Hidden Determinism, Probability, and Time’s Arrow

Hans Primas
ETH-Zentrum, CH–8092 Zürich (Switzerland)

Abstract

In present-day physics the fundamental dynamical laws are taken as a time-translation-invariant and time-reversal-invariant one-parameter groups of automorphisms of the underlying mathematical structure. In this context-independent and empirically inaccessible description there is no past, present or future, hence no distinction between cause and effect.

To get the familiar description in terms of causes and effects, the time-reversal symmetry of the fundamental dynamics has to be broken. Thereby one gets two representations, one satisfying the generally accepted rules of retarded causality (“no effect can precede its cause”). The other one describes the strange rules of advanced causality. For entangled (but not necessarily interacting) quantum systems the arrow of time must have the same direction for all subsystems. But for classical systems, or for classical subsystems of quantum systems, this argument does not hold. As a consequence, classical systems allow the conceptual possibility of advanced causality in addition to retarded causality.

Every mathematically formulated dynamics of statistically reproducible events can be extended to a description in terms of a one-parameter group of automorphisms of an enlarged mathematical structure which describes a fictitious hidden determinism. Consequently, randomness in the sense of mathematical probability theory is only a weak generalization of determinism. The popular ideas that in quantum theory there are gaps in the causal chain which allow the accommodation of the freedom of human action are fantasies which have no basis in present-day quantum mechanics. Quantum events are governed by strict statistical laws.

Freedom of action is a constitutive necessity of all experimental science which requires a violation of the statistical predictions of physics. We conclude that the presently adopted first principles of theoretical physics can neither explain the autonomy of the psyche nor account for the freedom of action necessary for experimental science.
1 Determinism Does not Deal with Predictions

A large part of the difficulty with the “determinism versus indeterminism” debate lies in the failure to define the terms of discussion clearly. Many physicists and philosophers do not make the important distinction between determinism and the concept of predictability, and thereby commit a category mistake by claiming that determinism implies the possibility of prediction the future course of the universe.

For example, MAX BORN considers the distinction between determinism and predictability as hairsplitting and completely superfluous.¹ He maintains that classical point mechanics is not deterministic since there are unstable mechanical systems which are epistemically not predictable.

I think the following remarks by JOHN EARMAN are appropriate:²

“The history of philosophy is littered with examples where ontology and epistemology have been stirred together into a confused and confusing brew. . . . Producing an ‘epistemological sense’ of determinism is an abuse of language since we already have a perfectly adequate and more accurate term – prediction – and it also invites potentially misleading argumentation – e.g., in such-an-such a case prediction is not possible and, therefore, determinism fails.”

That is, determinism does not deal with predictions. Nevertheless, ontic descriptions are often confused with epistemic ones.³ In the philosophical literature there are many examples for such category mistakes. For example RUDOLF CARNAP says: “Causal relation means predictability.”⁴ Likewise, KARL POPPER maintains: “Scientific determinism is the doctrine that the state of any closed physical system at any future instant can be predicted.” ⁵ Physicists drop similar careless assertions. For example LEON BRILLOUIN: “The Poissonian discontinuities correspond to conditions where prediction is actually impossible and determinism cannot exist.”⁶

2 Terminology and Basic Concepts

In order to set the stage for my discussion, I introduce some definitions and notions I will use in the following.

**Determinism:** First of all, determinism will be taken to refer exclusively to ontic descriptions, and should not be confused with statements concerning our knowledge or beliefs. In particular, according to the definition adopted here,

¹Compare for example MAX BORN (BORN (1955a); BORN (1955b), BORN (1958)). For a critique of Born’s view compare VON LAUE (1955).
³For more details on the ontic/epistemic distinction compare the contribution by HARALD ATMANSPACHER in this volume.
⁴CARNAP (1966), p.192
⁵POPPER (1962), p.36.
determinism does not to imply predictability. Both, predictability and retrodictability have their proper place only in the framework of epistemic descriptions.

**Causal relations:** A process within which one event is a necessary condition for another event is described by a *causal relation*. The producing event is known as the *cause* and the event produced as its *effect*. A causal relationship is an irreflexive, antisymmetric and transitive binary relation between two events. That is:
- no event can be the cause of itself;
- if $a$ is the cause of $b$, then $b$ cannot be the cause of $a$;
- if $a$ is the cause of $b$, and $b$ the cause of $c$, then $a$ is the cause of $c$.

**Causal ordering:** A causal nexus requires some universal order. A fundamental issue is the relation of causality to time. According to David Hume causal relations have three components: contiguity of time and place, temporal priority of the cause, and constant conjunction.\(^7\) For Hume “all inferences from experience ... are effects of custom, not of reasoning”,\(^8\) so that according to Hume’s view the idea of cause and effect is not a matter of fact but a mental habit of association, that is, essentially subjectively fabricated. Hume’s characterization implies that causal and temporal arrows are related *by definition*. Yet, this merging of the two very different ideas of causal order and temporal order is conceptually not sound. Moreover, it precludes many logical possibilities, like a backward causation, or a time-independent ordering of the causal nexus.

**Arrowless time in physics:** If one wants to characterize causal ordering by temporal ordering, then one has first to introduce *temporal direction*. Yet, the generally adopted first principles of physics do not distinguish the future from the past. First principles are characterized by high symmetries. A corresponding physical law is said to be fundamental if it is as independent as possible of any particular context. For example, we assume that the laws of nature are the same all the time and everywhere.

