

## OPEN ACCESS

# The criterion for time symmetry of probabilistic theories and the reversibility of quantum mechanics

To cite this article: A T Holster 2003 *New J. Phys.* **5** 130

View the [article online](#) for updates and enhancements.

## Related content

- [From physical assumptions to classical and quantum Hamiltonian and Lagrangian particle mechanics](#)  
Gabriele Carcassi, Christine A Aidala, David J Baker et al.
- [‘The concept of information in physics’: an interdisciplinary topical lecture](#)  
T Dittrich
- [Quantum pasts and the utility of history](#)  
James B Hartle

## Recent citations

- [The Species Problem and its Logic: Inescapable Ambiguity and Framework-Relativity](#)  
Steven James Bartlett and Ph.D.
- [Operational formulation of time reversal in quantum theory](#)  
Ognyan Oreshkov and Nicolas J. Cerf
- [Pre- and Post-Selected Ensembles and Time-Symmetry in Quantum Mechanics](#)  
D. J. Miller

## The criterion for time symmetry of probabilistic theories and the reversibility of quantum mechanics

**A T Holster**

12 Pa Road, Pukerua Bay, Wellington, New Zealand

E-mail: [ATASA@clear.net.nz](mailto:ATASA@clear.net.nz)

*New Journal of Physics* 5 (2003) 130.1–130.28 (<http://www.njp.org/>)

Received 12 June 2003

Published 9 October 2003

**Abstract.** Physicists routinely claim that the fundamental laws of physics are ‘time symmetric’ or ‘time reversal invariant’ or ‘reversible’. In particular, it is claimed that the theory of quantum mechanics is time symmetric. But it is shown in this paper that the orthodox analysis suffers from a fatal conceptual error, because the logical criterion for judging the time symmetry of probabilistic theories has been incorrectly formulated. The correct criterion requires symmetry between future-directed laws and past-directed laws. This criterion is formulated and proved in detail. The orthodox claim that quantum mechanics is reversible is re-evaluated. The property demonstrated in the orthodox analysis is shown to be quite distinct from time reversal invariance. The view of Satoru Watanabe that quantum mechanics is time asymmetric is verified, as well as his view that this feature does not merely show a *de facto* or ‘contingent’ asymmetry, as commonly supposed, but implies a genuine failure of time reversal invariance of the laws of quantum mechanics. The laws of quantum mechanics would be incompatible with a time-reversed version of our universe.

**Contents**

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>The meaning of time reversal invariance</b>	<b>3</b>
<b>3</b>	<b>Criterion for time reversal invariance of probabilistic theories</b>	<b>5</b>
<b>4</b>	<b>Proving the validity of the criterion</b>	<b>7</b>
<b>5</b>	<b>The failure of the orthodox criterion</b>	<b>11</b>
<b>6</b>	<b>Origin of the orthodox criterion and the deterministic case</b>	<b>15</b>
<b>7</b>	<b>The meaning of orthodox time reversal</b>	<b>17</b>
<b>8</b>	<b>The time asymmetry of quantum mechanics</b>	<b>20</b>
<b>9</b>	<b>The <i>de facto</i> irreversibility view and thermodynamic asymmetry</b>	<b>22</b>
<b>10</b>	<b>Summary</b>	<b>27</b>
	<b>Acknowledgments</b>	<b>27</b>
	<b>References</b>	<b>27</b>

**1. Introduction**

Physicists routinely state that the fundamental laws of physics are ‘time reversal invariant’ or ‘time symmetric’ or ‘reversible’, and in modern physics this claim stems largely from the commonly stated view that quantum mechanics (QM) is time symmetric or reversible. But Watanabe [1]–[5] argued convincingly that this orthodox view is mistaken, and based on a flawed assumption about the very concept of time symmetry. The orthodox result is that QM (with the exception of K-meson systems) satisfies a principle here called *orthodox reversal symmetry*. But Watanabe shows that orthodox reversal symmetry fails to capture time reversal invariance, because it only requires symmetry among future-directed probability laws, whereas true time symmetry requires symmetry between future- and past-directed laws.

A small number of writers have recognized the importance of this issue. Watanabe [2]–[5] is the most important; after publishing studies of time reversal of the probability theory for QED in the 1950s, he moved on to detailed examinations of time symmetry and retrodictive probabilities, showing that quantum mechanical probabilities are time asymmetric, and this paper is most indebted to his work, which I believe deserves to be re-examined. Healey [6] and Callender [7] have most clearly supported Watanabe’s conclusion that QM is time asymmetric, and a number of writers have been concerned with related points, such as the asymmetry between past- and future-directed probabilities, and the problem of retrodiction, including Reichenbach [8], Grunbaum [9], de Beaugard [10]–[13], Earman [14]–[17], McCall [18], Penrose [19], and Price [20]. Boltzmann [21], the Ehrenfests and others were already aware of the problem of introducing past-directed probabilities in the classical context.

However, there is still no consensus on the central conceptual issues, and there is a distinct lack of adequate proofs in the subject. The main aim of this paper is to define an adequate criterion for the time reversal invariance of probabilistic theories, analyse the concept of orthodox reversal symmetry, and show their logical relationship. Their logical independence is established through two simple models, *theories A and B*. These models prove that orthodox reversal symmetry is not time symmetry, and the conclusions reached here raise severe doubts about the analysis of probabilistic time symmetry currently accepted in physics and philosophy of physics.

The key application to QM is briefly summarized, and I conclude with some comments about the interpretation of this result, particularly its impact on the popular view that the ‘irreversibility’ of quantum measurement is merely a ‘*de facto*’ or ‘contingent’ time asymmetry. It is widely assumed that this form of ‘irreversibility’ does not reflect any ‘lawlike’ time asymmetry of QM. This popular view is rejected here, and Watanabe’s main conclusion is supported.

## 2. The meaning of time reversal invariance

In what follows, I will often use the short term *time symmetry* for the concept referred to synonymously in physics as *time reversal invariance*, *invariance under time reversal*, or *reversibility*. The second term is most literally correct. These terms are generally intended in the literature to refer to the same concept. However, some authors have made the distinction between the notion of *reversibility*, as commonly used to refer to the specific mathematical symmetry of QM, and the notion of *time reversal invariance*. For instance, after presenting his argument for the time asymmetry of QM, Watanabe [3, section 8, p 157] states that

This basic (time) asymmetry due to irretrodictability is compatible with reversibility or any other symmetry rules mentioned at the beginning of this section.

To make the distinction explicit, I use the term *orthodox reversal symmetry* in contrast to *time reversal invariance*, or *time symmetry*. The latter is defined as follows.

**Definition.** *A law or theory of physics is time reversal invariant just in case it is unchanged under the general transformation  $T : t \rightarrow -t$ , i.e. inversion of the time axis.*

We assume here that time is defined as a continuum of pointlike moments, infinite in both directions. This is easily adapted to include cases where time is bounded, but we do not need to consider that generalization. As a one-dimensional continuum, time has exactly two directions, called *earlier than* (or *past*) and *later than* (or *future*). These two directions are logically defined, and do not by themselves imply any asymmetry. However, in some situations, it is important to realize that there is only *one logically independent direction of time*. If we have two moments,  $t_1$  and  $t_2$ , where  $t_1 < t_2$  ( $t_1$  is earlier than  $t_2$ ), then by definition,  $t_2 > t_1$  ( $t_2$  is later than  $t_1$ ). That is, to define the time continuum, we need only appeal to one fundamental direction, say ‘later than’, and we define the opposite direction in terms of this.

This logical notion is quite distinct from the notion of ‘*the direction of time*’, by which is generally meant ‘the direction of time *flow*’, which involves the metaphysical concept of a flow from present to future. We will not consider the question of a unique ‘tensed direction of time’ here: it is additional to the simpler questions about time symmetry of physical theories or laws. We will only consider the question of whether various proposed laws or theories of physics are invariant w.r.t. the two directions presupposed on the time continuum.

When we write a law or theory of physics, we assume a coordinate representation of time in the usual way, with a variable  $t$  referring to moments. A *time reversal transformation* is a mapping  $T : t \rightarrow -t$ . Of course, a specific coordinate representation assumes some definite origin of time,  $t = 0$ , and, strictly speaking, there are innumerable distinct time reversal transformations, depending on what we choose for the origin. Any of these transformations, however, achieves the same effect of *exchanging the time directions* between any two moments, no matter where the origin is chosen. A very important point is that we normally assume that the theories we are dealing with are also *time translation invariant*: i.e. they are universally quantified over all times,

or, equivalently, they are invariant w.r.t. the choice of origin. On this preliminary assumption, which we normally take for granted in physics, invariance properties of a law under any specific time reversal (i.e. specific choice of origin for the mapping  $T : t \rightarrow -t$ ) are the same as under any alternative choice. Essentially, time translation symmetry allows us to give a *relational-vector* representation of time, where only the intervals and directions between moments count, and any reference to ‘absolute coordinates’ can be dismissed.

A *theory of physics* will be regarded here as the collection of all its *laws*. The class of all the laws or propositions entailed by a theory defines the theory. Laws or propositions of a theory are most simply interpreted as classes of *possible processes*, or more generally (if possible), as classes of *possible worlds*, with worlds being universal processes. In this view, a law or a theory states that certain logically possible processes or worlds are *physically possible*, and others are not. General symmetries of a theory reflect symmetries among the possible processes it allows.

This view of laws or theories is simple enough in a deterministic context, where we have only *possibilities* of processes, but it is trickier in probabilistic theories, when we also have to take into account *probabilities* of processes. The difficulty in representing probabilistic theories in a formal way is behind the key problems in formally deriving the conditions or criteria for symmetry properties of probabilistic theories. However, this difficulty can be circumvented, and a perfectly general criterion for time reversal symmetry of classes of probabilistic laws can be given.

In our basic scheme of things, at any rate, we begin with the following propositions:

**Proposition 1.** *For any sufficiently well defined physical theory,  $\mathbf{T}$ , the time-reversed image of the theory,  $T\mathbf{T}$ , is the image of  $\mathbf{T}$  under the general transformation  $T : t \rightarrow -t$ . Similarly for laws,  $L$ , of a theory (definition).*

**Proposition 2.**  *$\mathbf{T}$  is time reversal invariant just in case  $\mathbf{T} = T\mathbf{T}$ , i.e. just in case the theory and its time-reversed image are identical theories (definition).*

**Proposition 3.** *For any theory  $\mathbf{T}$ , and any law  $L$ ,  $\mathbf{T}$  entails  $L$  just in case  $T\mathbf{T}$  entails  $TL$  (see below).*

**Proposition 4.** *If  $\mathbf{T}$  is time reversal invariant, then for any law  $L$ ,  $\mathbf{T}$  entails  $L$  just in case  $T\mathbf{T}$  entails  $L$ . (Obvious since time reversal invariance means  $\mathbf{T} = T\mathbf{T}$ .)*

**Proposition 5.** *Conversely,  $\mathbf{T}$  is time reversal invariant just in case, for any law  $L$ ,  $\mathbf{T}$  entails  $L$  if  $T\mathbf{T}$  entails  $L$ . (Obvious if we assume that  $\mathbf{T}$  entails  $\mathbf{T}$ .)*

**Proposition 6.**  *$\mathbf{T}$  is time reversal invariant just in case, for any law  $L$ ,  $\mathbf{T}$  entails  $L$  just in case  $\mathbf{T}$  entails  $TL$ . (Obvious from previous propositions.)*

Proposition 3 is a logical result: it is a consequence of the fact that  $\mathbf{T}$  is equivalent to a class of processes, represented by its laws, and  $T\mathbf{T}$  is equivalent to the class of time-reversed processes, represented by the class of time-reversed laws.

