

# DETERMINISTIC VERSUS INDETERMINISTIC DESCRIPTIONS: NOT THAT DIFFERENT AFTER ALL?

CHARLOTTE WERNDL  
University of Cambridge

## 1 INTRODUCTION

The guiding question of this paper is: *can we simulate deterministic descriptions by indeterministic descriptions, and conversely?* By simulating a deterministic description by an indeterministic one, and conversely, we mean that the deterministic description, when observed, and the indeterministic description give the same predictions.

Answering this question is a way of finding out how different deterministic descriptions are from indeterministic ones. Of course, indeterministic descriptions and deterministic descriptions are different in the sense that for the former there is indeterminism in the future evolution and for the latter not. But another way of finding out how different they are is to answer the question whether they give the same predictions; and this question will concern us.

Since the language of deterministic and indeterministic descriptions is mathematics, we will rely on mathematics to answer our guiding question. In the first place, it is unclear how deterministic and indeterministic descriptions can be compared; this might be one reason for the often-held implicit belief that deterministic and indeterministic descriptions give very different predictions (cf. Weingartner and Schurz 1996, p.203). But we will see that they *can* often be compared. The deterministic and indeterministic descriptions which we will consider are measure-theoretic deterministic systems and stochastic processes, respectively; they are both ubiquitous in science. To the best of my knowledge, our guiding question has hardly been discussed in philosophy.

In this paper I will explain intuitively some mathematical results which show that, from a predictive viewpoint, measure-theoretic deterministic systems and stochastic processes are, perhaps surprisingly, very similar. I won't go into the technical details (ergodic theory and the mod-

ern theory of stochastic processes); they can be found in Werndl (2009c and 2009d).

This paper is organised as follows. In section 2 I will introduce measure-theoretic deterministic systems and stochastic processes. In section 3 I will explain how, at a basic level, deterministic systems and stochastic processes give the same predictions. Section 4 will be devoted to more advanced results on the predictive equivalence of deterministic and indeterministic descriptions. I will argue that the results show that from a predictive viewpoint, measure-theoretic deterministic systems and stochastic processes are very similar.

## 2 INDETERMINISTIC AND DETERMINISTIC DESCRIPTIONS

Let us introduce the indeterministic and deterministic descriptions which will concern us. There are two kinds of indeterministic and deterministic descriptions: for *discrete* descriptions the time increases in discrete steps, and *continuous* descriptions involve a continuous time-parameter. We confine ourselves to discrete descriptions. For continuous descriptions similar but also slightly different results to the ones discussed here hold (cf. Werndl 2009d).

### 2.1 Stochastic Processes

The indeterministic descriptions we are concerned with are *stochastic processes*. Basically all the indeterministic descriptions in science are stochastic processes, which are thus ubiquitous in science. A stochastic process is a process that evolves according to probabilistic laws. That is, for a stochastic process there usually is some indeterminism: even if the initial state of the process is known, there are many possibilities the process might go to, and these possibilities are measured by probabilities. A sequence describing a possible evolution of the stochastic process over time is called a realisation.

Let me introduce stochastic processes in more detail.<sup>1</sup> Probability theory is concerned with probabilistic experiments. The set of all possible outcomes of a probabilistic experiment is denoted by  $\Omega$  and called the *sample space*. It comes equipped with a *probability distribution*  $Q$  which

---

<sup>1</sup> All the seemingly mathematical definitions introduced in this paper are intuitive descriptions of measure-theoretic definitions which are given in Werndl (2009c, 2009d).

assigns probabilities to the subsets of  $\Omega$  you are interested in. In applied problems you are usually interested in functions of the outcomes of a probabilistic experiment. Hence the concept of a random variable is introduced. A *random variable* is a function  $Z$  from a sample space  $\Omega$  to a set  $E$ , called the *state space*, which assigns a value  $Z(\omega)$  to each  $\omega$  in  $\Omega$ . An *event*  $A$  is a subset of the set  $E$  which you are interested in, and the *probability distribution* of the random variable  $Z$  is given by looking back to  $Q$  on  $\Omega$ , i.e.  $P := Q(Z^{-1}(A))$  defined for all events  $A$  of  $E$ . Finally, a *stochastic process*  $(Z_t)$  with *state space*  $E$  is a collection of random variables  $Z_t$ , where  $t$  is in an integer, defined on the same sample space  $\Omega$  and taking values in  $E$ . The sequence  $r(\omega) := (\dots Z_{-1}(\omega), Z_0(\omega), Z_1(\omega) \dots)$  for any  $\omega$  in  $\Omega$  is called a *realisation* of the stochastic process. Intuitively,  $t$  represents time; so that  $\omega$  represents a possible history in all details, and  $r(\omega)$  represents the description of that history by giving the ‘score’ at each  $t$ .