The assumption that there is neither a favored point of the origin nor a preferred direction in time and space is a basic symmetry postulate required in all fundamental physical theories. Since a fundamental theorem by Emmy Noether theorem implies a deep connection between symmetries and conservation laws, the idea that fundamental laws should be characterized by high symmetries is not just an aesthetic concept.\(^9\) For example, Noether’s theorem requires that the time-translation symmetry implies and is implied by the

\(^{7}\text{Hume (1793), book 1, part 3, p.466.}\)
\(^{8}\text{Hume (1748), section V, part 1.}\)
\(^{9}\text{Noether (1918). Noether’s theorem says that if the action integral of a dynamical system is invariant with respect to a $n$-parameter continuous group of symmetry transformation, then the equations of motion have $n$ linearly independent conservation laws.}\)
conservation of energy. In fundamental physical theories the basic dynamical laws are not only taken as a time-translation-invariant but also as time-reversal-invariant. In the mathematical jargon we say that a fundamental dynamics is given by a *time-translation-invariant* and *time-reversal-invariant* one-parameter group of automorphisms of the underlying mathematical structure.

If we consider the time-reversal symmetry as primary, then there is no ordering so that we cannot use the concepts of cause and effect. In such a formulation of physics all reality is already pre-existent, and nothing new can come into existence. In order for time and causality to be genuinely active, some degree of freedom is necessary to provide a mechanism by which the events “come into being”. Without breaking the time-reversal symmetry nothing new can ever arise. Within special contexts a spontaneous breaking of this symmetry is possible, so that the direction of time has to be considered as *contextual*.

### 3 Breaking the Time-Reversal Symmetry

Every experiment requires nonanticipative measuring instruments, hence a distinction between past and future. In engineering physics the direction of causation is always assumed to go from past to future. That is, to derive experimental physics from first principles, the time-reversal symmetry of the fundamental laws has to be broken. The anisotropy of time is a *precondition* for any theory of irreversible processes.

The phenomenon of symmetry breaking is well-understood in modern physical theories. It is not an ad-hoc postulate, but follows from the first principles of theoretical physics. Nevertheless, there remains an important problem. The time-reversal symmetry is represented by a group of order two. If the time-reversal symmetry is broken one gets two representations, one satisfying the generally accepted rules of retarded causality (forward causation) and the other rule the strange rules of advanced causality (backward causation).

That is, if it is possible at all to derive the principle of retarded causality (“no effect can precede its cause”), then the very same methods allow the derivation of processes governed by advanced causality. The decision which of the two possibilities is appropriate can therefore not be derived from the first principles of physics. *So the conceptual problem is not the breaking of the time-reversal symmetry, but the proper selection of one or the other one-sided realization.*

### 4 Arrow of Time

Mechanical and electrical input–output systems are non-anticipative. In other words, they are characterized by the fact that any present output values do not depend on future input values. The success of such traditional phenomenological physical descriptions suggests that in epistemic descriptions the time-reversal symmetry is always broken. Also, conscious perception and cognition seems to presuppose the usual forward direction of time, implying a memory of the
past, but no anticipation of the future. The one-way property of time which we experience in our everyday life, and which we find in phenomenological physical laws has been called “time’s arrow” by Arthur Stanley Eddington.\textsuperscript{10}

In spite of these empirical facts, advanced causality is a conceptual possibility which is not ruled out by any fundamental physical law. So even on an ontic level we have to distinguish between the following two possibilities:

- **Forward determinism**, that is the thesis that, given the past of a physical system, there is a unique future (retarded realization).

- **Backward determinism**, that is the thesis that, given the future of a physical system, there is a unique past (advanced realization).

The nature and origin of a temporal asymmetry in the physical world is a perplexing problem for a theoretician. What is the origin of the arrow of time? Why do most (perhaps all) observable processes show the same arrow of time? What is puzzling about this temporal asymmetry is the universality of the direction of the arrow of time. The usual choice of retarded causality cannot be explained by a statistical mechanical formulation of the “second law” without an a priori postulate imposing an asymmetric evolution toward increasing time.

Backward causation is usually disregarded in science, but – as indicated above – should not be excluded rashly. From the viewpoint of physical first principles there are no reasons to select one of the two temporal directions. The question of whether epistemic descriptions in terms of backward causality are rewarding or not, should not be decided by some a-priori argument. It seems to be more reasonable not to hold any prejudices, but first to work out the theory in full mathematical detail and then discuss the consequences.\textsuperscript{11}

## 5 Indeterminism

The phrase “indeterminism” hides a number of different concepts and implications. It can refer to limits of knowledge, inherent or practical unpredictability, unpredictable mechanistic or non-mechanistic causation, uncaused action, or even lawlessness.

First we have to emphasize that noncausation has to be distinguished from indeterminism. On the other hand, there are ontically deterministic systems which produce epistemically irreducible random outcomes. In addition, the existence of strict statistical laws for certain random events suggests that there must be some causation. If a system produces epistemically irreducible random outcomes which fulfill the statistical laws of Kolmogorov’s probability theory,\textsuperscript{12} then we speak of statistical causation. It may be tempting to exclude indeterministic events which are not ruled by the laws of mathematical probability or any kind of statistical laws. But this would be premature without further arguments. The concept of noncausation will be used to refer to such events.

\textsuperscript{10}Eddington (1928), p.68.

\textsuperscript{11}Compare also the contribution by Phil Dowe in this volume.

\textsuperscript{12}Kolmogoroff (1933).
6 Statistical Causality

A broken time-reversal symmetry gives rise to two epistemic causal relations: prediction and retrodiction. Forward causality is logically independent from backward causality. Ontic determinism implies neither epistemic predictability nor epistemic retrodictability.

The usual causal explanations have predictive power and refer to a world to be acted on. Predictions require a memory of the past, they refer to probabilistic inferences of the future behavior of a system from empirically estimated past states. In a statistical description, predictive processes are entropy-increasing. Retrodictive explanations require a Carrollian “memory of the future”, they refer to probabilistic inferences of the past behavior of a system. In a statistical description retrodictive processes are entropy-decreasing.\(^{13}\)

Forward causal and backward causal descriptions have the same logical status a priori. Progress in science depends on the discovery of causal connections between events, regardless of whether this nexus is backward or forward. In spite of the fact that memories usually store past events only, we have to acknowledge that advanced causations are not forbidden by the first principles of physics.