The point of proposition 6 is that we can reduce the problem of determining the time reversal of a general theory to a problem about determining the time reversal of its *laws*, and when we have a complicated theory, we generally seek an account of the time-reversed images of the laws of a theory, rather than of the theory directly. Note also that it is *not* true (it is never true) that, if  $\mathbf{T}$  is time reversal invariant, then for any law  $L$ , if  $\mathbf{T}$  entails  $L$  then  $L \equiv T(L)$ . On the contrary,

every theory entails weaker laws that are *not* time reversal invariant when taken separately. This is why we need proposition 6.

In the following discussion, it is assumed that there is a unique time reversed image of theories of physics we are dealing with, although this is only strictly true for theories in which processes are defined precisely enough, so that a unique time reversed image of each process is determined. We expect such precision in fully developed fundamental theories, because they specify processes very precisely, in terms of fundamental variables. But it may be a problem if a theory is open to a number of competing interpretations, which is a genuine issue in QM. However, in such cases, the concept of *time reversal* itself is not ambiguous. Rather, the *theory* is ambiguous, because different interpretations effectively represent different theories, which may very well have different temporal properties. This is noted later in relation to deterministic interpretations of QM.

The criterion for time reversal invariance of probabilistic theories proposed here simply assumes that the theory in question provides a complete interpretation for the time reversals of states, and that the theory is completely defined by a class of probabilistic laws of the stated form. If this is problematic for a particular theory, then the *application* of the criterion to that theory may be problematic; but this is not a problem for the criterion itself. The orthodox account makes the same assumption that there is a correct criterion for the time reversal of such theories; the point here is that the criterion proposed in the orthodox account is simply wrong.

It should be stressed that the concept of time reversal invariance is not itself ambiguous: it means *invariance under the general transformation*  $T : t \rightarrow -t$ . The models defined in the following sections demonstrate that the orthodox criterion fails to capture this concept. We cannot maintain that the orthodox criterion simply represents an ‘alternative concept’ of time reversal invariance; on the contrary, it is shown here to represent a different kind of symmetry altogether, which is decisively different from symmetry under  $T$ .

### 3. Criterion for time reversal invariance of probabilistic theories

The critical problem is to determine *an adequate criterion for time reversal symmetry of a probabilistic theory*, and this is relevant to both QM and to statistical mechanics. We consider theories that are specified as *classes of probabilistic transition laws*, with the following basic form.

*Probability law L*

$$(\forall t)\text{prob}(s_2(t + \Delta t)|s_1(t)) = p.$$

This says that the probability of obtaining a system in the state  $s_2$  at the (later) time  $t + \Delta t$ , given that the earlier state is  $s_1$  at time  $t$ , is equal to  $p$ . Note that, in this law, the terms  $s_1(t)$  and  $s_2(t + \Delta t)$  represent propositions. Probabilities are introduced in QM through a version of the Born interpretation as follows.

*QM probability law*

$$\text{prob}(\psi_2(t + dt)|\psi_1(t)) = |(\psi_2, \psi_1)|^2.$$

In QM,  $\psi_1$  and  $\psi_2$  are the special ‘maximal quantum states’, which represent ‘complete information’ about the wavefunctions. I will briefly discuss the need to conditionalize these with respect to measurement interactions later. But we do not have to consider the detailed states in the law  $L$  yet: we first just want to find the general transformation under time reversal of a law of the form  $L$ .

Given that our theory is defined as a class of such laws, the time-reversed image of the theory is then obtained as the class of the time-reversed images of these laws, according to proposition 6.

What is the time reversal of the law  $L$ ? It is assumed in the orthodox analysis to be the following corresponding law.

*Probability law  $L^*$*

$$(\forall t) \text{prob}(T s_1(t + \Delta t) | T s_2(t)) = p.$$

$T$  here defines the time reversal operator on the states. The law  $L^*$  says that *the probability of obtaining a later state  $T s_1$  given earlier state  $T s_2$  equals  $p$* . Making this assumption gives the following criterion for time reversal of a probabilistic theory,  $\mathbf{T}$ .

**Definition (the orthodox criterion for reversal symmetry).**  $\mathbf{T}$  is a time reversal invariant theory just in case, for any law of the form  $L : (\forall t) \text{prob}(s_2(t + \Delta t) | s_1(t)) = p$  entailed by  $\mathbf{T}$ , the law  $L^*$ ,  $(\forall t) \text{prob}(T s_1(t + \Delta t) | T s_2(t)) = p$ , is also entailed by  $\mathbf{T}$ .

Or more simply,

$$\mathbf{T} \Rightarrow [\text{prob}(s_2(t + \Delta t) | s_1(t)) = \text{prob}(T s_1(t + \Delta t) | T s_2(t))].$$

This principle is generally summarized in QM as *the probabilistic principle of micro-reversibility*, that  $\omega = \omega_{\text{rev}}$ , where  $\omega$  is the transition probability  $|\langle \psi, \phi \rangle|^2$ , and  $\omega_{\text{rev}} = |\langle T\phi, T\psi \rangle|^2$ , e.g. [22].

Quantum mechanics satisfies the orthodox criterion, and this is the reason it is judged to be a time symmetric or reversible theory. The fact that QM satisfies the orthodox criterion is straightforward, but *the justification of the orthodox criterion* is not. Is there a good proof that the orthodox criterion *correctly defines time reversal invariance*, or that  $L^*$  is the time reversal of  $L$ ? There cannot be, because the time reversal of  $L$  is actually the following law, which will be shown to be independent of  $L^*$ .

*Probability law,  $TL$*

$$(\forall t) \text{prob}(T s_2(t - \Delta t) | T s_1(t)) = p.$$

The peculiarity of  $TL$  lies in the fact that the probabilities in  $TL$  are *past-directed conditional probabilities*, as opposed to the future-directed probabilities found in  $L$  and in  $L^*$ . This is the theme that Watanabe stressed: for time reversal invariance, we need a symmetry between future directed laws and corresponding past-directed laws, and not merely a symmetry between future directed laws and other future directed laws<sup>1</sup>.

Supposing that  $TL$  is the real time reversal of  $L$ , the (correct) criterion for time reversal invariance is the following.

<sup>1</sup> Note that Watanabe uses the alternative terms *retrodictive probabilities* and *predictive probabilities*, but those terms have an epistemological connotation, whereas the terms *past-directed* and *future-directed probabilities* refer more explicitly to objective probabilities in nature, rather than human abilities to actually predict things. Thanks to Graham Oddie for suggesting this terminology.

**Definition (the (correct) criterion for time reversal invariance).**  $\mathbf{T}$  is a time reversal invariant (TRI) theory just in case, for any law of the form  $L: (\forall t) \text{prob}(s_2(t + \Delta t)|s_1(t)) = p$  entailed by  $\mathbf{T}$ , the law  $T(L)^*, (\forall t) \text{prob}(Ts_2(t - \Delta t)|Ts_1(t)) = p$  is also entailed by  $\mathbf{T}$ .

Or more simply,

$$\mathbf{T} \Rightarrow [\text{prob}(s_2(t + \Delta t)|s_1(t)) = \text{prob}(Ts_2(t - \Delta t)|Ts_1(t))].$$

This is the key principle of time symmetry for probabilistic theories.

#### 4. Proving the validity of the criterion

To establish this, it must be *proved* that  $TL$ , as defined above, correctly represents the time-reversed image of  $L$ , and it must also be shown how and why the orthodox criterion fails. In view of the controversies that have dogged this subject for the last hundred years or more, the need for *proofs* of such claims is paramount<sup>2</sup>.

There are a number of ways of approaching a proof, which need to be seen in conjunction with each other. The first approach is a *syntactic derivation*: for this, we use rules for operating syntactically on the complex equations or statements that represent the laws, to derive their time-reversed images. Physicists typically appeal to such rules, the basic idea being that, to obtain the time reversal of a statement, we replace occurrences of the time variable, ‘ $t$ ’, in the statement with its reversal, ‘ $-t$ ’, and similarly we replace any state or property variables, ‘ $s$ ’, with their time reversals, ‘ $Ts$ ’, and so on. For example, see [1] for a system of such rules for quantum electrodynamics. But any such system is only justified by semantic arguments.

A *semantic proof* requires considering what the complex expressions *mean*, and deducing what the time-reversed images must be, from elementary considerations about the definition of time reversal. It is these semantic considerations that must be used to validate any syntactically based derivation, and we find that we are inevitably drawn into such a treatment in a proper proof.

A third approach is a kind of ‘pragmatic semantics’, where we consider hypothetical models to illustrate and check results, and to explain, for instance, why past-directed probabilities are necessary for time symmetry. This is necessary to test and confirm the conclusions we arrive at using the other approaches. A number of pertinent arguments put forward by writers including Watanabe [1]–[5], Healey [6], Earman [14]–[17], Callender [7], Penrose [19], Schrödinger [23], von Neumann [24], and others are of this kind. This is how Watanabe and others discovered the flaw in the orthodox criterion. But more careful proofs of the first kinds are required.

<sup>2</sup> Some authorities claim that there are no objective proofs of such results, and that the choice of a criterion for time reversal is partly arbitrary or conventional. In this view, we simply adopt one or other ‘interesting’ type of symmetry which has something to do with time, and give it the name ‘time reversal’, or ‘reversibility’ or ‘motion reversal’ or ‘irreducibility’, with the choice of name being a terminological convention. I strongly reject this view. The fact that we have defined time reversal as induced by the transformation:  $t \rightarrow -t$  means that there must be mathematical proofs about its application to things such as classes of probabilistic laws. On the other hand, given that the orthodox reversal symmetry is not time symmetry, we need to examine what it really is, and define its underlying transformation.



*Syntactic derivation*

For the most basic syntactic derivation, we just replace the time variables  $t$  and  $\Delta t$  in  $L$  with  $-t$  and  $-\Delta t$ , and we replace the states  $s_1$  and  $s_2$  with their time reversals,  $Ts_1$  and  $Ts_2$ . This gives

$$(\forall -t) \text{prob}(Ts_2(-t - \Delta t)|Ts_1(-t)) = p.$$

Because the term  $-t$  is universally quantified, and the variable  $t$  ranges over positive and negative time coordinates equally, the replacement of  $t$  by  $-t$  actually makes no difference whatsoever. The only effective changes are to the sign of  $\Delta t$  (which is a constant, not a quantified variable), and the states involved. Thus, we obtain  $TL$  as claimed above.

What may seem surprising is that the time reversal operation does not swap the order of ‘conditional implication’ between the states in the probability law. Instead, *it reverses the temporal ‘direction’ of the implication*. It generates a *past-directed conditional probability*, implying that there are probabilistic connections *from present states to past states*. This is indeed the natural reversal of the *future-directed conditional implication*, represented by the original law,  $L$ . Reflection on this led Watanabe to realize that  $L^*$  cannot represent the time reversal of a *future-directed conditional probability*, since it is also a future-directed conditional, whereas the time-reversed law must surely involve a *past-directed conditional probability*.