For instance, a *Bernoulli process* is a stochastic process where at each time point an  $N$ -sided die is tossed and the tosses are independent. That is, it is a stochastic process  $(Z_t)$  where (i) the state space is  $\{e_1, \dots, e_N\}$ , and for any  $t$  the probability for outcome  $e_i$  is  $p_i$  where  $p_1 + \dots + p_N = 1$ ; and (ii)  $P(Z_{i_1} = e_{j_1}, \dots, Z_{i_n} = e_{j_n}) = P(Z_{i_1} = e_{j_1}) \dots P(Z_{i_n} = e_{j_n})$  for any  $e_{j_1}, \dots, e_{j_n}$  in  $E$  and any integer-valued  $i_1, \dots, i_n$ .

## 2.2 Deterministic Systems

A description is *deterministic* exactly if any two solutions which agree at one time agree at all times (Butterfield 2005). Hence for deterministic descriptions there is no indeterminism about the future evolution. A sequence describing the evolution of a deterministic description over time is called a *solution*.

We are concerned with measure-theoretic deterministic descriptions, in short *deterministic systems*. They are denoted by  $(X, T, \mu)$  and consist of three components: a set  $X$ , called the *phase space*, which represents all possible states of the system; a bijective map  $T: X \rightarrow X$  which describes how states evolve, namely every  $x$  in  $X$  evolves to  $T(x)$  in one time unit and so to  $T^t(x)$  in  $t$  time units. And  $\mu$  is a *normalised measure*; intuitively, the subsets of  $X$  you are interested in are called *events*, and the measure assigns to all events a nonnegative size with total size one, i.e.  $\mu(X) = 1$ . As we will see, for connecting deterministic with indeterministic descriptions it is crucial that a measure is given on the phase space.

The *solutions* of a deterministic system are  $(\dots T^{-1}(x), x, T(x) \dots)$  for any  $x$  in  $X$ . We will often be concerned with measure-preserving deterministic systems, which are deterministic systems where the measure is *invariant* under time-evolution, i.e. where  $\mu(T(A)) = \mu(A)$  for all events  $A$ . There are various interpretations of invariant measures. For instance, according to the *time-average interpretation* the measure of an event  $A$  is the long-run proportion of time a solution spends in  $A$  and thus can be interpreted as probability. In what follows we make the common assumption that invariant measures can be interpreted as probability (cf. Werndl 2009a).

Deterministic systems, and measure-preserving deterministic systems in particular, are among the most important deterministic descriptions in science. To require that the measure is invariant is not very restrictive because for wide classes of deterministic systems invariant measures exist (cf. Werndl 2009c). For instance, all deterministic descriptions in Newtonian and statistical mechanics are measure-preserving deterministic systems.

When you observe a deterministic system a value is observed which is dependent on, but maybe does not uniquely encode the actual state. Hence observing the system can often be modelled by an *observation function*, i.e. a function  $\Phi: X \rightarrow M$  where  $M$  contains all possible observed values.

The deterministic system of the *baker's system* will accompany us throughout this paper. States  $(x, y)$  in the unit square  $X := [0, 1] \times [0, 1]$  are mapped to  $T(x, y) := (2x, y/2)$  if  $0 \leq x < 1/2$  and  $(2x-1, (y+1)/2)$  otherwise. *Figure 1* shows the dynamics of the baker's system. For the standard volume (the Lebesgue measure) we obtain a measure-preserving deterministic system. It describes a particle bouncing on mirrors (Pitowsky 1995).



