

Copyright is owned by the Author of the thesis. Permission is given for a copy to be downloaded by an individual for the purpose of research and private study only. The thesis may not be reproduced elsewhere without the permission of the Author.

**Time Flow And Reversibility
in a Probabilistic Universe**

**A thesis presented in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy
in Philosophy
at Massey University.**

Andrew Thomas Holster

1990

ABSTRACT

A fundamental problem in understanding the nature of time is to explain its 'directionality'. The commonplace view is that this directionality is provided by the 'flow of time'. Unfortunately this concept of 'time flow', which seems to make perfect sense to us in our everyday lives, has resisted philosophical and scientific analysis so well that today it is widely regarded as having no place in the scientific account of the world. Instead, various alternative physical concepts of the directionality of time have been developed, principally the notions of the time reversibility of physical laws or theories, and of the time asymmetry of physical processes. It is frequently argued by philosophers of physics that the scientific account of the directionality of time must be framed entirely in terms of these physical notions.

The thesis of the present work is that this conclusion has been reached far too hastily. It is argued that the concept of time flow is a legitimate physical concept, and furthermore, that time flow plays a real part in quantum theory.

A number of conceptual investigations are necessary to support this argument. Firstly, it is necessary to give an analysis

of what a physical theory of time flow might be like, and how it might be empirically established. This is given in Chapter One, which at the same time is an overview of the results of later chapters. It is found in Chapter One that the concept of physical time flow has an important connection with the concept of time reversibility, which makes it necessary to give an analysis of this notion. A detailed discussion of reversibility and time symmetry is given in Chapters Two to Five. Here it is demonstrated that the orthodox analysis of the reversibility of probabilistic theories is flawed. This conclusion allows it to be shown, in Chapter Six, that, contrary to current scientific belief, quantum theory is profoundly irreversible.

This result, together with the argument of Chapter One, allows a strong *prima facie* case for an interpretation of quantum probabilities as involving time flow to be given. However, because of the traditional problems with the notion of time flow, for this interpretation to become respectable it needs to be demonstrated that it is possible to construct a formal model of a physical ontology in which time flow can be represented. This is undertaken in Chapter Seven. In Chapter Eight, various points about the role of probabilities in quantum theory are discussed. Finally, in Chapter Nine, the implications of relativity theory for the proposed theory of time flow are considered. It is found that relativity theory poses a serious problem for a physical theory of time flow, but the implications of relativity theory for the proposed interpretation of quantum probabilities is not clear because of deeper foundational problems with quantum theory.

ACKNOWLEDGEMENTS

It is first of all a great pleasure to thank my supervisor, Graham Oddie, for all his help over the past three years. His many acts of kindness, hours of discussion of my research, and his patient criticism and advice throughout the preparation of this thesis have been invaluable. Special thanks also go to Pavel Tichy, of the University of Otago, who helped to supervise this project in its first year. It is also a pleasure to thank Kirsten McKay, Peter Milne, Roy Perrett, Bruce Small, and Jack Smart for valuable discussions and encouragement. Acknowledgement is also made of receipt of a UGC Postgraduate Scholarship.

CONTENTS

| | Page |
|---|----------|
| Abstract | ii |
| Acknowledgements | iv |
| | |
| Chapter One An Overview. | 1 |
| | |
| 1.1 The commonplace conception of time: IT FLOWS. | 2 |
| 1.2 The commonplace conception II: time flow in the physical world. | 3 |
| 1.3 Phenomenological directedness. | 4 |
| 1.4 Terminology. | 6 |
| 1.5 Against time flow: Illusionism. | 7 |
| 1.6 The scientific conception of time. | 8 |
| 1.7 Does scientific time flow? | 11 |
| 1.8 Reasons for [1]: inspection of the theory. | 13 |
| 1.9 Reasons for [1]: reversibility of F. | 16 |
| 1.10 Four major views on physical time flow. | 18 |
| 1.11 Do scientific theories prohibit time flow? | 21 |
| 1.12 A theory of objective physical time flow. | 24 |
| 1.13 The irreversibility of quantum theory. | 26 |
| 1.14 Quantum probabilism. | 27 |
| 1.15 The irreversible feature of quantum theory. | 28 |
| 1.16 The criterion for the reversibility of probabilistic laws. | 30 |
| 1.17 The correct criterion for reversibility: [CPR]. | 32 |
| 1.18 The lack of past-directed generic probabilities. | 33 |
| 1.19 Reversibility, thermodynamics, and phenomenological directedness. | 39 |
| 1.20 A proposal: time flow in quantum theory. | 42 |

| | | |
|----------------------|---|------------|
| 1.21 | A contingent identity theory of time flow. | 44 |
| 1.22 | The idea of a dynamic theory. | 47 |
| 1.23 | Dynamic probability. | 49 |
| 1.24 | The motivation for realism about time flow. | 56 |
| 1.25 | The concept of existence 1. | 59 |
| 1.26 | The concept of existence 2. | 63 |
| 1.27 | An argument for the reality of time flow. | 65 |
| 1.28 | Realism vs. Illusionism about the physical world. | 68 |
| 1.29 | The analogous argument for time flow. | 75 |
| 1.30 | Conclusion. | 86 |
| | | |
| Chapter Two | Phenomenological Directedness | 87 |
| 2.1 | Phenomenological directedness. | 87 |
| 2.2 | Time reversibility. | 90 |
| 2.3 | Explaining phenomenological directedness. | 91 |
| 2.4 | Explaining phenomenological directedness in the context of a reversible fundamental theory. | 92 |
| 2.5 | Explaining phenomenological directedness in the context of an irreversible fundamental theory. | 99 |
| 2.6. | General form of the explanation of temporal directedness. | 101 |
| | | |
| Chapter Three | Reversibility and Time Symmetry | 107 |
| 3.1 | Directional symmetries. | 109 |
| 3.2 | The directional symmetry of time. | 113 |
| 3.3 | Equivalence of the two definitions of reversibility. | 115 |

| | | |
|---------------------|---|------------|
| 3.4 | Reversibility and time symmetry. | 118 |
| 3.5 | Time flow and Newton's scholium. | 121 |
| 3.6 | A logical puzzle about reversibility. | 123 |
| 3.7 | Time flow and time symmetry 2. | 126 |
| 3.8 | The direction of time flow cannot be contingent. | 130 |
| 3.9 | Some basic theorems. | 135 |
| 3.10 | Summary. | 137 |
| 3.11 | M. Bunge on reversibility. | 138 |
| Chapter Four | Time Reversal Operators | 148 |
| 4.1 | The physicist's definition and the syntactic reversal operator. | 149 |
| 4.2 | Terminology: tokens and types of states and processes. | 154 |
| 4.3 | The metrisation of time. | 156 |
| 4.4 | Symmetries. | 160 |
| 4.5 | The time-reversal transformation. | 162 |
| 4.6 | Time reversal of states. | 163 |
| 4.7 | Return to the syntactic reversal operator. | 168 |
| 4.8 | Time reversals. | 170 |
| 4.9 | The syntactic time reversal operation: some examples. | 174 |

| | | |
|---------------------|---|------------|
| Chapter Five | The Criterion for Probabilistic Reversal. | 178 |
| 5.1 | The CPR | 179 |
| 5.2 | The PPMR | 180 |
| 5.3 | CPR 1: A way of picturing the time reversal of probabilities. | 182 |
| 5.4 | CPR 2: Model-theoretic representation of probabilities. | 183 |
| 5.5 | CPR 3: A statistical picture. | 185 |
| 5.6 | Previous recognition of the CPR | 186 |
| 5.7 | Failure of the PPMR as a criterion for reversibility. | 199 |
| 5.8 | A flaw in the interpretation of reversibility for deterministic laws. | 201 |
| 5.9 | Failure of the PPMR. | 208 |
| | | |
| Chapter Six | The Irreversibility of Quantum Theory | 214 |
| 6.1 | The lack of nomological past-directed probabilities. | 215 |
| 6.2 | Objection 1: an accident. | 224 |
| 6.3 | Objection 2: A biased sample? | 226 |
| 6.4 | Is Experiment 1 of the wrong type? | 228 |
| 6.5 | Objection 3: anthropomorphic bias? | 231 |
| 6.6 | Objection 4: long-term equilibrium? | 232 |
| 6.7 | Epistemic past-directed probabilities. | 235 |

| | | |
|----------------------|--|------------|
| Chapter Seven | A Dynamic Model of Time. | 238 |
| 7.1 | Preliminary sketch of the model. | 239 |
| 7.2 | McCall's dynamic model. | 247 |
| 7.3 | Past/future in the temporal sequence versus earlier/later in the universe-tree. | 255 |
| 7.4 | Temporal sequences of universe-trees. | 258 |
| 7.5 | Semantics of tensed propositions. | 263 |
| 7.6 | The present moment. | 266 |
| 7.7 | Modification of McCall's theory: the possibilistic universe-tree. | 272 |
| 7.8 | The semantics of untensed propositions. | 278 |
| 7.9 | The probabilistic universe-tree. | 285 |
| 7.10 | The need for temporally-extended universe-trees. | 291 |
| 7.11 | Some concluding comments. | 293 |
| Chapter Eight | The Role of Probabilities in Quantum Theory. | 296 |
| 8.1 | The objectivity of quantum probabilities. | 297 |
| 8.2 | The nature of quantum probabilities. | 305 |

| | | |
|---------------------|--|------------|
| Chapter Nine | Problems with Relativity Theory. | 310 |
| 9.1 | The 'metaphysical' postulation of simultaneity relations. | 311 |
| 9.2 | Is the denial of simultaneity relations compatible with quantum theory? | 313 |
| 9.3 | Are simultaneity relations necessary for a dynamic model? | 316 |
| 9.4 | Summary. | 322 |
| Appendix 1.1 | The Use of Spatial Diagrams of Time. | 324 |
| Appendix 4.1 | The Dependence of Reversibility on Interpretation. | 334 |
| Appendix 4.2 | Time Reversal for Quantum States. | 343 |
| Bibliography | | 348 |

CHAPTER ONE

AN OVERVIEW

The key feature of time that most needs to be explained is its *directionality*. This directionality strikes us immediately in everyday life as its most elemental feature, and it is something that we must get clear about before we can hope to gain much understanding of what time is. A number of different concepts of the directionality of time have been explored by philosophers and physicists. They fall into two main groups, corresponding to two distinct sources for our conception of time. I will call these the *commonplace conception*, and the *scientific conception* of time. Let us review these in turn.

1.1 The commonplace conception of time: IT FLOWS.

The commonplace conception is gained, or so we tend to think, from the direct experience or sensation of 'time passing' that we all feel we have. This experience gives us the idea that there is something commonly known as the 'passage of time', or 'time flow'. This passage seems to us in our everyday lives to be the very key to time.

This idea that 'time flows' is surely among the most basic ideas we have, and is fundamental to our everyday metaphysics. It is fundamental to the understanding of such undeniable human realities as *experience* and *action*. These central elements of human existence are both fundamentally structured by what we call the 'passage of time'.

This is most obvious with experience. Experience seems to be intrinsically dynamic. The flow or movement of experience seems intrinsically necessary to it. We certainly all seem to be immediately aware of such a 'flow'. What does this awareness amount to?

We are aware, even if only fleetingly, that there are experiences which we *have had*, which are over; that there is a current experience which we *are having*, which has a special reality for us; and that there are experiences which we *will later have*, which we busy ourselves contemplating and planning for. Thus we understand experiences as changing, or as coming into, and going out of, existence. This is what the 'flow of experience' implies for us, in our normal lives.

be an intrinsically temporal kind of existence. The experiences that one has in the course of a life do not all exist at once: rather, experiences *happen*, moment by moment, they come into existence, and go out of existence. There are three modes of temporal existence: what will be, what is now, and what has been. Or future, present, and past existence. We all distinguish in practical life the pastness, presentness, and futurity of various experiences with great objectivity. For instance, some half an hour ago I had the experience of tasting a cup of coffee – but that experience is now definitely over; presently, I have the experience of sounding these words to myself in my mind as I write; and I anticipate a future experience, in another hour or so, of the taste of some baked beans which I intend eating for my dinner.

It is implicit in this that our experiences *change*: that present experiences cease, and new experiences become present. It is *change* which links the three modes of temporal existence together, and gives them the structure which makes them what they are, namely modes of existence.

1.2 The commonplace conception II: time flow in the physical world.

We also quickly generalise from the flow of subjective experience to the idea of time flowing for the physical world. We believe that physical events, or states of affairs, have the same temporal modes of existence (past, present, future) as we attribute to subjective experiences. We believe that the states of physical objects *change*.

This is forced on us by our commonsense understanding that we are in a physical world, that our experiences provide us with perceptions of it, and that these perceptions are accurate. Changes in percepts must be taken to register changes in perceived objects for any sensible understanding of perception. Into the bargain our experience seems to tell us that we are partly or wholly physical things or processes ourselves, and form a part of the physical world; thus we naturally take the physical world to be caught up in the flow of time that we are aware of in experience. This forms a central part of our commonsense (and scientific) metaphysics.¹

Thus the commonplace view is that there is a flow or movement of time for the whole world, physical and mental.

1.3 Phenomenological directedness.

The physical world is also observed to be highly directional in time in another way: there is an overwhelming directedness of phenomenological processes. The processes which most catch our attention in life only ever occur in one temporal direction. Eggs break but never mend. Fires burn but never restore. Beans are digested by the intestines, not reconstituted.

We no doubt develop a very strong sense of this directedness of processes in time. This sense is easy to confirm by watching

¹Materialists (provided they are not eliminativists) must suppose that the mental and the physical share the same ontological features, so that they can hardly hold that there is a (real) flow of experience but no corresponding real flow of physical time. Similarly idealists, or monists of any kind. The dualist might possibly avoid this conclusion, but I see no reason for wanting to.

a film of any normal human activities played in reverse - the physical directionality inherent in even simple bodily movements is quite noticeable, let alone in such activities as spitting and diving into swimming pools.

It may be thought that this physical directionality further confirms the idea of time flow, but there seems no good argument for that. It seems unlikely that the mere directedness of phenomenological physical processes yields the *metaphysical* idea of time flow. This directedness just consists in the fact that there are types of processes which commonly occur but the 'reversals' of which never occur, (and cannot be made to occur). But this kind of directedness seems to indicate nothing about the *temporal modes of existence*. I.e. it hardly entails that *existence is structured by the three temporal modes, past existence, present existence, and future existence*.

It seems much more plausible that we gain the idea of time flow from the more fundamental 'sensation of flow' that we seem to have, since this sensation seems to be exactly the sensation of certain things (namely, experiences) coming in to, and going out of, *existence*. Of course, it remains a good question whether this sensation of the 'flow of existence' provides good evidence for the metaphysical thesis either: this is a question to which I will return at the end of the chapter.

The first main concept of the directionality of time is, then, this metaphysical idea of *time flow*. We have noted two different instances of the idea: first, the idea of a flow of experiential time, second, the idea of the flow of physical time. Both of these involve the same metaphysical idea: that the

existence of things, in the first case, experiences, in the second, physical states of affairs, comes with temporal modalities, structured by real change. The commonsense view is undoubtedly that physical and experiential time flow go hand in hand, and that there is a general flow of time which catches up all of what exists in its inexorable movement².

1.4 Terminology.

The use of the term 'time flow' requires some comment. It obviously arose as a metaphor, which was, no doubt, felt to nicely evoke the phenomenon being talked of. But some writers assume that if we use the term, we should take it more or less literally, as meaning that time in some literal sense *flows*. Thus questions like the following commonly arise: At what rate does it flow? Relative to what does it flow? and so forth.

But this presupposes that the metaphor, 'time flow', can be interpreted as an accurate structural map of the phenomenon it evokes. As many writers have argued, the term 'time flow' cannot be taken literally. We do not get a viable theory of what 'time flow' is by trying to plumb the literal meaning of this metaphor.

My aim is to give a completely literal account of what time flow amounts to, but I do not take the term 'time flow' at all literally. I will use it to *refer* to the phenomena that it evokes, but I will not take it as implying that therefore the phenomena referred to must consist in time 'flowing'.³

² With the possible exception of God, who is often believed to transcend time.

³ Broad [1938] rejects 'time flow' as implying an incoherent idea of a movement of time, while maintaining that there is still the phenomenon of 'Absolute

1.5 Against time flow: Illusionism

Is the commonplace view correct? Does time really flow? This is the central question I am concerned with. I will eventually defend a controversial answer to it: YES.

The majority of scientific writers on the subject have reached the opposite conclusion. They believe there is no real, objective flow of time. They usually wish to discard the whole idea of an objective flow of time as an illusion. I will call them *Illusionists*.

The Illusionists feel that there is no real evidence for time flow, merely a subjective feeling or sensation the meaning of which we misinterpret. All we need to do, say the Illusionists, is to explain away the sensations that lead us to postulate time flow. If we can explain, in a good scientific fashion, the existence of all these sensations, without ever appealing to the existence of time flow in our explanation, then the sensations would seem to offer no good evidence for time flow.

They are convinced that we can give such an explanation for two reasons. The first is that they are invariably materialists, and believe that all real phenomena including mental phenomena are ultimately physical phenomena, and have purely physical explanations. Thus they think that whatever counts as the 'sensation of time flow' consists ultimately in certain kinds of physical process.

Becoming'. But in Broad's place we would just consider that the term 'time flow' refers to the phenomenon of Absolute Becoming, but that it must not be taken literally.

Secondly, they believe that the world of physics admits no such thing as time flow, that the concept of time flow plays no part in any physical explanation, and therefore that it plays no part in the explanation of our sensations of time flow.

If these two theses were true, then it would be possible to explain away all the evidence for time flow without ever appealing to time flow, and there would seem to be no further empirical reason for believing in time flow.⁴

This notion that time flow plays no role in physics turns us to the second conception of time: the scientific conception.

1.6 The scientific conception of time.

As well as playing a role in our direct experience, time figures in all our key physical theories, from Newtonian mechanics, to relativity theory, to the latest quantum theory.

Such theories as these are often called *fundamental theories* because they are attempts to offer an account of the fundamental nature of physical existence. There is a fair bit of metaphysics involved in the idea of a fundamental theory, but I will not try to go into that here. The idea of fundamental physics is well enough grasped by everyone who understands the enterprise of modern physics, and I will assume an intuitive understanding of why

⁴Other kinds of explanations are also offered by the Illusionists: e.g. see Christensen [1976], and Smart [1987a] p.86-88 for 'linguistic' and 'psychological' explanations. But these are really explanations of the *effectiveness* of the 'illusion', after it has been decided that there *is* only an illusion and not real time flow. The possibility of the kind of explanation described above is what is fundamental in deciding that the appearance of time flow *is* an illusion.

these theories are taken to have a special authority in deciding questions about physical reality.

In the ontologies offered by fundamental theories of physics, time always plays a key role. It is normal to objectify time as a one-dimensional continuum of moments, mathematically identical to a spatial dimension. In the normal interpretation of special relativity, in particular, time and space are mixed up together in an inseparable way to form the 'space-time manifold', and thus time is objectified exactly as space is.

A few words should be said about this objectification of time (and of space) for it is frequently objected to, on the grounds that all we really need for science are events with temporal (or spatio-temporal) *distances* between them. It is claimed that we do not also need substantial items called 'moments' (or space-time points). This gives the *relativist* view of time and space first developed by Leibniz in his famous correspondence with Clarke.

On the relativist view, time does not exist as an *object* (the continuum of moments). But time still figures in this way: events remain set in an intrinsic temporal (or spatio-temporal) metric. There are physically real temporal distances between events.⁵

The introduction of a substantial time or space-time is often, from this point of view, reinterpreted as a convenient mathematical apparatus for representing the metric on events - we can say that 'e occurs at t_1 and f occurs at t_2 ', rather than: 'f occurs at a temporal duration $t_2 - t_1$ from e'. On the relativist's view, the events themselves and their temporal distances remain

⁵Or ratios between pairs of events.

perfectly real, and we may say that events remain 'set in time' in the sense that they remain set in an intrinsic temporal metric.

The absolutist/relativist debate is interesting, but it is not of much relevance to anything I will have to say. The arguments I will be concerned with are not sensitive to the distinction being made. Like all physicists in their practical work I will generally adopt the absolutist picture of time as the most mathematically convenient, but everything of substance that will be argued for could be reformulated instead in relativist terms. (Or rather: if the relativists are correct, and relativist time *is* all that is needed, then everything could be so reformulated.)