A theme recurring again and again is the obsession that the second law of thermodynamics is related with the arrow of time, and that it precludes the existence of entropy-decreasing systems. However, a system showing a globally entropy-increasing behavior can nevertheless possess open subsystems with locally retrodictive and entropy-decreasing behavior. There are many everyday physical examples for entropy-decreasing systems. A successful retrodictive teleological description of entropy-decreasing living systems is just as legitimate as the predictive description of an entropy-increasing system.

7 Why Can There Be Laws of Chance?

Mathematical probability theory refers to events or processes whose outcomes are individually random, but are governed by strict statistical laws. This theory has a rich mathematical structure so we have to ask under which conditions the usual “laws of chance” are valid. Long ago, MARIAN VON SMOLUCHOWSKY pointed out that the concept of probability can be defined, and the laws of probability can be derived from the theory of strictly deterministic but non-robust classical dynamical systems.\(^{14}\) So it may be tempting to presume that chance events which satisfy the axioms of classical mathematical probability theory always result from the deterministic behavior of an underlying physical system. Such a claim cannot be demonstrated though. What can be proven is the weaker statement that every probabilistic system which fulfills the axioms


\(^{14}\)VON SMOLUCHOWSKY (1918).
of classical mathematical probability theory can be embedded into a fictitious larger deterministic description.

To understand this claim we recall that the logic of classical probability theory is given by a (usually nonatomic) Boolean $\sigma$-algebra. Mathematical probability is defined as a $\sigma$-additive norm on a Boolean $\sigma$-algebra. A theorem by Lynn H. Loomis and Roman Sikorski implies that every probabilistic system has a (nonunique) extension to an atomic Boolean algebra, such that every probability can be represented as a mean over two-valued states. It follows that random variables fulfill the laws of chance if and only if it is possible to find a larger Boolean system such that every expectation value is the mean of a (possibly hidden) variable with respect to a two-valued state of the enlarged system. Such deterministic embeddings are usually not constructive and nothing substantial can be said about a possible ontic interpretation of hidden variables of the enlarged deterministic system. The important point is that the strict statistical regularities are due to the mere possibility of a deterministic mathematical description.

Mathematical details on probability and hidden determinism

Mathematical probabilistic theory\textsuperscript{15} can be described in terms of a probability algebra $(\mathcal{B}, p)$ where the probability $p$ is a strictly positive $\sigma$-additive normed measure on the elements of a Boolean $\sigma$-algebra $\mathcal{B}$ of events. According to the fundamental Loomis–Sikorski representation theorem every Boolean $\sigma$-algebra $\mathcal{B}$ is isomorphic to a $\sigma$-algebra $\Omega/\Delta$, where $\Omega$ is some point set and $\Delta$ is a $\sigma$-ideal $\Sigma$, $\mathcal{B} \sim \Omega/\Delta$.\textsuperscript{16} Every statistical state can be represented by a probability $p$ on $\mathcal{B}$, or equivalently, by the restriction of a probability measure $\mu$ on the measurable space $(\Omega, \Sigma)$ to the Boolean algebra $\Omega/\Delta$. Equivalently, a statistical state can be represented by the restriction of a probability measure $\mu$ on the measurable space $(\Omega, \Sigma)$ to the Boolean algebra $\Omega/\Delta$. In this representation the nonnegative number $\mu(\mathcal{B})$ is the probability of the event $\mathcal{B} \in \Sigma$ in this statistical state.

The power set $\Pi(\Omega)$ of all subsets of the set $\Omega$ is a complete and atomic Boolean algebra which can be considered as an extension of the Boolean $\sigma$-algebra $\mathcal{B}$ of events. Every element $\omega \in \Omega$ defines a dispersion-free individual state $\chi_\omega$ by

$$\chi_\omega(\mathcal{B}) = \begin{cases} 1 & \text{if } \omega \in \mathcal{B} \\ 0 & \text{if } \omega \notin \mathcal{B} \end{cases}, \quad \mathcal{B} \in \Pi(\Omega).$$

The classical system characterized by the Boolean algebra $\Pi(\Omega)$ is one of the many possible hidden-variable extension of the probability algebra $(\mathcal{B}, p)$, or equivalently, of the Kolmogorov probability space $(\Omega, \Sigma, \mu)$. Every probability measure $\nu$ on $(\Omega, \Sigma)$ is a mean over a family $\{\chi_\omega | \omega \in \Omega\}$ of two-valued states of the extended system with the Boolean algebra $\Pi(\Omega)$:\textsuperscript{17}

$$\nu(\mathcal{B}) = \int_\Sigma \chi_\omega \nu(\omega) d\omega, \quad \mathcal{B} \in \Sigma.$$

Usually the sample space $\Omega$ is uncountable so that an atom $\omega \in \Omega$ cannot represent an experimental proposition. In spite of the fact that the points of an uncountable phase space have no epistemic relevance whatsoever, the introduction of such “hidden variables” is a convenient mathematical tool. This result

\textsuperscript{15}For an outline of the “point-free” approach to probability, compare Halmos (1944); Kolmogoroff (1948); Los (1955), Kappos (1969).


\textsuperscript{17}Compare Kamber (1964), §7; Kamber (1965), §14.
implies that random variables fulfill the laws of chance if and only if they can formally be reduced to hidden two-valued variables.

Since the requirement $\Omega/\Delta \sim B$ does not determine the set $\Omega$ uniquely, any interpretation of individual points of the Kolmogorov probability space $(\Omega, \Sigma, \mu)$ can be misleading and should be avoided.

The possibility of a fictitious deterministic embedding of the statistical laws of mathematical probability theory explains the possibility of a frequency interpretation of probabilistic physics. Similarly, the modern consistent formulations of subjective probabilities postulate that a rational man acts as if he had a deterministic model compatible with his preknowledge.$^{18}$ Accordingly, these theories also presuppose a hidden Boolean determinism as an underpinning for a coherent rational behavior.