This ‘derivation’ is limited, of course, by our confidence in the idea that this syntactic manipulation really produces the correct result: in particular, should we also modify the *conditional probability function*,  $\text{prob}(\cdot|\cdot)$ , into a ‘reversed’ form, which we might write as  $T \text{prob}(\cdot|\cdot)$ , which also swaps the order of the implication? That is, should we write  $T \text{prob}(a|b) = \text{prob}(b|a)$ ? To answer this, we can further analyse the meaning of the conditional probability function in the usual way, as a function of absolute probabilities. By the usual definition,

$$\text{prob}(s_2(t + \Delta t)|s_1(t)) = \frac{\text{prob}(s_1(t) \text{ and } s_2(t + \Delta t))}{\text{prob}(s_1(t))}.$$

(The interpretation of such absolute probabilities may be controversial; but all we really require to show the main point here is that the conditional probabilities reflect meaningful *ratios of expected frequencies of the relevant events or processes*, as considered below. A detailed discussion of probability theory itself is beyond the scope of this paper, but some further comments about QM probabilities are given in section 8.)

We can now transform the simpler absolute probability terms on the right-hand side instead of the conditional probability directly. Taking the syntactic transformation of the terms on the right-hand side gives

$$T[\text{prob}(s_1(t) \text{ and } s_2(t + \Delta t))] = \text{prob}(Ts_1(-t) \text{ and } Ts_2(-t - \Delta t))$$

and assuming

$$T[\text{prob}(s_1(t))] = \text{prob}(Ts_1(-t))$$

then

$$\frac{\text{prob}(Ts_1(-t) \text{ and } Ts_2(-t - \Delta t))}{\text{prob}(Ts_1(-t))} = \text{prob}(Ts_2(-t - \Delta t)|Ts_1(-t))$$

and we get the same result. There is no reason to think that the order of implication is reversed. But we may still not be sure that the *absolute probability function*,  $\text{prob}(\cdot)$ , is not itself modified by time reversal: i.e., perhaps we should write  $T \text{prob}(\cdot)$ , and distinguish this from  $\text{prob}(\cdot)$ ? We will dismiss this possibility next: but at any rate, there is no plausible way that  $T$  can reverse the order of implication in the conditional probability law, as the orthodox view requires.

*Semantic proof*

To make the role of the probabilities more explicit, we can give a more concrete statistical interpretation, in terms of expected frequencies of states in a world governed by the probability laws in question. This leads us to introduce the semantic notion of *worlds*, and the semantic view of the meaning of the law. The conditional probability law,  $L$ , is interpreted as requiring the following relative *lawlike expected frequencies of states* in the actual world:

$$\frac{\text{Exp Freq}(s_1(t) \text{ and } s_2(t + \Delta t))}{\text{Exp Freq}(s_1(t))} = p.$$

The modern *intensional semantic* approach (in its simplest form) explicates *propositions* as true or false at *worlds*. Thus, we may say that a proposition,  $L$ , is true at some possible worlds and false at others. A proposition is thus naturally identified as a *mapping from worlds to truth values*, or more simply, as a class of possible worlds: the worlds at which it is true. This called an *intensional view of propositions*: it treats propositions in an objectual way.

The notion of *time reversal of propositions* is then explicated as follows. If a proposition  $L$  is true at a world  $W$ , then, by definition, the time-reversed image  $TL$  of  $L$  is true at the world  $TW$ , i.e. at the time-reversed image of the world  $W$ . If we write  $L[W]$  to explicitly represent the proposition  $L$  as a *function on worlds*, then the value of  $L[W]$  is a truth value<sup>3</sup>.

**Semantic definition of time reversal of propositions.**

*$L[W]$  is true if and only if  $TL[TW]$  is true.*

We first consider a world,  $W$ , that satisfies the expected frequencies, and then derive the time-reversed image of that world,  $TW$ . Then we know that the time-reversed law,  $TL$ , must be satisfied by the world  $TW$ . Thus we can directly obtain the meaning of the law  $TL$  and verify the syntactic derivation assumed in the previous section.

In the world  $W$ , we assume that there is a certain frequency of systems that display the sequence of states  $s_1(t)$  and  $s_2(t + \Delta t)$ , at any time  $t$ . This frequency reappears, in the time-reversed image of that world, as the frequency of  $Ts_1(t)$  and  $Ts_2(t - \Delta t)$ , for any time  $t$ . The latter is the frequency of cases where  $Ts_1$ , at some  $t$ , is *preceded* by  $Ts_2$  at  $t - \Delta t$ .

Similarly, the original frequency of  $s_1(t)$ , for any time  $t$ , appears, in the time-reversed world, as the frequency of:  $Ts_1(t)$ , for any time  $t$ . Thus the frequencies in the original world determine the corresponding frequencies in the time-reversed world.

This means the following lawlike relation must hold in the time-reversed image of the original world:

$$\frac{\text{Exp Freq}(Ts_1(t) \text{ and } Ts_2(t - \Delta t))}{\text{Exp Freq}(Ts_1(t))} = p.$$

And this corresponds to the probability law  $TL:\text{prob}(Ts_1(t - \Delta t)|Ts_2(t)) = p$ .

<sup>3</sup> Formal intensional semantics was originally developed independently by Montague [25] and Tichý [26]. See [27] and [28] for a summary of some of the main concepts. The techniques of intensional semantics are recognized as an indispensable tool in modern conceptual analysis, but there has been little recognition of it in the philosophy of physics, where naive positivist or instrumentalist principles remain rife, despite being generally discredited by professional semanticists. There is some difficulty applying intensional semantics to QM because the notion of the *world* as a whole is unclear, but I am convinced it must be applied in some form if we are to define an adequate compositional semantics for general transformations in physics.

If the original frequency law is judged to be satisfied in the world  $W$ , as a reflection of the probability law  $L$ , then this shows that the derived frequency law above must be judged to be satisfied in the world  $TW$ , and this reflects the reversed probability law,  $TL$ . Note that this does not depend on the criterion we use to *judge* that the expected frequencies are satisfied by the actual frequencies: rather, *if the initial expected frequencies are judged to be satisfied in  $W$ , then the time-reversed expected frequencies must equally be judged to be satisfied in  $TW$* , since they are identical (allowing us to circumvent any disputes about the criteria for judging such things).

Thus we confirm that the probability law,  $L$ , is directly transformed to the law  $TL$  by the time reversal transformation. And indeed, we can assume that the general probability functions  $\text{prob}(\cdot|\cdot)$  or  $\text{prob}(\cdot)$ , no matter how they are explicated in the final analysis, are nonetheless *world invariant functions*, which must be invariant under time reversal.

### *Syntactic proof fails*

The semantic proof should be enough to convince us that the syntactic rule we have found (the criterion) is correct. But this does not mean that the syntactic derivation we gave initially is adequate as a *formal deductive proof*. A proper deductive proof needs to provide syntactic rules for applying the time reversal operator over full symbolic representations of laws and propositions. For instance, we used have the symbolism  $s_i(t)$  to represent a complex construction, which is more adequately written as  $S(t) = s_i$ , a proposition stating that the system state at time  $t$  is  $s_i$ . More generally we can write  $S(n, t) = s_i$ , to mean that system  $n$  at time  $t$  is  $s_i$ . Similarly, the state  $s_i$  may be defined by more fundamental variables; and the law itself generates probability values through a general function on the  $s_i$  and  $\Delta t$ , i.e.  $p_{i,j} = f(s_i, s_j, \Delta t)$ . Should we not also take a transformation of probabilities as  $Tp_{i,j} = Tf(s_i, s_j, \Delta t)$ ? We cannot establish an adequate deductive proof unless we have given a full account of the fundamental constructions of the language, with deductive rules for applying  $T$  as an operator over all complex expressions.

I believe this cannot be done successfully in *ordinary physics object languages*, because it requires an extension of the ordinary (extensional) symbolism to include explicit formal recognition of *intensional variables*, such as the world variable,  $W$ , that appears in the usual intensional representation of *propositions*. For instance, the probability laws relate together propositions  $\text{prob}(S(t + \Delta t) = s_j | S(t) = s_i) = p_{i,j}$ . Any operator on this complex symbol must operate on the propositions; but propositions are logically higher-order objects, e.g. mappings from worlds to truth values, or world-times to truth values. An adequate symbolism must expand the representation to something like  $S(n, t, W) = s_i$ , meaning that *the state of system  $n$  at time  $t$  in world  $W$  is  $s_i$* . I maintain that an explicit representation of *world variables* is necessary before any adequate deductive rules defining the  $T$ -operator on expressions can be defined (see footnote 3).

There is no use therefore in pressing for a more precise syntactic derivation, without overhauling the extensional formalism of physics, and providing an intensional formalism. This involves us in the complex issues of interpretation of the *types of world* used to represent theories, and is a larger project in the foundations of physics. This needs to be undertaken to properly clarify the time reversal operator.

But we can still be satisfied by our semantic arguments, and they are confirmed next, when we turn to showing the logical relations between the different symmetries through their exemplification in simple models.

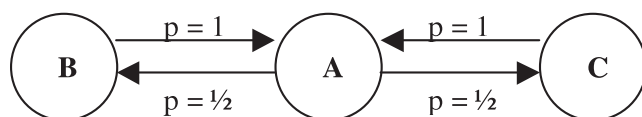


Figure 1. Probabilities for theory A.

## 5. The failure of the orthodox criterion

I will now show directly that the law  $L^*$  cannot be the time reversal of  $L$ , and that the orthodox criterion is invalid. This is done by defining two model theories that show that the orthodox criterion fails to provide either (i) a necessary condition, or (ii) a sufficient condition for time symmetry. The first is very simple.

*Theory A: orthodox reversal invariance is not necessary for time symmetry*

The following model theory, theory A, is obviously time symmetric but fails to satisfy the orthodox criterion. The theory defines three states,  $A$ ,  $B$ , and  $C$ , of a world. Each state is by definition its own time-reversed image: i.e.  $A = TA$ ,  $B = TB$ ,  $C = TC$ . There are four probability laws, illustrated in figure 1.

*Theory A.*

$$\begin{aligned} \text{prob}(B(t+1)|A(t)) &= 1/2 & \text{prob}(A(t+1)|B(t)) &= 1 \\ \text{prob}(C(t+1)|A(t)) &= 1/2 & \text{prob}(A(t+1)|C(t)) &= 1 \\ \text{Other future-directed transition probabilities} & & &= 0. \end{aligned}$$

The state jumps every second between  $A$ ,  $B$  and  $C$ , with  $A$  between  $B$  and  $C$ ; e.g. a typical sequence might look like

$$\dots \rightarrow A \rightarrow C \rightarrow A \rightarrow C \rightarrow A \rightarrow B \rightarrow A \rightarrow C \rightarrow A \rightarrow B \rightarrow A \rightarrow \dots$$

It must give rise to a process where the state  $A$  occurs deterministically in every second interval of time, and where  $B$  or  $C$  occur randomly in every other interval, with probabilities of  $1/2$  on each occasion. The time reversal of this sequence is

$$\dots \rightarrow A \rightarrow B \rightarrow A \rightarrow C \rightarrow A \rightarrow B \rightarrow A \rightarrow C \rightarrow A \rightarrow C \rightarrow A \rightarrow \dots$$

It is obvious that this theory is time symmetric, because the probability of any given sequence is the same as the probability of its time-reversed image. Conditional probability laws for the class of ordinary processes of theory A are exactly the same as those for the class of time reversals of those processes.