For instance, I will speak of two *directions of time*, 'earlier' and 'later'. These will be treated as directions between *moments* (which may be construed as properties of temporal vectors, constructed from classes of moments). But they could be introduced just as easily as directions between events.⁶ Or again, a certain view of existence called the *bloc universe* view will be presented as the view that the whole historical universe already exists as an unchangeable entity. But it doesn't matter whether the bloc universe is taken to be a collection of events placed in a substantial space-time, or as the same collection of events with just the appropriate spatio-temporal distances between them.

It will normally be assumed that in the scientific conception, time is a linear continuum of moments. Some solutions of the equations of general relativity give time topological properties

⁶Using the distance, $d(e,f)$ between two events, e and f , as a primitive notion, the direction from event e to event f is the direction from event f to event g just in case: $|d(e,f)| + |d(f,g)| \leq |d(e,g)|$.

which conflict with this, but I will be concerned with very much simpler matters.

1.7 Does scientific time flow?

The main question in reconciling the scientific to the commonplace conception of time is this: *does scientific time have any feature which corresponds to the commonplace view that time flows?*

We may note first of all that in most respects commonplace and scientific time are the same thing. Or more accurately, they are slightly different conceptions of a single thing, namely *time*. The commonplace conception attributes the same kind of intrinsic metric to time as the scientific view. We refer, in common language, to temporal distances between events, and these are real temporal distances between physical events. There is no great gap to be bridged at all between commonplace and physical time in respect of most of their features. The problem is time flow.

It is a problem because most present day philosophers and physicists who have studied the question have concluded that there is no flow of physical time, in the metaphysical sense. Time flow is regarded as an unscientific idea. There are varying forms of this conclusion, and the discussion really needs to be set in the context of our best current physical theories. Calling the group of our best current theories F , consider:

[1.1] F does not require any physical time flow.

[1.2] F requires that there is no physical time flow.

Or much stronger versions, which are not relative to F:

[1.3] No proper physical theory requires time flow.

[1.4] Every proper physical theory requires that there is no time flow.

Let us begin by considering the possible reasons for holding [1.1], which we will find to be the key to the debate.

[1.1] does not rule time flow out conclusively, it merely states that our present theories do not require it. If one takes these theories as giving a complete picture of physical ontology, or at least of the general framework for physical ontology, this is still a powerful thesis. For if F is complete, i.e. represents everything about the physical ontology, and if time flow is extraneous to F, then it has no place in the physical ontology, which is just to say that time flow is not physically real.

But what are the reasons for holding [1.1]?

1.8 Reasons for [1.1]: Inspection of the theory.

The simplest reason for [1.1] might be the failure to find any mention of time flow by simple *inspection* of the theories in question. For instance, we could make a list of the primitive concepts of the theories, and check that 'time flow' is not among them. The concept of time flow does not appear explicitly in either relativity theory or quantum theory.

There are at least two kinds of problems that might be raised against this argument. The first is that, although 'time flow' might not appear in the theory itself, it seems always to be presumed in the meta-theory. Any book which you care to open about either relativity theory or quantum theory (or any other scientific theory) will use a meta-theoretic language to present the theory itself. This meta-theoretical language is rather close to commonplace language, describing experimental procedures, and a variety of everyday activities. Implicit in the meta-theoretical concepts is the commonplace view that time flows. It seems quite impossible to escape from time flow in the meta-theoretical presentation of any substantial theory, and it may be thought that this means that science, after all, supposes or presupposes time flow.

In fact this objection, if it was successful, would probably establish the existence of physical time flow rather generally; but unfortunately it misses the point. The point is that the meta-theoretical presentation of the theory is exactly an attempt to bridge the gap between the theoretical world, and the commonplace or phenomenological world with which we are in

reasonably direct contact. Thus it necessarily uses commonplace concepts, such as that of deliberate human action, which naturally presuppose the commonplace metaphysics of time flow. Indeed, this is all that the observation that the meta-theoretical presentation of the theory presupposes time flow amounts to: that our *commonplace* concepts presuppose time flow. But we already know this. What is not established is that there is time flow in the *theoretical* ontology. That is the question.

This first objection will therefore be ignored. A second, very simple, objection is more important. It is that the concept of time flow might be *implicit* in the theoretical concepts, and hence ultimately required in the ontology, even though it is not mentioned explicitly. (In fact, this is exactly what I will eventually argue: I will argue that the idea of time flow is implicit in the concept of objective probability, which is primitive to quantum mechanics.) Because of this possibility, it seems that mere inspection of the explicit concepts of a theory is not enough to decide [1.1]: we must also check that time flow is not implicit in the theory, not hidden away in some dark conceptual closet. How can we do this?

A more systematic way of checking whether time flow enters the ontology is desired. A second method might be thought to do better at this task. This is to leave the list of theoretical concepts behind, and instead look directly at what the physical ontology required by the theory *is*, and to check it for time flow.

This is more or less what most writers attempt to do. Some way of representing or picturing the physical ontology required by the theory is adopted, and it is then shown that there is no time

flow involved in this ontology. For instance, a very common procedure is to draw a Minkowski diagram to represent some typical physical process, and then to observe that there is no place for time flow in the thing pictured in this diagram⁷

But there is really not much logical clarity to this idea, and indeed something rather odd about thinking that it does any better than the first method of mere inspection. The problem is that the 'direct inspection of the physical ontology' is not such at all, but really just inspection of another *representation* of the theory. What people do when they try to check out the ontology for time flow is to give a *secondary* representation of it, such as the Minkowski diagram. This is taken to represent the fundamental features of the ontology, and it is then observed that there is no place for time flow in this representation. But the problem now is whether the diagram is really an accurate representation. It may be, indeed, that in the secondary representation there is no place for time flow: but then that representation is *accurate* only so long as time flow really does not figure in the physical ontology. How have we become sure of that? By inspection?⁸

⁷ See D.Park [1970].

⁸See Appendix 1.1 for an extended criticism of the way spatial diagrams are used to represent time.

1.9 Reasons for [1.1]: Reversibility of F.

It seems that our first two methods of checking for time flow in a theoretical ontology are not as conclusive as they need to be. And since I will soon be arguing for the very controversial thesis that quantum theory *does* require time flow, having a conclusive method will be important. Fortunately there is a method which is far more conclusive. It makes use of a second concept about the directionality of time, which is known as *physical time symmetry*, or *time reversibility*.

These are two slightly different concepts, but they have such a close relationship they are almost identical. *Theories* are said to be *time reversible* (or irreversible). It is an objective mathematical fact, following from the formal structure of a theory, whether a given theory is reversible or not. If a theory is reversible, then the *time* postulated by the theory is said to be *symmetric*. It will be shown in detail in Chapter Three exactly to what time symmetry amounts, but it may be taken here at face value.

It will also be shown in Chapter Three that:

[1.5] Time flow requires an intrinsic *asymmetry* of time.

This is because it confers intrinsic asymmetric directional properties on time. If there was no intrinsic asymmetry of time, there could hardly be a flow of time. Hence, if a theory conferred no asymmetry on time, it could hardly confer a flow of time.

Since a theory confers asymmetry on time only if it is an

irreversible theory (shown in Chapter Three), we have the important thesis:

[1.6] If T is reversible then T does not require time flow.

On checking our best current theory F for time reversibility (an objective mathematical procedure) we must find either:

- (i) F is reversible, or else:
- (ii) F is irreversible.

In the first case we can decisively state that:

[1.1] F does not require physical time flow.

Thus at least one kind of decisive answer is possible for [1.1].

In the second case, where F is irreversible, we can draw no immediate conclusion. Irreversibility is compatible with time flow, but it in no way requires it. The kind of irreversibility that a theory suffers from might or might not be relevant to time flow. However once we isolate what the irreversible feature of a theory is, we have made considerable progress, because we know that if the theory does require time flow, then it must be closely connected to the way the theory is irreversible. We can check the specific irreversible feature of the theory, and it seems hopeful that mere inspection will now be a good guide as to whether this feature has anything to do with time flow.

This is the primary reason that I will investigate the

concepts of reversibility and time symmetry, which are the main concepts concerning the directionality of physical (scientific) time. It also turns out that these concepts are very interesting in their own right, and are necessary for exploring further questions and puzzles about physical time, particularly the puzzle of phenomenological directedness. These topics will occupy Chapters Two to Six. The discussion of time flow will be taken up again in Chapter Seven, after a clear theory of the physical directionality of time has been established.

1.10 Four major views on physical time flow.

Let us continue here with the logical relations between time flow and scientific theory. We have seen in some detail how the truth of [1.1] is to be investigated. There are two possible conclusions: (i) That F requires time flow, (ii) that F does not require time flow.

In the first case we would have the remarkable fact that *physical time flows*⁹ - and that it is a physical fact that it flows, delivered by the physical laws. I will call this the thesis of *objective physical time flow*. If it were established, then we could simply get on with spelling out how and why our commonsense belief in time flow connects with the reality of physical time flow, and we could probably expect a fairly direct answer. At any rate, consistency would have been established between our commonplace and our scientific conceptions of time over the most knotty issue.

In the second case, (ii), we would be faced with a problem -

⁹On the assumption, of course, that F is correct, or near enough correct.

indeed, with what is considered the most distinctive scientific problem about time flow, for it is presently thought that (ii) is correct. The problem is how to deal with the commonplace belief in time flow, given that pure physics leaves it out. There are three main choices:

(a) The main position is the *illusionist* position already noted earlier. This is the thesis that there is no time flow at all, merely the illusion of it. (Typical advocates: Grünbaum [1973 Ch.10], Smart [1954, 1987a].)

(b) Time flow might be postulated in a kind of metaphysical way, as a non-empirical feature of the world which empirical science simply does not and cannot get to grips with. I will call this the thesis of *objective non-empirical time flow*. (Typical advocates: perhaps Prior [1959], and Broad [1938], and perhaps a number of contemporary metaphysicians, who find they cannot do without time flow, but regard it as a non-empirical postulate.)

(c) If, because of (ii), we give up the idea of any objective flow of *physical* time, we might still maintain a real flow of 'psychological time' (a real flow of experience). We would thereby be supposing that subjective experience has a feature which distinguishes it essentially from physical events or processes, namely that experience exists *temporally*, with the temporal modes of existence (past, present, future). In contrast, physical events would simply exist, there would be no real temporal categories to physical existence. This would require a form of mind-body dualism, at least with respect to the kinds of existence that mental and physical things have. I will call this *objective mental time flow*. (I know of no advocates of this, but

a dualist might adopt it. It is rather suggested, though hardly deliberately, by Weyl's famous description of "consciousness crawling upward along the life-line of my body." [1949].)

A briefer list may be useful:

| <u>Thesis</u> | <u>Physical time flow.</u> | <u>Typical</u> <u>Metaphysics.</u> | <u>Typical</u> <u>Epistemology.</u> |
|-------------------------------|----------------------------|--|--|
| Illusionism | No. | Materialist. | Empiricist. |
| Objective physical time flow. | Yes. | Materialist or Dualist. | Empiricist. |
| Objective non-empirical flow. | Yes. | Materialist or Dualist. ¹⁰ | Rationalist. |
| Objective mental flow. | No. | Dualist. | - |

There are other consistent possibilities, for instance Idealism with or without time flow, or dualism without time flow (a dualist version of Illusionism). But the four views here described appear to be the important options.

It is worth relating this classification of views to a more

¹⁰The difference between objective physical time flow, and objective non-empirical time flow, is that in the first case time flow is a postulate of the physical laws, while in the second it is introduced as a 'metaphysical' thesis, over and above anything in the purely physical laws. The difference between these two theses will be discussed at greater length in later chapters.

common one. It is common to classify views on time in terms of belief in the existence in what is called the 'A series' and the 'B series', terms coined by McTaggart [1908]. Belief in the A-series is simply what we are calling belief in time flow. Belief in the B series is the belief, which we are assuming throughout, that there is an 'earlier-than' relation among moments of time (or if you are a relativist, among events). The classification offered above is really a detailed way of stratifying views on the reality of the A-series.

1.11 Do scientific theories prohibit time flow?

We have now seen in some detail how the proposition [1.1] may be investigated, and what kinds of views it might provoke if found to be true or false. Let us now go on to consider the much stronger proposition [1.2], that *F requires that there is no physical time flow*. Of course establishing [1.2] would be somewhat more useful for the Illusionist's cause than merely establishing [1.1], and it would mean the decisive failure of the thesis that I have already hinted I will eventually defend, namely a version of *objective physical time flow*.

There have been two distinct approaches to [1.2]. The first and best came with relativity theory. The specific idea is that relativity theory denies any reality to what we call the 'now', or the 'present state of the world'; that time flow requires a well-defined 'now'; and therefore that relativity therefore entails the impossibility of time flow.

The more general form of this argument is that: (i) time flow requires some general feature, G, while (ii) our good theory, F,

entails the failure or absence of G, thus (iii) F requires that there is no time flow.

This is a good argument, and in fact I regard the problem raised by relativity theory as the most formidable for any theory of time flow to face. I devote a chapter to this problem after I have made an attempt to establish a substantial theory of time flow in a non-relativistic quantum theoretic setting. Until then I will ignore this particular problem. (The apparent incompatibility of time flow with the relativistic denial of the existence of simultaneity relations is the only substantial argument of this kind that I know of.)

There is a second group of arguments for [1.2] which are also arguments for the much stronger: [1.3] *No proper physical theory requires time flow*, and [1.4] *Every physical theory requires that there is no time flow*.

These are *a priori* arguments against either the very coherence of the idea of time flow, or against its coherence in the setting of any physical ontology.¹¹ Famous arguments of the first kind are given by Parmenides in his 'Way of Truth', and McTaggart [1908], while notable arguments of the second kind are given by Grünbaum [1973], Smart [1954], Park [1970]. All these arguments purport in one way or another to demonstrate an inconsistency in the notion of change, or in the notion of the temporal structure of existence, which time flow entails.

I believe that all such *a priori* proofs against time flow are ultimately circular, roughly because they must rely on some *presupposition* that the only coherent concept of existence is an

¹¹These are the only kinds of arguments for [3] and [4].

a-temporal one. I will illustrate this point in some detail in the process of presenting a substantial theory of time flow. But it would not be practicable or useful to rebut every important argument that has been given against time flow in detail. Instead, I will let the coherence of the formal model of time flow that I present in Chapter Seven speak largely for itself. I believe this model demonstrates the consistency of the concept of time flow.¹²

It might be thought that I am overlooking another way of establishing [1.2], namely, simply to establish the reversibility of the theory F , from which it follows that time flow cannot be found in the ontology of F (by [1.6]). But it is a mistake to think that the reversibility of a theory F entails that *F requires that there is no time flow*. The reason is that F may be true, but only be a partial theory.¹³ The complete theory may be a stronger theory, F^* , which is *irreversible* and which adds time flow to the static ontology of F . Since F^* is compatible with F , time flow is compatible with F , and hence the mere reversibility of F does not mean that *F requires that there is no time flow*. This will be shown more fully in Chapter Three.

Having sketched the general setting of the debate over time flow, I will now summarize the main arguments that I will present.

¹²In fact the formal consistency of the tense logics developed by Prior [1968] and others, plus the fact that we all do understand and operate perfectly well with the concepts of 'past', 'present', 'future' and 'change' in everyday life seems to me enough to show the consistency of the idea.

¹³It would be true that: F is reversible and F is the complete fundamental theory entails that there is no time flow.

1.12 A theory of objective physical time flow.

With few enough exceptions (most notably Reichenbach [1953] and Capek [1961]) major recent commentators on physical time have generally felt that [1.1], [1.2], [1.3] and [1.4] are all true, and that there are pretty good arguments to show it¹⁴. It is thought, at any rate, that where F consists of quantum theory plus relativity, [1.1] is entirely certain, and [1.2] is certain barring some reinterpretation of relativity theory. [1.3] and [1.4] are perhaps more complex and dubious, but without worrying about these, [1.1] and [1.2] seem to put paid to the idea of physical time flow rather decisively.

I will defend the opposite view: that physical time flow is real. I will defend a theory which renders time flow as an *empirical* fact, for which we have a lot of evidence. If this is so then time flow ought to be reflected, somehow or other, in fundamental physics, and indeed it is: the rest of this chapter is an extended argument that *time flow is required in quantum theory*.

Hence I will be arguing for a version of what I have called *objective physical time flow*. This of course puts me at loggerheads with the illusionists, who represent the mainstream opinion of the day.

The first thing I must do is to contest the truth of both [1.1] and [1.2], since either of these would be fatal to my project. I have already said that the main argument for [1.2] is the

¹⁴E.g. Grünbaum[1969,1973], Smart[1954, 1987], Mehlberg [1980], Park [1970], and many others.

relativistic denial of an objective 'now', and I regard this as a substantial and real problem. However I will set this aside until Chapter 9, for this problem will be a lot more tractable after a number of other details have been discussed.

If [1.2] is removed, then the way would be open for what I have called *objective non-empirical time flow*. But for the kind of theory I will argue for, i.e. objective physical time flow, [1.1] must be removed as well.

I will contest [1.1] in the context of quantum theory. I will first argue that quantum theory is *irreversible*, thus removing any conclusive demonstration of [1.1] through a demonstration that quantum theory is reversible. I will subsequently argue that the kind of irreversibility the quantum theory suffers from yields, on the most natural interpretation, time flow in the quantum universe.

1.13 The irreversibility of quantum theory.

This claim that quantum theory is irreversible will probably make any experienced quantum physicist wince. For it is thought that (with a rather minor and still controversial exception with which my argument has nothing to do) all the known quantum theories are conclusively *reversible*. This is thought to be an indisputable mathematical fact. Demonstrations of the reversibility of quantum theories can be found in numerous textbooks on the subject, and it is a result accepted by practically everyone¹⁵. The only possible escape from the conclusion, it will be thought, would be through some radical reinterpretation of quantum theory, which would be bound to be highly controversial.

I dispute this result. I claim that a mistake has been made in the orthodox analysis of the reversibility of quantum theory, and I offer an essentially mathematical demonstration that quantum theory is irreversible in a very deep and inescapable way. And I do this without using any controversial interpretation of quantum theory. Indeed, my argument is insensitive to almost all issues of interpretation of quantum theory. It is certainly as insensitive to interpretive issues as the orthodox analysis that purports to show the reversibility of quantum theory.

The only important interpretive thesis on which my argument does depend is that quantum theory postulates probabilities in nature.

¹⁵Davies [1975], Elliott and Dawber [1979], Reichenbach [1956], Mehlberg [1980], Watanabe [1955a,b,c].

1.14 Quantum probabilism.

In every known version, quantum theory postulates probabilistic laws, normally through some version of the 'projection postulate'. There is no apparent way of formulating quantum theory without probabilistic laws. This should not be taken as the claim that there is no possibility of a deterministic interpretation of quantum theory.¹⁶ It is merely the claim that probabilistic laws are intrinsic to quantum theory, which seems so obvious that it is hardly worth arguing for.¹⁷ This does not mean that nature cannot be deterministic: all it means is that if nature *is* deterministic, then quantum theory is not the fundamental theory we take it to be. Any escape from quantum indeterminacy is bound to be into a new theory which is hardly to be called *quantum theory*.

This matter will be discussed at greater length in Chapter Eight. However, the reader skeptical about the thesis that quantum theory is a probabilistic theory can take my argument to demonstrate the conditional, that *if* quantum theory is a probabilistic theory, then quantum theory is irreversible.

¹⁶Although this is a respectable thesis in itself, for many hold that Bell's theorem rules out that quantum theory could be modelled in a deterministic ontology. I am not convinced of this, but this question is beside the point here.

¹⁷The 'many worlds' interpretation seems to be deterministic, but in fact it must recognise probabilities of some kind if it is to be adequate. This is discussed in Chapter Eight.

1.15 The irreversible feature of quantum theory.

What feature of quantum theory makes it irreversible? It is a feature which, in itself, has actually been recognised for many years. The physicist Watanabe seems to have been the first to see its real importance.¹⁸ Very simply, it is that *objective probabilities in quantum theory are future-directed, but never past-directed*. This is demonstrable from some very general observations, as will be shown in Chapter Six.

