

# Evidence for the Deterministic or the Indeterministic Description? – A Critique of the Literature about Classical Dynamical Systems

Charlotte Werndl, Lecturer, c.s.werndl@lse.ac.uk  
Department of Philosophy, Logic and Scientific Method  
London School of Economics

Forthcoming in: Journal for General Philosophy of Science

## **Abstract**

It can be shown that certain kinds of classical deterministic descriptions and indeterministic descriptions are observationally equivalent. In these cases there is a choice between deterministic and indeterministic descriptions. Therefore, the question arises of which description, if any, is preferable relative to evidence. This paper looks at the main argument in the literature (by Winnie, 1998) that the deterministic description is preferable. It will be shown that this argument yields the desired conclusion relative to in principle possible observations where there are no limits, in principle, on observational accuracy. Yet relative to the currently possible observations (the kind of choice of relevance in practice), relative to the actual observations, or relative to in principle observations where there are limits, in principle, on observational accuracy the argument fails because it also applies to situations where the indeterministic description is preferable. Then the paper comments on Winnie's (1998) analogy between his argument for the deterministic description and his argument against the prevalence of Bernoulli randomness in deterministic descriptions. It is argued that while there is indeed an analogy, it is also important to see that the arguments are disanalogous in another sense.

# Contents

1	Introduction	3
2	Deterministic and Indeterministic Descriptions	4
3	The Results on Observational Equivalence	10
4	Preference and Underdetermination	14
5	The Nesting Argument	16
6	The Analogy to Bernoulli Randomness	20
7	Conclusion and Outlook	23
	References	26

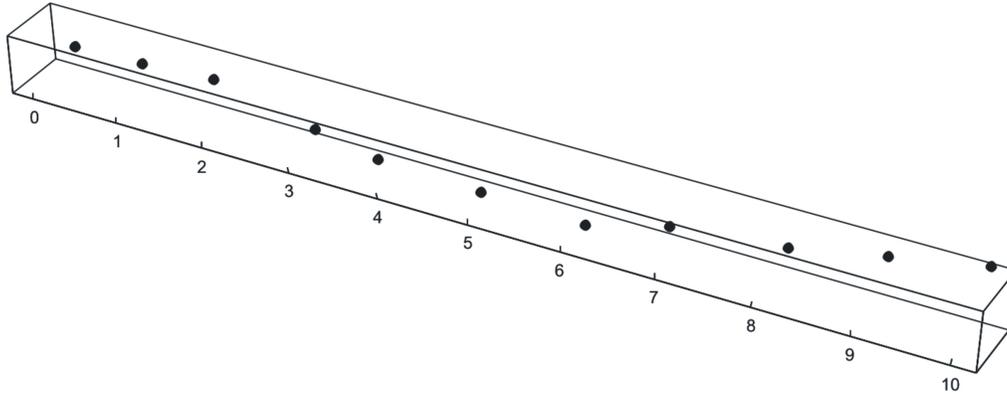


Figure 1: Observed positions of the particle at time 0, 1, . . . , 9, 10

## 1 Introduction

Consider the motion of a particle on a two-dimensional square. When one observes this motion, one obtains a sequence of observations. Figure 1 shows such a sequence, namely it represents the observed positions of the particle at time 0, 1, . . . , 9, 10. Science aims to find a model or description which reproduces these observations and which enables one to make predictions about the relevant phenomenon (in our example the position of the particle). One of the questions which arise here is whether the phenomenon is best described by a deterministic or an indeterministic model. Intuitively speaking, one might think that the data only allow for a deterministic or an indeterministic description. However, as we will see, in general, this is not so. In several cases, including our example of the position of the particle, the data allow for a deterministic as well as an indeterministic description. This means that there is a choice between a deterministic and an indeterministic description. Hence the question arises: Is the deterministic or the indeterministic description preferable? More specifically, *does the evidence favour the deterministic or the indeterministic description?* If none of the descriptions is favoured by evidence, there will be underdetermination.

This paper will critically discuss the main argument which purports to show that the deterministic description is preferable. More specifically, Section 1 will introduce the kinds of deterministic and indeterministic descriptions we will be dealing with, namely measure-theoretic deterministic descriptions and stochastic descriptions. Section 2 will present recent results

which show that, in several cases, the data allow for a deterministic and indeterministic description, i.e., that there is observational equivalence between deterministic and indeterministic descriptions. Section 3 will point out that this raises the question of which description is preferable relative to evidence, and several kinds of choice and underdetermination will be distinguished. Section 4 will criticise Winnie’s (1998) argument for the deterministic description. It will be shown that this argument yields the desired conclusion relative to in principle possible observations where there are no limits, in principle, on observational accuracy. Yet relative to the currently possible observations (the kind of choice of relevance in practice), the actual observations, or the in principle possible observations where there are limits, in principle, on observational accuracy the argument fails because it also applies to situations where the indeterministic description is preferable. Section 5 will comment on Winnie’s (1998) analogy between his argument for the deterministic description and his argument against the prevalence of Bernoulli randomness in deterministic descriptions. It will be argued that while there is indeed an analogy, it is also important to see that the arguments are disanalogous in another sense. Finally, Section 7 will summarise the discussion and provides an outlook; as a part of this, an argument will be sketched that, in certain cases, one of the descriptions is preferable and there is no underdetermination.

## **2 Deterministic and Indeterministic Descriptions**

This paper will only deal with certain kinds of deterministic and indeterministic descriptions – namely measure-theoretic deterministic descriptions and stochastic descriptions. These descriptions can be divided into two classes. Namely, either the time parameter varies in discrete steps (discrete-time descriptions), or the time parameter varies continuously (continuous-time descriptions). For reasons of simplicity, this paper concentrates on discrete-time descriptions. However, all that will be said holds analogously also for continuous-time descriptions. An intuitive characterisation of the discrete-time descriptions will be given; for the technical details the reader is referred to Werndl (2009a, 2011).

## Deterministic Descriptions

According to the canonical definition of determinism, a description is *deterministic* iff (if and only if) the state of the description at one time determines the state of the description at all (i.e., future and past) times. Similarly, a real-world system is *deterministic* iff the state of the system at one time determines the state of the system at all (i.e., future and past) times.

This paper deals with measure-theoretic deterministic descriptions, which are among the most important descriptions in science, and, for instance, include all deterministic descriptions of Newtonian theory. They are defined as follows. A *deterministic description* consists of a triple  $(X, f_t, P)$ , where the set  $X$  (called the *phase space*) represents all possible states of the system;  $f_t(x)$  (called the *evolution functions*), where  $t$  ranges over the integers, are functions, which tell one that the state  $x$  evolves to  $f_t(x)$  in  $t$  time steps; and  $P(A)$  is a probability measure<sup>1</sup> assigning a probability to subsets  $A$  of  $X$ . The *solution* through  $x$  represents the path of the system through  $x$ . Formally, it is the bi-infinite sequence  $(\dots, f_{-2}(x), f_{-1}(x), f_0(x), f_1(x), f_2(x), \dots)$ ;  $f_t(x)$  is called the  $t$ -th *iterate* of  $x$ . Because the  $f_t$ 's are functions, deterministic descriptions thus defined are deterministic according to the canonical definition given above.

When one observes a state  $x$  of a deterministic system, a value is observed which is dependent on  $x$  but is usually different from it (because one cannot observe with infinite precision). Hence an observation is modeled by a function  $\Phi(x) : X \rightarrow O$  (called the *observation function*) where  $O$  represents the set of all possible observed values. This means that when one observes a deterministic system over time, one sees the iterates coarse-grained by the observation function. For instance, if one observes the system starting in state  $x$  at times 0, 1, ..., 9, 10 one obtains the string  $(\Phi(f_0(x)), \Phi(f_1(x)), \dots, \Phi(f_9(x)), \Phi(f_{10}(x)))$ . A *finite-valued* observation function takes only finitely many values. In practice, observations are finite-valued; thus, in what follows, it will be assumed that observation functions are finite-valued. Given two observation functions  $\Phi$  and  $\Psi$ ,  $\Psi$  is called *finer* than  $\Phi$  and  $\Phi$  is called *coarser* than  $\Psi$  iff for any value  $o$  of  $\Psi$  there is a value  $v$  of  $\Phi$  such that for all  $x \in X$  if  $\Psi(x) = o$ , then  $\Phi(x) = v$ , and there are distinct values  $o_1, o_2$  of  $\Psi$  and a value  $w$  of  $\Phi$  such that for all  $x \in X$  if

---

<sup>1</sup>There are various possibilities of interpreting this probability measure. For instance, according to the popular time-average interpretation, the probability of  $A$  is defined as the (long-run) proportion of time that a solution spends in  $A$  (see Werndl 2009b).