Figure 1: Dynamics of the baker's system

### 3. BASIC INDISTINGUISHABILITY OF DETERMINISTIC SYSTEMS AND STOCHASTIC PROCESSES

This section is about indistinguishability results which are *basic* in the sense that they answer the question whether, given a deterministic system, we can find *any* stochastic process which simulates the system, and conversely. In section 3.1 I will explain how stochastic processes can be simulated by deterministic systems; these results seem to be unknown to philosophers. In section 3.2 I will show how deterministic systems can be simulated by stochastic processes; here the basic idea is known to philosophers of physics (Butterfield 2005).

#### *3.1 Deterministic Systems Simulated by Stochastic Processes*

How can you simulate deterministic systems by stochastic processes? Contrary to deterministic systems, for stochastic processes there is usually indeterminism in the future evolution. But when you *observe* a deterministic system, you see only how the system moves from one observed value to another; and because the observation function can map two states to the same value, the same current observed value can lead to different future observed values. Hence only if you make the assumption that a deterministic system is observed with an observation function can you hope to simulate it by a stochastic process. But this assumption is unproblematic: deterministic systems in science typically have an uncountably large phase space, and scientists can only observe finitely many different values.

A deterministic system comes equipped with a probability measure. Hence when observing the system, you see the probability distributions of sequences of possible observations of the deterministic system; and the predictions you obtain are these probability distributions. A stochastic process is defined by the probability distributions over its realisations, and these probability distributions are the predictions you obtain. Consequently, when saying *that a stochastic process  $(Z_t)$  and a deterministic system, as observed with  $\Phi$ , give the same predictions* we mean the following: (i) *the possible states  $E$  of the stochastic process are the possible observed values of the deterministic system, and (ii) the realisations of the stochastic process have the same probability distribution as the solutions of the deterministic system observed with the observation function.* Recall that we say that a deterministic system is simulated by a stochastic process, and con-

versely, exactly if the deterministic system, when observed, and the stochastic process give the same predictions.

Formally, for any  $(X, T, \mu)$  and any observation function  $\Phi$ ,  $(Z_t) := \Phi(T^t)$  is a stochastic process and hence simulates the deterministic system observed with  $\Phi$ . But the question is, of course, whether this stochastic process is *nontrivial*. To stress the point, if  $\Phi$  is the identity function, then although  $\Phi(T^t)$  is formally a stochastic process, it is evidently equivalent to the original deterministic system. But several results show that we often obtain nontrivial stochastic processes. Let me mention one result which is shown in Werndl (2009c).

**Proposition 1.** *Assume that a measure-preserving deterministic system is weakly mixing<sup>2</sup>. Then for any observation function with finitely many different values<sup>3</sup> the stochastic process  $(Z_t) := \Phi(T^t)$  is nontrivial: there are observed values, that is, elements  $e_i, e_j$  of the state space  $E$  of the stochastic process such that  $0 < P(Z_{t+1} = e_i \text{ given that } Z_t = e_j) < 1$ .*

Let us assume that only finitely many different values can be observed. Most measure-preserving deterministic systems, including many systems in science, are weakly mixing (Halmos 1944). Hence *Proposition 1* says that for most measure-preserving deterministic systems, regardless how you observe them, you always obtain nontrivial stochastic processes. For instance, the baker's system is weakly mixing. Therefore, regardless of how you observe it, you obtain a nontrivial stochastic process.

### 3.2 Stochastic Processes Simulated by Deterministic Systems

How can a given stochastic process be simulated by a deterministic system? Again, simulating the behaviour of a stochastic process by a deterministic system will only work if you allow that the deterministic system is observed. As argued, this assumption is unproblematic.

Assume that a stochastic process  $(Z_t)$  is given. We can construct a deterministic system as follows:  $X$  is the set of all imaginable realisations of

---

<sup>2</sup> A measure-preserving deterministic system  $(X, T, \mu)$  is weakly mixing iff for all events  $A, B$

$$\lim_{n \rightarrow \infty} 1/n (|\mu(A \cap B) - \mu(A)\mu(B)| + |\mu(T(A) \cap B) - \mu(A)\mu(B)| + \dots + |\mu(T^{n-1}(A) \cap B) - \mu(A)\mu(B)|) = 0.$$

<sup>3</sup> We assume that the observation function does not only take one value.