But this theory does not satisfy the orthodox criterion for time reversal invariance. Consider the law

$$\text{prob}(A(t+1)|B(t)) = 1.$$

The orthodox criterion requires that

$$\text{prob}(A(t+1)|B(t)) = \text{prob}(TB(t+1)|TA(t)).$$

But instead theory A entails

$$\text{prob}(TB(t+1)|TA(t)) = \text{prob}(B(t+1)|A(t)) = 1/2 \neq 1.$$

Thus, *the orthodox criterion is not a necessary condition for time reversal invariance.*

It is obvious enough that this theory entails symmetric past-directed probabilities that satisfy the correct criterion, confirming its validity in this case.

We see that a model to show (i) is easy to construct, but a model to show (ii) is more difficult. Yet this is the most important condition to prove: for we know that QM satisfies the orthodox criterion, and if this were a *sufficient condition* to guarantee the correct criterion, then we could conclude that QM is time symmetric after all, and we would not need to revise the orthodox results.

It should be noted that there are two distinct types of theory that provide counter-examples for the claim that the orthodox criterion is a *sufficient condition* for TRI.

- (i) QM type. *Theories with future-directed probabilities, and time translation symmetry, which satisfy the orthodox criterion, but without any lawlike past-directed probabilities specified at all.*
- (ii) Theory B type. *Theories with future-directed probabilities and past-directed probabilities defined, with time translation symmetry, which satisfy the orthodox criterion, but with a directional asymmetry in time.*

While ordinary QM itself is a counter-example, we do not want to use QM to show the general result here: we want to define an independent theory, which has probabilities in both directions, with an unambiguous time asymmetry evident, along with the other appropriate properties.

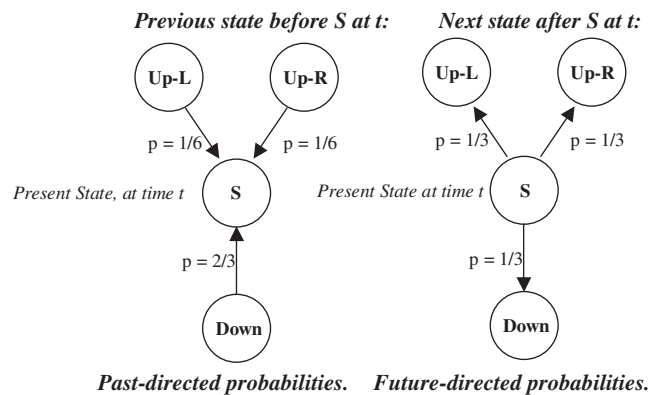
*Theory B: the orthodox criterion is not a sufficient condition for time symmetry*

The following theory B satisfies the orthodox criterion, but is explicitly *not time symmetric*. The theory defines a class of states connected in an infinite, open graph, called an infinite three-tree (see figure 2). This is a special type of graph, where each state (vertex or node) is directly connected to three others (edges); there are no loops; and any two states are connected by a finite path. In addition, two relative *directions* are defined between any two connected states: *Up* and *Down*. This is what makes it a ‘treelike’ structure, with more ‘branches’ in the *Up* direction than ‘trunks’ in the *Down* direction. Each state has two states connected in the *Up* direction, and one in the *Down* direction. (This is an asymmetric extension of the concept of ‘trees’ as usually defined in graph theory; see diagram.)

Connections in the *Up* direction are further distinguished as *Up-R (right)*, or *Up-L (left)*, although the distinction of left and right branches is essentially redundant. Thus, if we start at a given state, *S*, and go directly *Up*, then the number of higher states doubles every time we go up another level.

The theory defines each state to be its own time reversal. The probability laws for the theory are defined in two sets, *past directed*, and *future directed*. The past-directed probabilities are the probabilities that a given state *preceded the present state*; the future-directed probabilities are the probabilities that a given state *will follow the present state*.

For a given state, *S*, we will label the connected states as  $S_{\text{down}}$ ,  $S_{\text{up-left}}$ , and  $S_{\text{up-right}}$ . The proposition that the state at time *t* is *S* is represented  $S(t)$ , and similarly,  $S_{\text{down}}(t - 1)$  means the



**Figure 2.** Theory B: probabilities for transitions between the three states directly connected to a state  $S$ .

state at time  $t - 1$  is  $S_{\text{down}}$ ,  $S_{\text{up-left}}(t + 1)$  means the state at time  $t + 1$  is  $S_{\text{up-left}}$ , and so on in the obvious way.

The probability laws illustrated in figure 2 are defined as follows.

**Past-directed probabilities**

$$\text{prob}(S_{\text{down}}(t - 1)|S(t)) = 2/3$$

$$\text{prob}(S_{\text{up-left}}(t - 1)|S(t)) = 1/6$$

$$\text{prob}(S_{\text{up-right}}(t - 1)|S(t)) = 1/6$$

Other probabilities equal 0.

**Future-directed probabilities**

$$\text{prob}(S_{\text{down}}(t + 1)|S(t)) = 1/3$$

$$\text{prob}(S_{\text{up-left}}(t + 1)|S(t)) = 1/3$$

$$\text{prob}(S_{\text{up-right}}(t + 1)|S(t)) = 1/3$$

The *future-directed probabilities* are symmetric, in that a system has an equal chance of  $1/3$  of moving from a given state to any one of its three connected states in the next moment of time, no matter what direction the new state is in. However, the *past-directed probabilities*, of the system having *moved from* a given state to a given present state, depends on the direction of the earlier state relative to the present state. The probability that the system came from the *Down* direction (i.e. from the single lower state) is  $2/3$ , whereas the probability the system came from either of the states in the *Up* direction is only  $1/6$  each.

The process is Markov in the future direction, i.e. transition probabilities depend only on the present state, and are not contingent on the past sequence. The past probabilities can be considered as lawlike epistemic conditional probabilities, and have the reversed-Markov property, being independent of future states. The consistency of this theory is shown next, and assuming this, it serves our purpose, as follows.

- (i) Theory B is clearly time asymmetric or irreversible because it predicts that the system moves asymmetrically in time: if we look towards the future, we find that the system moves probabilistically into higher and higher states (at an average rate of climb of  $1/3$  node per unit of time). If we look towards the past, we find that the system has moved probabilistically from lower states to higher states, again, with an average rate of climb of  $1/3$  node per unit of time.

Note that this asymmetry holds equally *at any point of time, whether in the past or future, independently of the moment we start from*. That is, it is not a special feature of the special ‘present moment’ of time. The probability laws are *time translation symmetric*. They predict that the system has an average rate of climb of  $1/3$  throughout its history.

- (ii) On the other hand, a quick inspection shows that *the probabilities in this theory satisfy the orthodox criterion of time reversal invariance*. For instance, the orthodox criterion requires that

$$\text{prob}(S_{\text{down}}(t+1)|S(t)) = \text{prob}(S(t+1)|S_{\text{down}}(t))$$

which is true, because both of these probabilities are equal to  $1/3$ . (Other future-directed probabilities are exactly similar.)

We can conclude that *the orthodox criterion does not provide a sufficient condition for time reversal invariance of a theory*.

We also note that the correct criterion gives the right result: it is not satisfied by this theory, and it correctly identifies that the theory is irreversible. For example:

$$\text{prob}(S_{\text{down}}(t-1)|S(t)) \neq \text{prob}(S_{\text{down}}(t+1)|S(t)).$$

### Consistency of theory B

We must show that the probabilities defined in theory B are mutually consistent. It should be noted that the probability laws are only possible because of the special labelling of *directions in the tree*, and the directions are themselves topologically asymmetric. In this respect, the structure of states resembles a set of thermodynamic states, for an open system that has no upper or lower equilibrium point. We should strictly speaking give a constructive proof that such a network is possible first, but assuming this, we can derive the past-directed laws directly from the future-directed laws using a key feature of theory B: its *time translation symmetry*.

The predicted behaviour of the system towards the future is for the state to be *rising steadily*, at an average rate of  $1/3$  nodes  $s^{-1}$ , which is obvious from the fact that the future-directed probabilities are randomly  $1/3$  going in any direction, and there are two *Up* directions and one *Down* direction at each point.

The fact that the system must rise *Up* at an average rate of  $1/3$  node  $s^{-1}$ , plus time translation symmetry, logically corresponds to the average motion of going  $1/3$  node *Down* per unit of time towards the past. This rate towards the past is given by

$$[\text{prob}(S_{\text{down}}(t-1)|S(t))] - [\text{prob}(S_{\text{up-left}}(t-1)|S(t))] - [\text{prob}(S_{\text{up-right}}(t-1)|S(t))] = 1/3$$

and the sum of probabilities is unity:

$$\text{prob}(S_{\text{down}}(t-1)|S(t)) + \text{prob}(S_{\text{up-left}}(t-1)|S(t)) + \text{prob}(S_{\text{up-right}}(t-1)|S(t)) = 1.$$

By symmetry of *left* and *right*, the two probabilities on the right are equal, and the solution is just the solution to  $a - 2b = 1/3$ , and  $a + 2b = 1$ ; hence  $2a = 4/3$  or  $a = 2/3$ , and  $b = 1/6$ . Hence the past-directed probability of *Down* must be  $2/3$ .

This shows the consistency of the past-directed probabilities with the motion from the future-directed probabilities<sup>4</sup>.

<sup>4</sup> The structure of theory B should be contrasted with the alternative ‘symmetric’ system, where we assign past-directed probabilities as  $1/3$  like the future laws. In this case, there is no need to define any ‘down’ direction in the network to define the probability laws; and the state  $S$  at  $t = 0$  is a special ‘low-entropy’ state. It is now overwhelmingly probable that the system will display a *maximal low-point*, the lowest state in the whole history of the system, which should occur quite close to  $t = 0$ . But although the probability laws for this system are *time reversal symmetric* around this special initial moment, they are not symmetric around any others. And this symmetric theory is not *time translation* invariant. If we start at a later time,  $t > 0$ , and use the state at that time,  $S_t$ , in the probability laws, then the probability laws in the immediate past are altered.

## 6. Origin of the orthodox criterion and the deterministic case

We now turn to the meaning of orthodox reversibility symmetry, and observe how the mistake of identifying it with time reversal invariance arose. A typical explanation of time reversal invariance is given by Sklar.

If the laws of nature are time reversal invariant, then for any process that occurs in the world, the process that consists in starting with the time-reversed final state of the original evolution and ending up with the time-reversed initial state of the original process is equally compossible with the laws of nature. . . . Remember, of course, that in the quantum-theoretic context it is the transition probabilities between reversed states which must equal the probabilities of the unreversed states taken in opposite temporal order for the laws to be time reversal invariant. The physics of time reversal invariance, then, is quite simple. [29, p 368].

A slightly more technical explanation is given by Earman:

The laws  $L$  governing physical system  $\{S\}$  are said to be invariant under time reversal, or equivalently, the temporal processes of  $\{S\}$  are said to be nomically reversible if and only if (a) for any nomically possible states  $s_i$  and  $s_f$  of  $S \in \{S\}$ ,  $Ts_i$  and  $Ts_f$  are also possible states of  $S$  and (b) if the laws  $L$  imply that whenever  $S \in \{S\}$  is in state  $s_i$  it is state  $s_f$  after an interval of  $\tau$ , then they also imply that whenever  $S$  is in state  $Ts_f$  it is in state  $Ts_i$  after an interval  $\tau$ . . . . If the equations of motion contain a stochastic term, then the present characterization would have to be modified. The most obvious extension would be to require that the transition probability from  $s_i$  to  $s_f$  equal the transition probability from  $Ts_f$  to  $Ts_i$ . [14, p 281].