While this feature is well enough established, what has not been recognised is that it represents a conclusive *irreversibility* of quantum theory. In fact this is a simplification: Watanabe, who discussed this feature a great deal, recognised it as a kind of 'irreversibility'. Unfortunately he did not recognise the full sense in which it makes quantum theory irreversible. He considered it a *kind* of irreversibility, but without rejecting the standard sense in which quantum theory is said to be reversible. (He was a pioneer in establishing the orthodox reversibility of quantum electrodynamics.)

I will demonstrate, however, there is no such thing as a 'kind or irreversibility'. There are not many concepts of reversibility, but only one. There is a single objective sense in which a theory is be said to be reversible, and if it is not reversible it is irreversible. In this sense, the temporal directedness of quantum probabilities renders quantum theory irreversible.

This point needs stressing, as I have found through having presented my argument in a number of seminars. A common

¹⁸Watanabe [1955, 1965, 1966, 1970]

response to my argument has been: "Well, you have shown an interesting *kind* of 'irreversibility' of quantum theory. But it is only of a kind. The physicists show that, in another sense, quantum theory is perfectly reversible. Now why shouldn't we take their sense of reversibility as the best one?"

But this is specifically what I deny. The first part of my argument is to *demonstrate* that there is only one concept of reversibility, and to formulate this concept precisely. I go on to demonstrate that quantum theory is irreversible. It cannot be said to be irreversible in one sense, but not in another: it is simply irreversible.

What, therefore, is the mistake that the physicists have made? They have demonstrated a very powerful symmetry of quantum theory, which they have *called* reversibility (or time symmetry, or invariance under time reversal); but this is a mistake, because the symmetry in question does not represent reversibility at all. The mistake has been made because *the criterion for the reversibility of probabilistic theories* has been incorrectly formulated.

1.16 The criterion for the reversibility of probabilistic laws.

For the sake of analysing the reversibility of quantum theory, the orthodox analysis divides quantum processes into two types: deterministic processes, and probabilistic processes. The first is the deterministic evolution of the state vector, governed by the Schrodinger time-dependent equation. The second is represented by some form of the 'projection postulate', which tells us the probability of a system being found in a certain state given its initial state.

There is a controversy about this division of processes, but I will not discuss that here. The orthodox analysis assumes the division, and for the moment I will merely follow its lead. (See Chapter Eight for more comments on this.)

The treatment of reversibility falls into two corresponding parts: the treatment of the reversibility of the deterministic processes, and the treatment of the reversibility of the probabilistic processes. On the orthodox analysis, both types of process are found to be reversible (with the minor and still controversial exception of systems involving K^0 mesons, which will be of no concern here.)

With the first part of the result, the reversibility of the deterministic evolution of the state vector, I have no quarrel. I fully agree this process is reversible. What I claim is wrong is the analysis of reversibility for probabilistic laws. I believe that a systematic mistake has been made, because *the wrong criterion for the reversibility of probabilistic laws has been adopted.*

It is obvious enough what the reversibility of deterministic laws requires; but it is somewhat less clear what reversibility *means* for probabilistic laws. The criterion that has been adopted is sometimes called the *probabilistic principle of micro-reversibility*, which I will abbreviate as [PPMR]. Very roughly the [PPMR] states that *the transition probability from a state s_1 to a state s_2 equals the transition probability from the time-reversal of state s_2 to the time-reversal of state s_1 (for all states s_1 and s_2)*. It is thought that the [PPMR] is a necessary and sufficient condition for the reversibility of probabilistic laws, i.e. where T is a class of probabilistic laws:

[*] [PPMR] holds of T \Leftrightarrow T is a time reversible theory.

Indeed, the [PPMR] is probably taken as the very *meaning* of reversibility for probabilistic theories by many physicists. [*] seems convincing, but I argue that it is wrong. I will demonstrate that *the [PPMR] is neither a necessary nor a sufficient condition for reversibility*.

This is really a general result about the concept of the reversibility of probabilistic theories, and not a matter specific to quantum theory. Hence the main part of my argument is not specifically about quantum theory, but about the meaning of reversibility for probabilistic systems. In fact, only the slightest grasp of quantum theory is needed to follow my main argument.

To appreciate my argument, it must first be understood that the [PPMR] does not count as a definition of the *meaning* of

reversibility for probabilistic theories. There is a perfectly good and objective meaning of 'reversibility', which can be decided upon quite independently of the [PPMR]. I begin by analysing this meaning in an exact way. Having a precise definition of reversibility, I then demonstrate that the [PPMR] is neither sufficient nor necessary for reversibility. This result is illustrated with two hypothetical examples of probabilistic theories in Chapter Five. The first is a theory which fails the [PPMR], but is nevertheless reversible (in an obvious way, and in a way which satisfies the formal definition of reversibility that is developed). This demonstrates that the [PPMR] is not a necessary condition for reversibility. The second example is a theory which satisfies the [PPMR], but is nonetheless irreversible (again, in an obvious way.) This demonstrates that the [PPMR] is not a sufficient condition for reversibility.

1.17 The correct criterion for reversibility: [CPR].

Having rejected the [PPMR], I propose a new criterion for reversibility. I call it the Criterion for Probabilistic Reversibility, abbreviated to [CPR]. It is easy to show that the [CPR] is at least a *necessary condition* for reversibility. It is also shown that it is a sufficient condition for the reversibility of a theory consisting of any class of probabilistic transition laws.

It is important for my argument, however, only that the [CPR] is a necessary condition for reversibility, because it can be shown from very general facts that it fails of quantum theory. This is already evident from the results of Watanabe [1955,

1965,1966,1970]. Thus, quantum theory is shown irreversible, and on very general grounds.

This is the general form of my argument that quantum theory is irreversible. It will be useful if I try to say briefly what feature of quantum theory is behind the irreversibility. It has, of course, to do with probability. Very simply it is *the existence of future-directed probabilities, but the lack of symmetric past-directed probabilities.* (It is an appropriate symmetry between future-directed and past-directed probabilities that the [CPR] directly demands.)

It is on this feature that I will base a theory of time flow. Even if one decides to ignore the question of whether it represents an irreversibility of the theory, this feature can be considered directly for its possible relevance to time flow. To have any relevance, it must represent at least a structural feature of the theory which provides an intrinsic temporal asymmetry. In the next section I outline why this is so.

1.18 The lack of past-directed generic probabilities.

The asymmetry involves a probabilistic link from earlier to later states of quantum systems. There is a corresponding statistical link from earlier to later. This link is generic in nature. But there is no corresponding link in the other direction, from later states to earlier ones. Probabilities directed from later to earlier states do not exist generically in nature.

To illustrate the situation very simply, consider a very simple theory, which recognises four states, s_1, \dots, s_4 , of a certain kind of

system, with the probabilistic laws:

$$[L] \text{ PROB}(s_i(t+\Delta t), s_j(t)) = 1/4, \text{ for all } i, j.$$

[L] states that the probability of finding the system in any state s_i at a moment Δt later than t , given that it is in any other state s_j at t , is equal to $1/4$. There is an unspoken assumption that the system is to remain closed between t and Δt , for if it is open to other influences, then there might be some way of controlling the final state and contradicting the probabilities. (This assumption of closure is common to all dynamic laws.)

Now if we took a large ensemble of these systems, started them all in some initial states, and looked at their final states, we would expect the distribution of final states to include close to 25% of each type of state, s_1, \dots, s_4 . This is a statistical relation based on the objective probabilities. If this relationship failed in a substantial way, then we would have to concede that the probabilities given by the law [L] do not exist after all. But let us suppose they exist. Thus in our ensemble of systems, we have a distribution of *final* states in the proportions more or less of 25% s_1 , 25% s_2 , 25% s_3 , and 25% s_4 . This indicates a probabilistic relationship from earlier to later states¹⁹. We may say that the distribution of initial states *probabilistically controls* the distribution of final states.

Now is there a corresponding relationship in the reverse

¹⁹Because the distribution of final states is a *function* of the distribution of initial states. A boring function, since *every* distribution of initial states gives the same distribution of final states, viz. 25% s_i for all i ; but being boring doesn't matter.

direction in time? i.e. does the distribution of final states probabilistically control the distribution of initial states? If it does, then there will be generic probabilities of the form:

$$[M] \text{ PROB}(s_j(t)|s_i(t+\Delta t))= p_{i,j}$$

This states that the probability of the earlier state, s_j , given the occurrence of the later state, s_i , is equal to some number $p_{i,j}$. What $p_{i,j}$ needs to be for there to be complete temporal symmetry won't matter here, because we can show that there simply are no such reverse probabilities at all.

I am assuming a real-life-like situation, where it is under our control to choose the initial states of the systems in any ensemble. Thus we could choose four different ensembles: one where the initial states are all s_1 , a second where they are all s_2 , a third where they are all s_3 , and a fourth where they are all s_4 . And we are assuming of course that the distribution of final states in each ensemble is roughly 25% for each state s_j , since this is determined by [L].

Now this situation is clearly incompatible with the existence of any past-directed probabilities, of the form [M]. For no matter how the numbers $p_{i,j}$ are chosen, one or other of our ensembles will contradict the supposed probabilities. The fact is that the distribution of final states of each ensemble is the same, but the distribution of initial states varies wildly²⁰. This variation means that there is no generic probabilistic or statistical

²⁰The distribution of initial states is *not* a function of the distribution of final states. For the mapping is *one to many* from final to initial distributions of states.

relation from later states to earlier ones.

It may be thought that some kind of cheating is involved in the fact that the distribution of initial states has been *chosen*, while the distribution of later states has only been allowed to occur probabilistically *from* the chosen initial states. So, it might be said, of course this asymmetry will be apparent: but this is only because we have put it in there deliberately, and it only shows that we have *chosen* our ensembles so that they conflict with the hypothesis of past-directed probabilities. But this does not show that those probabilities do not, in normal circumstances, exist.

This criticism misses the point. It is the fact that the initial states *can be chosen*, independently of the final states, that leads us to reject the idea of a probabilistic connection from final to initial states. By contrast, the final states *cannot* be chosen independently of the initial states, and this leads us to postulate the future-directed probabilities.

Here is another way of putting the same point. Because of the forward-probabilities and the closure of the systems, the final states *cannot be controlled*. If the backwards probabilities postulated in [M] were really generic to nature, then similarly these probabilities plus the closure of the system would mean that the initial states could not be controlled either. (Remember that the systems are closed from final to initial state just as much as from initial to final state. There is no temporal asymmetry about closure.)

The simple fact is that the wild variation in the distribution of initial states for a given distribution of final states shows

that the backwards-probabilities, [M], do not exist.

And this is what it appears to be like in the real world: there are (if the fundamental theory is probabilistic) generic probabilities directed forwards in time, but there are no generic probabilities directed backwards in time.²¹

In real life, this is naturally interpreted as the *causal dependence* of later states upon earlier ones, and a lack of causal dependence of earlier states on later ones²². For in real systems, later states are perceived to *depend upon* earlier states in a lawlike probabilistic way; but earlier states do not depend probabilistically upon later states.

Probabilities must be carefully distinguished from mere possibilities here. Of course, the final state of a closed system determines the *possible* initial states of the system, by the condition that there must be a physically possible path from any initial state to the final state. In a fully deterministic system²³, for instance, there is only one such path, and hence only one possible initial state, thus ensuring a functional dependence of initial on final states, reflecting the dependence of

²¹It might be thought that there is another possibility: perhaps there really are such backwards-directed probabilities, and the kinds of phenomena that seem to show that there are not are just flukes. For where probabilities hold, almost anything is *possible*, and we cannot conclusively disprove any probabilistic hypothesis. This seems an obvious kind of fallacy, but I will have some more to say about it in Chapter Six.

²²A probabilistic conception of causality is intended here. i.e. A could be the cause of B in a certain case even though there is only a probability $p < 1$ of B given A.

²³In which both earlier states determine later states, and later states determine earlier states.

final on initial states.

In a probabilistic system, there may be many possible paths to reach a given final state, and hence many initial states consistent with a given final state. The *possibilities* of initial states for a given final state are thereby determined, but there is no extra lawlike *probabilistic* relation from final states to initial states. A final state entails no generic probability distribution of initial states. Probabilities look forwards in time, but not backwards.

1.19 Reversibility, thermodynamics, and phenomenological directedness.

The *evidence* for the lack of past-directed probabilities in the real world will be seen to be ultimately provided through the prevalence of disequilibrium processes in the real world (processes which involve large gains in entropy). The reason for this can be easily pictured. Consider a process where some ink and water are added together, and are observed to steadily mix. Suppose that the theory governing the system is probabilistic, in the sense that there are intrinsic probabilities of transitions from one state to another in the combined system. Viewed in the normal direction of time, this spontaneous mixing process will seem perfectly natural, roughly because (a) there are far more possible transitions to 'mixed' states than to 'unmixed' states on any occasion, and (b) the *nomological probabilities* of state transitions consequently determine a very high nomological probability of long-term evolution to a mixed state. But now consider the same process viewed in the reverse direction of time. It now consists of a system in a thoroughly mixed state evolving spontaneously into an extraordinary unmixed state. This appears an extremely unnatural process - in fact it appears to go decidedly against the *probabilistic laws of nature*. For the extremely high *nomological* probability that any given state will, in the long term, evolve to a mixed state has been contradicted.

What is especially important about this example is that the probability that has been contradicted is truly *nomological*, entailed by the fundamental laws of nature. This represents a

vital difference from the case where the system is governed by reversible *deterministic* laws. For in the deterministic case, the reversed evolution is nomologically possible, given the very special initial state of the reversed process. This special reversed initial state will, if the theory is really reversible, take the system deterministically back through the exact (reversed) original sequence of states.

It can be seen that the initial state of the reversed deterministic process represents very special *correlations* necessary to allow the special reversed process to unfold.²⁴ But where state transitions are intrinsically *probabilistic*, any 'special correlations' in the initial state of the reversed process are of no effect. For to follow the reversed path exactly will require an incredible sequence of *intrinsically probabilistic transitions*, whatever the initial state. This is just another way of saying that in the probabilistic system, the probability of long-term increase (or at least maintenance) of entropy is generally *nomological*, as opposed to the deterministic case, where the probabilities determining entropy increase are really epistemic, and the what actually happens is contingent only upon actual initial states.

Where the fundamental theory is probabilistic, therefore, the actual reversal of common disequilibrium processes such as the mixing of ink and water (the burning of petrol, the flowing of rivers, the radiating of the sun,...), would truly go against the laws of nature. The fact that these disequilibrium processes actually occur in our world shows that the probabilistic laws

²⁴See Bohm [1980,ch.6,7] for the concept of 'implicate order', which is the special sort of order in the reversed final state of the deterministic system.

governing processes in our world do not apply to our actual world in reverse. It is the existence of these processes that ultimately provides the evidence for the irreversibility of quantum theory. (This is discussed further in Chapter Six).

Now it happens that the processes that appear *phenomenologically directed* to us are of the same kind, i.e. they are disequilibrium processes.²⁵ So it may be said accurately enough that it is the predominance of phenomenologically directed processes that provides that evidence of irreversibility. But it is important to realise that although the phenomenologically directed processes provide this evidence, their commonness in our environment is not *explained* by the irreversibility of the laws. What is characteristic of these processes is that they are all disequilibrium processes: but quantum theory does not *require* that disequilibrium rather than equilibrium processes should occur. It would be equally consistent with quantum theory that the whole universe is in a kind of thermal equilibrium, and no phenomenological directedness is displayed at all.²⁶

²⁵Popper's points in his [1956] indicate that not all phenomenologically directed processes need directly involve entropy increases, since our inference of a common causal ancestor of a set of correlated events allows us to infer a causal direction in the process that caused the set of correlated events, without there necessarily being any entropy increase in this process. But it seems that there must instead have been an entropy increase earlier in the chain of events which gave rise to the 'causal ancestor', hence that the temporal directedness of the process in question relies on a disequilibrium process at some stage.

²⁶In this case there would be no evidence that the past-directed probabilities do not exist, hence no empirical *evidence* that the quantum world was in fact irreversible. Here - and only here - would the quantum world in fact *seem*

Thus the irreversibility of quantum theory that I am arguing for does not explain why the universe is in the kind of disequilibrium that it is in, or the phenomenological directedness of the environment. Physicists rather seem to expect that any serious irreversibility should provide for an explanation of this feature of the universe, but this expectation is wrong. This is discussed further in Chapter Two.

1.20 A proposal: time flow in quantum theory.

Having appreciated the temporal asymmetry of quantum probabilities, it is natural to wonder whether it has anything to do with the flow of time. My second main argument is that a convincing theory of time flow can be based upon this asymmetry.²⁷

It is important to make clear the kind of theory of time flow I am arguing for. In the broadest sense, I am arguing that *quantum theory requires time flow for the proper representation of quantum probabilities*. But this claim needs considerable explanation. Let me put forward a series of more explicit propositions:

properly reversible. For there would be the same actual statistical relations from future to past as there are from past to future, and it would not be possible to tell that there are in fact no generic probabilities from future to past.

²⁷There are a number of suggestions, most notably by Reichenbach [1953], Capek [1961], and McCall [1976], that a theory of time flow be based upon *indeterminism* in nature. I am in sympathy with the instinct behind these suggestions. McCall [1976] is discussed in detail in Chapter 7.

[1.7] Quantum probabilities *could* be interpreted in such a way as to require time flow.

[1.8] The theory that results from this is natural, conceptually unified, and empirically adequate to our experience of time flow.

[1.9] No other interpretation does justice to time flow (nor has advantages in other areas that could compensate for this lack). In other words the suggested interpretation should be regarded as giving rise to the *correct* version of quantum theory.

[1.10] This interpretation captures the implicit understanding of probabilities that physicists adopt in their ordinary use of quantum theory. Hence it can be seen as the normal interpretation.

To avert one confusion, what cannot be shown is that the mere quantum *formalism* requires time flow. The formalism concerns a kind of mathematical structure, and as such does not require anything physical at all.

For instance, in the formalism there may be variables x, y, z , normally interpreted to refer to physical space. But without the interpretation, there is no reference to physical space, and no implication about its existence. Once interpreted, a *theory* arises from the formalism, and (on the normal interpretation) this theory requires there to be physical space. We must decide upon an interpretation before we get a substantial theory: it is only then that we can claim that 'the theory requires physical space', or 'the theory requires time flow', or whatever else.

Thus what I am going to argue is that *the quantum theory*

that we get by taking a certain interpretation of quantum probability requires there to be time flow. Of course, what really needs to be shown is that the interpretation proposed is correct, or at least preferable to its rivals. That is the real problem, and that is why propositions [1.7] to [1.10] need to be argued for.

1.21 A contingent identity theory of time flow.

Let me first try to give a simple and rather picturesque account of the fundamental idea to be developed, and then return to consider [1.7] to [1.10] in detail.

Suppose that we take the idea of time flow seriously, and we wish to make a serious effort to find something in the quantum ontology that could correspond to it. That is: if time flow *were* real, where would it be located in the quantum world?

This is a normal kind of scientific question. We ask a similar question of water, for instance. Given the *prima facie* phenomenological evidence that there is this stuff we call 'water', we try to locate what the stuff could be in the quantum ontology. The only good candidate is a complicated thing called 'quantum-theoretical-H₂O-molecules', and it is in fact a very good candidate, so we adopt a contingent identity: phenomenological-water = quantum-theoretical-H₂O. What we mean is that the stuff which we call 'water' is actually stuff of a certain fundamental physical type, namely, the type quantum-theoretical-H₂O. This theory of water is contingent in the sense that if we came to reject the quantum theory, and to believe that

it is not a correct picture of fundamental physical ontology, then we would also reject the identity thesis.

Let us consider an analogous kind of contingent identity theory for time flow. Given the quantum ontology as the fundamental ontology for microscopic processes, with what could we identify time flow? The idea that suggests itself, given the time asymmetry of probabilities, is that *the actualisation of quantum probabilities* underlies what we call time flow.