the stochastic process, i.e. the set of sequences  $e = (\dots e_{-1}, e_0, e_1 \dots)$  where  $e_i$  are elements of the state space  $E$  of  $(Z_t)$ . Let  $T$  be the map on  $X$  which shifts each sequence to the left, i.e.  $T((\dots e_{-1}, e_0, e_1 \dots)) = (\dots e_0, e_1, e_2 \dots)$ . Finally,  $(Z_t)$  determines a probability distribution on the realisations and hence a measure  $\mu$  on  $X$ . Then  $(X, T, \mu)$  is a deterministic system. Now assume you observe the  $0^{\text{th}}$  coordinate, i.e. you employ the observation function  $\Phi(e) = e_0$ . Clearly, the deterministic system  $(X, T, \mu)$  observed with  $\Phi$  simulates the process  $(Z_t)$ .  $(X, T, \mu, \Phi)$  is called the *deterministic representation* of  $(Z_t)$ . Thus we can conclude that *every stochastic process can be simulated by at least one deterministic system which, when observed, gives the same predictions, namely its deterministic representation.*

The deterministic representation is, from a philosophical perspective, a cheat because its states are constructed to encode the future and past evolution of the indeterministic process; also, the observation function is unnatural because scientists usually do not observe the  $0^{\text{th}}$  coordinate of a bi-infinite sequence. Still, it is important to see how, mathematically, stochastic processes can be simulated by the deterministic representation. There remains the question whether there are also other deterministic systems that simulate a stochastic process; we will return to it later (in section 4.1).

For instance, for a Bernoulli process with  $N$  different outcomes the deterministic representation is the following: the phase space is the set of all possible bi-infinite sequences of the  $N$  outcomes, the evolution equation shifts each sequence to the left, and the observed value is the  $0^{\text{th}}$  coordinate of the sequence. Generally, – and this definition will be important later – a *deterministic Bernoulli system* is a deterministic system which is probabilistically equivalent to the deterministic representation of a Bernoulli process, and hence (observed with specific observation functions) simulates a Bernoulli process.

Let me briefly comment on the philosophical significance of the above results. The guiding question of this article is: can we simulate deterministic systems by indeterministic processes which give the same predictions, and conversely? Our results show that stochastic processes can always be simulated by deterministic systems, and that many deterministic systems, when observed, can be simulated by nontrivial stochastic processes. *Therefore, from the perspective of predictive power, stochastic processes and deterministic systems are similar.*

#### 4. ADVANCED INDISTINGUISHABILITY OF STOCHASTIC PROCESSES AND DETERMINISTIC SYSTEMS

This section is about indistinguishability results which are *advanced* in the sense that they concern the question whether it is possible to separate deterministic systems *in science* from stochastic processes *in science*.

##### *4.1 Deterministic Systems Simulating Stochastic Processes in Science*

We have seen that any stochastic process can be simulated by at least one deterministic system, namely its deterministic representation. However, the deterministic representation is artificial in the sense that scientists do not encounter it when modelling phenomena. So you might still guess that *the deterministic systems which simulate the stochastic processes in science, such as Bernoulli processes, are very different from the deterministic systems in science, such as Newtonian systems.*

Until the late 1950s it was widely believed that stochastic processes and deterministic systems are different in this way. More specifically, Kolmogorov conjectured that while the deterministic systems which simulate stochastic processes in science produce positive information, the deterministic systems in science produce zero information. Kolmogorov and Sinai introduced the *Kolmogorov-Sinai entropy* to capture the property of producing positive information; and this property was expected to be able to separate stochastic processes from deterministic systems. It was a big surprise when it was found that many deterministic systems in science, among them Newtonian systems, have positive Kolmogorov-Sinai entropy and thus produce positive information (Frigg and Werndl 2009; Sinai 1989, Werndl 2009b).

What is more, Bernoulli processes are often regarded as the most random stochastic processes because past outcomes are independent of future outcomes. And there are even deterministic systems in science which are deterministic Bernoulli systems; hence *there are deterministic systems in science which, when observed with specific observation functions, simulate a Bernoulli process.* For instance, to mention a few, dispersive billiard systems, the logistic map (for some parameter values) which models population dynamics and the climate, and the Hènon map which models weather phenomena (Lorenz 1964; Lyubich 2002; May 1976; Ornstein and Weiss 1991; Young 1997; see also Winnie 1998). Also the baker's system produces a Bernoulli process. Assume that you only see whether the state



of the baker's system is to the left or to the right of the unit square, i.e.  $\Phi((x, y)) := 0$  if  $x < 1/2$  and 1 otherwise. This makes the system equivalent to a fair coin toss, viz. a Bernoulli process with outcomes 0 and 1 which have both probability  $1/2$ .