**Conclusion:** A probabilistic system fulfills the axioms of classical mathematical probability theory if and only if it can be embedded into a larger deterministic system. The question whether physical theories are deterministic or merely probabilistic in the sense of mathematical probability theory is empirically undecidable.

8  Are There Statistically Irreproducible Events?

From a logical point of view one can neither exclude the occurrence of unique irreproducible events nor the existence of chance events for which the traditional “laws of chance” do not apply. Empirically this question is not decidable. Within the framework of the usual frequency interpretation of probability one encounters the well-known difficulty of “small probabilities”. Already in 1866 John Venn tried to define a probability explicitly in terms of relative frequencies of occurrence of events “in the long run”. He added that “the run must be supposed to be very long, in fact never to stop.” Yet, without additional assumptions nothing can be inferred about the value of the limiting frequency of a finite segment, no matter how long it may be. Supplementary decision rules that allow to decide which probability statements we should accept are notoriously difficult to formulate. Even Kolmogorov$^{19}$ had to adopt the working rule by Antoine Augustine Cournot:

> If the probability of an event is sufficiently small, one should act in a way as if this event will not occur at a solitary realization.

Yet, the theory gives no criterion for deciding what is “sufficiently small”. As emphasized by Wolfgang Pauli, no frequency interpretation can avoid a subjective factor.$^{21}$

---

18 Compare for example Savage (1954), Savage (1962), Good (1965), Jeffrey (1965), De Finetti (1972), De Finetti (1974), De Finetti (1975). For a convenient collection of the most important papers on the modern subjective interpretation, compare Kyburg & Smokler (1964).

19 Kolmogoroff (1933), p.4, postulate B.

20 Cournot (1843).

In this purely mathematical form, Bernoulli’s theorem is thus not as yet susceptible to empirical test. For this purpose it is necessary somewhere or other to include a rule for the attitude in practice of the human observer, or in particular the scientist, which takes account of the subjective factor as well, namely that the realisation, even on a single occasion, of a very unlikely event is regarded from a certain point on as impossible in practice. Theoretically it must be conceded that there is still a chance, different from zero, of error; but in practice actual decisions are arrived at in this way, in particular also decisions about the empirical correctness of the statistical assertions of the theories of physics or natural science. At this point one finally reaches the limits which are set in principle to the possibility of carrying out the original programme of the rational objectivation of the unique subjective expectation.

Wolfgang Pauli made the inspiring proposal to characterize unique events by the absence of any type of statistical regularity: 22

“The synchronicity phenomena considered by [Jung] elude being captured by “laws” of nature, since they are not reproducible, that is to say; they are unique and blurred by the statistics of large numbers. In physics, on the other hand, ‘acausalities’ are captured just by statistical laws of large numbers.”

9 Hadamard’s Principle of Scientific Determinism

For a mathematical formulation the traditional philosophical doctrine that each event is the necessary and unique consequence of prior events has to be sharpened by introducing the concept of the state of a physical system. What changes in a dynamical system is called the state of the system.

In modern system theory a state of a nonanticipative input-output system is represented by a kind of memory. It represents “the minimal amount of information about the past history of the system which suffices to predict the effect of the past upon the future”. 23 Intuitively a system-theoretical state can be considered as a kind of memory, representing the relevant history of the system. Under very general conditions there exists a state-space realization for every nonanticipative input-output system. The most compact description is given by the so-called Nerode state at time \( t \), defined as the equivalence class of all histories for \( t<0 \) of the system which give rise to the same output for all conceivable future. 24 No particular ontic or epistemic interpretation is implied by this definition.

A well-posed nonanticipative dynamical system requires a rule for determining the Nerode state at a given future time from a given present state.


24For the first time this idea was clearly expressed by Nerode (1958) in the context of automata theory. Compare also Kalman, Falb & Arbib (1969), chapters 1.3, 6.3, and 7.2.
Hadamard’s principle of scientific determinism requests that in a well-posed forward-deterministic dynamical system every initial state determines all future states uniquely. Hadamard’s principle of scientific determinism is not a natural law but a regulative principle which leads to an appropriate choice of the state space. If Hadamard’s principle is not fulfilled then it can often be enforced by choosing a larger state space.

Example: Hadamard’s principle in classical point mechanics

Newton’s second law \( m\ddot{\mathbf{q}}_t = \mathbf{F}(\mathbf{q}_t, t) \), \( \mathbf{q}_t \in \mathbb{R}^3 \), for a point particle of mass \( m \) under the influence of a force \( \mathbf{F} \) does not fulfill Hadamard’s principle for the configuration space \( \mathbb{R}^3 \). All the information about the past is given by the specification of the position and the velocity of every point particle at a particular time. The Lagrangian formulation considers the positions \( \mathbf{q} \) and the velocity \( \mathbf{v} := \dot{\mathbf{q}} \) as independent variables. This move allows us to rewrite Newton’s second law in a system-theoretic form as a system of two first order differential equations, \( \ddot{\mathbf{q}}_t = \mathbf{v}_t, m\ddot{\mathbf{v}}_t = \mathbf{F}(\mathbf{q}_t, t) \), \( \mathbf{q}_t, \mathbf{v}_t \in \mathbb{R}^3 \). This Lagrangian formulation fulfills Hadamard’s principle in the state space \( \mathbb{R}^6 \). The pair \( \{\mathbf{q}_0, \mathbf{v}_0\} \in \mathbb{R}^6 \) represents the Nerode state specifying the initial conditions necessary for the unique determination of the solution of the equations of motion for \( t > 0 \).

Since a stochastic process is nothing else but a family of random variables, every stochastic process can be dilated to a family of two-valued variables. Moreover, under appropriate continuity conditions, every stochastic process can be dilated to a Hadamard-deterministic process. Nevertheless, without further assumptions no ontological implication can be attributed to such a dilation.