By the 'actualisation of probabilities' is meant the occurrence of one probabilistic event out of all the possibilities. Since the probabilities are always directed towards the future, the *occurrence* of probabilistic events always takes us into the future, so to speak. These probabilities give the universe a method, so to speak, for transforming itself into a new state from the present state. I suggest we regard these probabilities as providing, in Nerlich's phrase, the 'engine of time' [1979,p.3.]²⁸

Identifying time flow with the actualisation of probabilities in this way achieves a number of desiderata:

(i) It provides for a *physical* definition of the future direction of time: viz. the future is the direction in which the quantum probabilities are directed. Provision for such a physical direction is the main structural feature that a physical theory of time flow requires. (If a direction was not provided for by the physical theory, then we would have to postulate it as an extra 'metaphysical' feature. We would then fail in our attempt to get time flow purely out of the physical ontology, as [1.7]

²⁸ Although Nerlich introduces the phrase only to call it an 'inchoate notion'. But I think it is a very nice metaphor.

requires.²⁹)

(ii) It is a theory that seems to have a good chance of being empirically adequate to our usual perception of time flow. A main constraint on any such theory is that it must give the flow of time a fine enough grain to satisfy our perception of time flow (change) as continuous. It must also make time 'flow' consistently 'forwards', flow in all of the universe and not just in special regions, and so forth. The theory to be developed will achieve all this.

(iii) It is also important that the theory be conceptually reasonable. By identifying time flow with some physical concept (the actualisation of probabilities) we affect our understanding of that concept. Is it conceptually reasonable to understand probabilities in the required way?

While (i) and (ii) are not so hard to defend, and go a long way towards establishing [1.7] and [1.8], at least three key questions remain. The first is the question of the conceptual coherence of the theory, raised in (iii). The second is the concern of [1.10]: does the suggested theory really capture the implicit understanding of practicing physicists? These two questions are quite closely connected, and I will discuss them in the following two sections, where I will argue that the required interpretation of probability is the natural one in the context of quantum theory.

The third problem is the concern of [1.9]: is there really a need

²⁹For instance, if we tried to introduce time flow into Newtonian mechanics, which is a reversible theory, we would, explicitly or implicitly, be adding an extra postulate rendering the theory *irreversible*. In fact Newton effectively introduced such a postulate himself in his famous definition of time in his *Principia*. This example is discussed at length in Chapter 3.

to give such an interpretation? Is a theory which postulates physical time flow better than an otherwise identical theory which does not? Is there adequate motivation for taking time flow seriously in the first place? This is a difficult problem, which I will discuss in the remaining sections of the chapter.

1.22 The idea of a dynamic theory.

First is a very general point about the notion of dynamic laws. Laws of physics are fundamentally *dynamic* laws, about how states of systems change with time. Dynamic laws are the obvious and the only place to look for time flow. Indeed, it is an intuitive view, which physicists have until they are taught better, that dynamic laws are exactly about *what happens to the world as time flows*.

This natural view is usually reformed when they meet relativity theory, and are taught to conceive of time as though it were all but another spatial dimension. It is still recognised that time differs from space, but it is generally said to differ only in the functional role it plays in the theory. For instance, the invariant interval in Minkowski space-time is: $dx^2+dy^2+dz^2-dt^2$, the temporal term having a negative sign in this equation, unlike the spatial terms. Similarly, in the Schrodinger equation in quantum theory, time and space play different functional roles.³⁰

But this explanation of the difference between time and space

³⁰There are many arguments that time and space are essentially identical. E.g. Taylor [1955], Gödel [1949], Quine [1960], Costa de Beauregard [1966], Webb [1977].

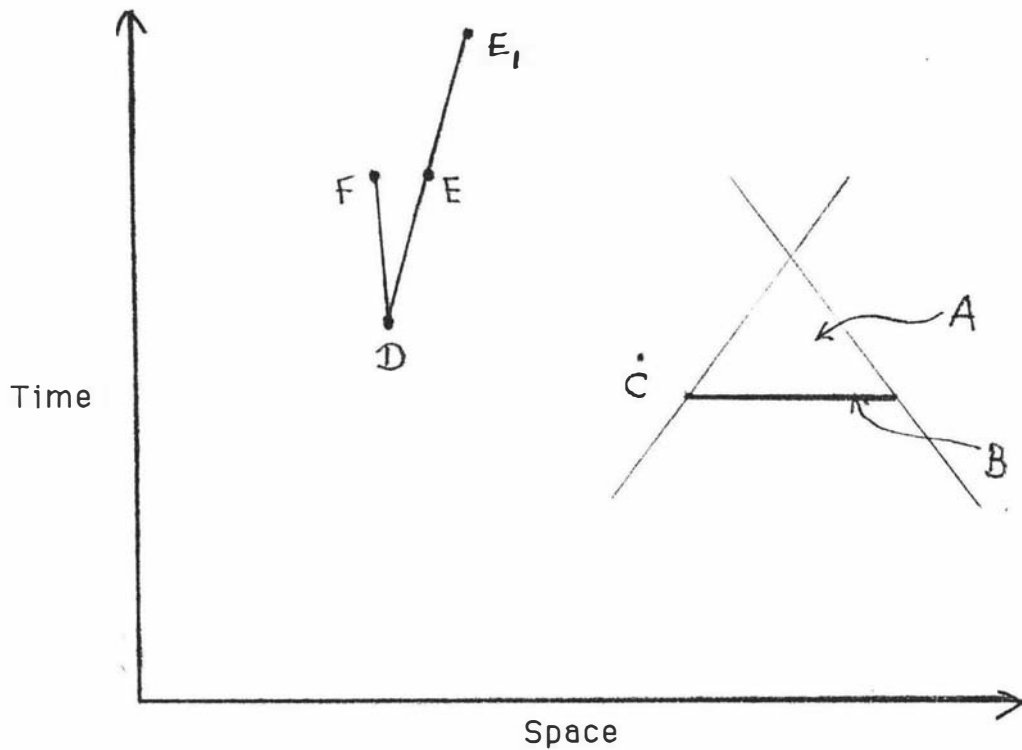


FIG. 1.1

A is the region within the lightcone that converges from the space-like hypersurface B. Relativity theory requires that the events in region A depend solely upon the events at the hypersurface B, since no influence from an event such as C outside of B could propagate to A. E and F are space-like separated events which are highly correlated. (For instance, they may represent the correlated contents of two copies of a newspaper). The correlation is 'across space', but it results from a common causal ancestor, D. E and E₁ are likewise highly correlated (for instance, they might represent the correlated contents of the same physical newspaper at different moment in time), but in this case the correlation results from the lawlike temporal evolution of the system involved. There is no spatial analogy to temporal evolution.

makes us overlook a far more obvious difference. The obvious difference is that there is a connectivity of objects through time that makes events in a later space-time region *depend upon* the events in earlier regions. For any space-time region, there is a future region which depends *only* on the earlier region. (See Fig. 1.1).

There is no such connection between spatially adjacent regions. What is the case in a given region generally has no bearing on what is the case in neighboring spatial regions. (See Fig. 1.1. If there is a relation across space, it is a contingent one. E.g. there may be a single cause in an earlier region of two correlated effects in spatially adjacent later regions. But the correlation is the result of the accidental contents of the world, and is generated in the first place by the temporal dependence that each region has on the earlier region.)

Now it is this *connectivity of systems through time* that is at the heart of all dynamic laws. That is why the laws of physics are dynamic, and time has a unique role in them. That is why Hume's problem of induction concerns induction to facts about later times, not to facts about left or right regions of space. The assumption of this connectivity through time is an overtly metaphysical foundation of physics, as well as of everyday understanding.

I have said that what happens in one region of time *depends upon* what happens in adjacent past regions. The irreversibility of quantum theory that I argue for means that the relation is not symmetric. I.e. the future depends (probabilistically) on the past, but what happens in the past is *not* dependent on what happens in

the future. This is what the asymmetry of probabilistic dynamics amounts to. So we have temporally asymmetric dependence.

This could naturally enough be interpreted as *causal dependence*.³¹ It need hardly be argued that a belief in real causation goes hand in hand with belief in time flow. And surely physicists in their informal moments appeal to the idea of causation all the time - it is only when they become philosophically sophisticated, and perhaps forget some of the most obvious lessons of their lives, that they are inclined to cast aside the ideas of causation, and time flow.

So we can say at least this much: if we are to interpret quantum theory as involving time flow, then the probabilistic dynamics is the obvious and natural arena for time flow. If there is to be any quantum theory of time flow, then the theory proposed here, or some close variation, is the only natural theory. And it does appeal in a natural way to the untutored instinct.

1.23 Dynamic probability.

Let us turn to a more specific point, the understanding of probability. In the ordinary understanding of quantum theory, probability is taken in a very strong way. This is expressed by saying that the result of a probabilistic quantum theoretic experiment (in general a measurement) is *indeterminate* until the experiment is finished. Or alternatively, that the *truth* about the result of an objectively probabilistic future event does not exist until the event happens.

³¹Obviously with a probabilistic sense of causation.

For instance, say we are going to measure the spin on the z-axis of an electron which has been prepared with spin up on the x-axis. Quantum mechanics tells us that there are two possible outcomes, *spin-up* or *spin-down*, and that there is an objective 0.5 chance of either result. Using e_+ and e_- respectively to denote the events of observing spin-up and spin-down, we have: $\text{PROB}(e_+) = \text{PROB}(e_-) = 0.5$.³²

Now it is normally thought that there is simply no answer to the question of whether e_+ or e_- is the result of the experiment *until* the time comes around and one or other event *does occur*. The natural picture is that when the time comes around, Nature must make a probabilistic choice, and, bingo, one or other event occurs. Until the 'bingo', there is simply no reality to either result.

This kind of probability may be contrasted with another sort of probability, which is epistemologically based. Imagine that, depending on the result of this first experiment, either a white or a black counter is placed in a sealed box. If e_+ is the result, a white counter is inserted, if e_- is the result, a black counter. This may be done mechanically, so that no one has any knowledge of which counter is in the box.

After the counter is inserted, we take the box, and rattle it to demonstrate that there is a counter in it, and ask ourselves what colour we would guess the counter to be.

Let us denote the event of finding that it is white (when we

³²These probabilities strictly need to be further conditionalised on the probability of performing the experiment at the future time. For simplicity it may be assumed that it is already physically determined that the experiment will be performed at the future time.

look at it sometime later) by e_W , and of finding that it is black by e_B . We would clearly assign probabilities of 0.5 to each event, i.e. hold that: $\text{PROB}(e_W) = \text{PROB}(e_B) = 0.5$. For we would certainly assign these probabilities before the result of the first (e_+/e_-) experiment was determined, since until then, the probabilities of e_W and e_B would be fully objective, being identical to the probabilities of e_+ and e_- . After the e_+/e_- experiment is completed, there is no effective change in our knowledge concerning e_W/e_B , so we must retain the 0.5 probabilities.

These probabilities have clearly been assigned by an objective, rational procedure. But although they are objective in a clear sense, they are not like the probabilities governing the events e_+ and e_- , because they imply no *indeterminateness* in the events e_W and e_B . There is no lack of an actual truth of the matter about whether e_W or e_B : only a lack of knowledge of what the truth is. We know that there is a definite counter of a definite colour in the box, that Nature has already made its choice about which of e_W and e_B will happen in the future.³³

For this reason, the second kind of probability is called 'epistemic' or 'subjective', because it depends upon a lack of knowledge, not upon a real indeterminateness in the facts. The first kind of probability is *physical* in a stronger sense because it involves a real physical indeterminateness. There seems to be really no truth of the matter in the first case, until Nature makes a real, probabilistic decision, and either e_+ or e_- is actualised.

³³Note that e_W and e_B are the (future) events of *observing the counter to be white/black* - not merely the event of it *being white/black*, which is a present event.

In this way of describing the physical probabilities of quantum theory, there is an obvious and intimate connection with the flow of time. This can be seen in the appeal to the temporal categories of existence - future, present and past - in the description of the *actualisation* of these probabilities. e_+ and e_- were first of all mere *future possibilities*; neither was yet actual. Time had to move on, one or other result has to be *actualized*, before the result became a present reality.

It is not merely that e_+ and e_- pertained to moments *later than the present* that generated this, for before the second experiment was completed, e_W and e_D (which are the events of *observing* one or other colour of the ball) pertained to later moments, but they were nevertheless already physically determined in the present. Clearly there can be presently determinate facts about later (or earlier) moments of time. We talk about such things all the time: "Hitler died in 1945"; "The sun will come up tomorrow morning", and so on, believing these to be determinate facts *now*.

We have been talking of 'determinateness' without any decent definition of it. For the moment let us take 'is determinate' to mean 'has the *present* modality of existence'. What was felt about e_+ and e_- was that for a time, neither had any *present reality* - until something happened, and one or other was realised in the world. This is not in conflict with the idea that one or other of e_+ and e_- was always the result which *would* be actualised. Either: *e_+ will occur* or: *e_- will occur* can be taken to have been always true. Say for instance that e_- turns out in fact to be the result of the experiment, realised at a moment t_2 . Then at an earlier moment t_1 , it was true to say that

e_- will occur at t_2 (and e_+ will not occur). This is tantamount to the modal claim: e_- -at- t_2 has the future modality of existence.

But this does not require that e_- -at- t_2 has the present modality of existence. We cannot, in other words, go from x will exist (or occur) to x does exist, or from x will be true to x is true.

In fact this is an issue that goes very deep, as will be evident when the matter is fully aired in Chapter Seven, and the few comments above do not serve to clear any of the real problems up. But they are meant to establish for the moment only that ordinary intuitions about quantum probabilities seem to rely on the concepts of temporal modalities of existence, or in other words, on the supposition of time flow.

I will call this ordinary understanding of probabilities the *dynamic* interpretation of probability. I am not arguing that this is the only way that objective probabilities can be interpreted. In fact I will later contrast the 'dynamic' view of objective probabilities with a 'static' or 'bloc universe' interpretation, to make the competition clear. All I am observing is that the dynamic interpretation is *prima facie* attractive to physicists, and could plausibly be claimed to lie behind the implicit understanding of probabilities used most of the time in practical quantum theory. In other words I am showing that proposition [1.10] ("This interpretation captures the implicit understanding of probabilities that physicists adopt in their ordinary use of quantum theory.") is plausible.

Another feature of the ordinary view that strongly reinforces

this conclusion is the *prima facie unimagability* of past-directed objective probabilities. Above, I have described quantum theory as lacking past-directed probabilities. Now this claim may have struck some readers as bizarre in itself, because it may seem unimaginable what *past-directed probabilities* could be. It may be thought that I really do not need to take much trouble to empirically disprove the existence of past-directed probabilities, because such things are a conceptual impossibility. They cannot even be imagined.

Why should they seem impossible or unimaginable? I think it is because the kinds of objective probabilities we are talking of are assumed to be *dynamic*, and hence to depend upon time flowing forwards into the future for the probabilistic outcomes to be actualised and become real. The idea of past-directed probabilities of this dynamic variety is certainly inchoate, since there is no way that such probabilities could be actualised.

But this is only because probabilities are imagined as dynamic, for on the alternative 'static' or 'bloc universe' interpretation, there is no problem at all with imagining past-directed probabilities. Such probabilities may be manifested in certain objective statistical relations from later to earlier events. (These relations are empirically found to be absent, and are not required by quantum theory, which is why quantum theory is irreversible whatever interpretation of probability is adopted.) Since on the bloc universe conception, past-directed probabilities are at least *imaginable*, whereas in the popular conception they are not, the bloc universe view cannot provide the popular conception of probability.

I have made an attempt to show the plausibility of [1.10] (that the dynamic interpretation captures the implicit understanding of physicists.) This is about all I will say in defence of it. To show [1.10] conclusively would not achieve much even if it was possible to do so. Even if every living physicist voted in favor of dynamic probabilities, opponents of this interpretation of physical probabilities need not take any notice, if they have good objective arguments against it. They need only say: Yes, but the physicists are just philosophically unsophisticated. So while it is of some polemically importance that [1.10] is plausible, it is hardly a key issue: what is crucial is the quality of the arguments that the proposed interpretation is *correct*.

A vital step is to show the conceptual coherence of my proposal. What has been already been said goes some way towards this, showing that the idea that quantum probabilities are *dynamic* probabilities fits in very naturally with ordinary ideas. It may still be objected, though, that the very idea of *dynamic* probabilities is internally incoherent. The source of such objections are mainly arguments or presuppositions that the concept of time flow is already incoherent. Since dynamic probability already presupposes time flow, if one rejects the latter as an incoherent notion, so will one reject the former. It is vital that I overcome such objections.

I will argue for the coherence of time flow and dynamic probability by showing how a formal model of them can be constructed. This is done in Chapter Seven.

1.24 The motivation for realism about time flow.

I turn to the last key piece of my argument that still needs to be filled in: the motivation for introducing time flow in the first place. This must be made clear if we are to establish [1.9], that no other interpretation does justice to time flow (nor has advantages in other areas that could compensate for this lack).

The theory of dynamic probabilities deals with time flow by making it a real feature of the physical world. The main alternative, the bloc universe view of the illusionists, deals with it by saying that there is no real time flow, and the impression or sensation of it can be explained away in an effective way. Obviously there is only an advantage in the dynamic view if there are good reasons for wanting to take time flow realistically in the first place.

These reasons cannot come from physics, of course: they are required to provide our motivation for taking time flow seriously *in* physics. They must rest instead on pretty direct phenomenological grounds.

It is worth stressing that this is a normal kind of situation in science. An exactly comparable situation would arise from the question of whether the existence of *physical objects* should be reflected in the interpretation of a certain theory. It is no doubt possible to reinterpret almost any physical theory in an idealist way, or even a solipsistic way, and say that it does not imply the existence of any external objects. We could hold that there is no real external physical world at all, just sensations, or some such mental things. A solipsist might reinterpret quantum theory or

Newtonian mechanics to support this view. On the reinterpretation, the main body of theory would become a mere formalism, serving only to express complex relationships between 'observations' or some such thing. This is indeed the program of some instrumentalists.

But such a radical interpretation is normally rejected, because the phenomenological evidence appears so overwhelmingly to favour the postulate of an external, physical world. There is no *logical necessity* to do this, but the ordinary canons of informed common sense convince most of us. Thus we ordinarily interpret our scientific formalism in a specific way: we take certain kinds of terms to denote physical objects.

Reasons of the same sort for taking time flow seriously enough in the first place are required. In fact we need not look far. The apparent usefulness of the time-flow concept in interpreting so much about our world, and the thoroughness with which tenses infiltrate all our natural language, would probably be enough to convince most people that if there is a good scientific way of incorporating time flow, it should be had. *Prima facie*, at least, there seems somewhat more of a need to justify abandoning the idea of time flow, than a need to justify taking it seriously.³⁴

There is, then, a sufficient *prima facie* motivation for the project of dynamic probability, without the need to say very much at all. But while the initial motivation for the project is hardly under question, it would not be very satisfactory to leave matters

³⁴It must be remembered that, if a good scientific way of having time flow is found, the Illusionist's arguments against time flow are dismissed, since they are essentially arguments that we cannot have such a scientific account of time flow.

here. It seems that there must be reasons why we are convinced that there is time flow, and it will always be necessary, in the end, to analyse them, and to decide whether they are good reasons or bad ones. If they really are bad, then, despite all our strong feelings that time flow is desirable, there would be no ultimate need or justification for bringing it into the scientific ontology.

So this introduction concludes with an exploration of what these reasons could be. A positive argument from phenomenological evidence for the reality of time flow is given. This argument is intended, at the same time, to be a diagnosis of our deeper reasons for believing in time flow.

I must acknowledge that this final argument is far from complete. It is really only a sketch that needs a much fuller development. However, the weakness of this particular part of the case is hardly crucial for the project as a whole, and it seems better to give a partial, and no doubt faulty, analysis, than to avoid the issue altogether.

1.25 The concept of existence 1.

The view that time flows is really a view about the nature of *existence*, as has been sufficiently emphasized. The dynamic view of existence is that there are three modes of existence, past, present and future. The alternative view of the Illusionists is that there is only one mode of existence: existence pure and simple. Whatever exists for the Illusionist exists timelessly. This is often called the *bloc universe* concept of existence. The 'bloc universe' consists of the whole collection of events or states of affairs that occur throughout the actual course of history. This is an object extended in time as in space. There is no special moment within the universe which is the 'present' (just as there is no special place which is 'here'), and the categories of past, present and future are not real according to the Illusionist. They can be thought of only as perspectival effects (or illusions) apparent from particular points of view within the universe.

Let me first dispel an attempt to deflate the dispute into a mere verbal disagreement. It might be objected that the dispute is merely over two different possible meanings for the term 'existence'. Why not be diplomatic, allow the Illusionists their concept of existence and the proponents of time flow theirs, and merely observe that they are using the term 'existence' in different senses. So long as they can each adequately translate the other's concepts into their own terms, there will be adequate agreement between them.