These comments show how the intuition behind the orthodox criterion arises: it seems the natural extension of the simpler concept of reversibility given for deterministic transition laws. This is claimed to be the ‘*most obvious extension*’ by Earman, and Sklar tells us that we should ‘*Remember, of course, . . .*’ and that ‘*the physics of time reversal invariance, then, is quite simple*’. But no proofs are provided, and the conclusion is a mistake. In fact, Earman goes on shortly afterwards to observe that this symmetry is not equivalent to symmetry under inversion of the time axis, and there is some conceptual problem here; but he fails to give a deeper analysis.

In an interesting recent account, Callender [7] provides a definition of time reversal invariance (TRI), and it seems clear from the context (see his section 4) that he *intends* to define the correct criterion. But his definition is stated as follows (using superscripted  $T$  to indicate state reversal):

A theory is TRI just in case given a lawful sequence of states of a system from an initial state  $S_i$  to final state  $S_f$  with chance equal to  $r$ , the sequence from the temporally reflected final state  $S_f^T$  to the temporally reflected initial state  $S_i^T$ , also has chance equal to  $r$ , i.e.  $\text{prob}(S_i \rightarrow S_f) = \text{prob}(S_f^T \rightarrow S_i^T)$  [7, p 254].

However, this seems to be just the definition of orthodox reversal symmetry. The problem is caused partly by the opacity of his symbolism: the use of the arrow, ‘ $\rightarrow$ ’ and the suppression of explicit reference to the times at which the states occur can only lead to confusion. This flaw is common to practically all discussions. How are we supposed to perceive the temporal relations involved when references to times are suppressed in the formalism? It is not clear whether he



intends this arrow to reflect a direction of *inference*, or a temporal direction—but in any case, it seems impossible to interpret his definition correctly. This confusion is evident in a subsequent paragraph, where he proposes that an alternative account might be that

[A]n indeterministic process is TRI just in case if  $S_i \rightarrow S_f$  is compatible with the laws then so is  $S_i^T \rightarrow S_f^T$ . This is the probabilistic version of what is sometimes called ‘motion reversal invariance’. Motion reversal invariance holds that if state  $S_i$  evolves to state  $S_f$ , then it is also possible for the state  $S_f^T$  to evolve to state  $S_i^T$  [7, p 254].

The final sentence makes sense as the orthodox reversal symmetry: but I am mystified as to how the condition that ‘it is also possible for the state  $S_f^T$  to evolve to state  $S_i^T$ ’ is meant to be symbolized by  $S_i^T \rightarrow S_f^T$ . I think that the use of arrows to symbolize processes, and the habit of suppressing the explicit reference to *time*, should be avoided.

It is also wrong to characterize processes as TRI, as Callender does here: it is theories or laws that have this property. This reflects a failure in Callender’s account to give direct account of time reversal as a transformation on *theories*, or to prove the adequacy of his criterion for TRI from a definition of what time reversal means. But his article nonetheless raises essential points overlooked in most accounts.

It is also noteworthy that most leading writers appeal to the intuitive notion of *reversing a movie of a process* to explicate the meaning of the time reversal transformation. For instance, after describing the concept of space reversal (through the intuitive notion of a mirror), Sachs explains

A similar treatment of time reversal requires introducing a ‘mirror’ of the time variable. A perfectly practical mirror exists for classical mechanical systems. We simply take a motion picture (‘movie’) of the system with high enough resolution (in both space and time) to make measurements of the motion in the projected film. Then by running the film backward we can observe and make measurements on the motion as a function of the reversed time variable,  $t'$  [30, p 21].

It might be thought that this is just an intuitive explication of time reversal, to be backed up by a more rigorous analysis: but Sachs provides nothing more rigorous. And the inadequacy of this explanation is clear. For a start, it is circular, because the concept of reversing a movie *already* requires the notion of time reversal of a physical process—i.e. of *running a movie backwards in time*. It is also a severe exaggeration to claim that making a movie represents a ‘perfectly practical mirror’; first, the resolution of a movie camera is not sufficient to reveal motions on the atomic scale; and how do we make a movie of a *theory*, as opposed to a mere *process*? The ‘movie’ analogy is really just an intuitive ploy to help visualize the meaning of time reversal: but it seems to be used by all writers who support the orthodox analysis.

### *The deterministic criterion*

In fact, close examination shows that even the simple criterion used to define the meaning of time symmetry for *deterministic laws* is incorrect. The principle is usually explained by first defining a simple deterministic law in a form such as *if*  $S_1(t)$  then  $S_2(t + \Delta t)$ . This law means that a system that started in state  $S_1$  at any time  $t$  will evolve to state  $S_2$  at  $t + \Delta t$ . It is then observed that the reversal of such a process will involve a system in the reversed final state  $T S_2$  evolving to the reversed initial state  $T S_1$  at a time  $\Delta t$  later. The corresponding law for this process is *if*  $T S_2(t)$

then  $TS_1(t + \Delta t)$ . Therefore, it is claimed, *time symmetry requires the second deterministic law in symmetry with the first.*

But this intuitive reasoning is already proved wrong by theory A above. This is a time symmetric theory, defined with three states,  $A$ ,  $B$  and  $C$ , and the two *deterministic laws*: *if  $B(t)$  then  $A(t + 1)$* , and *if  $C(t)$  then  $A(t + 1)$* . But these laws are *not* matched by the corresponding laws  *$TA(t)$  then  $TB(t + 1)$* , or  *$TA(t)$  then  $TC(t + 1)$* . On the contrary, theory A has the laws  $\text{prob}(TB(t + 1)|TA(t)) = \text{prob}(TB(t + 1)|TA(t)) = 1/2$ .

The correct criterion for time symmetry shows what is wrong: namely, the time-reversed image of the deterministic *future-directed* law *if  $B(t)$  then  $A(t + 1)$* , is the deterministic *past-directed* law *if  $TB(t)$  then  $TA(t - 1)$* , which in indeed a law in theory A. The time reversal of deterministic transition laws appears as the appropriate special case of the reversal of probabilistic laws with probability of one.

**Necessary condition for deterministic time reversal symmetry.** *A theory that entails a deterministic transition law of the form if  $S_1(t)$  then  $S_2(t + \Delta t)$ , or equally  $\text{prob}(S_2(t + \Delta t)|S_1(t)) = 1$ , can be time reversal symmetric only if it also entails the time-reversed law if  $TS_1(t)$  then  $TS_2(t - \Delta t)$ , or equivalently  $\text{prob}(TS_2(t - \Delta t)|TS_1(t)) = 1$ .*

However, for theories entailing *bilateral determinism*, i.e. full determinism towards both past and future, where every transition probability between fundamental states is either zero or unity, the orthodox criterion and the correct criterion for reversibility are equivalent. Bilateral determinism holds in fundamental theories of classical physics, so that the conventional picture of reversibility is used there without material error. But this does not extend to classical statistical mechanics, or laws treating macro-states, and it is simply not true, as so commonly assumed, that the time reversal of the future-directed law *if  $s_1(t)$  then  $s_2(t + \Delta t)$*  is the future-directed law *if  $TS_2(t)$  then  $TS_1(t + \Delta t)$* .

There is a clear psychological impulse behind the conceptual fallacy described above: this is our penchant to think in *causal* terms, where we always imagine causality going forwards in time. We intuitively associate *cause* and *effect* with *earlier* and *later*, on the one hand, and with the *condition* and the *consequence* in a conditional statement, such as a transition law, on the other hand. We readily describe deterministic laws (which are conditional statements) in terms of future-directed causality (which are cause-and-effects): ‘if we set up such-and-such a system in some state,  $s$ , it must lead to a future state,  $s'$ .’ It is not possible to visualize past-directed *causes* in such a way—‘if we set up such-and-such a system in some state,  $s'$ , it must have originated from a certain past state,  $s$ .’ Past-directed conditional probabilities have no ready visualization in terms of causes, but they may be taken as epistemic lawlike probabilities.

## 7. The meaning of orthodox time reversal

The orthodox criterion for time reversal really represents a combined *time reversal and cause-and-effect exchange*. This is seen by considering the operation that turns the law  $L$  into the law  $L^*$ , or more simply, the deterministic law *if  $S_1(t)$  then  $S_2(t + \Delta t)$* , into the law *if  $TS_2(t)$  then  $TS_1(t + \Delta t)$* . This operation consists of

- (i) applying the (correct) time reversal transformation,  $T$ , to *if  $S_1(t)$  then  $S_2(t + \Delta t)$* , to obtain the law *if  $TS_1(t)$  then  $TS_2(t - \Delta t)$* , and then
- (ii) applying a second *cause-and-effect exchange*, to obtain *if  $TS_2(t - \Delta t)$  then  $TS_1(t)$* .

I will symbolize this second cause-and-effect exchange operation by an operator  $E$ . More generally,  $E$  maps a conditional probability law  $L$  to a new law,  $EL$ , as follows.

**Cause–effect exchange operator,  $E$ .** If  $L$  is the probability law  $(\forall t)\text{prob}(s_2(t + \Delta t)|s_1(t)) = p$ , then  $EL$  is the probability law  $(\forall t)\text{prob}(s_1(t)|s_2(t + \Delta t)) = p$ .

This is not a ‘physically valid transformation’, in the sense that  $E$  does not hold for real laws, and it is not recognized as a general symmetry in physics. (It does hold for deterministic transition laws for classical micro-states, but classical physics is not true, and these are not the real laws of physics. It may hold also for specially defined classes of laws governing systems in thermodynamic equilibrium, e.g. [31], but these are also not the real laws of physics.)

The law  $EL$ , like  $TL$ , is a past-directed probability law, and we will see shortly that  $E$  is what Watanabe calls *retrodictibility*. Most importantly, it must be stressed that  $E$  is not a general transformation, because (unlike  $T$ ) it is not based on any general mapping (automorphism) between fundamental elements of the models. The operator  $E$  is a very specific mapping between special types of proposition. It maps conditional statements in the form  $L$  to the form  $EL$ . But it does not make sense when applied to entities other than the special propositions for which it is defined. Thus we define the  $E$ -image of a probabilistic theory  $\mathbf{T}$  as follows.

**$E$ -image of a theory  $\mathbf{T}$ .** If  $\mathbf{T}$  is a theory defined by a class of fundamental transition laws,  $\{L\}$ , then the image  $E\mathbf{T}$  of  $\mathbf{T}$  is the theory defined by the class of  $E$ -images of those laws: i.e.  $E\mathbf{T} = \{EL: L \text{ is a fundamental law of } \mathbf{T}\}$ .

The theory  $\mathbf{T}$  has  $E$ -symmetry (or is  $E$ -invariant) just in case  $\mathbf{T} = E\mathbf{T}$ .

We find that invariance under the combined operation  $ET$  is orthodox reversal symmetry. But  $ET$  makes sense as a symmetry only when applied to laws involving maximal states, i.e. to transition laws for the fundamental micro-states of a theory.

**$ET$  symmetry is orthodox reversal symmetry.** If  $L$  is any fundamental transition law of a theory, then the theory has orthodox time symmetry, or the orthodox reversibility property, just in case the law  $ET(L)$  is also a law of the theory.

It is critical to realize that, if  $L$  is a law entailed by an orthodox reversible theory, but it is not a fundamental law, then  $ET(L)$  is not generally entailed by the theory, and there is no analogue of proposition 6 for this symmetry. This is evident if we consider applying  $ET$  to laws for macro-states defined as classes of micro-states (coarse-graining).