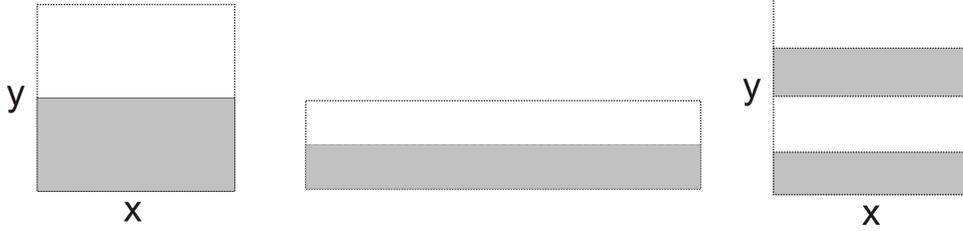


Figure 2: The map  $f((x, y))$  of the baker's transformation

	$\mathbf{e}_{13}$	$\mathbf{e}_{14}$	$\mathbf{e}_{15}$	$\mathbf{e}_{16}$
$\mathbf{y}$	$\mathbf{e}_9$	$\mathbf{e}_{10}$	$\mathbf{e}_{11}$	$\mathbf{e}_{12}$
	$\mathbf{e}_5$	$\mathbf{e}_6$	$\mathbf{e}_7$	$\mathbf{e}_8$
	$\mathbf{e}_1$	$\mathbf{e}_2$	$\mathbf{e}_3$	$\mathbf{e}_4$
	$\mathbf{x}$			

Figure 3: The observation function  $\Phi_{16}$

$\Psi(x) = o_1$  or  $\Psi(x) = o_2$ , then  $\Phi(x) = w$ .

The following two examples of deterministic descriptions will accompany us throughout this paper.

*Example 1: The baker's transformation.*

On the unit square  $X = [0, 1] \times [0, 1]$  consider the map:

$$f((x, y)) = \left(2x, \frac{y}{2}\right) \text{ if } 0 \leq x < \frac{1}{2}; \left(2x - 1, \frac{y+1}{2}\right) \text{ if } \frac{1}{2} \leq x \leq 1. \quad (1)$$

As shown in Figure 2, the map first stretches the unit square to twice its length and half its width; the rectangle obtained is then cut in half and the right half is stacked on top of the left. For  $f_t((x, y)) := f^t((x, y))$  ( $f$  is applied  $t$  times) and for the uniform probability measure  $P$  on the unit square, one obtains the deterministic description  $(X, f_t, P)$  called the baker's transformation. This description models the motion of a particle subject to Newtonian mechanics which bounces on several mirrors. More specifically, the description models a particle which moves with constant speed in the direction perpendicular to the unit square. The particle starts in initial position  $(x, y)$ , then it bounces on several mirrors, which cause it to return to the unit square at  $f((x, y))$  (for more on this, see Pitowsky

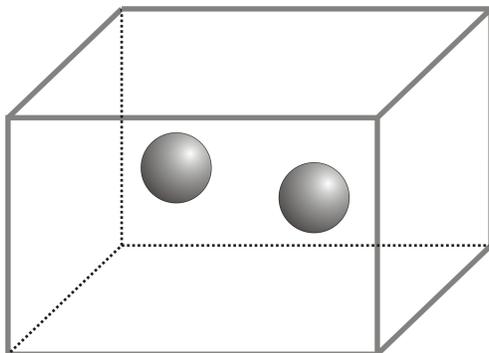


Figure 4: Two hard balls moving in a box

1996, 166). Now consider the observation function  $\Phi_{16}$  with sixteen values  $e_1, e_2, \dots, e_{16}$  assigned as follows:  $\Phi_{16}((x, y)) = (\frac{2i+1}{8}, \frac{2j+1}{8})$  for  $\frac{i}{4} \leq x < \frac{i+1}{4}$  and  $\frac{j}{4} \leq y < \frac{j+1}{4}$ ,  $i, j \in \mathbb{N}$ ,  $0 \leq i, j \leq 3$ . Figure 3 shows this observation function. Suppose that the baker's transformation is initially in  $(0.824, 0.4125)$ . Then the first 11 iterates coarse-grained by the observation function  $\Phi_{16}$  are:  $(\Phi_{16}(f_0((0.824, 0.4125))), \Phi_{16}(f_1((0.824, 0.4125))), \dots, \Phi_{16}(f_{10}((0.824, 0.4125))))$   
 $= ((\frac{7}{8}, \frac{3}{8}), (\frac{5}{8}, \frac{5}{8}), (\frac{3}{8}, \frac{7}{8}), (\frac{5}{8}, \frac{3}{8}), (\frac{1}{8}, \frac{5}{8}), (\frac{3}{8}, \frac{3}{8}), (\frac{5}{8}, \frac{1}{8}), (\frac{3}{8}, \frac{5}{8}), (\frac{7}{8}, \frac{3}{8}), (\frac{7}{8}, \frac{5}{8}), (\frac{7}{8}, \frac{7}{8}))$ . This is exactly the sequence shown in Figure 1. Hence the time series shown in Figure 1 can arise by observing the baker's system.

*Example 2: Description of two hard balls in a box.*

Consider the continuous-time Newtonian mechanical description of two hard balls with a finite radius and no rotational motion; the balls move in a three dimensional box and interact by elastic collisions (see Simányi 1999). Figure 4 visualises such a hard ball system. These hard ball descriptions are of some importance in statistical mechanics because they are the mathematically most tractable models of a gas. Because the hard ball potential involves strong idealising assumptions, it does not accurately describe the potential of real gases. Despite this, in some contexts hard ball models are expected to share properties of real gases, and for this reason they are relevant, e.g., in the context of the H-theorem and the foundations of statistical mechanics. Now suppose that one considers this continuous-time description at discrete time-intervals, i.e., at time points  $ns_0$ ,  $s_0 \in \mathbb{R}^+$  arbitrary,  $n \in \mathbb{Z}$ . Then one obtains what will be our second main example, namely a discrete-time description  $(X, f_t, P)$  of two hard balls in a box. The phase space  $X$  represents all possi-

ble states of the system and is the set of all vectors consisting of the possible position and velocity coordinates of the two balls. If the system is initially in state  $x$ , then the evolution functions tell one that the system will evolve to  $f_t(x)$  after  $t.s_0$  time steps. The meaning of the probability measure  $P$  is as follows: For subsets  $A$  of the phase space,  $P(A)$  is the probability that the two hard balls are in one of the states which are represented by  $A$ . Finally, the solution through  $x$  is the sequence  $(\dots, f_{-2}(x), f_{-1}(x), f_0(x), f_1(x), f_2(x), \dots)$ , which represents the path of the hard ball system which is initially in  $x$  (over the time points  $ns_0, n \in \mathbb{Z}$ ).

## Stochastic Descriptions

A description of a real-world process is *indeterministic* iff it is not deterministic. The indeterministic descriptions of concern in this paper are stochastic descriptions. They model probabilistic processes and are the most important indeterministic descriptions in science. A *stochastic description* is denoted by  $\{Z_t\}$ . It consists of a family of functions  $Z_t : \Omega \rightarrow E$ , where  $t \in \mathbb{Z}$ ,  $E$  is a set (called the *outcome space*) representing all possible outcomes of the process, and  $\Omega$  is a set which encodes all possible paths of the process (i.e., each  $\omega \in \Omega$  encodes a specific path in all its details but is usually unknown in practice). The outcome of the process at time  $t$  is given by  $Z_t(\omega)$ . Since stochastic descriptions model probabilistic processes, there is a probability distribution  $P(Z_t \in A)$ , giving the probability that the outcome is in  $A$  at time  $t$  (for subsets  $A$  of  $E$  and any  $t \in \mathbb{Z}$ ). Similarly, there is a probability distribution  $P(Z_t \in A \text{ and } Z_r \in B)$ , giving the probability that the outcome is in  $A$  at time  $t$  and in  $B$  at time  $r$  (for subsets  $A, B$  of  $E$  and any  $t, r \in \mathbb{Z}$ ); and there is a probability distribution  $P(Z_t \in A \text{ given } Z_r \in B)$ , assigning a probability to the event that the outcome is in  $A$  at time  $t$  given that it is in  $B$  at time  $r$  (for subsets  $A, B$  of  $E$  and any  $t, r \in \mathbb{Z}$ ).