The proponents of time flow can easily translate the

Illusionist's 'existence' into their terms. *x exists* in the bloc universe lingo means *x has, does, or will exist*. But can the Illusionist translate the concepts of *x has existed*, *x presently exists*, and *x will exist* into bloc universe terms?

Reichenbach [1947], Goodman [1951], and Quine [1960] have supported such a theory of translation, which has proved very popular. The statement *x is past* is taken to mean *x exists earlier than this token utterance*. Similarly, *x is present* is taken to mean *x exists simultaneously with this token utterance*, and *x is future* to mean *x exists later than this token utterance*.

In these translations, the term 'exists' must be taken as signifying timeless, bloc universe existence, while 'is earlier than' and 'is later than' are merely temporal relations, similar to spatial relations like 'is north of' and 'is south of', and have no modal significance. (Perhaps it is worth saying that 'x exists earlier than...' means: 'x exists *and* x is earlier than...' as opposed to: 'x exists-earlier-than', with 'earlier than' modifying the type of existence.)

The argument now goes that these translations are sufficient for the conveyance of all substantial items of knowledge. For instance, to say that *x is past* conveys no more nor less than that *x exists at an earlier moment than the moment of utterance of this token*. Similarly, *x is future* conveys that *x exists at a later moment than the moment of utterance of this token*. Thus, what the disagreement is over is nothing substantial but merely the use of the term 'existence'. The proponents of time flow take there to be three distinct kinds of existence; the Illusionists take

there to be just one kind, which can be stratified into three groups relative to any moment in time. Each can say in his own language everything that can be said in his opponent's.

But there is an obvious objection to this. It must simply be denied that *x is future* means the same as *x exists later than the moment of this utterance*. When we say that *x is future* we intend to say something about *x*'s existence: viz, that it does not yet exist, but when some time has passed, it will exist. (This is the very meaning of the term 'is future'.) This implies, for instance, the reality of time flow. But the statement *x exists later than the moment of this utterance* has no such implication.

The Illusionist must change tack a little. A much more reasonable claim is that the statement *x is future* conveys at least the information that *x exists later than the moment of this utterance*; and that although it may seem to convey something more, something about 'time flow', in reality it doesn't convey any more real information because there simply is no time flow. (The whole idea of it is incoherent.) Thus the information that is effectively conveyed by *x is future* is just that *x exists later than the moment of this utterance*.³⁵

But of course this reply takes us away from the claim of a mere 'verbal disagreement', and back into a substantial dispute about whether time flow is real or not. For the proponent of time flow holds that time flow *is* real, and that the Illusionist is simply wrong in saying it isn't. And it is this dispute that must be resolved before any agreement can be reached.

The Illusionist might try a second tack to switch the debate

³⁵E.g. according to Goodman, a token sentence of the type *World War II is past* "tells us simply... that World War II is prior to the sentence in question."

to one about meanings. Let us say that it is admitted that there is no translation from tensed language *as it is intended to be understood* into the non-tensed language of the Illusionist.³⁶ This is because tensed language presupposes time flow, and time flow requires at least one primitive concept that is not available to the Illusionist (e.g. the concept of 'real change').

Nevertheless, it might be claimed that the term 'existence' just has the meaning that the Illusionists ascribe to it, namely bloc universe existence. The terms 'is past', 'is present', and 'is future' do not denote modes of existence at all, because it is a conceptual truth that *existence* means bloc universe existence, not 'temporal existence'.

The trouble is that this is totally implausible. For practical purposes, the core sense of *existence* is *temporal existence*. For instance, if you say 'Napoleon Bonapart exists', or 'My experience of the taste of baked beans exists', you are taken naturally to be saying that these things exist *presently*. 'Napoleon Bonapart exists' is false according to common sense and common usage: 'Napoleon Bonapart *did* exist' is true. Common sense takes existence to be temporal existence: it seems very unlikely that Illusionism could be established on purely conceptual grounds.

³⁶I reserve the term *tensed proposition* for propositions containing real tenses (besides the present tense), for instance: that *Polly was eating a lolly*, that *Polly will eat a lolly*, that *Polly ate a lolly in the past*. Propositions may also be indexed with the times, e.g. that *Polly is eating a lolly* may be modified to that *Polly is eating a lolly at 12 o'clock*. I will always call the latter a *time indexed* proposition, not a tensed proposition.

1.26 The concept of existence 2.

It is interesting to consider the concept of existence further. I will begin the positive case for time flow with some comments about how the concept of existence is learnt.

Everyone has a concept of existence, and it is obviously a pretty fundamental idea. How do we learn it? It would seem that we learn it by being able to contrast things that *don't* exist with things that do exist. After all, if there was nothing that we could contrast with the things (objects, experiences, states-of-affairs) that do exist, what practical use would the concept be?

The Illusionist can easily allow that there are plenty of things that don't exist - all the nomologically possible but unactualized worlds. But we are not in perceptual contact with these non-existent things, and surely we don't *learn* our concept of existence by contrasting the actual world that does exist with other worlds that do not.

Rather, we learn about existence because things (or states of affairs) come in and out of existence within the actual world. Coming in and out of existence happens in time. In our normal experience we are all certain of this fact, that things come in and out of existence in the world.

Experiences themselves are the primary candidates. Babies undoubtedly get hungry, and they must come to recognise that the experience of hunger comes and goes quite regularly. How they recognise this is of little concern here, and what I am saying does not suppose some kind of 'private language' theory of terms for internal states or anything like that. All I am saying, which is

obvious, is that babies learn to recognise the experience of hunger, and as they develop memories, they must recognise that hunger comes and goes. They come to recognise that all sorts of experiences come and go, that indeed there is a flux of experience or sensation, not a single unchanging experience. If there *was* but a single unchanging experience it is doubtful that the concept of existence itself would be a part of it. (It would certainly not be a concept that would *develop*, since nothing would change).

All I wish to say is that *we would not acquire the concept of existence unless things did at least seem to come in and out of existence to us*. And therefore our concept of existence allows for things coming in and out of existence, and is thus a concept of *changing existence*, or *temporal existence*. This is the primary and natural concept of existence.

I make this point to try to bring the Illusionists down to earth somewhat from the heights of abstract speculation. The Illusionist's view of existence is a highly abstract one, gained from attempts to make mathematical or logical models of the world. These models are made mainly for the sake of giving precise semantic theories. In mathematics, as opposed to real life, a static concept of 'existence' is certainly common. Mathematical Platonists believe that there are realms of eternal mathematical objects, which exist in the static, bloc universe way. I believe the Illusionists come to transfer this notion of existence from their mathematical models of the world, to the real world itself (the fallacy discussed in Appendix 1.1) But the physical world seems to be an entirely different thing from the realm of mathematical objects. It is different particularly

because it exists *in* time, because it is not static but dynamic.

1.27 A positive argument for the reality of time flow.

So far the points made have been largely negative, against the Illusionist's objections to time flow. It is time for a positive argument for the reality of time flow. In a general sense the argument is an empirical one, i.e. it is argued that there is empirical evidence supporting time flow. But the inference is a rather abstract one. The strategy of the argument is sketched in this section, details are given in following sections .

The argument is an empirical one, taking certain empirical observations as premises, and concluding with the proposition that time flows. The empirical premises consist, roughly, of the observation of the *continuity through time of one's personal experience, and of the world of physical objects*. The meaning of this will be spelt out in detail shortly, but for the moment it will be symbolised by the term *CONTINUITY*. The conclusion is symbolised as *TIME FLOW*, and hence the inference as: *CONTINUITY* \Rightarrow *TIME FLOW*.

The key problem, not surprisingly, will be to justify the inference. The inference depends upon the fact that the conclusion (*TIME FLOW*) *explains* the premise (*CONTINUITY*). There is rather more to it than that, but let us begin with this idea of an inference which is legitimated by the fact that the conclusion explains the premise. This kind of inference is not unusual, as we can see by considering a concrete example like the following. Suppose that we observe a crater, and scattered around it fragments of metallic rock, and other features which all suggest

that a meteorite had crashed. Let us call this evidence *DEBRIS*. Clearly, in ordinary circumstances we would feel justified in inferring from *DEBRIS* that a meteor had crashed on the site, that is we would make an inference appropriately symbolised: *DEBRIS* \Rightarrow *METEOR*. What justifies this inference the fact that the hypothesis *METEOR* provides such a good *explanation* of the premise *METEOR*. Of course, other features are also necessary: for instance, there must be an initial plausibility to the idea that a meteor has crashed there, and also there must be a lack of plausible competing explanatory hypotheses. However it is clear that inferences of this kind are commonly taken to be legitimate.

The inference: *CONTINUITY* \Rightarrow *TIME FLOW* is not as straightforward as the: *DEBRIS* \Rightarrow *METEOR* inference, however, because the explanation it is based upon is not a simple causal explanation of one contingent event from another, as it is in the latter case. It is rather a very high-level theoretical (I am inclined to say 'metaphysical') explanation. Rather than trying to offer a precise analytical justification of the *CONTINUITY* \Rightarrow *TIME FLOW* inference, I will justify it instead by analogy. The analogy I will draw is with an argument for a similarly 'metaphysical' conclusion. It is an argument which, in fact is controversial in philosophical circles exactly because no satisfactory analysis of the inference has been given, but it is nevertheless an argument which is widely accepted, and forms the basis of scientific realism. The argument is for the thesis that there is a real, concrete external world which exists independently of our experiences of it. I will signify this conclusion by the term *REALISM*. The assumption of *REALISM* is a kind of metaphysical

underpinning of both our commonplace understanding of the world, and of the normal scientific view. Do we have any good reasons for REALISM? I think we do: the reasons are roughly that REALISM *explains* so many features of our experience – what I will summarize in the following section as *the coherence of experience*, and will symbolise here as COHERENCE. COHERENCE is broadly speaking an empirical premise. Hence I think that: COHERENCE \Rightarrow REALISM is a good argument, and that what justifies the inference is the fact that the conclusion REALISM *explains* the premise COHERENCE. We will see the details of this in the following section, but the point here is the analogy between the inference: COHERENCE \Rightarrow REALISM, and the inference: CONTINUITY \Rightarrow TIME FLOW. I will represent these as inferences of a very similar kind: both have broadly empirical premises and rather 'metaphysical' conclusions, and the inference is legitimated through the fact that the conclusion provides an explanation for the premise. Thus I will effectively argue that we have the same kind of empirical reasons for believing in TIME FLOW as we have for believing in REALISM.

Of course many philosophers object to REALISM, and perhaps the argument for it has no ultimate justification at all. But this is a question I will not try to decide. If one wishes to reject REALISM, then one has taken up a position of radical skepticism about the very existence of the physical world, and from this point of view the whole project of trying to establish a realistic interpretation of time flow within physics is doomed before the Table of Contents has been reached.³⁷ However, if it is accepted

³⁷But the Illusionist's view is undercut in exactly the same way of course, so this kind of skepticism does not represent not an Illusionist objection

that experience offers us good reason for REALISM, then I will argue that experience offers us reasons of a similar force for TIME FLOW. This argument represents a challenge to the presumption of the Illusionists that there can be no empirical reasons for belief in time flow.

1.28 Realism vs. Illusionism about the physical world.

That the external world is real means, I take it, that there are durable physical objects which exist independently of our own perceptions of them. Now there is a well-known species of 'Illusionism' about this view, normally called Idealism or Phenomenalism. It arises from asking how we *know* that there is an external world. Apparently, through our perceptions or sensations of it. But do these perceptions or sensations really count for anything? What if we could explain away all the perceptions as illusory, without ever appealing to the reality of external objects in our explanation? Then they would seem to provide no good reason for believing in the external world after all.

The Idealist does exactly that. Let us suppose that all that exist are the percepts themselves. They are not caused by anything external, and provide no information about external objects.

Of course we still need to account for the structure of percepts, for they are obviously highly structured. Indeed they cohere perfectly, as far as we can tell, with the predictions of our complex scientific theories. But this structure is not hard to

account for. We simply take a thoroughly Idealist interpretation of the scientific theories. We say that they provide true accounts of the structure of our percepts or observations. They do this in a complex way, using a rather indirect formalism to generate the predictions. People commonly misinterpret this formalism, and take it to be about something real (the subatomic world of microprocesses, for instance), but that doesn't matter. Our Idealist theory does just as well as the traditional Realist theory, so far as pure experience is any guide.

So why don't we (most of us) take this kind of Idealism seriously? It seems that it is because we feel that the postulate of durable physical objects has some kind of powerful *explanatory* value, which the Idealist view cannot provide. Take, for instance, my present experience of looking across the meadows at some cows. I look at the cows, and the cow-percepts arise, then I look away, and the cow-percepts cease. Whenever I choose to look back, the cow-percepts reappear. This is a fundamental kind of regularity in my percepts, which obtains in a thoroughly systematic way among all my senses. My percepts cohere into a highly structured whole. What is the nature of this coherence? The simplest way we have of describing it is to say that it is *exactly as though* I have real perceptions of durable, external objects, which exist in a three-dimensional space, and move about it in a continuous way, and so on.

The postulate of external-objects-which-are-perceived (which I will call the Realist postulate) seems to explain the coherence of percepts. How does it explain them? Or simpler: what *difference* does this postulate make? Believing in it seems

to take away the mystery that is otherwise felt about the coherence of percepts: but how is the mystery removed?

Obviously because, if it is true, and the kind of world we are in is indeed a physical world of durable objects then the coherence of percepts is more or less *necessary*. The possibility of things being otherwise does not exist. There is much less mystery why I see the cow again when I look back at it if it *is* a durable object which I perceive.

The postulate of external objects is an ontological postulate about the *kind* of world the actual world is. In possible world terms, this is equivalent to placing a constraint on the space of (nomologically) possible worlds to which our world belongs. Not everything is physically possible that at first seems to be possible: in particular, there is no real possibility for our percepts *not* to cohere more or less as they do (except under special conditions, when we dream, hallucinate, etc.)³⁸

³⁸The idea that our world comes with a space of real physical possibility is implicit in the project of fundamental physics. The project is to plumb the nature of the fundamental ontology, or in other words, to determine what the space of *real* possibility that comes with our world is. This is easily enough seen by imagining an experiment that is performed to determine whether or not a certain kind of fundamental particle 'exists'. The question is really whether a particle of a certain kind is *possible*. Suppose that it requires very special conditions for a particle of this kind to exist, even conditions which have never been realised in the universe before. We experiment, to decide the question, by generating these special conditions: if the particle is a physical possibility, it will be produced in the experiment, if it is not a physical possibility, it will not be produced. Now surely the universe already 'knows' whether the particle in question is possible or not before we do the experiment. It seems only natural to consider it an objective fact about the universe that the possibility existed all the time, or that it never existed. In doing the experiment we are only discovering the truth. This at least is

On the Idealist account, the alternative possibilities are not ruled out. It is *possible* that I first have cow-percepts, and then spontaneously have dragon-percepts, and then the sky rains with pink angels for three minutes, and then.... There is nothing in the idealist ontology to rule out these possibilities. Thus it is a mystery why they do seem to be ruled out in fact.

The Idealist might object to this last judgement, and say that the inappropriate possibilities have been ruled out after all. They are ruled out by the further postulate that *percepts in fact cohere as though they were really perceptions of the external objects proposed by the Realist*. By adding this second postulate, the Idealist effectively gets all the predictions about the structure of percepts that the Realist gets.

The problem is whether the postulate really provides an *explanation* of the phenomenon. It is a statement to the effect that percepts *do* have such-and-such a structure. But we do not explain something by merely stating that it is true.

We are not dealing with a causal explanation of a particular event, but a theoretical explanation of a law-like regularity. Other examples would be the explanation of the regularity expressed by Boyle's law by appeal to the molecular behaviour of gasses; or the explanation of the electrical conductivity of iron in terms of the electron structure of iron atoms. A simpler and more picturesque example is given in Hung [1978], which concerns the explanation of the apish behaviour of the people that we see in mirrors. We all know that when we look into a mirror, we see a person in there doing various things. Strangely enough, the person the presumption of physics, and it seems to be confirmed by the discovery of real structure in the universe.

apes exactly what we ourselves do. Mirror people do not seem to have the freedom to do anything else.

We explain this by changing our ontological view of what the 'people-in-the-mirror' really are. We postulate that they are not *people*, with their own wills and thoughts, as they first seemed, but merely images formed by the geometrical behaviour of light. It then becomes clear why the regularities exist. Changing the ontology in this way makes it impossible for the 'mirror people' to do anything else but ape our own behaviour. The freedom that appeared to be there at first when we thought there were real people in the mirror – the normal freedom of people to move about at whim – does not really exist.

This kind of explanation is sometimes called a theoretical reductive explanation of a regularity, as opposed to a simple causal explanation of a particular event. It seems likely, however, that a theoretical reductive explanation might be regarded as just providing a general schema for giving particular causal explanations. For instance, having the reductive explanation of mirror-behaviour, we are in a position to provide a causal explanation of any particular instance of mirror-behaviour. If a child asks us on some occasion why the other child in the mirror is waving her hand, we explain that it is because she herself is waving her hand, and the light reflects from the mirror and lets her watch herself doing this. This is a causal explanation of a particular event. The form of the explanation is provided by the larger theoretical reductive explanation, which says effectively that we can always explain particular mirror-behaviour in terms of the behaviour of physical objects and the

reflection of light.

Now let us return to the explanation of the coherence of percepts. The Realist postulate provides a theoretical reductive explanation of the coherence. For any particular instance of a coherence of percepts, we can identify the real objects in the world, and the observers causal relations to the objects (e.g. receiving light rays from the objects into the retina...) , and we construct a good causal explanation of the particular behaviour of the percepts. The Realist postulate provides a general schema for causal explanations of particular events, and it seems a good theoretical reductive explanation.

Does the Idealist explanation achieve the same? Not at all. The key postulate in the Idealist explanation is that *percepts in fact cohere as though they were really perceptions of the external objects proposed by the Realist*. This allows *predictions* of the coherence of percepts in particular cases, but it provides no scheme for finding causal explanations for the particular cases. If I ask: "Why do my cow-percepts cohere so well in *this* case?", there is no causal explanation forthcoming from the Idealist. There is just the observation that percepts *do* cohere (entailed by: "Percepts in fact cohere as though they were really perceptions of the external objects proposed by the Realist"). That is not a causal explanation, for no *causes* have been established.

The key point made above remains: the Realist postulate is about the ontological nature of the world, and it is such that it limits the possibilities of the behaviour of percepts so that there is no (nomological) possibility of them not cohering. What

seemed like a host of other possibilities - for instance, all the bizarre behaviors of percepts that are possible when one is dreaming, and for which the Realist postulate palpably fails - are not really possibilities at all.

The Idealist theory on the other hand leaves the bizarre possibilities as real, and leaves us in mystery about why the observed regularities hold.

Of course there are many who would object to this account, because they think that there is no real causation in the world, and consequently no genuine causal explanations. (They are illusionists about causation). These people may think that the Realist account does no better than the Idealist account, because the apparent 'causal explanations' of the Realist are illusory.

But my point here is merely to try to bring out why it is commonly *felt* that Realism is better than Illusionism, not whether it really is better. Why are people so inclined to realism about external objects? My answer is: Realism is *recognised* to provide for the explanation of the coherence of experience. Idealism is not.

Whether the proffered Realist explanation is really any good, or ultimately any better than some Idealist explanation, is another question, since my concern will only be to draw a parallel between the reasons for believing in the external world, and the reasons for believing in time flow. If there are good reasons for believing in the external world, then there are similar kinds of good reasons for believing in time flow. If one rejects belief in the external world as unjustified, then the argument for time flow will have little impact, for such radical skepticism undermines the very foundations the project.

1.29 The analogous argument for time flow.

I will now argue that the postulate of time flow explains a great deal that the Illusionist theory does not. Indeed, it does more than just explain: it provides the metaphysical basis for our whole interpretation of the world. It provides the metaphysical basis for the whole idea of *causal explanation* itself, for instance. It provides the metaphysical basis for the view that *present experience informs us about happenings at other times*, without which we would have no justification at all for believing in time (as a dimension of the world). In brief, it provides for a whole spectrum of interpretive principles we adopt quite universally and unthinkingly, and without which we simply would not be able to begin to make sense of the world. Hence the metaphysical view that time flows really has immense implications. The static view of time, I will argue, if seriously adopted, would undermine our normal interpretation of the world in a radical way. The Illusionists are content with it only because they have not perceived what a radical metaphysical view it is.

