, M. (2003). *Bigger than chaos* (Cambridge (Mass.): Harvard University Press).

Strevens, M. (2011). "Probability out of determinism." In *Probabilities in physics*, edited by C. Beisbart and S. Hartmann, pp- 339-364. Oxford: Oxford

University Press, 2011.

Suppes, P. (1993). "The transcendental character of determinism." *Midwest Studies in Philosophy* 18: 242–257.

Suppes, P. and de Barros, A. "Photons, billiards and chaos." In: *Law and Prediction in the Light of Chaos Research*, edited by P. Weingartner and G. Schurz, pp. 189-201. Springer, Berlin, 1996.

Szabò, L. E. (2007). "Objective probability-like things with and without indeterminism." *Studies in History and Philosophy of Modern Physics* 38: 626-634.

von Plato, J. (1983). "The method of arbitrary functions." *The British Journal for the Philosophy of Science* 34: 37-47.

Van Strien, M. (2014). "The Norton dome and the nineteenth century foundations of determinism." *Journal for General Philosophy of Science* 45 (1): 167-185.

Werndl, C. (2009a). "Are Deterministic Descriptions and Indeterministic Descriptions Observationally Equivalent?" *Studies in History and Philosophy of Modern Physics* 40: 232-242.

Werndl, C. (2009b). "What are the New Implications of Chaos for Unpredictability?" *The British Journal for the Philosophy of Science* 60: 195-220.

Werndl, C. (2010). "The Simple Behaviour of Complex Systems Explained? Review of 'Bigger Than Chaos, Understanding Complexity Through Probability' (Michael Strevens)." *The British Journal for the Philosophy of Science* 61 (4): 875-882.

Werndl, C. (2011). "On the Observational Equivalence of Continuous-Time Deterministic and Indeterministic Descriptions." *European Journal for the Philosophy of Science* 1 (2): 193-225.

Werndl, C. (2013a). "On Choosing Between Deterministic and Indeterministic Models: Underdetermination and Indirect Evidence." *Synthese* 190 (12): 2243-2265.

Werndl, C. (2013b). "Justifying Typicality Measures of Boltzmannian Statistical Mechanics and Dynamical Systems." *Studies in History and Philosophy*

of Modern Physics 44 (4): 470-479.

Wilson, M. (2009). "Determinism and the mystery of the missing physics." *The British Journal for the Philosophy of Science* 60 (1): 173-193.

Winnie, J. "Deterministic Chaos and the Nature of Chance." In *The Cosmos of Science – Essays of Exploration*, edited by J. Erman and J. Norton, pp. 299-324. Pittsburgh: Pittsburgh University Press, 1998.

Wuethrich, C. "Can the world be shown to be indeterministic after all?" In *Probabilities in Physics*, edited by C. Beisbart and S. Hartmann, pp. 365-389. Oxford: Oxford University Press, 2011.

Xia, Z. (1992). "The existence of noncollision singularities in Newtonian systems." *Annals of Mathematics* 135: 411-468.

Zinkernagel, H. "Causal fundamentalism in physics" In *EPSA Philosophical Issues in the Sciences: Launch of the European Philosophy of Science Association*, edited by M. Suarez, M. Dorato, and M. Redei, pp. 311-322. Dordrecht: Springer, 2010.