The typical situation is that given the present outcome of a stochastic description, several possible outcomes can follow, and the likelihood of these possible outcomes is measured by the probability distribution. Because several possible outcomes can follow, the description is indeterministic. A *realisation* of the stochastic description represents a possible path of the stochastic process over time. Formally, it is a bi-infinite sequence  $(\dots, Z_{-2}(\omega), Z_{-1}(\omega), Z_0(\omega), Z_1(\omega), Z_2(\omega), \dots)$  for an arbitrary  $\omega \in \Omega$ .

Two stochastic descriptions will accompany us throughout this paper.

*Example 3: Bernoulli descriptions.*

The well-known tossing of an unbiased coin over time is a special case of a Bernoulli description. Generally, for Bernoulli descriptions at each point of time a (possibly biased)  $N$ -sided die is tossed, and each toss is independent of the other ones. Formally, a Bernoulli description  $\{Z_t\}$  is defined as follows: (i) the outcome space consists of  $N$  ( $N \in \mathbb{N}$ ) symbols  $E = \{e_1, \dots, e_N\}$ ; (ii) the probability for the side  $e_m$  is  $Q_m$ , i.e.,  $P(Z_t = e_m) = Q_m$ ,  $1 \leq m \leq N$ , for all times  $t \in \mathbb{Z}$ , where  $Q_1 + \dots + Q_N = 1$ ; (iii) and the condition of independence holds, i.e.,  $P(Z_{l_1} = e_{i_1}, \dots, Z_{l_h} = e_{i_h}) = P(Z_{l_1} = e_{i_1}) \dots P(Z_{l_h} = e_{i_h})$  for all  $e_{i_1}, \dots, e_{i_h}$  in  $E$  and all  $i_1, \dots, i_h \in \mathbb{Z}$ ,  $h \in \mathbb{N}$ . A realisation of a Bernoulli description represents the sequence of the outcomes obtained (e.g., for the fair coin the sequence of heads and tails).

*Example 4: Markov descriptions.*

Markov descriptions are probabilistic descriptions where the next outcome depends only on the previous outcome. So, unlike Bernoulli descriptions, Markov descriptions are history-dependent but only the previous outcome matters. Markov descriptions are among the most frequently encountered descriptions in science. Formally, a Markov description  $\{Z_t\}$  is defined as follows: (i) the outcome space consists of  $N$  ( $N \in \mathbb{N}$ ) symbols  $E = \{e_1, \dots, e_N\}$ ; (ii) and the next outcome only depends on the previous one, i.e.,  $P(Z_t = e_h \text{ given } Z_{t-1}, Z_{t-2}, \dots, Z_j) = P(Z_t = e_h \text{ given } Z_{t-1})$  for all  $t \in \mathbb{Z}$ , all  $j \leq t-1$  and all outcomes  $e_h$ .

An example of a Markov description is the description  $\{W_t\}$ : There are sixteen possible outcomes  $e_1 = (1/8, 1/8)$ ,  $e_2 = (3/8, 1/8)$ ,  $e_3 = (5/8, 1/8)$ ,  $e_4 = (7/8, 1/8)$ ,  $e_5 = (1/8, 3/8)$ ,  $e_6 = (3/8, 3/8)$ ,  $e_7 = (5/8, 3/8)$ ,  $e_8 = (7/8, 3/8)$ ,  $e_9 = (1/8, 5/8)$ ,  $e_{10} = (3/8, 5/8)$ ,  $e_{11} = (5/8, 5/8)$ ,  $e_{12} = (7/8, 5/8)$ ,  $e_{13} = (1/8, 7/8)$ ,  $e_{14} = (3/8, 7/8)$ ,  $e_{15} = (5/8, 7/8)$ ,  $e_{16} = (7/8, 7/8)$ .  $P(e_i) = 1/16$ ,  $1 \leq i \leq 16$ . Each present outcome can be followed by two other outcomes, where the probability that the present outcome is followed by any of these two outcomes is  $1/2$ :  $e_1$  is followed by  $e_1$  or  $e_2$ ,  $e_2$  by  $e_3$  or  $e_4$ ,  $e_3$  by  $e_9$  or  $e_{10}$ ,  $e_4$  by  $e_{11}$  or  $e_{12}$ ,  $e_5$  by  $e_1$  or  $e_2$ ,  $e_6$  by  $e_3$  or  $e_4$ ,  $e_7$  by  $e_9$  or  $e_{10}$ ,  $e_8$  by  $e_{11}$  or  $e_{12}$ ,  $e_9$  by  $e_5$  or  $e_6$ ,  $e_{10}$  by  $e_7$  or  $e_8$ ,  $e_{11}$  by  $e_{13}$  or  $e_{14}$ ,  $e_{12}$  by  $e_{15}$  or  $e_{16}$ ,  $e_{13}$  by  $e_5$  or  $e_6$ ,  $e_{14}$  by  $e_7$  or  $e_8$ ,  $e_{15}$  by  $e_{13}$  or  $e_{14}$ ,  $e_{16}$  by  $e_{15}$  or  $e_{16}$ . In one of the realisations of  $\{W_t\}$  the entries from time 0 to 10 are:  $(e_8, e_{11}, e_{14}, e_7, e_9, e_6, e_3, e_{10}, e_8, e_{12}, e_{16})$ . This is exactly the sequence shown in Figure 1. Hence the time series shown in Figure 1 can arise from a Markov description. Recall that the time series shown in Figure 1 can also arise from observing the deterministic baker's

transformation. This gives a first hint that there is observational equivalence between deterministic and stochastic descriptions. The next section will explain the results on observational equivalence between deterministic and stochastic descriptions.

### 3 The Results on Observational Equivalence

*Observational equivalence* between a stochastic description, a deterministic description and an observation function of the deterministic description as understood in this paper means that *the stochastic description gives the same predictions as the deterministic description relative to the observation function*. Note that observational equivalence in this sense is a relation between a stochastic description, a deterministic description, *and an observation function*. That is, this notion of observational equivalence is relative to a specific observation level. What does it mean that a stochastic description and a deterministic description, relative to an observation function, give the same predictions? The predictions obtained from the stochastic description are the probability distributions over the sequences of outcomes (i.e., over the realisations). Recall that there is a probability measure  $P$  defined on the phase space of the deterministic description. Therefore, relative to the observation function, the predictions derived from the deterministic description are the probability distributions over the sequences of observed values (i.e., over the solutions of the description coarse-grained by the observation function).

This means that a stochastic description  $\{Z_t\}$  *gives the same predictions* as a deterministic description  $(X, f_t, P)$  relative to the observation function  $\Phi$  iff the following holds: The set of possible outcomes  $E$  of  $\{Z_t\}$  is identical to the set of possible observed values of  $\Phi$ ; and the probability distributions over the sequences of outcomes of  $\{Z_t\}$  are the same as the probability distributions over the sequences of observed values of  $(X, f_t, P)$  relative to  $\Phi$ .<sup>2</sup>

Consider an observation function  $\Phi : X \rightarrow O$  of a deterministic de-

---

<sup>2</sup>Here the paper follows the extant literature Suppes (1993), Suppes and de Barros (1996) and Winnie (1998) in assuming that measure-theoretic deterministic and stochastic descriptions are tested and confirmed by deriving probabilistic predictions from them. That nonlinear models can be confirmed has been questioned (see Bishop 2008). This is just to mention the controversy about confirmation of nonlinear models; a thorough treatment would require another paper.

scription  $(X, f_t, P)$ . Then  $\{Z_t\} := \{\Phi(f_t)\}$  is a stochastic description. It arises by applying the observation function  $\Phi$  to the deterministic description. Therefore, the set of possible outcomes of  $\{Z_t\}$  is identical to the set of possible observed values of  $(X, f_t, P)$ , and the probability distributions over the sequences of outcomes of  $\{Z_t\}$  and the probability distributions over the sequences of observed values of  $(X, f_t, P)$  are the same. From this it follows that  $(X, f_t, P)$ , relative to  $\Phi$ , is observationally equivalent to  $\{\Phi(T_t)\}$ .