The idea that the hypothesis of time flow explains something serious is *prima facie* plausible, simply because of the overwhelmingly strong belief in time flow by almost everyone. Despite the view of the Illusionists, it actually seems very difficult to imagine that time does *not* flow, (that existence is not truly dynamic), and it is plausible that this is because the postulate of time flow accounts for something substantial.

What does the postulate of time flow account for? What does

it explain? What difference does it make to our conception of the world? Imagine that there really is no time flow: what features of the world become mysterious or strange or inexplicable? My suggestion is, very generally, that time flow explains the apparent *connectedness of the world through time*. If we do not adopt a dynamic metaphysics, then any connectedness of the world through time appears mysterious. Indeed, without a dynamic view of existence, we are not justified in interpreting our present experience as implying any kind of connectedness through time.

At least two kinds of 'connectedness through time' are very important to us: (i) the continuity of *experience* through time; and (ii) the continuity of *physical objects* through time. Consider these in turn.

Experience has a definite temporal structure. In the first place, the *content* of experience seems to come in 'time slices'. That is, there seems to be a more or less definite *content of experience for each moment of time*.³⁹ This is most evident in

³⁹'Content of experience' is left as an unanalyzed primitive here. A more thorough analysis would require an extended treatment of the idea. Since we are considering the *phenomenological* evidence for time flow, a certain amount of description of the phenomenological world, couched in such phenomenological terms as 'the content of experience', is unavoidable. Note that the claim that *there is a distinct content of experience for each moment of time* is an empirical claim. The 'content of experience' of an agent is not merely *defined* to be the instantaneous, or momentary, bundle of experience of that agent: the situation is that the 'content of experience' comes in instantaneous or momentary bundles *in fact*. Thus there is no circularity in the claim that the content of experience is momentary: it is a substantial empirical claim (which in fact might only be approximately true). Graham Oddie has pointed out that analysis here of the temporal structure of experience bears some close resemblances to Kant's analysis of the threefold synthesis in his transcendental deduction. See Kant [1970] pp.131-138.

the part of experience we consider to be direct perception of the external physical world. For instance, at 11.45 a.m., I am sitting at my desk, and my visual field is filled with a representation of a number of books, papers, coffee cups, a brown table surface, a green wall, and so forth. Let this be called visual field F_1 . At 12.15 p.m., some time later, I am sitting at a patio table outside and my visual field is filled with representations of green grass, a variety of trees, some cows across a meadow, and a plate of baked beans within my reach. Let this visual field be called F_2 . The two visual fields, F_1 and F_2 , seem to *exist* quite distinctly and independently: in particular, they do not both belong to any *single present content of (my) experience*. They belong to quite distinct contents of experience.

Present experience does have internal parts; for instance the visual field F_1 contains a flat vertical green part above a flat horizontal brown part, and some colorful rectangles amidst the green. But the present content of experience does not have distinct *temporal* parts. Experiences at distinct moments never belong to a single content of experience. This seems to be a fundamental phenomenological feature of experience, and is what I intend to convey by the expression *experience comes in time-slices*⁴⁰.

Now the remarkable thing is that, although the particular content of experience is momentary, or near enough to being

⁴⁰It may be that a given content of an agent's present experience actually has a certain thickness or duration in time, rather than existing at a point-like moment. (See Capek [1970]) But if there is such an extension, it is certainly very small - of the order of a tenth of a second at most - and experiences such as F_1 and F_2 which are separated in time by a sizable duration clearly do not intersect in any single content of experience.

momentary, experience is nevertheless taken to be *continuous through time*. Most of us believe that our present experience has developed, in a fairly continuous way, through a sequence of previous experiences. We take it that we have had other experiences at moments earlier in time. And we expect that the present experience will itself soon be past, and an experience at a later moment become our present experience.

The key function in constructing this view seems to be *memory*. A certain portion of present experience is identified quite unmistakably by most people as *memory experience*. This portion of present experience is interpreted as *providing information about experiences earlier in time*. This realistic interpretation of memory, as *referring to events at earlier moments*, is absolutely central to our view of the world. It is also central to our understanding of the internal coherence of our percepts.

There is a wide variety of memory experience, from very short-term perceptual memory, to long-term propositional memory. Very short term memory of the changing field of our perceptual experience seems certain to have a close connection with our 'sensation of time flow'. For instance, we often continue to be 'aware' of a sound for a short time after the original sensation has passed, or of certain features of a visual impression after the original sight is over. Perhaps this awareness is generally awareness of a *sequence of experience* - a process - rather than just an instantaneous experience. We clearly organize what we *remember* we have experienced into a linear temporal sequence, stretching backwards in time from the

momentary present. We are compelled to assume this linear sequence of earlier experiences through what I will call the *nested structure* of memories. In our memories of past experience we can sometimes identify *earlier memory experiences*, i.e. we can remember what we were remembering. E.g. at t_0 we might see a bright yellow flash in the sky. At t_1 we might remember that we saw a bright yellow flash in the sky, and simultaneously see a blue flash in the sky; at t_2 we might remember the original experience of seeing a bright yellow flash, and also remember the second experience of *remembering* seeing the yellow flash and of simultaneously seeing a blue flash. Thus at t_2 we can order the experiences we remember having at t_0 and t_1 as being in that causal order and not the reverse order. This 'nested structure' of memories (which might be perceived in a more or less unconscious way) generally determines an intrinsic linear ordering of remembered experiences. This is taken to the extreme in the very short-term memory of the perceptual field: this very short-term memory seems to provide images of *continuous processes through time*. It seems clear that the human mind can survey many of its own processes, and perhaps the 'sensation of time flow' arises from some higher-order observation on the *process of constructing memories*, so that this sensation really does represent a *perception* of the passage of time.

But the precise details of the mechanisms of memory and the 'perception' of time flow are not important here. The crucial point is that we interpret certain kinds of *present experiences*, namely memories, to *refer across time*, to other experiences

which have occurred earlier in time. The question to be asked is: *what justifies this interpretation of present experience?*

Let us consider the different implications for this question of the two competing views about time, the dynamic view, and the bloc universe view.

On the *dynamic view*, the idea of the *continuity of experience through time* is perfectly natural, since the dynamic view is exactly that what exists (in this case the content of experience) *changes* as time goes on. We may thus choose the hypothesis about the way things have been in the past which best explains or fits the internal coherence of present experience. This seems to be provided by the normal interpretation of 'memories' as providing information about past experiences.

On this hypothesis, there is a good *causal* explanation of the coherence of present experience. The memory portion of present experience has been *caused* by the storage and access of information about a real previous sequence of experience. The coherence of memory with present perception is explained.

But on the contrary, bloc universe, metaphysics, I think we are left without any such explanation, or equally, without any justification for *interpreting memory as we do*.

On the bloc universe view, a time-slice of experience exists as an eternal (or timeless) object. Consider my previous visual field, F_1 . If there really was such a visual field of mine at an earlier moment, then according to the bloc universe model, that visual field continues to exist exactly as it always has - only, of course, at a different place in time from my present experience as I write this. The memory awareness of the occurrence of F_1 which I presently have is rather like a *perception* of something

which exists in a different temporal place. Memories supply information about states-of-affairs which co-exist at different temporal places. The problem is why present experience should be taken to have any relation at all to something that exists at a different moment. Why interpret present memory to *refer* to anything? The answer cannot be that the postulate of a sequence of earlier experiences provides a causal explanation for the content of present memory or experience. For the present experience (including memory) is something that exists eternally. It has not been brought into being, and it does not go out of being; it suffers no change of any kind. On the bloc universe view, it is simply an object that is. In explaining its existence, what use is it to appeal to earlier events (or indeed anything) as having 'causally generated' it? Quite simply, nothing *has* generated it.

Perhaps the Illusionist will say: present experience has an internal coherence exactly *as though* it is generated in part by memories of earlier sequences of experience. Let us therefore postulate that the appropriate earlier experiences exist, and that they (and the functions of memory storage and recall) should be viewed as the cause the coherence of present experience.

But this is already incoherent in the Illusionist's own terms, since it appeals to a notion of real causation. To be consistent, the Illusionist must do entirely without a concept of real causation. For the Illusionist, the world comes as a temporal whole: no part of it is *caused* by any other part of it. But then how can we make sense of the idea that occurrences at one moment in time help explain occurrences at another moment?

Many Illusionists would no doubt be happy to grasp this nettle,

and accept that there are no such explanations. In the first place, this simply contradicts the ordinary conception of the world, in the same sort of radical way as Idealism contradicts it. In the second place, if it is accepted that present experience requires no earlier *cause*, then what legitimates the inference from present experience to past experience? What legitimates the normal interpretation of memory?

The Illusionist might claim that, although he has banished real causation from the world, 'explanations' of states of affairs, or inferences from one state-of-affairs to another, are legitimated by true universal generalisations. For instance, he might summarize the generalisations he believes in as follows: *there is in fact a continuity of experience through time exactly as would be implied by the content of present experience given real time flow.*

But it is difficult to see how this has any *explanatory value*. It should be compared to the Idealist's postulate (see p.67) that "percepts in fact cohere as though they were really perceptions of the external objects proposed by the Realist." Superficially this seems to do the same job as the Realist postulate: but all the same, few of us are prepared to seriously accept Idealism. But what reason have we for rejecting Idealism that we do not have in greater measure for rejecting the bloc universe view of time?

This argument is hardly conclusive, and it is hardly as strong as the ordinary conviction of time flow. I expect that this means that the argument, even if it is on the right general lines, fails to bring out the full force of the reasons for belief in time flow. The Illusionist would probably prefer to conclude that it is a sign that

the ordinary conviction is stronger than it ought to be. But note again the analogy with the Phenomenalism/Realism debate: the theoretical arguments for Realism are much weaker than the ordinary conviction of Realism. One would similarly conclude that the theoretical arguments do not bring out the full force of the reasons for belief in Realism.

To conclude I will sketch a second argument for time flow from the second kind of 'connectivity through time' mentioned above: the continuity of physical objects through time. We have already noted (Section 1.22) the feature of time that distinguishes it ontologically from space: that what happens *later* in a region of space depends upon what holds earlier in that region. This ontological dependence does not hold across spatial directions. That a certain region of space is filled with air has no bearing on whether the air continues in any given direction in space. But of course, that a region of space is filled with air at one moment or interval has all the bearing in the world on whether it is filled with air in the next moment.

There is clearly a commonplace belief in a connectivity of objects through time. This is represented in the four-dimensional Minkowski picture by the fact that objects appear as four-dimensional 'pipes' through time. The dynamic conception of objects is that they *persist through time*, and this, of course, is a central, commonplace metaphysical conception about the world. But why should there be such connectivity through time, but not through space?

If the Illusionists are right, there is no ontological necessity for any such connectivity through time, no need for persistence of objects through time. The four-dimensional Minkowski space-

time could as easily be filled with four-dimensional spheres or random blobs, giving a true symmetry between time and space. This would be reflected in 'experience' (if experience were possible in such a world) by the fact that objects would spontaneously begin and end in time, as they do in space. This would make predictions very difficult. We wouldn't know whether to trust the continued existence of the world at all.

This is very reminiscent of Hume's famous problem of induction from past to future events. How can we have any reasonable beliefs about what lies in the future? Well, if we adopt the bloc universe view, and make time really equivalent to space, I don't think we can. What is the case *here* in space is no guide to what is the case over there. Similarly, on the bloc universe view, what is the case *now* is no guide to what is the case later or earlier. There is no *intrinsic connectivity* of the world through time.

The metaphysical postulate of time flow solves this problem, by giving an intrinsic connectivity of the world through time. The dynamic view is that what exists is essentially a collection of objects in a three-dimensional space. The state of these objects *change* - it is because change occurs that we say that 'time flows'. The crucial ontological ingredient is that objects *persist through changes*. (Persisting through changes means persisting through time.) Thus it is no mystery about the ontological dependence of what is *later* on what is *earlier* when we have time flow, no mystery about the special connectivity of the world through time.

It may be thought that the Illusionist has a reply to this: why

not simply make the further postulate that the world (or its objects) *are* connected through time? This will be an extra feature of the Illusionist's account of the world - but surely such extra features are needed anyway, as when complex scientific laws are postulated to account for all sorts of phenomena.

The Illusionist in other words can postulate that: *There is in fact a continuity of objects through time exactly as there would be if time flowed.*

But once again, the kind of objection that the Realist raised against the corresponding Idealist postulate that "Percepts in fact cohere as though they were really perceptions of the external objects proposed by the Realist" (p.67) may be repeated. Namely, that this postulate does not *explain* anything. All it does is to state a general proposition that is believed to hold. In contrast, the metaphysical postulate of time flow *explains* the general proposition, just as the realist postulate of external objects explains the coherence of percepts.

1.30 Conclusion.

This overview of a rather long and difficult argument has itself been rather long and difficult. But though difficult in some of the details, the central idea is simple and intuitively appealing: that quantum theory entails that there are generic probabilities in nature, which are directed towards the future, and that in their actualisation we have the 'flow of time'.

The following chapters and appendices contain detailed treatments of concepts and claims referred to in this overview. These are provided with the intention of supporting the argument presented in this chapter, but their interest does not lie solely in that, for the questions and concepts they deal with are interesting in themselves.

CHAPTER TWO

PHENOMENOLOGICAL DIRECTEDNESS

The main topic of following chapters will be the concept of the *time reversibility of physical theories (or of natural laws)*. But I will begin the discussion of the physical directionality of time by considering the much more immediate temporal directionality evident in our environment, which I will call the *phenomenological directedness of processes in the environment*.

2.1 Phenomenological directedness.

At the phenomenological level, the physical processes in our environment appear to be highly directional in time. For instance, if we took a film of some ordinary kinds of processes, and played it in reverse, most people could easily tell that they were watching a reversed film and not the original. Why is this?

It is because there are many types of processes which are

perfectly common, but the 'reversals' of which never occur at all. For instance, intact eggs frequently go through processes in which they break, but broken eggs never go through processes in which they mend; rivers run down to the sea, never up to the mountains; wood burns and turns to ash, but ash never turns to wood.

In fact we feel that these reversed processes not only do not occur, but cannot occur. We cannot make eggs mend, rivers run uphill, or ash turn back into wood, no matter how hard we try to set up a situation in which these processes would occur. Such reversed processes would seem to go against nature itself.

Such processes will be called *phenomenologically directed*. Phenomenological directedness is a very common feature of processes in our environment.

A slightly more technical vocabulary will help before giving a definition. If P denotes some given type of process, then I will use the term: P^R to denote its time reversal¹. P^R is also a type of process. It is the type of process that would occur if a process of type P were to run 'backwards in time'. P^R is conveniently imagined as the kind of process we would see if we ran a film of the original process P backwards. E.g. if P involves an egg falling from the bench to the floor and breaking, then P^R would involve a broken egg on the floor reassembling itself, and leaping back onto the table.

Technically, the superscripted R represents an operator on process-types. It is called the *time reversal operator*.² It

¹I will frequently just talk of *processes* when I really mean *process types*. Context should make the intended meaning clear. See Chapter Four for a more detailed discussion.

maps the class of possible processes back onto itself. This operator will be defined precisely in following chapters, but the intuitive understanding of it will do here.

Phenomenological irreversibility then amounts to this:

[2.1] A type of process P is *phenomenologically directed* just in case instances of P are common, or can be brought about, while instances of P^R never occur, and cannot be brought about.

I will also talk of the *phenomenological directedness of the environment*, meaning that our environment is saturated with phenomenologically directed processes. Why is our environment so saturated with such processes? This is one of the key scientific questions about our universe. Many physicists think that it is the key scientific question about the nature of time. So far physics has not found a complete answer to it. I will not try to provide an answer either, but I will spend some time explaining the logical structure of the question, and particularly its relation to the second kind of reversibility, the time reversibility of theories.

² Or more accurately the *time reversal operator for processes*, since there are also *time reversal operators for processes, theories, and sentences*.

2.2 Time reversibility.

If Q is a type of phenomenologically directed process, then Q is common but Q^R never occurs at all³. This makes it appear that there is a natural law against Q^R occurring, while obviously there is no natural law against Q occurring. A natural question is therefore whether Q^R is *nomologically possible*. This brings us to concept of the *time reversibility of a physical theory*, which will now be defined.

Let T be the fundamental physical theory governing the domain of processes to which Q and Q^R belong. T tells us which *micro-physical processes* in its domain are nomologically possible. I will say that P is a *T-process* to mean that P is a type of process consistent with the theory T . Time reversibility is defined:

[2.2] A theory T is *time reversible* just in case, for every T -process P , P^R is also a T -process. Otherwise T is time irreversible.

This definition will be examined and justified in more detail in the three following chapters. It will be shown that time reversibility is an objective feature of a theory (once the interpretation of the theory has been settled).

Note how different phenomenological directedness and time reversibility are. Whether a process P is phenomenologically

³ The variable Q will be used to range over phenomenologically directed (types of) processes, while P ranges over processes in general.

directed or not is an empirical matter, a function of the particular environment or world of the observer. But whether a theory T is reversible is a fact about a theory, and not relative to the environment at all.

2.3 Explaining phenomenological directedness.

It is easy to assume that the explanation of phenomenological directedness must appeal to the irreversibility of the fundamental physical laws. In fact it is generally considered a puzzle and even a 'paradox' that our environment is phenomenologically directed, while (as it is thought) the fundamental theories are time reversible. But it is well recognised that there is no real paradox in this.⁴ That is: *phenomenological directedness (of the local environment, or the entire universe) does not require irreversibility for its explanation.*

It must equally be recognised that irreversibility of the laws of nature will not necessarily provide any explanation of phenomenological directedness, i.e. *irreversibility does not necessarily have anything to do with the explanation of phenomenological directedness.* These points will be demonstrated more generally later, but a practical illustration will be useful first.

⁴E.g. DeBeauregard [1970,1977], Davies [1975], Reichenbach [1956].

2.4 Explaining phenomenological directedness in the context of a reversible fundamental theory.

To explain phenomenological directedness, two questions need to be answered:

- (i) Why do the kinds of processes which are phenomenologically directed occur?
- (ii) Why do the reversals of these kinds of processes not occur?

If the fundamental laws were irreversible, then the reversals of phenomenologically directed processes *might* be nomologically ruled out, which would answer (ii). But they would not necessarily be ruled out. They would only be ruled out if they all happened to be 'irreversible' processes, but there is no *a priori* reason to think they will be. For instance, if the only real irreversible processes were those involving K^0 meson decay, then the reversals of most phenomenological processes would still be nomologically possible.

However, given a strong kind irreversibility, the reversals of phenomenologically directed processes could be ruled out, and (ii) would be answered - the reversals of phenomenologically directed processes never occur and cannot be made to occur because they are nomologically impossible. But the problem of (i) remains. Why do the kinds of processes which are phenomenologically directed occur at all? Why is our environment

saturated with them? The answer to this is likely to remain the *major* step in the explanation.

The irreversible feature of the natural laws might offer nothing towards the answer. Perhaps it would be unusual if it did. Irreversibility means only that the reversals of certain possible processes are nomologically impossible. So it rules out certain classes of processes. But it does not immediately say anything about the kinds of processes that *do* occur.

In particular, in the actual world, there is a large class of possible processes that do not appear phenomenologically directed at all - 'equilibrium processes'. By this I mean processes which do not involve any large-scale changes in entropy. The reversal of an equilibrium process generally looks, phenomenologically, just like the original equilibrium process. E.g. the *apparent* behaviour of a volume of well-mixed gasses in thermodynamic equilibrium is the same in either direction of time. (Nothing *appears* to happen at all.) Thus equilibrium processes are not phenomenologically directed, since, at the phenomenological level, their reversals appear as common as they are.

Since equilibrium processes are nomologically possible, it is a mystery why the environment is not filled with them, instead of with phenomenologically directed processes. The fundamental laws might not imply that any processes other than equilibrium processes need occur, whether or not they are irreversible.

To answer (i) purely by appeal to the laws of nature may not be possible, because there may be no nomological reason for the particular events that occur in our world. It is probably

consistent with known laws of physics that the entire world throughout its history was and will be in a state of thermodynamic equilibrium, and that no phenomenologically directed processes were ever common. In this case, the particular nature of our world might have to be blamed merely upon 'accident'. In particular, upon special boundary conditions for the universe, which are not compelled by the fundamental laws, but obtain only accidentally.