**Proof.** We define a theory with at least three distinct micro-states  $s_1, s_2$  and  $s_3$ , and their time reversals,  $Ts_1, Ts_2$  and  $Ts_3$ . (There are also other micro-states which we need not consider.) We define a system to be in the macro-state  $S$  just in case it is either in the state  $s_2$  or  $s_3$ . Thus it is in the time-reversed macro-state  $TS$  just in case it is either in the state  $Ts_2$  or  $Ts_3$ . The theory is defined with probability laws  $\text{prob}(s_1(t + 1)|s_2(t)) = \text{prob}(s_1(t + 1)|s_3(t)) = 1/2$ , and  $\text{prob}(Ts_2(t + 1)|Ts_1(t)) = \text{prob}(Ts_3(t + 1)|Ts_1(t)) = 1/2$ . This is consistent with  $ET$ -symmetry, i.e. orthodox reversibility, and we assume that the theory has this symmetry generally. For instance, transforming the first law  $TE(\text{prob}(s_1(t + 1)|s_2(t)) = 1/2)$  gives the law  $\text{prob}(Ts_2(t + 1)|Ts_1(t)) = 1/2$ , which is the third law. However, consider the  $ET$ -transformation on the macro-state law:  $\text{prob}(s_1(t + 1)|S(t)) = 1/2$ . This is also a law of the theory, since it is entailed by the first two probability laws. But the  $ET$ -transformation of this would be  $ET(\text{prob}(s_1(t + 1)|S(t)) = 1/2)$  giving  $\text{prob}(TS(t + 1)|Ts_1(t)) = 1/2$ . However,

this is not a law of the theory. On the contrary, it follows from the second two laws that  $\text{prob}(TS(t+1)|Ts_1(t)) = 1$ .

This shows that the  $ET$  symmetry really only applies sensibly to *micro-states*, and it is simply *not applicable to macro-states*. This point is essential when we consider the orthodox reversibility of thermodynamics, which is defined through such macro-state laws. It is evident that statistical mechanics defined on coarse-grained states *must fail the  $ET$  symmetry (orthodox reversibility)*, for this simple logical reason. But this has no bearing on the *time symmetry* of statistical mechanics. This is a point of common confusion in thermodynamics. Many authors claim that *coarse graining of micro-states* generates time asymmetry or irreversibility from a time-symmetric micro-theory (e.g. [32]). They are correct that coarse-graining generates a class of transition laws which fails *orthodox reversibility symmetry*: but it does not and cannot generate a *time asymmetric theory* from a time symmetric micro-theory.

It is also easily verified that time symmetry, i.e.  $T$ -symmetry, does not suffer this logical deficiency. That is, *if a theory is  $T$ -symmetric w.r.t. all its micro-state laws, then it is also  $T$ -symmetric w.r.t. all its macro-state laws based on the micro-state laws*. In the present example, for instance, the  $T$  transformations of the micro-state laws are  $T(\text{prob}(s_1(t+1)|s_2(t)))$  giving  $\text{prob}(Ts_1(t-1)|Ts_2(t))$  and  $T(\text{prob}(s_1(t+1)|s_3(t)))$  giving  $\text{prob}(Ts_1(t-1)|Ts_3(t))$ . Let us suppose that the theory has  $T$ -symmetry, so these two transformed laws are also entailed by the theory. The  $T$  transformation of the macro-state law is  $T(\text{prob}(s_1(t+1)|S(t)) = 1/2)$  giving  $\text{prob}(Ts_1(t-1)|TS(t)) = 1/2$ ; and this is entailed by the theory also, as a consequence of the  $T$ -symmetry of the micro-state laws. These points are easily proven in a completely general form.

#### *E*-symmetry is Watanabe retrodictibility

The symmetry  $E$  alone represents what Watanabe calls *retrodictibility*. The central point that Watanabe makes may be now reformulated in these terms. While  $ET$  is a physical symmetry of known theories (orthodox reversibility),  $E$  alone is not a physical symmetry of any theories, except for the class of deterministic transition laws for classical micro-states. Watanabe took  $E$  to be directly necessary for time symmetry; from our point of view, the situation is rather that, since  $ET$  symmetry holds, but  $E$  symmetry fails,  $T$  symmetry must also fail.

<b>Symmetry</b>	<b>Name</b>
<i>Invariance under <math>ET</math></i>	<i>Orthodox reversibility</i>
<i>Invariance under <math>T</math></i>	<i>Time reversal symmetry</i>
<i>Invariance under <math>E</math></i>	<i>Cause–effect exchange, or retrodictibility</i>

#### **Simple logical relations**

$EE = I$	<i>Identity</i>
$TT = I$	<i>Identity</i>
$ET = TE$	<i><math>E</math> and <math>T</math> commute</i>
$(ET)T = E(TT) = E$	<i>Associativity</i>
$E(ET) = (EE)T = T$	<i>Associativity</i>
<i><math>E</math> and <math>T</math> entails <math>ET</math></i>	<i>Entailment relations</i>
<i><math>ET</math> and <math>T</math> entails <math>E</math></i>	
<i><math>ET</math> and <math>\sim E</math> entails <math>\sim T</math></i>	

## 8. The time asymmetry of quantum mechanics

I now briefly summarize the application to QM on the assumption that we can treat its probability laws as conditional probabilities. I will comment on this assumption subsequently. To evaluate the time reversal invariance of QM, we must check it against the correct criterion for time symmetry, and it fails because it fails  $E$ , i.e. Watanabe retrodiction. We will take the QM state reversal operator, although denoted simply by  $T$ , to signify the usual Wigner operator<sup>5</sup>,  $T^*$ . We begin with the well known fact that QM (except for K-meson decay) has orthodox time reversal symmetry, represented by the orthodox criterion.

### $ET$ . Orthodox reversibility of QM (holds).

$$\text{prob}(\psi_2(t + dt)|\psi_1(t)) = \text{prob}(T\psi_1(t + dt)|T\psi_2(t)) = |(\psi_1, \psi_2)|^2.$$

This obtains because  $|(\psi_1, \psi_2)|^2 = |(T\psi_2, T\psi_1)|^2$ . But we must evaluate whether QM has the property:

### $T$ . Correct criterion for time symmetry of QM (fails).

$$\text{prob}(\psi_2(t + dt)|\psi_1(t)) = \text{prob}(T\psi_2(t - dt)|T\psi_1(t)) = |(\psi_1, \psi_2)|^2.$$

The fact that this property *does not obtain* is shown in detail by ([1, part 3], [3, 5]). He argues correctly that  $E$  is a necessary condition for time symmetry.

### $E$ . Watanabe's retrodictibility condition for QM (fails).

$$\text{prob}(\psi_2(t + dt)|\psi_1(t)) = \text{prob}(\psi_1(t - dt)|\psi_2(t)).$$

Watanabe calls this failure the *irretrodictability of QM*. Because  $ET$  symmetry holds in QM, if  $E$  symmetry fails, then  $T$  symmetry must also fail. We have already observed this logical relation, but it is derived directly in this case simply by substituting the central term in the orthodox criterion above for the left-hand term of the retrodictibility condition, to give

$$\text{prob}(T\psi_1(t + dt)|T\psi_2(t)) = \text{prob}(\psi_1(t - dt)|\psi_2(t)).$$

And this is just the correct criterion in a rearranged form.

Now QM as ordinarily formulated explicitly postulates future-directed probability laws, but does not postulate past-directed probabilities. Hence, *prima facie*, the retrodictibility condition is not satisfied by QM. But it may be thought that either QM implicitly entails the existence of past-directed probabilities, which satisfy retrodictibility, or that QM could be simply extended to entail such past-directed probabilities. Watanabe shows that the simple addition of past-directed probabilities in QM is inconsistent with empirical observations.

If the retrodictibility condition were true, then it would be valid to *retrodict* an earlier state  $\psi_1$  from a later state  $\psi_2$  with exactly the same probability that one could *predict*  $\psi_2$  from  $\psi_1$ . But while the latter probabilities are nomic or lawlike, and cannot be controlled in any way, the simple empirically well known possibility of *controlling the initial states of systems independently of*

<sup>5</sup> The usual claim that QM is time symmetric or reversible has two parts: first, that the probabilistic laws of QM form a time symmetric theory, and second, that the deterministic equation of motion of the wavefunction (or the Schrödinger time-dependent equation) is time symmetric. While I only consider the probabilistic laws here, the second claim is also problematic; e.g. see [7, 12]. This turns on the problem of the choice of time reversal operator for QM states. The orthodox view of this is assumed here.

*their final states*, as assumed in ordinary experiments, immediately rules out the possibility of lawlike retrodictive probabilities. Watanabe shows that the future-directed probabilities of QM reflect the *uncontrollability of future states w.r.t. initial states*, while the *controllability of initial states* reflects the failure of past-directed probabilities; see e.g. [3, section 8, p 156–7].

The same point is also made by observing that the past-directed probabilities in symmetry with future-directed probabilities entail boundary conditions on the absolute probabilities, amounting to a requirement of *general equilibrium*. We begin with the definition of conditional probabilities:  $\text{prob}(b|a) = \text{prob}(a\&b)/\text{prob}(a)$ , and  $\text{prob}(a|b) = \text{prob}(a\&b)/\text{prob}(b)$  (where  $a$  and  $b$  refer to micro-states). Hence, if  $\text{prob}(b|a) = \text{prob}(a|b)$ , then  $\text{prob}(a\&b)/\text{prob}(a) = \text{prob}(a\&b)/\text{prob}(b)$ , and it follows immediately that  $\text{prob}(a) = \text{prob}(b)$ . Thus it means that *every micro-state of a system is a priori equally likely*. But this can only be true for systems that are expected to be in *thermodynamic equilibrium*. Thus, maintaining time reversal invariance entails that *every physical system is expected to be in thermodynamic equilibrium*. In a disequilibrium universe like our own, this is openly contradicted.

The main problem that remains is that we have not yet taken into account an extra condition required for the future-directed QM probabilities to come into play. In fact, the assumption about this is partly sensitive to the interpretation of QM. The most basic view is that probabilistic transitions (or state vector reductions) are *conditional on appropriate types of measurement being made*. For instance, suppose we perform an experiment where we take 1000 systems in an initial state  $\psi_1$ , and then *measure the spin on the x-axis*. We find all the systems in the state  $\psi_1$ . This reflects the probability law that  $\text{prob}(\psi_1(t + dt)|\psi_1(t)) = 1$ . But this law only comes into application *given the extra condition that we measure the state  $\psi_1$*  (or at least, impose conditions that the state vector is reduced with this as an eigenstate).

We may tackle this by adding a *measurement condition* as an explicit part of the ordinary probability laws. We take the term *measure –  $\psi_1$  – at –  $t + dt/2$*  to indicate the condition that *the only interference on the QM system between time  $t$  and  $t + dt$  is a measurement having  $\psi_1$  as an eigenstate, which is made at the intermediate time  $t + dt/2$* . Our probability law is then filled out to

$$\text{prob}(\psi_1(t + dt)|\psi_1(t) \text{ and } [\text{measure} - \psi_1 - \text{at} - t + dt/2]) = 1.$$

And we should repeat the arguments using these elaborated probability laws. See [7] for a similar approach.