Let me explain this result with the example of the baker's transformation<sup>3</sup> (Example 1). Suppose that the observation function  $\Phi_{16}$  is applied to the baker's transformation (see Example 1 for the definition of  $\Phi_{16}$ ). Then the set of possible outcomes of the stochastic description  $\{\Phi_{16}(f_t)\}$  is the same as the set of possible values of  $\Phi_{16}$ , and the probability distributions of this stochastic description are determined by applying  $\Phi_{16}$  to the baker's transformation. Consequently, the deterministic baker's transformation, relative to  $\Phi_{16}$ , and the stochastic description  $\{\Phi_{16}(f_t)\}$  are observationally equivalent. This example is interesting because the stochastic description  $\{\Phi_{16}(f_t)\}$  is actually identical to the Markov description  $\{W_t\}$  (see Example 4 for the definition of  $\{W_t\}$ ).<sup>4</sup> Hence also *the Markov description  $\{W_t\}$  and the baker's transformation, relative to  $\Phi_{16}$ , are observationally equivalent*. This explains why the time series shown in Figure 1 can arise from both the baker's transformation, relative to  $\Phi_{16}$ , and the Markov description  $\{W_t\}$ . What is shown in Figure 1 is  $V = ((\frac{7}{8}, \frac{3}{8}), (\frac{5}{8}, \frac{5}{8}), (\frac{3}{8}, \frac{7}{8}), (\frac{5}{8}, \frac{3}{8}), (\frac{1}{8}, \frac{5}{8}), (\frac{3}{8}, \frac{3}{8}), (\frac{5}{8}, \frac{1}{8}), (\frac{3}{8}, \frac{5}{8}), (\frac{7}{8}, \frac{3}{8}), (\frac{7}{8}, \frac{5}{8}), (\frac{7}{8}, \frac{7}{8}))$  – the first eleven iterates of the baker's transformation starting in state

<sup>3</sup>The baker's transformation involves strong idealisations – in particular, that there is no friction. Consequently, scientists do not derive probabilistic predictions from it to compare it with the data. Because the example is so easy to understand, instead of referring to another example, the reader is asked to pretend that scientists indeed use it for this purpose.

<sup>4</sup>This is not hard to see. As shown in Werndl (2009a, Section 4.1), the baker's transformation  $(X, f_t, P)$  is isomorphic via  $\phi$  to the Bernoulli shift  $(Y, h_t, Q)$  with values 0 and 1 (the states are bi-infinite sequences  $y = \dots y_{-2}y_{-1}y_0y_1y_2\dots$ ,  $y_i \in \{0, 1\}$ ). For  $(Y, h_t, Q)$  define the observation function  $\Xi_{16}$  as follows. It takes the value  $(\frac{1}{8}, \frac{1}{8})$  if  $y_{-1}y_0y_1y_2 = 0000$ ,  $(\frac{3}{8}, \frac{1}{8})$  if  $y_{-1}y_0y_1y_2 = 0001$ ,  $(\frac{5}{8}, \frac{1}{8})$  if  $y_{-1}y_0y_1y_2 = 0010$ ,  $(\frac{7}{8}, \frac{1}{8})$  if  $y_{-1}y_0y_1y_2 = 0011$ ,  $(\frac{1}{8}, \frac{3}{8})$  if  $y_{-1}y_0y_1y_2 = 1000$ ,  $(\frac{3}{8}, \frac{3}{8})$  if  $y_{-1}y_0y_1y_2 = 1001$ ,  $(\frac{5}{8}, \frac{3}{8})$  if  $y_{-1}y_0y_1y_2 = 1010$ ,  $(\frac{7}{8}, \frac{3}{8})$  if  $y_{-1}y_0y_1y_2 = 1011$ ,  $(\frac{1}{8}, \frac{5}{8})$  if  $y_{-1}y_0y_1y_2 = 0100$ ,  $(\frac{3}{8}, \frac{5}{8})$  if  $y_{-1}y_0y_1y_2 = 0101$ ,  $(\frac{5}{8}, \frac{5}{8})$  if  $y_{-1}y_0y_1y_2 = 0110$ ,  $(\frac{7}{8}, \frac{5}{8})$  if  $y_{-1}y_0y_1y_2 = 0111$ ,  $(\frac{1}{8}, \frac{7}{8})$  if  $y_{-1}y_0y_1y_2 = 1100$ ,  $(\frac{3}{8}, \frac{7}{8})$  if  $y_{-1}y_0y_1y_2 = 1101$ ,  $(\frac{5}{8}, \frac{7}{8})$  if  $y_{-1}y_0y_1y_2 = 1110$ ,  $(\frac{7}{8}, \frac{7}{8})$  if  $y_{-1}y_0y_1y_2 = 1111$ . Since  $(Y, h_t, Q)$  is a Bernoulli shift,  $\{\Xi_{16}(h_t)\}$  is a Markov description, which is easily seen to be identical to  $\{W_t\}$ . Because  $(X, f_t, P)$  is isomorphic via  $\phi$  to  $(Y, h_t, Q)$  and  $\Phi_{16}(x) = \Xi_{16}(\phi(x))$ ,  $\{\Phi_{16}(f_t)\} = \{\Xi_{16}(\phi(f_t))\} = \{\Xi_{16}(h_t(\phi))\}$ ; thus  $\{\Phi_{16}(f_t)\}$  is also identical to  $\{W_t\}$ .

(0.824, 0.4125) coarse-grained by  $\Phi_{16}$ . But  $V = (e_8, e_{11}, e_{14}, e_7, e_9, e_6, e_3, e_{10}, e_8, e_{12}, e_{16})$ , which are the first eleven outcomes of a realisation of the Markov description  $\{W_t\}$ .

The Markov description  $\{\Phi_{16}(f_t)\}$ , which is identical to  $\{W_t\}$ , is clearly a nontrivial stochastic description (a stochastic description is *nontrivial* iff it is not trivial; for a *trivial* stochastic description the probability that any arbitrary outcome follows another arbitrary outcome is always 0 or 1). So there are examples of deterministic descriptions which, relative to some observation functions, yield nontrivial stochastic descriptions. The question arises whether there are *general* results to the effect that the stochastic description  $\{\Phi(f_t)\}$  is nontrivial. This is important: If  $\{\Phi(f_t)\}$  is trivial, then one might argue that  $\{\Phi(f_t)\}$  (although mathematically a stochastic description) is really a deterministic description; and hence that there is no observational equivalence between deterministic and indeterministic descriptions. Several general results show that  $\{\Phi(f_t)\}$  is *often nontrivial*. Let me state one result (Proposition 1 in Werndl 2009a).

**Theorem 1** *Given a deterministic description  $(X, f_t, P)$ , suppose that there is no  $m \in \mathbb{N}$  and  $G \subseteq X$ ,  $0 < P(G) < 1$ , such that  $f_m(G) = G$ . Let  $\Phi : X \rightarrow O$  be an arbitrary nontrivial finite-valued observation function<sup>5</sup>. Then the stochastic description  $\{Z_t\} = \{\Phi(f_t)\}$  is nontrivial in the following sense: For any arbitrary  $n \in \mathbb{N}$  and any  $t \in \mathbb{Z}$ , there exist an  $e_l \in O$  such that for all  $e_h \in O$ ,  $0 < P(Z_{t+n} = e_h \text{ given } Z_t = e_l) < 1$ .<sup>6</sup>*

Theorem 1 applies to Example 1 of the baker's transformation (see Werndl 2009a, Section 3.1) and to Example 2 of the description of two hard balls (see Simányi 1999). It also applies to many other descriptions which are important in science. For instance, it applies to descriptions of  $N$ ,  $N \geq 2$ , hard balls moving on a torus for almost all values  $(m_1, \dots, m_N, r)$ , where  $r$  is the radius of the balls and  $m_i$  is the mass of the  $i$ -th ball (see Simányi 1999 and

---

<sup>5</sup>A *nontrivial* finite-valued observation function has at least two outcomes with positive probability.