Mathematical models for deterministic motions

The mathematical description of any kind of motion uses an uninterpreted concept of time. In fundamental descriptions time is represented by the additive group \( \{t | t \in \mathbb{R}\} \) of real numbers. Our ability to distinguish between before and after requires that time intervals are oriented. If the time interval \( t_2 - t_1 \) between two instance is positive, the time \( t_2 \) is said to be later than \( t_1 \). The physically fundamental equations of motion are not only invariant under translations \( t \to t' = t + \tau, \tau \in \mathbb{R} \), but also under the time-reversal transformation, an involution which exchanges the time parameter \( t \) by \( -t \). If the time-reversal symmetry is broken, time is represented either by the additive semigroup \( \{t | t \geq 0\} \), or by the additive semigroup \( \{t | t \leq 0\} \).

In a mathematical description of a dynamical system a Nerode state at time \( t \) is represented by an element \( \mathbf{x}_t \) of some topological space, called the state space \( \mathcal{X} \). A motion of the system is represented by a function \( t \mapsto \mathbf{x}_t \in \mathcal{X} \) which maps each instant \( t \) on exactly one element \( \mathbf{x}_t \) of the state space.

For a Hadamard-deterministic system the motion \( t \mapsto \mathbf{x}_t \) of a system-theoretical state element \( \mathbf{x}_t \) is ruled by a family \( \{\varphi_{r,s} | 0 \leq s \leq r\} \) of continuous state transition maps \( \varphi_{t,s,t} \), fulfilling

\[
\varphi_{r,s} \circ \varphi_{s,t} = \varphi_{r,t}, \quad \varphi_{t,t} = \text{id}, \quad 0 \leq t \leq s \leq r.
\]

---


26Doob (1953), p.46.

27An involution is an operation whose square is the identity. The involution associated with time-reversal \( T (T^2 = 1) \) also involves the space reflection \( P (P^2 = 1) \) and the charge conjugation \( C (C^2 = 1) \) (PCT-theorem).
The dynamics of a Hadamard-deterministic system is then given by the functional relation \( x_s = \varphi_{s,t}(x_t) \) with \( s \geq t \).

A dynamical system is called autonomous if it is not subject to time-varying external influences. In this case the state transition map \( \varphi_{s,t} \) depends only on the difference \( s - t \) so that the time evolution is time-translation-invariant and given by the one-parameter semigroup \( \{ \varphi_t | t \geq 0 \} \) with \( \varphi_t = \varphi_{0,t} \) and \( \varphi_s \circ \varphi_t = \varphi_{s+t} \).

The semigroup of time-translation invariant state transition maps \( \varphi_t \) is called a semiflow. Under very general conditions, a non-autonomous dynamical system can be embedded into a larger autonomous dynamical system. \(^{28}\)

An autonomous dynamical system is said to be reversible if for each \( t \geq 0 \) the inverse state transition map \( \varphi_t^{-1} \) exists. If the family of state transition maps \( \varphi_t \) form a one-parameter group, \( \varphi_s \circ \varphi_t = \varphi_{s+t} \), \( s, t \in \mathbb{R} \), then \( \{ \varphi_t | t \in \mathbb{R} \} \) is called a flow. For every initial state element \( x_0 \) the state element \( x_t \) at a later instant \( t \) is given by \( x_t = \varphi_t(x_0) \). For a reversible time-translation-invariant dynamical system the inverse \( \varphi_t^{-1} \) is given by \( \varphi_{-t} \), so that the trajectory \( s \mapsto \varphi_s^{-1}(x_s) \), \( 0 \leq s \leq t \), arrives back at \( x_0 \) after the time \( t \). \( x_0 = \varphi_{-t}^{-1}(x_t) \). Since the dynamics can be uniquely reversed, we speak of a time-reversal-invariant dynamics. Note that a time-reversal-invariant dynamics entails the possibility of a process reversal, not the reversal of the direction of time.

### 10 Experimental Science Requires Freedom of Action

The premise that we often have a choice about what we are going to do is called the free-will thesis.\(^ {29}\) Some philosophers claim the laws of physics and the freedom of will are of a different type and operate at a different level so that one can without contradiction subscribe to both determinism and free will. An example for such a “compatibilist viewpoint” is the suggestion that we should abandon the view that the laws of nature act like inviolable prescriptions, and that we should adopt a descriptive view of natural laws so that the problem of free will does not even arise. This descriptive view claims that whatever happens in the world, there are true descriptions of those events, and that whatever you do, there is a true description of what you have done. Other philosophers and scientists opt for another variant, claiming that all behavior is determined, and that “free will” is a “meaningless concept”, or an “illusion”.

At present the problem of how free will relates to physics seems to be intractable since no known physical theory deals with consciousness or free will. Fortunately, the topic at issue here is a much simpler one. It is neither our experience of personal freedom, nor the question whether the idea of freedom could be an illusion, nor whether we are responsible for our actions. The topic here is that the framework of experimental science requires a freedom of action in the material world as a constitutive presupposition. In this way “freedom” refers to actions in a material domain which are not governed by deterministic first principles of physics.

\(^{28}\) Compare Howland (1974); Reed & Simon (1975), section X.12; Nickel (1996); Engel & Nagel (2000); section VI.9. Compare also the contribution by Gregor Nickel in this volume.

To get a clearer idea of what is essential in this argument we recall that the most consequential accomplishment by Isaac Newton was his insight that the laws of nature have to be separated from initial conditions. The initial conditions are not accounted for by first principles of physics, they are assumed to be “given”. In experimental physics it is always taken for granted that the experimenter has the freedom to choose these initial condition, and to repeat his experiment at any particular instant. To deny this freedom of action is to deny the possibility of experimental science.

In other words, we assume that the physical system under investigation is governed by strictly deterministic or probabilistic laws. On the other hand, we also have to assume that the experimentalist stands out of these natural laws. The traditional assumption of theoretical physics that the basic deterministic laws are universally and globally valid for all matter thus entails a pragmatic contradiction between theory and practice. A globally deterministic physics is impossible.