Hence even Bernoulli processes, often regarded as the most random stochastic processes, are simulated by deterministic systems in science when observed with specific observation functions. But then any attempt to separate such deterministic systems from the deterministic systems needed to simulate the stochastic processes in science must fail. The philosophical implication of this result is that the above guess is wrong, viz. *the deterministic systems which simulate the stochastic processes in science include deterministic systems in science. This indicates that from a predictive viewpoint, stochastic processes and deterministic systems are very similar.*

#### *4.2 Stochastic Processes Simulating Deterministic Systems in Science at Any Observation Level*

The deterministic systems which simulate the stochastic processes in science include deterministic systems in science. Still, you might conjecture that *only if you employ specific coarse observation functions* for such deterministic systems, do you obtain stochastic processes in science; and if you make observations of such deterministic systems which are precise enough, you will not any longer obtain a stochastic process in science or an observation of a stochastic process in science. Hence you might guess that stochastic and deterministic descriptions are different in the following way: *the stochastic processes which simulate deterministic systems in science at any observation level are very different from the stochastic processes in science.*

Let us introduce two natural ways of understanding the notion that stochastic processes of a certain type simulate deterministic systems at any observation level. In practice, for sufficiently small  $\alpha_1$ , you cannot distinguish states of a deterministic system which are less than  $\alpha_1$  apart; and, for sufficiently small  $\alpha_2$ , you won't be able to observe differences in probabilistic predictions of less than  $\alpha_2$ , and you can thus neglect events of probability less than  $\alpha_2$ . Assume that  $\alpha$  is smaller than  $\alpha_1$  and  $\alpha_2$ . Then a deterministic system and a stochastic process in practice give the same predictions if the solutions of the deterministic system can be put into one-to-one correspondence with the realisations of the stochastic process such that the state of the deterministic system and the corresponding outcome of the

stochastic process are less than  $\alpha$  apart except for states of probability less than  $\alpha$ . Formally, this idea is captured by the notion of  $\alpha$ -congruence (Ornstein and Weiss 1991). The deterministic system  $(X, T, \mu)$  and the stochastic process  $(Z_t)$  are  $\alpha$ -congruent exactly if there is a function  $\Psi: X \rightarrow X$  such that (i)  $(Z_t)$  is the stochastic process  $\Psi(T^t)$ , and the deterministic representation of  $(Z_t)$  is probabilistically equivalent to  $(X, T, \mu)$ , and (ii)  $x$  and  $\Psi(x)$  differ by less than  $\alpha$  for all  $x$  in  $X$  except for a set of measure smaller than  $\alpha$ . Note that the notion of  $\alpha$ -congruence *does not assume that the deterministic system is observed with an observation function*. But observation functions can, of course, be brought in: assume you make a coarse-grained observation of a deterministic system. Then the probabilistic predictions you obtain are essentially the same (differ at most by  $\alpha$ ) as the probabilistic predictions of the corresponding observed stochastic process, i.e. the process obtained by applying the observation function to the stochastic process which is  $\alpha$ -congruent to the deterministic system. So we arrive at that a plausible meaning of the notion that stochastic processes of a certain type simulate a deterministic system at any observation level is that for any  $\alpha > 0$  there is a stochastic process of this type which is  $\alpha$ -congruent to the system.