<sup>6</sup>Theorem 1 is equivalent to Proposition 1 in Werndl (2009a), but there are two minor differences in the statement of the theorems. First, unlike Werndl (2009a), it is not required that  $(X, f_t, P)$  is ergodic because this follows automatically from the premises (ergodicity is equivalent to the condition that there exists no  $G \subseteq X$ ,  $0 < P(G) < 1$ , such that  $f_1(G) = G$ ; cf. Petersen 1983, 42). Second, it is stated here that  $0 < P(Z_{t+n} = e_h \text{ given } Z_t = e_l) < 1$  for any arbitrary  $n \in \mathbb{N}$  and not only for  $n = 1$  as in Werndl (2009a). This stronger conclusion follows immediately from the weaker one because if the premises hold for  $(X, f_t, P)$ , they also hold for  $(X, f_t^n, P)$  where  $f_t^n$  is the  $n$ -th iterate of  $f_t$ .

2003; the motion of hard balls on a torus is mathematically more tractable than the motion of balls in a box); to many billiard descriptions such as billiards in a stadium (Chernov and Markarian 2006); and also to dissipative descriptions (i.e., where the Lebesgue-measure is not invariant under the dynamics) such as to the Hénon description, which models weather phenomena, for certain parameter values (Benedicks and Young 1993, Hénon 1976). It is mathematically extremely hard to prove that descriptions satisfy the assumption of Theorem 1. Therefore, for many descriptions it is conjectured, but not proven, that they satisfy this assumption, e.g., for any finite number of hard balls in a box or for the Hénon description for many other parameter values (see Benedicks and Young 1993; Berkovitz et al. 2006).

So far we have focused on how to obtain stochastic descriptions when deterministic descriptions are given. There are also converse results about how to obtain deterministic descriptions when stochastic descriptions are given. For what follows later it is not necessary to go into the details of these results. It suffices to say that the basic construction for these converse results is the same: *Given a stochastic description  $\{Z_t\}$ , one finds a deterministic description  $(X, f_t, P)$  and an observation function  $\Phi$  such that  $\{Z_t\}$  is identical to  $\{\Phi(f_t)\}$ .* Moreover, several results show how, given stochastic descriptions of the type used in science, one can find deterministic descriptions of the type used in science which, relative to certain observation functions, are observationally equivalent to these stochastic descriptions. For instance, given Markov descriptions (Example 4), one can find discrete-time Newtonian descriptions which are observationally equivalent to these Markov descriptions (see Werndl 2009a, 2011).<sup>7</sup>

---

<sup>7</sup>There are other results in the literature which are based on quite different assumptions but which could also be interpreted as results about observational equivalence between deterministic and stochastic descriptions. An example is Judd and Smith (2001, 2004); they present results under the ideal condition of infinite past observations of a deterministic system and the assumption that there is observational error modeled by a probability distribution around the true value. While these results are interesting, this paper will not discuss them. They are based on different assumptions and hence are of limited relevance for assessing the extant philosophical literature of concern here (i.e., Suppes 1993, Suppes and de Barros 1996, Winnie 1998).

## 4 Preference and Underdetermination

We have seen that, in certain cases, there is observational equivalence between a stochastic description  $\{Z_t\}$  and a deterministic description  $(X, f_t, P)$  relative to an observation function  $\Phi$ . So the question arises: *Is the stochastic description or the deterministic description preferable relative to evidence?* If evidence equally supports both descriptions, this is a case of *underdetermination*. This section will differentiate between different kinds of choice and underdetermination which could arise.

When asking which description is preferable, it is important to distinguish between the cases where (i) the descriptions are about the *same level of reality*; (ii) the descriptions are about *different levels of reality*. An example for (i) is when both descriptions model the evolution of the positions and velocities of two hard balls in a box. An example of (ii) is a Boltzmannian and Newtonian description of a gas. At the lower level there is the Newtonian description: The states are represented by the position and velocity coordinates of the gas particles and the evolution is deterministic. At the higher level there is the Boltzmannian description: The states are represented by values of variables such as temperature, pressure and volume, and the evolution is stochastic. The higher-level states are a function of the lower-level states (this means that they are related by an observation function) (cf. Frigg 2008, section 3.2). This paper will concentrate on case (i). The extant philosophical literature (Suppes 1993; Suppes and de Barros 1996; Winnie 1998) also focuses on this case. The case (ii) of the descriptions about different levels of reality is very different. For instance, here one might argue that there is no conflict – reality might well be deterministic at one level and stochastic at another level.

A first significant result is the following. Suppose one considers all *in principle possible observations where there are no limits, in principle, on observational accuracy*. (That is, although one will never be able to observe infinitely precisely, in principle possible observations allow one to come arbitrarily close to these infinitely precise values.) It is assumed that it is not the case that the deterministic description  $(X, f_t, P)$  and all of the stochastic descriptions  $\{\Phi(f_t)\}$  are disconfirmed by the in principle possible observations (here none of the descriptions will be acceptable). Then, *relative to these in principle possible observations, there is no underdetermination*: On the one hand, if given any arbitrary finite-valued observation, one can always make finer observations, the deterministic description will be preferable because

only the deterministic description allows that one can always make finer observations. On the other hand, suppose one can observe the values corresponding to the finite-valued observation function  $\Phi$  and observations show that there are no other states apart from these values. Then the stochastic description  $\{Z_t\} = \{\Phi(f_t)\}$  will be preferable because only this stochastic description does not have more states and thus agrees with the observations.

Yet other kinds of underdetermination are not as easily dismissed. Most importantly, first, consider the choice relative to *currently possible observations* (given the available technology). Here, if the observation function  $\Phi$  is fine enough, it will not be possible to find out whether there are more states than the values given by  $\Phi$ . Then the relevant predictions of the deterministic description and of the stochastic description  $\{Z_t\} = \{\Phi(f_t)\}$  are the same (the relevant predictions are those which can be tested by the currently possible observations). If these predictions are confirmed and other evidence does not favour a description, there is underdetermination relative to all currently possible observations. Arguably, of interest in practice is this underdetermination relative to all currently possible observations (cf. Laudan and Leplin 1991). A second case is the choice relative to the *actual observations* which have been made. Here, if  $\Phi$  corresponds to an observation which is finer (or the same) than any of the observations which have actually been made, the relevant predictions of the deterministic description and of the stochastic description  $\{Z_t\} = \{\Phi(f_t)\}$  are the same (the relevant predictions are those which can be tested by the actual observations which have been made). If these predictions are confirmed and other evidence does not favour a description, there is underdetermination relative to the actual observations. A third case is the choice relative to *in principle possible observations where there are limits, in principle, on observational accuracy*. Here, if  $\Phi$  corresponds to an observation which is finer (or the same) than any in principle possible observation, the relevant predictions of the deterministic description and of the stochastic description  $\{Z_t\} = \{\Phi(f_t)\}$  are the same (the relevant predictions are those which can be tested by the in principle possible observations). Again, if these predictions are confirmed and other evidence does not favour a description, there is underdetermination relative to in principle possible observations where there are limits, in principle, on observational accuracy.

In what follows, to avoid a trivial answer, it will be assumed that the relevant observations do not disconfirm the deterministic description and all of the stochastic descriptions (here none of the descriptions will be acceptable).

And whenever the concern is the choice relative to currently possible observations/actual observations/in principle possible observations where there are limits, in principle, on observational accuracy, it will be assumed that  $\Phi$  is fine-enough, viz. that it is currently not possible to find out/that the actual observations do not allow to find out/that in principle possible observations do not allow to find out whether there are more states than the ones given by the observation function  $\Phi$ . Let us now turn to the main argument in the literature against underdetermination.