11 Quantum Randomness

Often it is claimed that according to quantum mechanics “the basic constituents of matter behave in a fundamentally random way,” or that quantum mechanics allows “uncaused” events. Such statements are misleading since all quantum events are governed by strict statistical laws. Probability is an essential element in every complete epistemic description of quantum events. From an epistemic viewpoint, individual quantum events are in general irreducibly random. But this epistemic quantum randomness does neither imply ontic randomness nor that determinism has been refuted by quantum theory. As the following example illustrates, a link between randomness and determinism, analogous to that discussed in Section 7, holds also for quantum physics.

Example of a deterministic embedding for quantum measurements

In spite of the fact that no experiment can ever realize a measurement of the first kind, it is worthwhile to discuss this customary idealization in the framework of quantum theory. We consider the simple example of a first-kind measurement of the observable \( \sigma_3 \),

\[
\sigma_3 = |\alpha\rangle \langle \alpha| - |\beta\rangle \langle \beta|, \quad \alpha, \beta \in \mathbb{C}^2, \quad \langle \alpha | \alpha \rangle = \langle \beta | \beta \rangle = 1, \quad \langle \alpha | \beta \rangle = 0,
\]

of a two-level quantum object system. We assume that the initial state is pure. It can either be described by a normalized state vector \( \Psi_{\text{initial}} = c_1 |\alpha\rangle + c_2 |\beta\rangle \in \mathbb{C}^2 \), or by the density operator \( D_{\text{initial}} = |\Psi_{\text{initial}}\rangle \langle \Psi_{\text{initial}}| \).

The result of any measurement is irreversibly recorded by a Boolean device. Therefore a full statistical description of a first-kind measurement process has to include the object system, the measuring apparatus, and the classical recording device. In the framework of algebraic quantum mechanics such a full description is possible. Here we discuss only the resulting statistical state transition map \( D_{\text{initial}} \rightarrow D_{\text{final}} \) connecting the initial with the final state of the object system.

\[\text{Compare for example Barrett (1999), p.1}\]
\[\text{Margenau (1957), p.724.}\]
For a first-kind measurement of $\sigma_3$ the final density operator of the object system is given by

$$D_{\text{final}} = |c_1|^2 |\alpha\rangle\langle\alpha| + |c_2|^2 |\beta\rangle\langle\beta| = \frac{1}{2} \sigma_3 + \frac{i}{2} \text{tr}\{D_{\text{initial}} \sigma_3\} \sigma_3 .$$

Since in quantum mechanics the set of statistical state functionals is not a simplex, the final density operator $D_{\text{final}}$ allows infinitely many distinct convex decompositions into pure states. Therefore the linear map $D_{\text{initial}} \rightarrow D_{\text{final}}$ does not imply that the final state is a classical mixture of pure states described by the density operators $|\alpha\rangle\langle\alpha|$ and $|\beta\rangle\langle\beta|$. Nevertheless, the probabilities $|c_1|^2$ and $|c_2|^2$ are the correct conditional probabilities for a measurement of the first kind. The condition is that the classical measuring instrument has factually and irreversibly registered one of the eigenvalues of the observable $\sigma_3$. In other words, the density operator $D_{\text{final}}$ is the conditional expectation of $D_{\text{initial}}$, conditioned by the commutative $W^*$-algebra generated by the observable $\sigma_3$.

The simplest dynamical description of the state transition map $D_{\text{initial}} \rightarrow D_{\text{final}}$ is given by the linear dynamical semigroup

$$D(t) = \frac{\omega}{2i} [\sigma_3, D(t)] + \frac{\kappa}{4} [\sigma_3, [\sigma_3, D(t)]] ,$$

with $D_{\text{initial}} = D(0)$, $D_{\text{final}} = D(\infty)$. The constant $\omega \in \mathbb{R}$ is an angular frequency describing the Hamiltonian dynamics, and $\kappa > 0$ is a relaxation frequency describing the dissipative interaction with the environment. For an arbitrary initial density matrix $D(0)$ we get the asymptotic density matrix $D(\infty)$,

$$D(\infty) := \lim_{t \to \infty} D(t) = \frac{1}{2} \sigma_3 + \frac{i}{2} \text{tr}\{D(0) \sigma_3\} \sigma_3 .$$

The statistical map $D_{\text{initial}} \rightarrow D_{\text{final}}$ can be generated in many different ways by an individual Hadamard deterministic motion. Consider for example a solution $t \mapsto \Phi(t)$ of the following nonlinear stochastic Schrödinger equation in the sense of Stratonovich

$$i d\Phi(t) = \frac{1}{2} \omega \sigma_3 \Phi(t) dt + i \kappa \{\Phi(t) | \sigma_3 \Phi(t)\} \{\sigma_3 - \Phi(t) | \sigma_3 \Phi(t)\} \Phi(t) + i \sqrt{\kappa/2} \{\sigma_3 - \Phi(t) | \sigma_3 \Phi(t)\} \Phi(t) \circ dw(t) , \quad \Phi(0) = \Psi_{\text{initial}}.$$

Then the mean density operator $D_\Phi(t) = \mathbb{E}\{|\Phi(t)\rangle\langle\Phi(t)|\}$ satisfies the same dynamical semigroup as the density operator $D(t)$. Moreover, an individual trajectory $t \mapsto \Phi(t)$ describes the individual behavior of the object system within the hypothetical extended description. For an arbitrary initial state vector $\Psi_{\text{initial}}$, one gets asymptotically\(^{32}\)

$$\Psi_{\text{initial}} \quad \lim_{t \to \infty} \quad \begin{cases} \alpha & \text{with probability } |\langle\alpha | \Psi_{\text{initial}}\rangle|^2 \\ \beta & \text{with probability } |\langle\beta | \Psi_{\text{initial}}\rangle|^2 \end{cases} ,$$

where $\alpha$ and $\beta$ are the normalized eigenvectors of $\sigma_3$.