Some interpretations of QM assume that probabilistic transitions occur ‘autonomously’, without requiring explicit or deliberate ‘measurements’. For example, we may suppose that radioactive decays occur autonomously in nature, without any human or conscious agents to measure the results. This is the ‘Schrödinger cat’ problem, and involves the deep controversy about the conditions necessary for quantum probabilities to be ‘actualized’, or for wavefunction collapse to occur. We can hardly be entirely confident about our understanding of quantum probabilities while this issue is unresolved, and there remains a cloud over the argument given here because of this. But the result appears robust simply on the weak assumption that quantum probabilities are fundamental and irreducible, and *are actualized somehow*.

Callender [7] observes that if we take some deterministic interpretation of QM, such as Bohm proposes, then the QM probabilities cease to be fundamental, and there is a way out of this conclusion. But this shows that there are two different theories: Bohm’s deterministic version (as well as far from complete) is not standard QM, and has different time symmetry properties. In standard QM, probabilities are fundamental and irreducible, and on this assumption, by applying the correct criterion for time symmetry, I think we must conclude that QM *as we know it* is time

asymmetric in its probabilistic laws. We should reserve some scepticism about whether this is dependent on our interpretation of QM probabilities, but a more general discussion of that topic is beyond the scope of this paper.

### 9. The *de facto* irreversibility view and thermodynamic asymmetry

I think that the case that the orthodox criterion for reversibility does not represent TRI is indisputable, and the ‘correct criterion for TRI’ stated in section 3 is indeed the relevant condition. This means that the usual analysis of TRI of QM is mistaken. And Watanabe’s arguments show that ordinary QM as we know it fails the TRI condition, because ordinary QM fails to provide any past-directed probabilities.

However, the interpretation of this result still has a number of complications, and the ultimate source or cause of *irreversibility* remains deeply problematic. While it is beyond the scope of this paper to deal with the ramifications in detail, I will comment here on some important points, and try to pre-empt some likely misunderstandings of the situation.

The most important issue involves the orthodox explanation of irreversibility in the QM measurement process as a ‘merely *de facto*’ irreversibility, rather than a lawlike irreversibility. Many writers have noticed that the QM measurement process, or state vector reduction, appears irreversible, but most have concluded that QM is nonetheless time symmetric ‘in principle’. Their primary reason for this is that they have mistaken orthodox reversibility for TRI, but we have seen this is a mistake. However, they have also provided an explanation of the source of irreversibility, and this needs to be addressed, because it will appear to many writers that this explanation is still viable, and undercuts the failure of TRI that we have seen.

The orthodox explanation is that the irreversibility or time asymmetry of state vector reduction merely reflects a *de facto* thermodynamic asymmetry of the universe. The argument goes that ordinary state vector reduction is asymmetric in time, but this merely results from the thermodynamic disequilibrium of the ordinary systems we deal with. If we dealt with systems in equilibrium, we would find that a time symmetric version of QM applies. (For example, Aharonov *et al* [33] and Cocke [34] investigate a time symmetric formalism in response to Watanabe’s arguments.) But (the argument goes) the thermodynamic disequilibrium state of the universe is merely *de facto*, or contingent. An equilibrium universe is perfectly conceivable, and a completely time-reversed universe is also conceivable. Hence, the irreversibility seen in QM processes is merely a contingent or *de facto* feature of *actual* processes, not a lawlike feature, and it fails to reflect a lawlike time asymmetry of the fundamental laws of QM.

This kind of argument contains a number of mistakes that need to be dismissed, and I will summarize some important points.

- (1) First, it is true that *a version of QM constrained to apply only to systems in equilibrium may be time symmetric*. The argument in section 8 shows in fact that a time symmetric version of QM (formed by adding past-directed laws in symmetry with future-directed laws to satisfy TRI) can *only* apply to systems in equilibrium. But any such version of QM is severely constrained; it is not real QM. It does not describe real systems, which are generally not in equilibrium. This point is simply a *non sequitur*, because *it is not about real QM*, merely about an artificially constrained version. (It is like saying that thermodynamics is time symmetric because a version of thermodynamics for systems in equilibrium is time symmetric.)
- (2) Real QM fails the criterion for TRI, and this is the essential condition we must come back to. Given that the criterion for TRI formulated in this paper is correct, any positive case that

QM is TRI must be able to show that QM satisfies the correct criterion. The ‘*de facto* irreversibility’ explanation fails to show this. It is really just an informal sketch of an argument, and its inaccuracies are hidden by vagueness.

- (3) The notion that ‘a time-reversed universe is conceivable’ is misleading. Of course, such a universe is *conceivable*; but the question is whether such a universe would *obey the ordinary laws of QM*. It would not, because it would contradict the future-directed laws of ordinary QM (see below). By definition, it would obey *the time-reversed laws of QM*, and this is really what many writers seem to imagine: that we reverse the process of the universe, and this obeys a *time-reversed quantum theory*; and this shows that a time-reversed version of QM is coherent. But this conclusion is misguided.

Obviously, if we reverse the ordinary future-directed laws, and impose them towards the past instead, then the time-reversed universe would obey these, given that the original universe obeys ordinary QM. This is just a logical fact. But what is required for time symmetry of QM is that the time-reversed universe obeys the *original laws*, with the original future-directed probabilities. The fact that it does not is exactly what is proved by the arguments in section 8. It would require that QM probabilities apply towards the past in symmetry with the way they apply towards the future, but they do not.

- (4) This illustrates a typical kind of fallacy common in discussions of time reversal. Another example is to reason that we only define the ‘future direction’ of time by reference to irreversible processes, and then conclude that if we were in the time reversal of our universe, we would define ‘the future’ and ‘the past’ in the opposite directions that we currently define them, and then we would define a version of QM in this universe with the time directions exchanged, and this would be perfectly adequate. How then could we possibly have evidence of whether we are in the QM universe or its time reversal? But without evidence (the argument goes) the distinction between the two is meaningless.

This argument is a fallacy: it fails to show that QM is TRI, or that QM defined in our universe applies correctly to the time-reversed universe. On the contrary, it reveals the typical logical fallacy inherent in the ‘positivist–empiricist’ theory of meaning, which is seen as follows. Suppose that we have a time asymmetric theory, **T**, which applies to our universe. We can construct a weaker theory: **T or TT** i.e. the disjunction of **T** and its time reversal, **TT**. This weaker theory applies to both our universe and its time reversal, and it is obviously time symmetric. The argument above leads us to conclude that any theory, **T**, is ‘empirically equivalent’ to the disjunction: **T or TT**. Now they are obviously not logically equivalent, because (**T or TT**) does not entail **T**. And if we supposed that they are logically equivalent—or have identical meanings—this would remove any possibility of ever establishing a time asymmetric theory! Yet their equivalence appears justified by the positivist–empiricist line of reasoning. Certainly, **T** entails (**T or TT**), and any observational evidence for the theory **T** is also evidence for (**T or TT**). On the other hand, can we have evidence that **T** is true, whereas **TT** is not true, so that we can establish the stronger theory **T** by itself? The empiricist–positivist account makes **T** and **TT** appear indistinguishable: for how can we distinguish whether we are *really* in a **T**-universe, or in a **TT**-universe? Assuming we cannot, we are led to take **T** to be equivalent to **T or TT**. But this is not a valid argument: instead it is an illustration of a typical kind of fallacy inherent in positivist–empiricist conceptions of meaning. Such fallacies have deeply infected the subject, and they are only properly dispelled by tackling modern semantics seriously.



- (5) The cause of the early low entropy state of the universe following the Big Bang remains the major issue in the subject. This boundary condition appears to be the cause of the wealth of irreversible processes in the present universe. The most popular view is that this boundary condition is merely known as a *de facto* condition, and that it does not have a lawlike cause (although I will argue below that this is really unknown). However, even if we assume this is a merely *de facto* condition, this does not entail that the time asymmetry of QM is merely '*de facto*', and therefore not an asymmetry of the *laws of QM*. It does not prevent QM from being a *time asymmetric theory*. It can only mean that the time asymmetry of QM does not explain or entail the thermodynamic asymmetry. The point is clarified by comparison with the time asymmetry of classical thermodynamics. There is no doubt that thermodynamics is a *time asymmetric theory*: its laws describe irreversible processes, but not their time reversals. At the same time, in the classical context, the ultimate source of this asymmetry is normally inferred to be merely contingent, or to reflect a *de facto* condition, because the classical laws are firmly established to be time symmetric. Yet this does not mean that classical thermodynamics is consequently a *time symmetric theory*: its laws remain indisputably asymmetric! Rather, we infer that classical thermodynamics is not a *fundamental theory of the laws of nature*, because it contains some propositions which reflect merely 'contingent facts' about our particular universe. Similarly, we might well conclude that QM is not a complete theory, or that it is not a truly fundamental theory: but as long as it fails to satisfy the criterion for TRI, it is a time asymmetric theory, independently of the ultimate cause of irreversibility in our universe.
- (6) Despite this comparison, there is a critical difference between the situation in classical and quantum thermodynamics. Returning to point (3), let us visualize the universe reversed in time, with the low entropy boundary condition in the future (instead of the past), and ask whether QM *as we know it* would correctly describe this universe. It could not possibly describe it, because it would require systems to undergo bizarre state transitions which contradict QM probabilities based on present states. Now it might describe it correctly if we modified QM by adding the *future boundary condition* as an additional conditionalization of the usual QM probabilities. This is possible in classical thermodynamics because the statistical probabilities there are merely *epistemic conditional probabilities*, based on our ignorance of precise micro-states corresponding to observed macro-states, and we can add extra *de facto* conditions as we wish, without contradicting the original conditional probabilities.
- But it is not possible in QM, because QM transition probabilities between maximal states are irreducible and maximal. By definition of the quantum theory, they depend on *the present state of a system*, not on additional conditions w.r.t. future states. If we add future boundary conditions that *contradict the results we would expect on the basis of QM transition probabilities alone*, then we have contradicted those probabilities and contradicted QM.
- (7) This brings out a crucial difference between the theoretical possibility of reversing classical and quantum systems. In a classical universe, we can imagine taking the exact time reversal of the present micro-state of the universe, and reason that it will deterministically lead back to the original low entropy boundary condition. But in QM, if we start with the exact time reversal of the present micro-state, and let it develop via QM with probabilistic

transitions, we will *not* get back to the original state, because of the chaos introduced by the probabilities.

The reversed classical universe can retrace the original process *precisely* because it is *precisely arranged* to encode its past, and because it is deterministic. But in the reversed QM system, given there are probabilistic transitions, some of these will occur differently than in the original development, and this will rapidly upset the time-reversed process. Note that this does not hold if we demand a purely deterministic quantum mechanical evolution of the wavefunction for the universe, by the (reversible) Schrödinger equation: but as long as we maintain objective and irreducibly probabilistic events in QM, it becomes impossible in principle to generate the time reversals of complex processes.

- (8) The same consideration also shows that, even if we took the exact time reversal of the present state of the QM universe, we could never generate *time-reversed thermodynamics*, as we would do in a reversed classical universe. Instead, if we started from the time-reversed present QM state, and applied a probabilistic development: (i) thermodynamic processes would run ‘backwards’ for a short period, (e.g. rivers would run up-hill for a few moments); but then (ii) states would become probabilistically ‘scrambled’, and there would be a period of chaotic confusion, when these thermodynamic systems reorganized their behaviour (e.g. rivers would go into a turmoil, losing their coherent motion up-hill, and then begin to turn down-hill again); and then (iii) thermodynamic directionality would reassert itself in the future direction as normal (rivers would thereafter run down-hill as normal).
- (9) This kind of consideration supports Watanabe’s contention that the time asymmetry of QM is sufficient by itself to explain at least one major aspect of thermodynamic directionality: the fact that complex quantum systems in disequilibrium must become increasingly disordered towards the future. The future behaviour is necessitated by the intrinsic probabilities of QM, rather than depending on any additional assumptions that micro-states are ‘random’ w.r.t. the future, as in classical statistical mechanics. Watanabe held that the time asymmetry of QM makes a decisive difference to the problem of irreversibility, e.g.