## 5 The Nesting Argument

Winnie (1998) discusses the choice between Newtonian deterministic descriptions and the stochastic descriptions obtained by applying an observation function  $\Phi$  to these deterministic descriptions. He aims to identify sufficient conditions under which the deterministic description is preferable. Let us turn to his argument. For the deterministic description consider the observation functions of interest relative to the given kind of choice (i.e., the observation functions which, according to the deterministic description, are either in principle possible, or currently possible, or have actually been applied). Suppose that the observations corresponding to these observation functions of interest can be made (i.e., that it is either in principle possible to make them, or that it is currently possible to make them, or that they have been made). Further, suppose that one finds that finer observations lead to stochastic descriptions at a smaller scale (i.e., stochastic descriptions where there is at least one outcome of the stochastic description at a larger scale such that two or more outcomes of the stochastic description at a smaller scale correspond to one outcome of the stochastic description at a larger scale). Then this provides evidence that the phenomenon is deterministic and that the deterministic description is preferable. Winnie endorses this argument, which will be called the ‘nesting argument’:

To be sure, at any stage of the above process, the system may be modeled stochastically, but the successive stages of that modeling process provide ample—inductive—reason for believing that the deterministic model is correct. [Winnie 1998, 315]

For instance, assuming that the motion of the two hard spheres (Example 2) is really deterministic, the nesting argument establishes that the determinis-

tic description is preferable.

Winnie does not explicitly state which kind of choice he is concerned with. Suppose his concern is underdetermination relative to in principle possible observations where there are no limits, in principle, on observational accuracy. As argued in Section 4, here it is straightforward which description is preferable. Now consider the observation functions of interest for the deterministic description, namely the observation functions which, according to the deterministic description, are in principle possible. If the observations corresponding to these observation functions can be made, then the deterministic description is indeed preferable. This is so because only the deterministic description allows that always finer observations can be made, and hence only the deterministic description will agree with the possible observations. And then one automatically also finds that finer observations lead to stochastic descriptions at a smaller scale. Hence *the nesting argument yields the desired conclusion*.<sup>8</sup> Winnie may have well been concerned with the choice relative to in principle possible observations where there are no limits, in principle, on observational accuracy, and here his argument succeeds.<sup>9</sup>

Does the nesting argument deliver the desired conclusion for the other kinds of choice discussed in Section 4, viz. relative to currently possible observations, relative to actual observations, or relative to in principle possible observations where there are limits, in principle, on observational accuracy? It will now be argued that *the nesting argument fails in these cases* by giving an example where the premises of the nesting argument are true but where *not* the deterministic description is preferable. This is important because Winnie develops his argument as a criticism of Suppes (1993) and Suppes and de Barros (1996). These papers defend the claim that *there is underdetermination*: the stochastic description and the deterministic description are equally well supported by evidence. Suppes (1993) and Suppes and de

---

<sup>8</sup>Still, the nesting argument seems a bit complicated because the reason why there is no underdetermination is simply that only the deterministic description agrees with the in principle possible observations.

<sup>9</sup>For in principle possible observations where there are no limits, in principle, on observational accuracy, the nesting argument succeeds in giving sufficient conditions under which the deterministic description is preferable. Here it is also true that *if the premises of the nesting argument are not true, a stochastic description will be preferable*. Then there is an observation function  $\Phi$  such that the values corresponding to  $\Phi$  can be observed, but where observations show that there are no other states apart from these values. Hence only the stochastic description  $\{Z_t\} = \{\Phi(f_t)\}$  agrees with the observations and is thus preferable (cf. Section 4).

Barros (1996) do not state explicitly which kind of choice they are concerned with. Yet they say:

Remember you are not going to predict exactly. That's very important. It's different when you identify the point precisely. Then you have ideal observation points. You would not have a stochastic situation. (Suppes and de Barros 1996, 200)

Because there are several comments in this direction (highlighting that stochastic descriptions arise by coarse-graining the phase space), it is plausible that they were not concerned with in principle possible observations. Rather, it seems, they were concerned with choice relative to the currently possible observations (or something similar). *If so, then Winnie's criticism misses the target* because it fails for the kind of choice Suppes was concerned with.

This counterexample involves the idea of *indirect evidence*. This idea will first be explained with an example. Consider Darwin's theory about natural selection; this theory is only about selection which is natural and not artificial. Nevertheless, data from breeders about artificial selection are generally regarded as supporting Darwin's theory of natural selection. How is this possible? The data from breeders provide indirect evidence for natural selection: They support Darwin's general evolutionary theory; and by supporting the general evolutionary theory, they also support Darwin's theory about natural selection (even though the data are not derivable from this theory).

Laudan and Leplin (1991) and Laudan (1995) have argued that indirect evidence can provide an argument against underdetermination. Namely, suppose that the same predictions are derivable from the theory of natural selection than from another theory  $U$ , and that  $U$  is not derivable from Darwin's general evolutionary theory or any other more general theory. Because only the theory of natural selection is additionally supported by indirect evidence, the theory of natural selection, and not  $U$ , is preferable relative to evidence. Consequently, there is no underdetermination. Thus the main insight here is that even if two theories have the same observational consequences, one theory may still be preferable based on observations because only one theory is additionally supported by indirect evidence.

Okasha (1998) has criticised Laudan and Leplin as endorsing the view that indirect evidence amounts to being derivable from the same statement, which leads to the unacceptable consequence that any statement is confirmed by a

statement.<sup>10</sup> Okasha (1998) is certainly right that such a definition of indirect evidence is untenable. However, with a more sophisticated understanding of indirect evidence, Laudan and Leplin’s insight that indirect evidence can provide an argument against underdetermination holds (as also admitted by Okasha 1998, 2002). Namely, in this paper statement  $A$  is understood to provide indirect evidence for statement  $B$  (where  $A$  is not derivable from  $B$ ) when  $A$  is confirmed and  $A$  and  $B$  are unified by a well-confirmed theory. Clearly, for the example of natural selection this unifying theory is Darwin’s general evolutionary theory. With this understanding, the difficulty pointed out by Okasha (1998) is avoided and Laudan and Leplin’s main insight that indirect evidence can block underdetermination holds.

Now we can construct an argument based on indirect evidence for the stochastic description, exposing the weakness of the nesting argument. First of all, note that, relative to currently possible observations/actual observations/in principle possible observations where there are limits, in principle, on observational accuracy there is a choice between a deterministic description and a stochastic description obtained by applying an observation function  $\Phi$  to the deterministic description where  $\Phi$  corresponds to an observation finer or the same than any of the currently possible/actual/in principle possible observations. Because  $\Phi$  is so fine, *regardless* of whether the stochastic or deterministic description is correct, the following holds for the observation functions which, according to the deterministic description, are currently possible/have actually been applied/are in principle possible: The observations corresponding to these observation functions can be made, and for finer observations one obtains stochastic descriptions at a smaller scale. That is, the premises of the nesting argument are true.

Now suppose that the stochastic description  $\{Z_t\} = \{\Phi(f_t)\}$  can be derived from a well-confirmed theory  $T$ , but that the deterministic description  $(X, f_t, P)$  does not derive from any theory. Also, suppose that there are many descriptions which support  $T$ , which do not derive from  $(X, f_t, P)$  or  $\{Z_t\}$ , and which provide indirect evidence for  $\{Z_t\}$ . It follows that the stochastic description  $\{Z_t\}$  is preferable relative to evidence because it is additionally supported by indirect evidence.<sup>11</sup> Hence the conclusion of the nesting argu-

---

<sup>10</sup>Namely,  $A$  confirms  $A$ . Given any arbitrary statement  $B$ , both  $A$  and  $B$  are derivable from  $A \wedge B$ . From this it follows that  $A$  confirms  $B$  (see Okasha 1998).

<sup>11</sup>For example, consider the choice between a stochastic description in statistical mechanics of the evolution of the macrostates of a gas and the deterministic representation (see Werndl 2009a) of this stochastic description. Then, at the level of reality of the

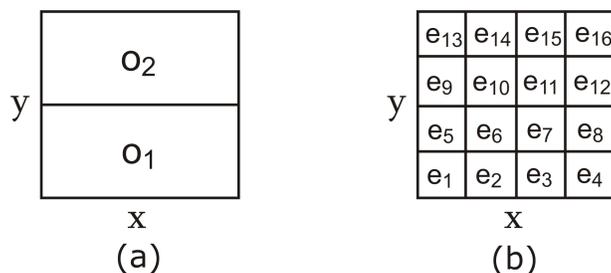


Figure 5: Observation functions of the baker's transformation: (a)  $\Phi_2$  yields a Bernoulli description; (b)  $\Phi_{16}$  yields a Markov description

ment is not true. This argument also shows that, in certain cases, there is no underdetermination. Thus Suppes' (1993) claim that there is underdetermination relative to the currently possible observations cannot be generally correct.