In this case irreversibility has little to do with the explanation of phenomenological directedness. The explanation of (i) would appeal crucially to an accidental feature of the world, and this would remain the major part of the explanation.

I do not want to say, of course, that the statement: "Phenomenological directedness is merely an accident" would be an adequate explanation. To be convincing, the 'accident' that the asymmetries are blamed upon must (1) have independent confirmation (over and above just the phenomenological directedness it explains), and (2) must provide a *comprehensive* account of the directed processes, explaining why they all have the same direction in time, why they are all so common, and why the directedness of phenomenological processes is for any practical purpose lawlike. But given that the explanation *did* appeal to such an accident, whether the fundamental laws were reversible or not might be very much beside the point.

In fact the explanation accepted at present does appeal to such an accident - specifically, to the special boundary condition of the universe produced by the 'big bang' event. What many physicists probably imagine is that the discovery of irreversibility is needed to provide an explanation for this

'accident'. To have an explanation of the big bang would be a great thing, but it is wrong to see it as necessarily connected with irreversibility. Firstly, there might be an explanation of the 'big bang' without there being any fundamental irreversibility at all. Secondly, there might be irreversibility without it affording any explanation of the big bang.

The actual explanation that is presently given for phenomenological directedness goes something like this. Sometime in the distant past (about 15 billion years ago by present calculations) the universe passed through, or was 'created' in, a state of incredibly low entropy. That it went through this post-big-bang state is known empirically, by the observed expansion of the universe, and the observed background radiation and overall material composition, in conjunction with quantum mechanics and general relativity which have been established on independent grounds. Given such a state of initial low entropy, it is statistically almost certain that entropy would increase rapidly, both in the universe as a whole and in all of its significant subsystems, for an extremely long period to come. The phenomenologically directed processes are essentially all processes involving enormous entropy increases, and they have been made factually likely in our era just because the universe passed through the original low-entropy state. The reversals of such processes would involve correspondingly enormous decreases in entropy, and this being astronomically unlikely to happen by chance, the reversals never occur.

This is the standard explanation, and still a pretty rough one perhaps, but it seems essentially correct. The main feature of it

is that there is no explanation of why the special boundary condition (the 'big bang') occurred. That is unexplained at present, left as an 'accident' which is merely *observed* to have occurred. Although this is a major gap in our understanding of the genesis of the universe, it does not invalidate the explanation. To see this we can note the distinction between two kinds of explanations:

- (i) Explanations of laws,
- (ii) Explanations of contingencies.

An example of the first is the explanation of the law that water has a latent heat of vaporization in terms of the fundamental processes that occur when water is vaporized. Or the explanation of the electroconductivity of iron in terms of the electronic structure of iron atoms. What is explained is not a *particular* event but the lawlike behaviour of certain types of things under certain types of conditions. The explanation works by showing that a certain observed generalisation is compelled to hold by the fundamental laws of nature.

An example of the second type would be the explanation of a rough radius of fragments of meteoric-type rock in a certain place, by the hypothesis that a meteor fell there and shattered. Or the explanation of the existence of a variety of distinct yet similar kinds of cats (tabbies, lions, lynxs, etc) by the hypothesis of a common ancestor from which all evolved. This is properly called *causal* explanation. It involves explaining a variety of effects as the arising from some particular antecedent cause. The effects are typically a group of apparently independent but at the same time unusually correlated events. The hypothesis of a

unifying causal ancestral event 'explains' the correlations.

Returning to the explanation of phenomenological directionality, we see that what has been offered is an explanation of the second type, not the first. This seems puzzling at first sight, because it seems at first sight that what we are explaining - the directionality of ordinary processes - is *lawlike*. Hence we feel that we need an explanation of the first kind, a derivation of the directionality of processes from the fundamental laws. But if the fundamental laws are time symmetric there can be no such explanation, and we must fall back upon an explanation of the second kind, which is what is presently done.

What is given, then, is a causal explanation, which explains why our environment is full of phenomenologically directed processes by appealing to a special event or condition that occurred in the past. A causal explanation can always be improved, in a sense, by pushing for an account of the contingent event/s that it appeals to. We could explain why the meteor fell in the first place, or why the cat ancestor arose in the first place. Of course we would only find ourselves with a new 'contingent event' which remains unexplained, but we would have understood more about the causal chain of events that gave rise to the effect. Sometimes there is little to be gained by pushing a causal explanation back any further - in the meteor case, for instance, it is sufficient just to observe that meteor crashes are reasonably common: to trace the causes of the crash of this *particular* meteor, apart from being practically impossible, wouldn't improve our understanding of the effect we are trying to understand, namely, the distribution of fragments of rock. But

sometimes a causal explanation does demand a further explanation, and the explanation offered for temporal asymmetries is like this. We are not very satisfied with the mere *observation* that there was a 'big bang' event: it is such a peculiar event that we feel a need to push for an further explanation of it. One might also feel that the big bang is so peculiar that there must be some lawlike explanation of it, not merely a causal explanation.

But although these kinds of feelings make the given explanation seem unsatisfying, it still remains a good explanation as far as it goes. The temporal directedness of processes in our environment *does* depend upon the universe having passed through the special boundary condition we call the 'big bang', and the explanation will probably always appeal to this fact. Whether or not we someday push the explanation further back, and explain the big bang event as well, we will not invalidate this explanation, but only deepen it.

We may conclude that there is no real paradox, indeed really no conflict at all, in the conjunction of a reversible fundamental theory with a phenomenologically directed environment. We have to explain the latter as a contingent feature, in the sense that it is generated by an earlier contingent event, and not entailed by the laws of nature.

It is worth noting that although the existence of phenomenologically directed processes is thereby seen as contingent, it doesn't follow that all features of the temporal directedness of the world are also contingent. Vital features can still be explained as necessary (or at least extremely probable) -

for instance, the fact that, throughout the whole universe, the directed processes occur in a consistent temporal direction. It is not explained why they all run in the direction they do run in, but it *is* explained why they all take the same direction. It is because they all have the same cause in the the big bang. This lets us predict, for instance, that anything comparable to eggs on planets in distant galaxies will commonly break just as they do on earth, and that no eggs will mend, just as they do not mend on earth. Hence, although the explanation of the phenomenological directedness of our world presently makes it a contingent feature of the world, important features of that directedness are still accounted for.

2.5 Explaining phenomenological directedness in the context of an irreversible fundamental theory.

While reversibility does not prevent us giving a good explanation of phenomenological directedness, irreversibility does not automatically provide us with any better explanation. This is seen by considering two proposed kinds of irreversibility: first the very weak 'irreversibility' that has been detected in the decay of K^0 mesons; second the rather strong kind of irreversibility that I claim to hold generally of quantum theory.

The first kind of irreversibility only pertains to a very small class of processes, and it neither rules out the reversals of most phenomenologically directed processes, nor explains the commonness of phenomenologically directed processes.⁵

⁵See Penrose [1979, sec.12.2.1].

The relation of the irreversibility that I claim for quantum theory to phenomenological directedness is more interesting. Firstly, the irreversibility I propose is universal: every quantum process in which anything probabilistic happens is irreversible. The reason is that in every such process, there are intrinsic probabilities directed towards the future, so to speak, but never such probabilities directed towards the past. (This will be a key subject of following chapters).

This does not mean that all quantum processes *appear* to be 'irreversible', or phenomenologically directed, however. Consider an equilibrium process, such as the continual random mixing of gasses in an already well-mixed state. This process appears the same in reverse, and thus cannot constitute a phenomenologically directed process. Nevertheless, on my account, the actual quantum micro-process underlying the phenomenological appearance is irreversible.⁶ In other words, *processes can be irreversible at the micro-level, but appear reversible at the phenomenological level.*

The irreversibility that I claim therefore does nothing immediate to explain (i) why the kinds of processes which are phenomenologically directed are common in our world. For our world might equally have been filled with equilibrium process, and although these would remain microscopically irreversible, they would not be phenomenologically directional since they would not *appear* irreversible.

What my proposed irreversibility does explain perfectly well is (ii), i.e. why the reversals of phenomenologically directed

⁶So long as it involves 'probabilistic' processes, and not a merely deterministic evolution of state.

processes do not occur.⁷ The reason is just that such reversed processes would be astronomically unlikely. And they would be unlikely *according to the laws of physics* - not merely on account of boundary conditions that hold or fail to hold of the actual universe.

Thus the irreversibility I propose does nothing to solve the major step in the explanation of phenomenological directedness, for it does nothing to explain (i). But this is no surprise.

2.6. General form of the explanation of temporal directedness.

What has just been illustrated with examples in the two preceding sections can be seen more generally. To explain the phenomenological directedness of our general environment we do something like this. First we identify the phenomenologically directed processes in our environment as being of a certain kind. Let us call it the kind G . The reversals of the phenomenologically directed processes must be of a contrary kind, which it is appropriate to denote as G^R . We then try to explain why the processes of kind G are common, while processes of the contrary kind G^R are non-existent.

For instance, in the real world, the phenomenologically directed processes are all processes which involve *large increases in entropy*. Their reversals (if they ever occurred) would involve *large decreases of entropy*. I will call these *entropy-climbing processes* and *entropy-falling processes*, respectively.⁸ So what we try to explain in practice is why, in

⁷Watanabe [1965,66,70] pointed this out in some detail.

⁸It should be remembered that *entropy-climbing processes* must involve

our environment, entropy-climbing processes (G processes) are common, while entropy-falling processes (G^R processes) are non-existent.

It may seem at first sight that this does not give a specific enough explanation of phenomenological directedness. For instance, showing why entropy-climbing processes are common doesn't show why any *specific* phenomenologically directed processes are common. The breaking of eggs, or the burning of wood are specific types of phenomenologically directed processes common in our environment. But merely explaining the commonness of entropy-climbing processes does not explain why these *specific* processes are common.

But in fact we do not want too specific an explanation. Obviously it is common for eggs to break in our world only because a long process of evolution has produced hens (and other animals) which lay eggs. Similarly it is common for wood to burn only because trees have evolved. But to explain the *general* predominance of phenomenologically directed processes we don't want to have to go into the detailed explanations of the origins of hens and trees, and the countless other strange things with which our particular environment is filled. These things might easily not have appeared at all, but we would still want an explanation of phenomenological directedness apparent in countless other phenomena. We only want to explain why the environment is saturated with phenomenologically directed processes in general, and if we could explain why it is saturated with

large increases in entropy, similarly *entropy-falling processes* must involve *large* decreases in entropy.

entropy-climbing processes we would be well on our way to a good explanation.⁹

Thus we pick out some characteristic universal features, G and G^R , of phenomenologically directed processes and their reversals respectively, and we try to explain the predominance of processes of kind G and the absence of processes of kind G^R . The question is: What has the reversibility of the fundamental laws got to do with this explanation?

It is useful to gain some grasp of the internal structure of the class of nomological processes. If the fundamental theory T is *irreversible*, then there are some T -processes P the reversals P^R of which are not T -processes. These *irreversible* T -processes form a subclass of all the T -processes. This subclass of irreversible processes will be denoted by T^1 , and the remaining T -processes by T^0 . Thus T^1 and T^0 are mutually exclusive classes whose union comprises all the T -processes.¹⁰ T^1 might be only a part of the class of T -processes, or it might comprise all the T -processes. For instance, if, in the real world, K^0 -meson decay was the only basic type of irreversible process, then processes involving K^0 -meson decays would be the only types of

⁹Even if this does not give a complete explanation, it is a necessary part of any good explanation. This is because *entropy-climbing* is a necessary feature of all the phenomenologically directed processes in our world. Hence to explain the commonness of phenomenologically directed processes we must, implicitly or explicitly, explain why entropy climbing processes are common. Entropy-climbing is *not* also a sufficient condition for phenomenologically directedness, so the proposed explanation would not necessarily be complete. But in practice it would almost certainly provide the basis for as complete an explanation as is possible.

¹⁰In fact from the next chapter on, T will normally be treated as being literally the class of all T -processes. This is adequate because specifying the class of T -processes specifies everything about the theory T . See Section 3.4.

process in T^1 , and T^1 would comprise a very small subset of the T-processes. By contrast, if my claim about the irreversibility of quantum theory is correct, then all probabilistic quantum processes are irreversible.

Each process P in T^1 has a reversal, P^R . The class of these reversed processes is $(T^1)^R$.¹¹ $(T^1)^R$ has no intersection with the class of T-processes (or with T^1). Similarly, $(T^0)^R$ is the class of reversals of the processes in T^0 . However each process in $(T^0)^R$ is a T-process, and T^0 and $(T^0)^R$ are the same class. I.e. $T^0 = (T^0)^R$.

Note also that every process of kind G is the reversal of a process of kind G^R , and vice versa. G and G^R are disjoint classes of processes (otherwise G has not been well chosen), but they do not necessarily exhaust the class of nomologically possible processes, since there may be processes which are neither G 's nor G^R 's. (e.g. equilibrium processes, in the example where G are entropy-climbing processes and G^R are entropy-falling processes).

G should be chosen in the first place so that all G processes are T-processes. (For insofar as G processes are not nomologically possible, G has not been well chosen as the characteristic feature of phenomenologically directed processes). Assuming this, there are two important groups of possible relationships:

¹¹In general, where Z is any class of processes, Z^R is the class of reversals of the Z -processes. Here the superscripted R is the *time reversal operator for theories*.

- a. All G's are T¹'s.
 - a'. No G's are T¹'s.
 - a''. Some G's are T¹'s and some are not.
-
- b. All T's are G's.
 - b'. Not all T's are G's.
 - b''. In any T-world, some processes must be G's.

If a holds, then G^R processes are nomologically impossible, which explains why they never occur. If a' holds instead, then G^R processes are all nomologically possible, and the explanation of why they do not occur must appeal to something beyond the laws of nature - probably to contingent events in the universe. If a'' holds, then some of the G^R's are rendered nomologically impossible and some are not. I will argue later that a holds of quantum theory.

If b holds, then all nomologically possible processes are necessarily phenomenologically directed, which would give a perfect explanation of phenomenological directedness. However it seems unrealistic that b should hold, since this would imply that all possible processes are phenomenologically directed. Instead, b' is almost certain. Given b', then the only hope for a nomological explanation of the existence of processes of kind G is if b'' holds. But there is no *a priori* necessity for scientific theories to guarantee b''. Whether there is any practical reason to expect it is quite unclear at present. It does not seem to hold where T is quantum theory. However it must be recognised that

the quantum theories so far proposed have limited application, and do not provide the complete fundamental theory for the whole universe. The complete fundamental theory must also incorporate what we presently know as gravitation. Various possibilities are being investigated, but it is difficult to tell what progress can be expected in this area. However, the uniqueness of the big bang seems to persuade us that there must be some nomological reason why it occurred, and if there is, then the complete theory will entail \underline{b} ". Perhaps the most obvious speculation is that 'big bangs' are actually periodic events, and the correct theory will entail their occurrence. This would explain the occurrence of the last big bang as just one in a long series of such events¹². Another possibility might be that our universe is not unique at all, but has been produced by some repeatable process, a kind of fundamental process which perhaps always produces universes from big bang events. In either case the nature of our universe in the present cosmic era would be seen as a necessary stage in the evolution of a universe. It would hardly be necessary that there are human beings, or even life. But it would be necessary that universes go through states which entail subsequent periods of great entropy increase, and this would provide the main premise in the explanation of the phenomenological directedness of time.

However despite our natural inclination to believe that there must be some such explanation of the origin of our universe, whether such an explanation exists is presently only a matter of speculation.

¹²E.g. Landsberg [1982b].

CHAPTER THREE

REVERSIBILITY AND TIME SYMMETRY

Reversibility was defined in the previous chapter as:

[3.1] A theory T is time reversible just in case, for every T -process P , P^R is also a T -process. Otherwise T is time irreversible.

This is the fundamental definition of reversibility (for theories), and is well agreed upon in the literature.¹ I will call it the primary definition of reversibility.

¹ More precisely, this is a common style of definition with philosophers or logicians, e.g. Earman [1969,74], Kroes [1986]. Most physicists use a different style of definition, but it achieves the same end, as shown in Section 3.3.

This definition presupposes the concept of the *time reversal of processes*, since it relies upon P^R being well-defined. The main problem in applying the definition in practice will be in interpreting the nature of process reversal, and this will be considered in detail in the next two chapters. But before going on to that, it is important to establish the motivation for the definition of the reversibility of theories.

What needs to be made clear in particular is the close relationship between *reversibility* and the *directional symmetry of time*. In this chapter I will offer a systematic analysis of this relationship, and of a cluster of related concepts. The systematization of concepts that I present here derives mainly from the work of Reichenbach [1956], Mehlberg [1980], Grunbaum [1973, ch.8], and Earman [1967,69,74], particularly the last two. But I differ from each of these writers on a number of points (as they also differ amongst themselves). Perhaps what is most obvious is that there is only a partial consensus among these writers (and others) on terminology. Insofar as this is merely a stylistic matter, I have tried to adopt the simplest, most systematic, and most orthodox system of terminology I could. But the problem runs beyond mere terminological conventions, into substantial disagreements over the proper analysis of various concepts. The main innovation I have made is to give a precise analysis of the notion of the *symmetry of time*, and of its relation to *reversibility* (or time reversal invariance, or T-invariance). This leads me to some divergences from the above writers. Where my analysis differs from those referred to

above, I hope it will prove to be correct, but I have not made detailed comparisons of the various views here.

While my treatment is essentially similar to those referred to above, it is very different to another analysis, given by Bunge [1970]. I have felt it illustrative to consider Bunge's views in some detail. This is done in the final section of the chapter, where I criticize and reject many of Bunge's conclusions.

Let us turn now to the main topic of this chapter, the concept of the *directional symmetry of time*.

3.1 Directional Symmetries.

There is an exact analogy between the concepts of the directional symmetry of time and the directional symmetry of space. As space is a more concrete entity than time, and easier to get an intuitive grasp of, I will begin by discussing the symmetry of space. The conclusions will transfer directly to our understanding of the symmetry of time.

Physics normally treats space as directionally symmetric.² What is meant by this cannot be that space is actually *homogeneous* in all directions – for it is obviously not. E.g. in one direction we find the sun and in another direction we do not, and this is a lack of spatial homogeneity.

But this is merely an asymmetry in the distribution of things in space, not an intrinsic asymmetry of space itself. When we

²In the context of quantum theory this no longer holds because of the violation of parity in K^0 -meson decay.

call space *directionally symmetric* what we intend is that there is no intrinsic asymmetry of space itself. What is meant by this 'intrinsic symmetry of space'?

A physicist might explain it by saying that the spatial variable x could be exchanged for the variable $-x$ in the fundamental laws without making any difference to them. This property is called *invariance under spatial reversal*.³ Denoting the spatial *directions* by $+x$ and $-x$, this invariance means that Nature is blind to any difference between these two spatial directions.

I wish to show how this can be understood in a formal way, as meaning that *the intrinsic properties of the spatial directions are identical*. That is, I will define directional symmetry of space as follows:

[3.2] A spatial axis x is *directionally symmetric* just in case the intrinsic properties of the spatial directions $+x$ and $-x$ are identical. Otherwise x is *directionally asymmetric*.

To understand this we must understand (i) the idea of 'intrinsic properties' of such a thing as space (and later time), and (ii) the idea of properties of *spatial directions*.

(i) Intrinsic Properties. The intrinsic properties of things will be taken to be just those properties entailed by the laws of Nature. (For 'intrinsic' you may substitute *essential* or *nomological*.) Thus the intrinsic properties of space are those

³To be consistent with this terminology, what we are calling *time reversibility* would be called *invariance under time reversal*.

properties that space has in all physically possible worlds. An example (according to present physics) is that space is a continuum.

Other properties of space are *extrinsic*, depending upon merely contingent features of a particular world. (For 'extrinsic' you may substitute *non-essential* or *contingent*.) For instance, it may be that if we look in the x direction of space we find substantially more galaxies than in the y direction. Thus the x and y directions differ in this extrinsic property. But it clearly does not pertain to the intrinsic nature of space, and is of little interest in discussing the *kind* of thing that space is.

There is little to say to justify this interpretation of 'intrinsic properties' of natural things, except to observe that the very notion of 'laws of Nature' involves the idea of features that are *intrinsic* to physical things, as opposed to merely contingent features. One might object to making any such distinction at all between intrinsic and extrinsic properties, but one would thereby be rejecting the whole idea of 'laws of Nature'. If we are to distinguish intrinsic from extrinsic properties at all, then the only natural way to make the distinction is in terms of natural laws.