The white noise process $t \mapsto dw(t)/dt$ is a generalized stationary Gaussian process. It is well known that the trajectories of every continuous Gaussian processes can be generated by a linear Hamiltonian system.\(^{33}\) A physically realistic stochastic differential equation is never driven by white noise processes but by some variant of band-limited smooth noise processes. A theorem by Wong and Zakai implies that the white-noise limit of a stochastic differential equation with an essentially band-limited noise process is the solution of the corresponding stochastic differential equations in the sense of Stratonovich.\(^{34}\) Therefore the stochastic Stratonovich equation mentioned above has a Hadamard deterministic

\(^{32}\)Gisin (1984).

\(^{33}\)Compare for example Picci (1988).

\(^{34}\)Wong & Zakai (1965a), Wong & Zakai (1965b).
dilation to a Hamiltonian equation of motion for the deterministic trajectories of the stochastic process \( t \mapsto \Phi(t) \).

Of course, the sketched construction of a deterministic representation of a quantum measurement process is quite ad hoc and implies nothing about an ontic interpretation. The salient point is to show that the strict statistical regularities of a quantum measuring process can be understood in terms of a hidden determinism. Note that the various theorems which show that in quantum theory it is impossible to introduce hidden variables only say that it is impossible to embed quantum theory into a deterministic Boolean theory.

The context-independent laws of an ontic description of quantum mechanics are strictly deterministic but refer to a non-Boolean logical structure. The epistemically irreducible probabilistic structure of quantum theory is due to the fact that every communication in terms of an unequivocal language requires a Boolean domain of discourse. The nonpredictable outcome of a quantum experiment is related to the projection of the non-Boolean lattice of the deterministic ontic description to the Boolean algebra of the epistemic description of a particular experiment. The associated epistemic quantum-theoretical probabilities cannot be attributed to the object system; they are conditional probabilities for state transitions induced by the interaction of the object system with a classical measuring apparatus. The epistemic probabilities depend on the experimental arrangement, but for a fixed context they are objective since the underlying ontic structure is deterministic.

12 Quantum Mechanics Cannot Explain Free Will

Since the Hadamard determinism of classical mechanics and the freedom of actions collide, some physicists and many philosophers find delight in the belief that the so-called “uncertainty principle” of quantum mechanics can solve the problem of freedom.\(^{35}\) For instance, the neurophysiologist JOHN ECCLES argued that in addition to the material world there is a nonmaterial and nonsubstantial mental world. He speculated that our mind acts on the brain at the quantum level by momentarily increasing the probability of exocytosis in selected cortical areas and thereby controlling “quantum jumps”, turning them into voluntary excitations of the neurons that account for body motion.\(^{36}\)

The main idea of all such fantasies seems to be that in quantum theory there are gaps in the causal chain which allow the accommodation of free will and corresponding action. Yet, quantum events are governed by the strict mathematical laws of Kolmogorov’s probability theory which lead to the empirically reproducible statistical rules of statistical physics. Consequently, quantum randomness is just a weak generalization of determinism.\(^{37}\) Human actions are, however, in general not reproducible, hence not subject to the mathematical

\(^{35}\)Compare for example JORDAN (1932); JORDAN (1934); JORDAN (1956), pp.114ff.; MARGENAUX (1957); ROHS (1996), pp.232–239.

\(^{36}\)BECK & ECCLES (1992); ECCLES (1994).

\(^{37}\)Similarly, WOLFGANG PAULI has stressed that “quantum mechanics is a very weak general-
laws of probability theory.\textsuperscript{38} Freedom of action requires a violation of the statistical predictions of quantum theory. Hence present-day quantum theory is neither in a position to explain the autonomy of the psyche nor the existence of free will.

13 Why Does Time’s Arrow Always Point in the Same Direction?

The presupposed possibility of breaking the time-reversal symmetry and defining the arrow of time locally is not yet sufficient to understand the seemingly global character of the direction of time. Why should the arrow of time in two noninteracting physical systems point to the same direction? Classical causality is local and cannot explain the \textit{universality} of the direction of the arrow of time. In quantum physics the omnipresence of Einstein–Podolsky–Rosen correlations even between noninteracting subsystems implies a global character of the time-reversal operation in quantum systems. In an entangled system it is impossible to define a \textit{local} time-reversal operation for one subsystem only. Consequently the arrow of time of \textit{entangled} (but not necessarily interacting) quantum systems must have the same direction. \textit{Quantum causality is holistic and requires the same direction of the arrow of time even for non-interacting entangled subsystems.}

\textbf{Representing time-reversal symmetry in quantum mechanics}

In the C*-algebraic description of arbitrary physical systems a symmetry is represented by a \textit{Jordan automorphism} of the underlying C*-algebra \( \mathfrak{A} \). A Jordan-automorphism is a linear \(*\)-preserving bijection which respects the symmetrized product,\textsuperscript{39}

\[ \alpha(A^*) = \alpha(A)^* , \quad \alpha(AB + BA) = \alpha(A)\alpha(B) + \alpha(B)\alpha(A) , \quad A, B \in \mathfrak{A} . \]

A C*-automorphism \( \alpha \) is a Jordan-automorphism which preserves the order of the ordinary product, \( \alpha(AB) = \alpha(A)\alpha(B) \). An anti-automorphism \( \alpha \), which reverses the order of the terms in the ordinary product, \( \alpha(AB) = \alpha(B)\alpha(A) \), is also a Jordan-\(*\)-isomorphism. Every Jordan-automorphism of a C*-algebra is the sum of a C*-automorphism and a C*-anti-automorphism.\textsuperscript{40}

\textsuperscript{38}Curiously, Warren Weaver (1948) (p.33) claims “that individual human decisions, like the individual events of physics, are not ruled by causality; while the statistical behavior of a man, like the statistics of a physical ensemble, is ruled by causality.” It may be that some human actions are ruled by statistical laws, but certainly there are also singular non-recurring human actions.

\textsuperscript{39}Compare for example Emch (1972), p.152.