To conclude, for the choice relative to in principle possible observations where there are no limits, in principle, on observational accuracy, the nesting argument delivers the desired conclusion. However, the nesting argument fails for the choice relative to currently possible observations (and this is the kind of choice of concern in practice) or relative to actual observations or relative to in principle possible observations where there are limits, in principle, on observational accuracy.

## 6 The Analogy to Bernoulli Randomness

One of Winnie's (1998) main topics is randomness. He claims that his argument for the deterministic description is analogous to an argument he gives against the prevalence of Bernoulli randomness. This section will critically discuss this claim. This is important because, while there is indeed an analogy, it will be argued that the arguments are also disanalogous in another sense.

First of all, Winnie's argument against the prevalence of Bernoulli randomness will be outlined. Winnie calls descriptions *Bernoulli random* when they are probabilistic and when the next outcome of the description is independent of the previous outcomes; i.e., when the history is irrelevant. So

---

macrostates, the stochastic description can be preferable because of indirect evidence.

Bernoulli random descriptions are simply Bernoulli descriptions (Example 3). Winnie starts by asking the question of whether deterministic descriptions  $(X, f_t, P)$  can be Bernoulli random relative to some observation functions  $\Phi$ . The answer is affirmative: There are deterministic descriptions, including deterministic descriptions of the kinds used in science, which give rise to Bernoulli descriptions. For instance, consider the baker's transformation (Example 1) and the observation function  $\Phi_2((x, y)) = o_1$  if  $0 \leq y \leq 1/2$  and  $\Phi_2((x, y)) = o_2$  if  $1/2 < y \leq 1$ . Figure 5(a) shows this observation function  $\Phi_2$ . It is not hard to see that the stochastic descriptions  $\{Z_t\} = \{\Phi_2(f_t)\}$  is the Bernoulli descriptions of a fair coin (two outcomes 'o<sub>1</sub>' and 'o<sub>2</sub>' with probability 1/2); hence there is Bernoulli randomness. Similarly, also for the deterministic description of two hard spheres in a box (Example 2) there are observation functions  $\Phi$  such that the resulting stochastic description  $\{Z_t\} = \{\Phi(f_t)\}$  is a Bernoulli description (see Simányi 1999).<sup>12</sup>

Winnie then goes on to emphasise that even if some observation functions yield Bernoulli descriptions, there are always finer observation functions which do not yield Bernoulli descriptions. For instance, for the baker's transformation recall the observation function  $\Phi_{16}$  (see Example 1 in Section 2), which is shown in Figure 5(b).  $\Phi_{16}$  is finer than  $\Phi_2$ , and we have already seen that the stochastic description  $\{W_t\} = \{\Phi_{16}(f_t)\}$  is not a Bernoulli description but a Markov description (see the discussion at the beginning of Section 3). Each outcome of this Markov description has equal probabilities  $P(Z_t = e_i) = 1/16$ ,  $1 \leq i \leq 16$ . There is history dependence because  $e_1$  is followed by  $e_1$  or  $e_2$ ,  $e_2$  is followed by  $e_3$  or  $e_4$ , etc.; as a consequence, for instance,  $P(Z_t = e_1 \text{ given } Z_{t-1} = e_2) = 1/2 \neq 1/256 = P(Z_t = e_1)P(Z_t = e_2)$ .

On this basis, Winnie criticises claims in the literature that deterministic descriptions can be Bernoulli random relative to any finite-valued observation function, regardless how fine.<sup>13</sup> Winnie is correct to point out that *deterministic descriptions are always only Bernoulli random relative to some finite-valued observation functions*.<sup>14</sup> But another question is left open.

---

<sup>12</sup>There are observation functions which lead to Bernoulli descriptions with equal probabilities (fair dice), and also observation functions which lead to Bernoulli description with non-equal probabilities (biased dice).

<sup>13</sup>It is questionable that the literature which Winnie criticises really claims this. But we can set this issue aside because it will not be relevant in what follows.

<sup>14</sup>This point, Winnie argues, also underlies a wrong interpretation of Brudno's theorem. There is no need to go here into Brudno's theorem because the basic conceptual point is exactly the same: Brudno's theorem does not tell us that all coarse-grainings lead to

Namely, suppose a deterministic description is not Bernoulli random relative to an observation function  $\Phi$ . Can there still be observation functions  $\Psi$  finer than  $\Phi$  such that the deterministic description is Bernoulli random relative to  $\Psi$ ? This question is interesting because if the answer is negative, then, for fine-enough observations of deterministic descriptions, there will never be Bernoulli randomness. As has been proven by Werndl (2009c), the answer is indeed negative: Whenever  $\{\Phi(f_t)\}$  is not a Bernoulli description, there exists no finer observation function  $\Psi$  such that  $\{\Psi(f_t)\}$  is a Bernoulli description.<sup>15</sup>

Now Winnie claims that there is the following analogy: *Suppes' argument for underdetermination fails for the same reason as the claim that there is Bernoulli randomness for all finite-valued observation functions.* The following quotes illustrate this.

I shall begin by discussing a similar interpretation of [the argument about Bernoulli randomness]<sup>16</sup> and argue that it is misleading. Then [...] I shall argue that Suppes' interpretation is similarly misleading. (Winnie 1998, 310)

As the earlier discussion of [the argument about Bernoulli randomness] has shown, the fact that deterministic systems generate stochastic behaviour under a given partitioning in no way undermines the fundamental determinism of such systems. (Winnie 1998, 317)

Winnie does not state in which exact sense the argument against the prevalence of Bernoulli randomness and the argument against underdetermination are analogous. He might have had the following analogy in mind: The argument about Bernoulli randomness is that even if a deterministic

---

Bernoulli random sequences; all it tells us is that there are some coarse-grainings which lead to Bernoulli random sequences.

<sup>15</sup> The basic idea of Werndl's (2009c) proof is as follows. By assumption, there is an observation function  $\Phi : X \rightarrow Q$  of the deterministic description such that  $\{\Phi(f_t)\}$  is no Bernoulli description. Now suppose that there exists an observation function  $\Psi : X \rightarrow O$  finer than  $\Phi$  such that  $\{\Psi(f_t)\}$  is a Bernoulli description. Because  $\Psi$  is finer than  $\Phi$ , there is a function  $\Gamma : O \rightarrow Q$  such that  $\Phi = \Gamma(\Psi)$ . But it is not hard to see that if  $\{\Psi(f_t)\}$  is a Bernoulli description, then  $\{\Phi(f_t)\} = \{\Gamma(\Psi(f_t))\}$  is also a Bernoulli description. But this contradicts the assumption that  $\{\Phi(f_t)\}$  is no Bernoulli description.

<sup>16</sup>The text in square brackets has been replaced by the terminology used in this paper. I will use this convention throughout.

description is Bernoulli random relative to some observation function  $\Phi$ , the deterministic description will cease to be Bernoulli random relative to some observation functions which are finer than  $\Phi$ . Analogously, even if the stochastic description  $\{\Phi(f_t)\}$  is observationally equivalent to the deterministic description  $(X, f_t, P)$  relative to the observation function  $\Phi$ , the stochastic description  $\{\Phi(f_t)\}$  is not observationally equivalent to  $(X, f_t, P)$  relative to some observation functions finer than  $\Phi$  (indeed all observation functions finer than  $\Phi$ ). Hence if, given any arbitrary finite-valued observation, it is in principle always possible to make finer observations, then the deterministic description is preferable (relative to the in principle possible observations where there are no limits, in principle, on observational accuracy). This is so because only the deterministic description allows one to always make finer observations. *So there is indeed an analogy.*

However, it is also important to see that the cases are disanalogous in another sense. Bernoulli randomness is not prevalent because even if deterministic descriptions are Bernoulli random relative to some observation functions, for fine-enough observation functions they will cease to be Bernoulli random. The analogous argument for the choice between deterministic and stochastic descriptions is as follows: Even if deterministic descriptions yield nontrivial stochastic descriptions relative to some finite-valued observation functions, for fine-enough finite-valued observation functions one would not obtain nontrivial stochastic descriptions any longer. However, this is wrong. As we have seen in Section 3, for many deterministic descriptions, relative to *any arbitrary* finite-valued observation function, there is observational equivalence to a nontrivial stochastic description. Hence, regardless of how fine the observation functions are, one obtains nontrivial stochastic descriptions. *So there is no analogy in this sense.* Because Winnie claims that the two cases are analogous and this is an obvious potential analogy, it is important to point out that there is no analogy in this sense.