Another approach is to start with the idea that for sufficiently small  $\alpha > 0$ , you won't be able to distinguish an observed deterministic system from a stochastic process if the stochastic process nearly simulates the deterministic system; i.e. if the deterministic system, when observed, and the stochastic process have the same outcomes, and their probabilistic predictions are either the same or differ by less than  $\alpha$ . The following notion captures this idea: a stochastic process  $(Z_t)$   $(\alpha, \Phi)$ -simulates a deterministic system  $(X, T, \mu)$  observed with  $\Phi: X \rightarrow M$  exactly if there is a function  $\Psi: X \rightarrow M$  such that (i)  $(Z_t)$  is the stochastic process  $\Psi(T^t)$ , and (ii)  $\Psi$  has the same ranges as  $\Phi$  and differs from  $\Phi$  only on a set of measure  $< \alpha$  (cf. Ornstein and Weiss 1991, p.95). Hence the notion of  $(\alpha, \Phi)$ -simulation captures the idea that in practice the observed deterministic system and the stochastic process give the same predictions. Unlike  $\alpha$ -congruence,  $(\alpha, \Phi)$ -simulation *assumes that the deterministic system is observed with an observation function*. By generalising over  $\alpha$  and  $\Phi$  we obtain a plausible meaning of the phrase that stochastic processes of a certain type simulate a deterministic system at any observation level, namely: for every  $\Phi$  and every  $\alpha$  there is a stochastic process of this type which  $(\alpha, \Phi)$ -simulates the deterministic system.

For Bernoulli processes future outcomes are independent of past ones. In contrast, for deterministic systems in science future states are strongly constrained by its previous state. Thus, intuitively, *such deterministic systems cannot be simulated at every observation level by Bernoulli processes*. This intuition is correct. First, as shown in Werndl (2009c), for any deterministic system in science there is a  $\alpha > 0$  such that no Bernoulli process is  $\alpha$ -congruent to the deterministic system. Second, for any deterministic system there is an observation function and a  $\alpha > 0$  such that no Bernoulli process  $(\alpha, \Phi)$ -simulates the deterministic system (Werndl 2009d). Despite these results, there remains the question whether deterministic systems in science can be simulated by other paradigm stochastic processes at every observation level.

To answer it, we have to introduce two paradigm stochastic processes, namely Markov processes and multi-step Markov processes. For a *Markov process* the probability distribution of the next state only depends on the previous state. That is,  $(Z_t)$  is a Markov processes exactly if (i) the state space consists of  $N$  states  $e_1, \dots, e_N$ , and (ii)  $P(Z_{t+1} = e_i \text{ given } Z_t, Z_{t-1}, \dots, Z_k) = P(Z_{t+1} = e_i \text{ given } Z_t)$  for any  $t$  and any  $k, k \leq t$ , in the integers and any  $i, 1 \leq i \leq N$ .<sup>4</sup> We will be concerned with *irreducible* and *aperiodic* Markov processes. These are Markov processes which cannot be split into two separate processes and which have no periodicities, and they are generally regarded as the most random Markov processes.

*Multi-step Markov processes* generalise Markov processes. They consist of all Markov processes of order  $n$ , where  $n$  is a natural number. For a Markov process of order  $n$  the probability distribution of the next state only depends on the previous  $n$  states. That is,  $(Z_t)$  is a Markov process of order  $n$  exactly if (i) the state space consists of  $N$  states  $e_1, \dots, e_N$ , and (ii)  $P(Z_{t+1} = e_i \text{ given } Z_t, Z_{t-1}, \dots, Z_k) = P(Z_{t+1} = e_i \text{ given } Z_t, \dots, Z_{t+n-1})$  for any  $t$  and  $k, k \leq t$ , in the integers and any  $i, 1 \leq i \leq N$ . Again, we will be concerned with the most random multi-step Markov processes: namely those which cannot be split into two separate processes and which have no periodicities, called *irreducible* and *aperiodic* multi-step Markov processes.

Recall that deterministic Bernoulli systems are deterministic systems which are probabilistically equivalent to the deterministic representation of a Bernoulli process (section 3.2), and that many deterministic systems in science are deterministic Bernoulli systems (section 3). *It holds that several deterministic systems in science, namely all deterministic Bernoulli*

---

<sup>4</sup> We assume that the probability distribution of Markov processes and of Markov process of order  $n$  is independent of time.

systems, can be simulated at any observation level by paradigm stochastic processes, namely by Markov processes or by multi-step Markov processes. More precisely, for the first meaning of the phrase that deterministic systems are simulated at every observation level it holds that every deterministic Bernoulli system can be simulated at any observation level by irreducible and aperiodic Markov processes. For the second meaning of the phrase it holds that every deterministic Bernoulli system can be simulated at any observation level by irreducible and aperiodic multi-step Markov processes. It is not hard to see that any Markov processes of order  $n$  is an observed Markov process, i.e. a process which is obtained by assigning to some of the states of the Markov process the same value (Werndl 2009d).<sup>5</sup> Hence we rediscover a result which we already arrived at for the first meaning of simulation at every observation level, viz. that regardless how fine we observe a Bernoulli system, the outcomes could have resulted from an observed Markov process:

**Theorem 1.** *Let  $(X, T, \mu)$  be a deterministic Bernoulli system. Then for every  $\alpha > 0$  there is an irreducible and aperiodic Markov process which is  $\alpha$ -congruent to the deterministic system (Werndl 2009c).*

**Theorem 2.** *Let  $(X, T, \mu)$  be a deterministic Bernoulli system. Then for every observation function  $\Phi$  and every  $\alpha > 0$ , there is an irreducible and aperiodic multi-step Markov process which  $(\alpha, \Phi)$ -simulates the deterministic system (cf. Werndl 2009d).*

Also, deterministic Bernoulli systems are the only systems in science which can be simulated at every observation level by multi-step Markov processes or by Markov processes. That is, a deterministic system where for every observation function and every  $\alpha$  there is an irreducible and aperiodic Markov process which is  $\alpha$ -congruent to the system is a deterministic Bernoulli system (Ornstein 1974, p.45). And a deterministic system in science where for every  $\Phi$  and  $\alpha > 0$  there is an irreducible and aperiodic multi-step Markov processes which  $(\alpha, \Phi)$ -simulates the system is a deterministic Bernoulli system (Werndl 2009d). For instance, the baker's system is a deterministic Bernoulli system. Thus it can be simulated by

---

<sup>5</sup> This follows because a multi-step Markov process can be reduced to a Markov process by considering the process whose state space consists of all possible  $n$ -tuples of the states of the Markov process of order  $n$ .

Markov processes or by multi-step Markov processes at any observation level.

Let me comment on the philosophical significance of these results. We guessed that the stochastic processes which simulate deterministic systems in science at any observation level are very different from the stochastic processes in science. *Clearly, the above results show that this guess is wrong because several deterministic systems in science can be simulated at any observation level by Markov processes or by multi-step Markov processes. This further underlines that from a predictive viewpoint deterministic systems and stochastic processes are very similar.*

Suppes (1993), Suppes and de Barros (1996) and Winnie (1998) are the only philosophical discussions about the results of this section I could find. They all discuss some results involving the notion of  $\alpha$ -congruence. Suppes and de Barros (1996) and Winnie (1998) claim that the philosophical significance of some  $\alpha$ -congruence results is that, at every observation level, you have a choice between a deterministic description in science and a stochastic description. However, this seems weak and hence misguided: that, at every observation level, you have a choice between a deterministic description in science and a stochastic description is already shown by the results in section 3.1, and these results were known much earlier. As argued, the real philosophical significance of the  $\alpha$ -congruence results is that the stochastic processes needed to simulate deterministic systems in science at any observation level include paradigm stochastic processes (cf. Werndl 2009c).<sup>6</sup>

There remains, of course, the very important question: if we have a choice between a deterministic and a stochastic description, which one is better? Winnie (1998) criticises Suppes (1993) for arguing that in the case of the  $\alpha$ -congruence results both descriptions are equally good. Winnie thinks that the deterministic description is preferable, and his argument goes as follows. Assume we have a Markov process which simulates the deterministic system for the current observation level. The observation level in the future may become so fine that we need to introduce another Markov process to simulate the deterministic system at the finer observation level. Hence the deterministic system is better. However, I think the situation is often not as clear as Winnie claims. For instance, suppose that in principle the states of the modelled phenomenon can never be distinguished beyond a certain level, e.g. because these states do not take infi-

---

<sup>6</sup> Furthermore, Suppes and de Barros (1996) and Winnie (1998) both do not really understand the meaning of  $\alpha$ -congruence (see Werndl 2009d).

nately precise values. Then it seems less clear what is preferable: the deterministic system or the stochastic process resulting from an observation function which assigns to any two states that can never be distinguished the same value. I think the question of whether the deterministic or stochastic description is better needs more careful consideration. For some thoughts see Werndl (2009d), and it is my hope that further work will be dedicated to this question.