The reason (for the phenomenological one-way-ness of temporal developments) is, as we shall presently see, that quantum physics is basically irretrodictable, both microscopically and macroscopically, whether or not it obeys reversibility or any other similar invariance law. It is precisely irretrodictability which is related to phenomenological one-way-ness [4, section 8, p 156].

Watanabe’s results show that the ‘reversibility paradox’ of classical thermodynamics is removed when we turn to QM thermodynamics. We no longer have the paradox of obtaining a time asymmetric macro-theory from a time symmetric micro-theory, because QM is simply not time symmetric. We obtain the future-directed laws of thermodynamics from QM without introducing any additional assumptions about ‘random distributions of micro-states’ employed by Boltzmann, because quantum probabilities provide an inescapable ‘chaos’ in the micro-states adequate to ensure future entropy increase.

- (10) But this does not mean that QM time asymmetry *explains irreversibility* in the broader sense of explaining the predominance of irreversible processes in the universe, because it does not explain the reason for the low entropy of the universe in the past. This is evident because QM as we know it is compatible with a universe in thermodynamic equilibrium, and hence does not *entail* the predominance of irreversible processes. No explanation of this feature seems possible without a more complete theory of the cosmological origins of the universe, including a theory combining gravity and QM. Different cosmological theories will give different kinds of explanation, and until we have a convincing general cosmological theory I think it is premature to judge whether irreversibility has a contingent cause or is a fundamental lawlike feature of the universe.
- (11) We can also note that the time asymmetry of QM does not necessarily show that the *fundamental laws of nature are time asymmetric*, because QM is not a complete theory of fundamental physics, and may not be a fundamental theory. It needs at least to be combined with theories of gravity and cosmology. A complete theory might still turn out to be time symmetric.
- (12) But we should be aware that orthodox reversibility symmetry does not represent this goal, and the time asymmetry of QM undermines the usual justification physicists have for maintaining that TRI is a fundamental symmetry of nature. The reason usually given is that TRI holds for (almost) all known theories of fundamental micro-physics, but this now turns out to be incorrect for QM, which is the most important theory. Judging by QM alone, TRI should now be expected to fail in fundamental physics. Physicists and philosophers may well continue to pursue time symmetry as a fundamental symmetry in seeking to develop a more complete theory. But the apparent empirical basis for this seems to have evaporated, and it would now seem to rest on a ‘metaphysical intuition’, of a kind generally rejected by empiricists in other contexts. On the other hand, the way should be open to contemplating *irreversible* cosmological theories, which distinguish the future from the past.
- (13) Finally, I note that the failure of QM to provide any past-directed probabilities threatens to leave us with a major gap in the laws of nature. I will call this the *QM retrodiction problem*: how can we consistently extend QM to include lawlike past-directed epistemic probabilities? How can we validly retrodict in a quantum mechanical world unless QM provides some lawlike past-directed epistemic probabilities?
- We surely need some lawlike basis to support ordinary retrodiction, if we are to take our ordinary information about the past seriously. For information is encoded in the *present* physical states of systems. We only *infer* facts about the past from these systems. Grünbaum, de Beaugregard, and others hold that we need to support retrodiction generally by assuming a *de facto* condition of a low entropy early state of the universe; but surely we cannot assume this to start with without some lawlike basis for inferring this early state from presently observed conditions. What is the physical basis for such inferences if QM alone provides the fundamental laws? We cannot use Bayesian probability theory with ‘blind retrodiction’, because this generates (time symmetric) past probabilities that the real past fails to satisfy—it is inconsistent with the *de facto* boundary condition we want to infer. The same problem also appears intractable in classical thermodynamics. I agree with Price [20] that this seems impossible to solve without some decisive advance in theoretical physics, and it represents a deep problem in the foundations of our physical account of the universe.

## 10. Summary

Time reversal invariance,  $T$ , for probabilistic theories, requires a strong symmetry between past- and future-directed laws. Orthodox reversal symmetry,  $ET$ , is a quite different symmetry, which is a combination of  $T$  and  $E$ , where  $E$  is a *cause-and-effect* reversal between two fundamental states.  $ET$  only pertains to future-directed laws.  $T$  is based on a general underlying transformation  $T : t \rightarrow -t$ , but  $E$  is not a true general symmetry, and is not based on an underlying general transformation. It is relative to a special class of states or propositions. It represents a *probabilistic cause-and-effect exchange*. The orthodox account begins by mistakenly adopting  $ET$  in the treatment of time reversal of deterministic transition laws.

This conceptual flaw in the orthodox account undermines our confidence in many of the accepted conclusions in the subject of physical time directionality. The formal  $T$ -operator in physics needs a closer logical examination, and an intensional semantic analysis is recommended to establish the foundations on a more secure footing.

Contrary to accepted opinion, QM is found to fail time reversal invariance in a very powerful way. This is shown through the failure of retrodictibility, and is a consequence of ‘irreversibility’, i.e. of the existence of the wealth of ‘irreversible processes’ evident in our universe, as Watanabe showed. However, the failure of TRI in QM does not explain irreversibility by itself. Irretrodictibility provides the evidence for the failure of time symmetry, and is a consequence of thermodynamic irreversibility. But the time asymmetry of QM does not provide any direct explanation for the peculiar entropy state of our universe in the first place, the physical feature responsible for the predominance of irreversible processes.

The key problem in this regard is to explain the apparent low entropy state of the early universe; and this is an open question, which requires advances in our understanding of cosmological theory, rather than being derivable directly from QM or the time asymmetry of micro-processes. In the meantime, the question of whether this is a contingent (or *de facto*), rather than a lawlike (or *nomically necessary*), feature of our universe remains open. But even if a *de facto* cause is eventually agreed, this does not mean that the time asymmetry of QM is a ‘*de facto*’ asymmetry. The failure of TRI for QM is an asymmetry of the probabilistic laws of QM.

We are also faced with the closely related problem of extending physics to provide lawlike past-directed epistemological probabilities that can serve as a realistic physical basis for retrodiction.

## Acknowledgments

Thanks go to Professor Graham Oddie and the late Pavel Tichý for many invaluable discussions of points of logic and semantics. Thanks to the *NJP* referees of earlier versions of this paper for a number of valuable comments and additional references.

## References

- [1] Watanabe S 1955 Symmetry of physical laws. Part 1 and 2. Q-number theory of space–time inversions and charge conjugation *Rev. Mod. Phys.* **27** 40–76
- [2] Watanabe S 1955 Symmetry of physical laws. Part 3. Prediction and retrodiction *Rev. Mod. Phys.* **27** 179–86
- [3] Watanabe S 1965 Conditional probability in physics *Prog. Theor. Phys. Suppl.* 135–67
- [4] Watanabe S 1966 Time and the probabilistic view of the world *The Voices of Time* ed J T Frazer (Berlin: Springer)

- [5] Watanabe S 1972 Creative time *The Study of Time* ed J T Frazer (Berlin: Springer)
- [6] Healey R 1981 Statistical theories, quantum mechanics and the directedness of time *Reduction, Time and Reality* ed R Healey (Cambridge: Cambridge University Press) pp 99–127
- [7] Callender C 2000 Is time handed in a quantum world? *Proc. Arist. Soc.* **121** 247–69
- [8] Reichenbach H 1957 *The Direction of Time* (Berkeley, CA: University of California Press)
- [9] Grunbaum A 1973 *Philosophical Problems of Space and Time* (Dordrecht: Reidel)
- [10] de Beaugregard O C 1972 No paradox in the theory of time asymmetry *The Study of Time* ed J T Frazer *et al* (Berlin: Springer)
- [11] de Beaugregard O C 1977 Two lectures on the direction of time *Synthese* **35** 129–54
- [12] de Beaugregard O C 1980 CPT invariance and interpretation of quantum mechanics *Found. Phys.* **10** 513–31
- [13] de Beaugregard O C 1987 *Time, the Physical Magnitude* (Dordrecht: Reidel)
- [14] Earman J 1969 The anisotropy of time *Aust. J. Phil.* **47** 273–95
- [15] Earman J 1974 An attempt to add a little direction to the problem of the direction of time *Phil. Sci.* **41** 15–47
- [16] Earman J 1967 Irreversibility and temporal asymmetry *J. Phil.* LXIV **18** 543–9
- [17] Earman J 1986 *A Primer on Determinism* (Dordrecht: Reidel)
- [18] McCall S 1976 Objective time flow *Phil. Sci.* **43** 337–62
- [19] Penrose R 1989 *The Emperor's New Mind* (Oxford: Oxford University Press)
- [20] Price H 1996 *Time's Arrow and Archimedes' Point* (Oxford: Oxford University Press)
- [21] Boltzmann L 1964 *Lectures on Gas Theory* (Berkeley, CA: University of California Press) pp 51–62, 441–52 (Transl. S G Brush)
- [22] Davies P C W 1974 *The Physics of Time Asymmetry* (Surrey University Press)
- [23] Schrödinger E 1950 Irreversibility *Proc. R. Ir. Acad.* **53** 189–95
- [24] von Neumann J 1932 *Mathematical Foundations of Quantum Mechanics* (Princeton, NJ: Princeton University Press) (Transl. 1955 by R T Beyer)
- [25] Montague R and Thomason R H (ed) 1974 *Selected Papers of Richard Montague* (New Haven, CT: Yale University Press)
- [26] Tichý P 1971 An approach to intensional analysis *Nous* **5** 273–97
- [27] Partee B H and Hendricks H L W 1997 Montague grammar *Handbook of Logic and Language* (Cambridge, MA: MIT Press)
- [28] Janssen T 1997 Compositionality *Handbook of Logic and Language* (Cambridge, MA: MIT Press)
- [29] Sklar L 1974 *Space, Time, and Spacetime* (Berkeley, CA: University of California Press)
- [30] Sachs R G 1987 *The Physics of Time Reversal* (Chicago, IL: University of Chicago Press)
- [31] Lewis G N 1932 The symmetry of time and physics *The Enigma of Time* ed Landsberg (Bristol: Hilger)
- [32] Hollinger H B and Zenzen M J 1985 *The Nature of Irreversibility* (Dordrecht: Reidel)
- [33] Aharonov Y, Bergmann P and Lebowitz J 1964 Time asymmetry in quantum process of measurement *Phys. Rev. B* **134** 1410–8
- [34] Cocke W J 1967 Statistical time symmetry and two-time boundary conditions in physics and cosmology *Phys. Rev.* **160** 1165–70
- [35] Frazer J T (ed) 1966 *The Voices of Time* (Berlin: Springer)
- [36] Frazer J T (ed) 1972 *The Study of Time* (Berlin: Springer)
- [37] Landsberg 1982 *The Enigma of Time* (Bristol: Hilger)
- [38] van Benthem J and ter Meulen A 1997 *Handbook of Logic and Language* (Cambridge, MA: MIT Press)