## 7 Conclusion and Outlook

At the beginning of the paper we looked at a time series of observations of the position of a particle moving on a square. Science aims to find a description which reproduces these observations and yields new predictions. One of the questions arising here is whether the phenomenon is better modeled by a deterministic or an indeterministic description. Intuitively, one might think

that the data only allow for a deterministic or an indeterministic description, but as we have seen in Section 3, this is not the case. In several cases, measure-theoretic deterministic descriptions are observationally equivalent to stochastic descriptions at any arbitrary observation level. Therefore, the question arises which description is preferable relative to evidence. If the evidence equally supports a deterministic and a stochastic description, there is underdetermination. Section 4 discussed the different kinds of choice these results pose, namely relative to in principle possible observations where there are no limits, in principle, on observational accuracy, relative to currently possible observations, relative to actual observations, or relative to currently possible observations where there are limits, in principle, on observational accuracy.

Section 5 discussed the following argument by Winnie (1998): Consider the possible observation functions which, according to the deterministic description, one should be able to apply. Suppose that it is actually possible to make these observations and that one obtains stochastic descriptions at a smaller scale when finer observations are made. Then, Winnie (1998) claims, the deterministic description is preferable. It was argued that Winnie's argument yields the desired conclusion relative to in principle possible observations where there are no limits, in principle, on observational accuracy. Yet relative to the currently possible observations (this is the kind of choice of relevance in practice), relative to the actual observations, and relative to in principle possible observations where there are limits, in principle, on observational accuracy the argument fails because it also applies to situations where the stochastic description is preferable. Section 6 was about Winnie's claim that his argument against underdetermination is analogous to the argument that even if a deterministic description is Bernoulli random at some observation level, it will fail to be Bernoulli random relative to finer observation levels. There is indeed an analogy – if a stochastic description is observationally equivalent to a deterministic description at some observation level, it will fail to be observationally equivalent for finer observation levels. However, it was argued that these cases are disanalogous in an important sense. While the deterministic description will cease to be Bernoulli random for fine-enough observation levels, the deterministic description will be observationally equivalent to stochastic descriptions at any arbitrary observation level.

Returning to the question of which description is preferable, we have seen that none of the arguments discussed above is tenable relative to the

currently possible observations (the kind of choice arising in practice), relative to the actual observations or relative to the in principle possible observations where there are limits, in principle, on observational accuracy. So which description is preferable relative to evidence in these cases? An answer will now be proposed when one has to choose between a deterministic description  $(X, f_t, P)$  derivable from Newtonian theory and a stochastic description  $\{Z_t\}$  which is not derivable from any theory (cf. Werndl 2009c and 2011). This includes the particle that bounces off several mirrors (Example 1), the two hard spheres moving in a box (Example 2) as well as the examples discussed by the extant literature (Suppes 1993, Suppes and de Barros 1996, Winnie 1998).

This answer involves the idea of indirect evidence as discussed in Section 5. The relevant predictions derived from the deterministic description  $(X, f_t, P)$  and the stochastic description  $\{Z_t\}$  are the same (the relevant predictions are those which can be tested by the currently possible observations, the actual observations or the in principle possible observations). Still, there are many deterministic descriptions from Newtonian theory which do not follow from  $(X, f_t, P)$  or  $\{Z_t\}$ , which are supported by evidence, and which provide indirect evidence for  $(X, f_t, P)$  but not for  $\{Z_t\}$  (because these deterministic descriptions and  $(X, f_t, P)$  are unified by Newtonian mechanics). Thus the deterministic description  $(X, f_t, P)$  is preferable because it is additionally supported by indirect evidence, and there is no underdetermination. This means that the deterministic description is preferable for the particle that bounces off several mirrors (Example 1) and the two hard spheres moving in a box (Example 2). Generally, because the relevant predictions derived from  $(X, f_t, P)$  and  $\{Z_t\}$  are the same, one can escape the conclusion of underdetermination only by appealing to other predictions; here these other predictions are provided by the other Newtonian descriptions.

This argument can be generalised: Suppose that one of the descriptions is additionally supported by indirect evidence because it is derivable from a well-confirmed theory and that the other description is not derivable from any well-confirmed theory. Then the description which is additionally supported by indirect evidence is preferable. It is important to note that the preferred description may be *deterministic or stochastic*. In the case of Newtonian mechanical descriptions, indirect evidence leads one to prefer the deterministic description. Yet in other cases the stochastic description might be preferable. To conclude, the main point is that even if the relevant predictions derived from the deterministic and stochastic description are the same,

indirect evidence may still make one of the descriptions preferable.

## References

- Benedicks, M. and Young, L.-S. (1993). ‘Sinai-Bowen-Ruelle-measures for certain Hénon maps’. *Inventiones Mathematicae* **112**, 541–567.
- Berkovitz, J., Frigg, R. and Kronz, F. (2006). ‘The ergodic hierarchy, randomness and Hamiltonian chaos’. *Studies in History and Philosophy of Modern Physics* **37**, 661–691.
- Bishop, R. (2008). Chaos. In E. Zalta, ed., ‘The Stanford Encyclopedia of Philosophy (Fall 2009 Edition)’. URL = <http://plato.stanford.edu/archives/fall2009/entries/chaos/>.
- Chernov, N. and Markarian, R. (2006). *Chaotic Billiards*. American Mathematical Society, Providence.
- Frigg, R. (2008). A field guide to recent work on the foundations of statistical mechanics. In D. Rickles, ed., ‘The Ashgate Companion to Contemporary Philosophy of Physics’. Ashgate, London, 99–196.
- Hénon, M. (1976). ‘A two dimensional mapping with a strange attractor’. *Communications in Mathematical Physics* **50**, 69–77.
- Judd, K. and Smith, L.A. (2001). ‘Indistinguishable States I: Perfect Model Scenario’. *Physica D* **151**, 125–141.
- Judd, K. and Smith, L.A. (2004). ‘Indistinguishable States II: Imperfect Model Scenario’. *Physica D* **196**, 224–242.
- Laudan, L. (1995). ‘Damn the Consequences!’ *Proceedings and Addresses of the American Philosophical Association* **69**, 27–34.
- Laudan, L. and Leplin, J. (1991). ‘Empirical equivalence and underdetermination’. *The Journal of Philosophy* **88**, 449–472.
- Okasha, S. (1998). ‘Laudan and Leplin on empirical equivalence’. *The British Journal for the Philosophy of Science* **48**, 251–256.

- Okasha, S. (2002). ‘Underdetermination, holism and the theory/data distinction’. *The Philosophical Quarterly* **208**, 303–319.
- Petersen, K. (1983). *Ergodic Theory*. Cambridge University Press, Cambridge.
- Pitowsky, I. (1996). ‘Laplace’s Demon Consults an Oracle: The Computational Complexity of Predictions’. *Studies in the History and Philosophy of Modern Physics* **27**, 161–180.
- Simányi, N. (1999). ‘Ergodicity of hard spheres in a box’. *Ergodic Theory and Dynamical Systems* **19**, 741–766.
- Simányi, N. (2003) ‘Proof of the Boltzmann-Sinai ergodic hypothesis for typical hard disk systems’. *Inventiones Mathematicae* **154**, 123–178.
- Suppes, P. (1993). ‘The transcendental character of determinism’. *Midwest Studies in Philosophy* **18**, 242–257.
- Suppes, P. and de Barros, A. (1996). Photons, billiards and chaos. In P. Weingartner and G. Schurz, eds, ‘Law and Prediction in the Light of Chaos Research’. Springer, Berlin, 190–201.
- 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. (2009c). *Philosophical Aspects of Chaos: Definitions in Mathematics, Unpredictability and the Observational Equivalence of Deterministic and Indeterministic Descriptions*. Ph.D. thesis. University of Cambridge.
- Werndl, C. (2011). ‘On the observational equivalence of continuous-time deterministic and indeterministic descriptions’. *European Journal for the Philosophy of Science* **1**, 193–225.
- Winnie, J. (1998). Deterministic chaos and the nature of chance. In J. Earman and J. Norton, eds, ‘The Cosmos of Science – Essays of Exploration’. Pittsburgh University Press, Pittsburgh, 299–324.