## 5. CONCLUSION

Our guiding question has been: can we simulate deterministic systems by stochastic processes which give the same predictions, and conversely? I have first explained that many deterministic systems, when observed, can be simulated by nontrivial stochastic processes, and that every stochastic process can be simulated by at least one deterministic system. Yet you might still guess that the deterministic systems which simulate the stochastic processes in science are very different from the deterministic systems in science. I have shown this to be false since there is a class of deterministic systems, including several paradigm deterministic systems, which even produce the most random stochastic process of tossing a die. Given this, you might still guess that the stochastic processes which simulate the deterministic systems in science at every observation level are very different from the stochastic processes in science. I have explained that also this guess is wrong because there are many deterministic systems in science which can be simulated at every observation level by paradigm stochastic processes, namely by Markov processes or by multi-step Markov processes. I have argued that all these results show that from a predictive viewpoint, deterministic systems and stochastic processes are very similar.

## ACKNOWLEDGMENTS

I am indebted to Jeremy Butterfield. Many thanks also to Cymra Haskell, Franz Huber, Thomas Müller, Amy Radunskaya, Peter Smith and the audience at the Wittgenstein Symposium 08 for valuable discussions. I am grateful to St John's College, Cambridge University, for financial support.

## REFERENCES

- Butterfield, J.N. (2005). 'Determinism and Indeterminism'. *Routledge Encyclopaedia of Philosophy*.
- Frigg, R. and Werndl, C. (2009). 'Entropy and Probabilities'. Forthcoming in: C. Beisbart and S. Hartmann (eds). *Probabilities in Physics*. Oxford: Oxford University Press.
- Halmos, P.R. (1944). 'In General a Measure-preserving Transformation is Mixing'. *The Annals of Mathematics* 45, pp.786–792.
- Lorenz, E.N. (1964). 'The Problem of Deducing the Climate from the Governing Equations'. *Tellus* XVI, pp.1–11.
- Lyubich, M. (2002). 'Almost Every Regular Quadratic Map is Either Regular or Stochastic'. *Annals of Mathematics* 156, pp.1–78.
- May, R.M. (1976). 'Simple Mathematical Models with Very Complicated Dynamics'. *Nature* 261, pp.459–467.
- Ornstein, D. (1974). *Ergodic Theory, Randomness, and Dynamical Systems*. New Haven: Yale University Press.
- Ornstein, D. and Weiss, B. (1991). 'Statistical Properties of Chaotic Systems'. *Bulletin of the American Mathematical Society* 24, pp.11–116.
- Pitowsky, I. (1995). 'Laplace's Demon Consults an Oracle: The Computational Complexity of Prediction'. *Studies in History and Philosophy of Modern Physics* 27, pp.161–180.
- Sinai, Y.G. (1989). 'Kolmogorov's Work on Ergodic Theory'. *The Annals of Probability* 17, pp.833–839.
- Suppes, P. (1993). 'The Transcendental Character of Determinism', *Midwest Studies in Philosophy* 18, pp.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*. Berlin: Springer, pp.189–201.
- Weingartner, P. and Schurz, G., eds. (1996). *Law and Prediction in the Light of Chaos Research*. Berlin: Springer.
- Werndl, C. (2009a). 'What Are the New Implications of Chaos for Unpredictability?'. *The British Journal for the Philosophy of Science* 60, pp.195–220.
- Werndl, C. (2009b). 'Justifying Definitions in Mathematics – Going Beyond Lakatos'. *Philosophia Mathematica*, doi: 10.1093/phimat/nkp006.
- Werndl, C. (2009c). 'Are Deterministic and Indeterministic Descriptions Observationally Equivalent?'. Submitted to: *Studies in History and Philosophy of Modern Physics*.
- Werndl, C. (2009d). *Philosophical Aspects of Chaos: Definitions in Mathematics, Unpredictability and the Observational Equivalence of Deterministic and Indeterministic Descriptions*. Ph.D. thesis. University of Cambridge.
- Winnie, J.A. (1998). 'Deterministic Chaos and the Nature of Chance'. In: Earman, J. and Norton, J., eds., *The Cosmos of Science – Essays of Exploration*. Pittsburgh: University of Pittsburgh Press, pp.299–324.

Young, L.-S. (1997). 'Ergodic Theory and Chaotic Dynamical Systems'. *XII-th International Congress of Mathematical Physics, Brisbane*. Cambridge, MA: International Press, pp.311–319.