\textsuperscript{40}Kadison (1951), theorem 10.
In the irreducible representation of quantum mechanics on a Hilbert space $\mathcal{H}$, the time reversal is represented by an anti-unitary operator on the algebra $\mathfrak{B}(\mathcal{H})$, so that it implements a $C^*$-anti-automorphism on the algebra $\mathfrak{B}(\mathcal{H})$. More generally, for an arbitrary physical system, the time reversal is represented by an involutary anti-automorphism $\tau$ of the underlying $C^*$-algebra $\mathfrak{A}$,

\[ \tau(A^*) = \tau(A)^* , \quad \tau(AB) = \tau(B)\tau(A) , \quad \tau(\tau(A)) = A , \quad A, B \in \mathfrak{A} . \]

The time-reversal map $\tau$ is positive, but for quantum systems (where $\mathfrak{A}$ is non-commutative) the map $\tau$ is not completely positive. That is, if $(\mathfrak{A}_1, \tau_1)$ and $(\mathfrak{A}_2, \tau_2)$ describe two noninteracting quantum systems, then the local map $\tau_1 \otimes \tau_2$ on the minimal $C^*$-tensor product $\mathfrak{A}_1 \otimes \mathfrak{A}_2$ are not positive, hence they are not time-reversal maps. The time-reversal map for the composite quantum system is given by the global positive map

\[ (\tau_1 \otimes 1_2)(1_1 \otimes \tau_2) = \tau_1 \otimes \tau_2 . \]

This fact implies that even when the two quantum systems do not interact in any way, the time-reversal for the first quantum system is not given by $\tau_1 \otimes 1_2$. The map $\tau_1 \otimes 1_2$ represents the time-reversal operator for the first subsystem if and only if the two systems are not entangled with Einstein–Podolsky–Rosen correlations. This is the case for every state of the combined system if and only if at least one of the subsystems is classical (in the sense that either $\mathfrak{A}_1$ or $\mathfrak{A}_2$ is commutative). In an entangled system with broken time-reversal symmetry the direction of the arrow of time has to point to the same direction for all (even noninteracting) subsystems. This requirement of global consistency distinguishes quantum causality from classical causality.

14 Hadamard Determinism Cannot Be Globally Valid

Complex systems can sometimes establish truly novel emergent properties, new properties that the component parts do not have. Often the doctrine of determinism is opposed by the principle of emergence. Yet, if the underlying physical system is deterministic in the sense of Hadamard, then all consistent higher-level theories can be extended to deterministic systems in the sense of Hadamard.

From the point of view adopted here freedom refers to actions in the material world which are not determined by the presently adopted first principles of physics. Such a view violates the usually assumed universal validity of the first principles of physics. Yet, such a violation is not surprising. The postulate that the underlying fundamental time evolution should be an automorphism is a queer assumption indeed. An automorphism is by definition a map which does not change any physically relevant feature. From a physical point of view the dynamics of interacting systems cannot simply be postulated to be automorphic, but has to be derived from first principles. Yet, a mathematically consistent theory of interactions does not yet exist.

\[ Wigner (1932). \text{ Compare also Wigner (1959), chapter 26.)}\]

A linear map $\tau : \mathfrak{A} \rightarrow \mathfrak{A}$ is said to be completely positive if the linear map $\tau \otimes 1_n : \mathfrak{A} \otimes \mathfrak{B}(\mathbb{C}^n) \rightarrow \mathfrak{A} \otimes \mathfrak{B}(\mathbb{C}^n)$ is positive for all $n \geq 1$, where $\mathfrak{B}(\mathbb{C}^n)$ is the $C^*$-algebra of all complex $n \times n$-matrices and $1_n$ is the identity transformation of $\mathfrak{B}(\mathbb{C}^n)$ onto itself.

This is a consequence of theorem 4.14 in Takesaki (1979), p.211.
Remarks on the derivation of the dynamics from first principles

Nowadays it is generally accepted that all known interactions of matter are due to gauge fields. The dynamics of so-called bare elementary systems is known to be automorphic. Here, a bare elementary system is defined as an indecomposable representation of a kinematical group (like the Galilei or Lorentz group), disregarding interactions intermediated by gauge fields. Such elementary systems are called bare since they only characterize the transformation properties under the actions of the kinematical group. Yet, elementary systems with non-zero mass or non-zero electrical charge are inevitably coupled to the gravitational or electromagnetic gauge fields. Due to this interaction, a bare elementary system acquires a complicated structure. This process is referred to as dressing.

The derivation of the dynamics of a dressed system from the group-theoretically known automorphic dynamics of the bare elementary systems and the bare gauge fields in a mathematically rigorous way involves many difficulties. In particular our understanding of radiation effects and the infinities of self-interactions is far from satisfactory. It is an open question whether the resulting dynamic is automorphic. In spite of the fact that present-day quantum field theory is quite successful, it has only the status of a phenomenological theory.

Notwithstanding the problematic character of automorphic time evolutions, mounting evidence shows that for practically all concrete applications the time evolution of any open subsystem of the material world can be described by the restriction of an automorphic dynamics of a larger system consisting of the object system and its environment. So we have to ask why a hypothetical universal automorphic dynamics leads to contradictions. A possible answer derives from the remark that not everything that is locally true is necessarily true globally. So it may be worthwhile to contemplate about theories which only locally, but not globally correspond to the contemporary theories. At present we have no idea how to achieve such a goal. Besides, it is questionable whether a solution of this problem could contribute to an understanding of the problem of freedom of action in the material world.

It is disquieting that we have no idea where precisely in the material domain the presently adopted first principles of physics do not apply. Presumably new physical principles will need to be discovered. Since in physics time is used in a rather ad-hoc manner, one could think about a more physical concept of time, say as time operator which represents time as a dynamical observable. But it is very much an open question what form a physical theory including the freedom of action will take.

Acknowledgments

I am grateful to Harald Atmanpacher for helpful comments on an earlier version of this contribution.

---

44 Compare for example the ideas discussed by Misra (1995).
References