(ij) Properties of the spatial directions. For any spatial axis, x, there are two directions, which will be denoted $+x$ and $-x$ respectively. These directions may be considered as entities with their own properties.⁴ This may sound a little unusual to the physicist, but it sounds perfectly natural to the logician.

⁴The directions are properties of spatial vectors.

What are the properties of $+x$ and $-x$?

Such properties may be obtained by a simple process of abstraction. For instance, let us suppose that there are two stars in the universe, A and B, which are the biggest and the smallest stars, respectively. Let us suppose that *the direction from the biggest to the smallest star is $+x$* . This tells us a property that the direction $+x$ has. Namely, the property of *being the direction from the biggest star to the smallest star*.

If $+x$ has this is a property in our world then no other spatial direction, including $-x$, has it in our world. Thus the spatial directions $+x$ and $-x$ differ by this property in our world. We can thus say that the spatial axis x is *directionally asymmetric with respect to this property*. The asymmetry in this case is only an extrinsic or contingent one, because the property which distinguishes the two directions $+x$ and $-x$ is only a contingent property of spatial directions.

For *directional asymmetry* proper, according to [3.2], the spatial directions must be distinguished nomologically. That is, the laws of nature must distinguish the spatial directions. It can be shown that this holds just in case the laws are non-invariant under spatial reversal.

Invariance under spatial reversal means, roughly, that the laws remain the same when the spatial variable x is replaced with $-x$. This means roughly that it doesn't matter whether the spatial directions $+x$ and $-x$ are exchanged in the laws. Let the laws be represented by a class of sentences T . Let $T[v]$ be T with the essential occurrences of $+x$ replaced by any other variable v .⁵

That is, $T[v]$ would say everything about v that T says about $+x$. Thus $T[+x]$ is just T .

Now $T[.]$ represents *all the nomological properties of $+x$* . $T[.]$ is a predicate denoting the complete intrinsic (or nomological) property that T attributes to $+x$. Invariance of T under spatial reversal just means that $T[+x]$ is identical to $T[-x]$ – which is to say, that $+x$ and $-x$ have exactly the same intrinsic properties. Thus *invariance (of T) under spatial reversal* is logically equivalent to the *directional symmetry of space (in all T worlds)*.

3.2 The Directional Symmetry of Time.

The concept of *directional symmetry of time* is defined in exact analogy with that of space:

[3.3] Time is *directionally symmetric* just in case the intrinsic properties of the temporal directions $+t$ and $-t$ are identical. Otherwise time is *directionally asymmetric*.

What needs to be shown is that *time reversibility* (of a theory T) is logically equivalent to the *directional symmetry of time* (where T is the fundamental theory).

In fact this brings out two different ways of defining time reversibility. Physicists commonly define it in the same kind of

⁵See Quine [1960] for the concept of essential occurrence.

way as they define invariance under spatial reversal, viz: T is time reversible just in case T is invariant when all occurrences of t (the temporal variable) are replaced by $-t$. This is rather imprecise as it stands, but it conveys the intended idea. The actual 'reversal operation' will be considered very closely in following chapters.

If this is our adopted definition of time reversibility, then what has been said in the previous section about space would hold equally for time. I.e. Time reversibility would be logically equivalent to the directional symmetry of time, just as space reversibility is equivalent to the directional symmetry of space.

However the definition of time reversibility that has already been adopted seems rather different: *a theory T is time reversible just in case, for every T -process P , P^R is also a T -process.* This definition is commonly found in philosophical discussions of reversibility but not so often in scientific discussions. It needs to be shown either that this definition is equivalent to the physicist's conception of time reversibility, or alternatively, that this definition also makes time reversibility logically equivalent to the directional symmetry of time. It is instructive to show both.

3.3 Equivalence of the Two Definitions of Reversibility.

The physicist's view is that we find the time-reversed version of a theory by replacing occurrences of t with $-t$. This talk of 'replacing occurrences of t with $-t$ ' is deceptively simple, and there is a lot left unsaid, but what this 'time reversal operation' must do is to generate a new theory, which I will denote by T^R , from the original theory T . The superscripted R once again denotes a time reversal operator - in this case a time reversal operator for theories. But to make clear exactly what this operator is we first need to clear up an ambiguity in our idea of what a *theory* is.

Traditionally a theory would be taken as a syntactic item - a class of sentences. Alternatively, in line with more modern logical techniques, a theory can be taken as the *proposition* which is expressed or denoted by the conjunction of the sentences. For most of our purposes it will be more useful to explicate a theory in this second way, as the *proposition*.

This proposition would normally be explicated in possible-world semantics as the class of (types of) worlds for which the theory obtains - i.e. T would be the class of T -worlds.

It is more convenient for some purposes, however, to take it as the class of *processes* for which the theory obtains - i.e. as the class of T -processes. (Adopting this 'possible-process semantics' instead of 'possible-world semantics' does not make a great difference, but some advantages will be pointed out later.)

Generally theories will be understood in this second, realistic

way. The term 'T' will represent a theory in this sense. I will use the italicized term '*T*' to refer to a syntactic item which denotes the theory T. The realistic understanding of theories simply makes the discussion much more direct.

When the theory is taken as a syntactic item, then the time reversal operator *R* is a *syntactic* operator. This is clearly what the operation of 'replacing occurrences of *t* with *-t*' would be - a syntactic operation, generating a new class of sentences out of the original class.

When the theory is taken as the proposition (the class of processes) then the time reversal operator for theories is an operation on that class. I will call this the *semantic time reversal operator*. This operator is easy to define (provided process-reversal has already been defined):

[3.4] Where *T* is any class of process-types, then T^R is the class of the time-reversals of the *T*-processes. (I.e. for any *P* in *T*, P^R is in T^R).

The alternative definitions of reversibility can now be recast in an analogous form. Compare:

[3.5] *T* is time reversible just in case $T \equiv T^R$ (Physicist's definition).

[3.6] *T* is time reversible just in case $T = T^R$. (Primary definition).

[3.5] is equivalent to the physicists' definition if we understand the operator R to be the *syntactic* operation of 'replacing occurrences of t with occurrences of $-t$ '.

[3.6] is equivalent to the primary definition if we understand the operator R as just defined above, since $T = T^R$ just in case every T -process is a T^R -process, which is just to say, in case the reversal P^R of every T -process P is also a T -process.

What must be demonstrated, therefore, is that [3.5] and [3.6] are logically equivalent. They will be equivalent just in case a certain relationship holds between the syntactic operator R and the semantic operator R . The necessary relationship is this: Where T denotes T , T^R must denote T^R .

In fact we cannot verify that this is so until we see in more detail how the syntactic time reversal operator R is defined. The recipe 'replace t with $-t$ ' is really just a rule of thumb, and needs to be further developed before it becomes precise⁶. But the point is that R is *intended* to be defined exactly to render the above equivalence true. The whole point of the syntactic operator is to make this equivalence true. The necessity to render the equivalence true provides the main guide for interpreting what the syntactic reversibility operator should be. In the meantime we will assume that the physicist's definition of reversibility is indeed logically equivalent to the primary definition.

⁶A precise rule is developed in the following chapter.

3.4 Reversibility and time symmetry.

The second problem is to show the equivalence of reversibility and time symmetry. What is meant precisely by this equivalence is that: *A theory T is reversible just in case T renders time directionally symmetric.*

What has been said in the last two sections already effectively shows why this equivalence holds. But it is enlightening to consider it directly.

First the definition of the directional symmetry of time:

[3.7] Time is *directionally symmetric* just in case the intrinsic properties of the temporal directions $+t$ and $-t$ are identical.

What are the 'intrinsic properties' of the temporal directions? They could be explicated in the same way as the intrinsic properties of spatial directions were explicated, by abstracting the temporal directions from the fundamental theory T and thus forming the complete 'nomological property' that T attributes to each of them. However they can be treated in a more direct and intuitive way.

It can be shown quite formally that the nomological properties of the temporal directions are basically *the possibilities for processes to occur in directions of time.*⁷ In other words, the

⁷It is assumed that other 'geometric' features of time, e.g. the cardinality of moments, and various features of the order relation among moments do not suffice

fact that *a process of type P is physically possible in the $+t$ direction of time* directly represents an intrinsic property of the temporal direction $+t$. Equivalently, the fact that *a process of type P is physically possible in the $-t$ direction of time* directly represents an intrinsic property of the temporal direction $-t$. Now by definition of R, *a process of type P is physically possible in the $-t$ direction of time* just in case *a process of type P^R is physically possible in the $+t$ direction of time*. (Process reversal is defined precisely to make this equivalence hold.) It follows immediately that, if there is some T-process P such that P^R is not a T-process, then the intrinsic properties of the temporal directions $+t$ and $-t$ are distinct. Thus:

[3.8] Irreversibility entails time asymmetry.

To show also that time reversibility entails time symmetry, and thus make the equivalence complete, what must be shown is that the *possibilities for processes to occur in a temporal direction* represent *all* the (nomological) properties of that temporal direction. To make this claim plausible, it is worth first dispelling what may seem a counterexample to it.

Imagine that T postulates time which has a beginning but no end. Obviously time in this theory should be judged asymmetric. The asymmetry may be brought out like this. Taking $-t$ as the to distinguish the directions of time, and these are ignored. If they do then asymmetry will be immediate. E.g. it might be that time has the 'geometric' structure of a ray rather than a line, having a 'first' moment but no last moment.

direction towards the beginning of time, we have that: $-t$ is the temporal direction from any given moment towards the first moment. Of course $+t$ does not have this property - in fact it has the contrary property that: $+t$ is the temporal direction from any given moment away from the first moment.

Thus the temporal directions differ by an intrinsic property, and so time is asymmetric on our definition. But is this asymmetry reflected by *irreversibility of processes*? That is to say, are there processes possible in the $+t$ direction which are impossible in the $-t$ direction?

Yes: any infinitely long process which has a definite beginning (initial state) but no final state is impossible in the $-t$ direction, since there is not infinite time available in the $-t$ direction. But there is always at least one such infinite process possible in the $+t$ direction, even if it is just the 'null process', where nothing happens at all.

It is also worth noting that the trivial temporal 'sequence' of a single state: $[s(t)]$ is here taken to be a process. Hence if s were a nomologically possible state, but s^R were not (which would render time asymmetric), the process: $P=[s(t)]$ would be nomologically possible, but: $P^R=[s^R(t)]$ would not be possible, rendering the laws of nature irreversible.

Thus the asymmetry of the intrinsic properties of the directions of time is reflected in process irreversibility. What has been said in the last three sections shows clearly enough for our purposes that time asymmetry always entails irreversibility.

3.5 Time flow and time symmetry 1.

I will now argue for [1.5]: that a theory that entails time flow must render time asymmetric, and consequently must be an irreversible theory. I will begin with a simple argument for this thesis, illustrated through the example of Newton's famous scholium on time. In the two following sections I will note a flaw in this argument, and give a deeper analysis of the problem.

A convenient example of a theory which entails time flow was practically offered by Newton himself: consider Newtonian mechanics with Newton's scholium on absolute time included as a *postulate* of time flow. In Newton's famous words: "Absolute, true, and mathematical time, of itself and from its own nature, flows equably without relation to anything external."⁸ Let us understand this as a postulate which states that *there is a flow of time*, and add it to the rest of the theory of Newtonian mechanics. (This was hardly Newton's intention, but I am only interested in concocting a useful example.)

By the 'flow of time' we understand more precisely that the world suffers real change, and thus that real properties of being *past, present, or future* attach to events, or to the moments at which events occur. (These properties of course change with time).

A key feature of time is therefore this: there is one *direction* of time, which I will denote $+t$, which is the direction into the future, and which is not the direction into the past. The opposite

⁸Newton [1962, p.6]

direction, $-t$, is the direction into the past and is not the direction into the future.

Hence the theory entails that the two directions of time differ by a property: $+t$ has the property of being the *direction into the future*, $-t$ does not have this property. Since this asymmetry of properties is nomological, assigned by the theory, this comprises an intrinsic asymmetry of time.

It is useful to emphasize why the *ad hoc* addition of a postulate of time flow of this kind is unsatisfactory. The problem is that, in the theory that arises, no relation is established between it and any other physical concepts. In particular, the *ad hoc* postulate of time flow has *no implications for any other observations we might make*. Equivalently, there is no possibility of physical evidence for time flow: no inference from physical observations to the existence of time flow has any basis. One would expect to be able to appeal to a set of physical observations, O , for evidence that time flows in the $+t$ direction. But so far as the laws of nature establish any connection between time flow and physical facts, any observation O would count just as much as evidence that time flows in the opposite direction, $-t$. So what is the reason for proposing time flow in the $+t$ direction rather than the $-t$ direction?

One might imagine that what counts as evidence for time flow is provided by a 'common sense' understanding of time flow, rather than by the physical theory. For example, perhaps the direction in which time is *experienced* to be passing is the only plausible direction for time flow, and thus our experience

provides evidence of time flow after all. But this could be so only if time flow has relations (e.g. to kinds of experiences) that are simply not allowed for by the physical theory. This is to imagine that the physical theory is *incomplete*. To be completed it would have to be strengthened to imply further relations between time flow and physical (or mental?) situations: but adding these relations is exactly what would turn it into an empirical, rather than an *ad hoc* theory of time flow. For instance, one might propose, instead of simply that 'time flows in the + t direction', that 'time flows in the direction of major entropy increase'. The theory would now be an empirical one.⁹ Similarly, the proposal that I make - that time flows as physical probabilities are actualised - generates an empirical theory. It means, for instance, that the direction of time flow can be *observed* - by observing the temporal direction of physical probabilities, which in a world like ours is very easy to do.

3.6 A logical puzzle about reversibility.

The argument given above is plausible, but there is a flaw in it. To bring this flaw out, it is useful to first consider an apparent logical puzzle for the concept of reversibility. The puzzle is seen most easily if we take a *possible-worlds* instead of a *possible-process* semantics for theories. I will use an alternative script (\mathbf{T} , \mathbf{T}^R , etc) where a possible-worlds semantics is intended.

⁹Reichenbach [1956] seems to propose something like this as a *conceptual truth*, but that is not what is intended here.

Take a theory T to be represented as a class of possible worlds, viz. the class of worlds in which T holds. Imagine also that T is irreversible. This means that $T \neq T^R$. I.e. there are T -worlds which are not T^R -worlds.

But consider now a new theory, T^* , formed as the *disjunction* of T and T^R . I.e. $T^* = (T \text{ or } T^R)$. T^* is obviously reversible.¹⁰ As well as being reversible, it might be argued that T^* is just as good in every practical respect as T , for the following reasons. Firstly, the *confirmation* we have of T is also confirmation of T^* (since T entails T^* .) Secondly, the chances of *our disconfirming* T are just the chances of *our disconfirming* T^* . This is because it can be presumed that T^R has already been disconfirmed (otherwise we are premature in holding to T rather than to T^R in the first place), and thus any disconfirmation of T^* must be through the disconfirmation of T . Thirdly, supposing that T^R has been disconfirmed of *our* world already, T^* then implies that T must be true of our world, and in this way T^* supplies us with all the predictions about our actual world as does T .

Hence it might be argued that the only difference between T and T^* is a 'metaphysical' kind of affair: T rules out a certain class of worlds as being nomologically impossible (the 'irreversible' T -worlds), which T^* does not rule out. Thus according to T^* , the class of nomologically possible worlds is larger than according to T . But if one is within a T -world, there is no test to tell which theory is correct (which is another way of

¹⁰Since T^* is the union of T and T^R , every w in T^* is in T or in T^R ; if w is in T , then w^R is in T^R , hence in T^* ; similarly, if w is in T^R , then w^R is in T , hence in T^* . Hence for every w in T^* , w^R is also in T^* , and hence T^* is reversible.

saying that there is no practical difference for us between T and T*.) It seems therefore that the difference between T and T* is 'metaphysical', in the positivistic sense of being observationally undecidable.¹¹

If this is so, then we seem to have a simple way of overcoming the irreversibility of any theory. If T is irreversible, then just adopt T* instead. We lose nothing of practical value: we gain reversibility. This possibility would seem to rob reversibility of practical importance.

An adequate solution to this problem must involve a denial that the confirmation we have of T* is necessarily as good as the confirmation we have of T. It must be possible to have stronger reasons for holding to the latter theory than to the former. I will just observe that this is a result that must hold of a satisfactory theory of confirmation for more general reasons. For otherwise, the same argument would apply where T* was the disjunction of T and any other theory; rather than holding to T, we would end up holding to: (T or T₁ or T₂ or...), where T₁, ... are all the theories we could think of. This would make science a little absurd.¹²

An adequate theory of confirmation can, therefore, be expected to provide a solution to this problem; but the idea on which it is based raises another problem, this time for the thesis of [1.5] that a theory which entails time flow must render time asymmetric.

¹¹This emphasises the fact that reversibility is a property of the *space of possible worlds*, not of any individual world.

¹²This solution was pointed out to me by Graham Oddie.

3.7 Time flow and time symmetry 2.

Imagine that the irreversible theory T considered in the previous section is a theory which entails time flow. Let us suppose that it directly entails that the direction of the future is $+t$. In this case its reversal, T^R , will also entail time flow, but it will give time flow the opposite direction, i.e. it will entail that the direction of the future is $-t$. Now the theory T^* is both reversible, as described above, and entails time flow. For T^* entails that for any world, either T is true of it, or T^R is true of it: in the first case time 'flows towards $+t$ ', in the second case, time 'flows towards $-t$ '. T^* therefore seems to be a counterexample to [1.5].

In the first place it should be noted that if this argument was accepted, and [1.5] rejected, it could only be to the advantage of the programme of finding time flow in physics. For it would mean that reversibility does not rule out time flow, and hence raise the prospect that even a reversible theory might entail time flow. This would just widen the class of respectable physical theories which entail time flow.

Nevertheless it seems to me that the argument against [1.5] is mistaken, and that [1.5] is essentially correct. I will try to show why this is, because it is important for the understanding of what a physical theory of time flow involves.

Firstly, if the artificial device of adopting the theory: T^* (= T or T^R) was the only way of getting a reversible theory which

entailed time flow, then the problem might be dismissed as academic. For T^* would be *logically possible*, but that doesn't mean that it has any respectability as a physical theory. And indeed, as we have seen, a satisfactory theory of confirmation should judge T^* an inadequate theory.

However T^* is not the only possibility. Consider another kind of theory, T , concocted in the following way from a respectable reversible theory, such as classical mechanics. T represents all the usual postulates of classical mechanics, but into the bargain, it represents a postulate of the following form: that there is time flow in a world w in the temporal direction $+t$ just in case w has a certain property P ; and there is time flow in a world w in the temporal direction $-t$ just in case w lacks the property P . This means that time flows in each T -world, but that the direction of flow depends upon a contingent feature (P) of that world.

We may suppose moreover that the property P partitions T in a 'time symmetric' way, so that for any world w , if w has the property P then w^R does not, and if w does not have the property P , then w^R does.¹³ In this case, T will be reversible. Hence T is a reversible theory which entails time flow.

Before discussing the problem which T represents, some comments on its general structure will be helpful. It can be seen that T represents the general structure of any reversible theory which entails time flow in every world. For (i) since T entails

¹³For instance, P might be a property, roughly, that *greatest entropy increase occurs in the direction $+t$* . This property would have to be made far more precise, since as stated it would be undefined for most worlds, but the general idea of identifying the 'future' with the direction of entropy increase has been seriously proposed by Reichenbach [1956].

time flow in every world, in any T -world, w , either $+t$ must be the direction of the future, or $-t$ must be the direction of the future. (ii) If $+t$ is the direction of the future in w , then $-t$ is the direction of the future in w^R (by the definition of time reversal), and since T is reversible, w^R must be in T . (iii) Similarly, if $-t$ is the direction of the future in w , then $+t$ is the direction of the future in w^R , and since T is reversible w^R must be in T . (iv) Hence T is partitioned into two classes, call them T_+ and T_- , such that time flows in the $+t$ direction in each world w in T_+ , time flows in the $-t$ direction in each world w in T_- , and T_+ is exactly the set of reversals of the worlds in T_- .

A more general kind of reversible theory, which entails time flow in only *some* worlds, can be constructed by adding to T a further (reversible) class of worlds, in none of which is there time flow. But everything in the discussion below which is said about theories of the form T also applies to theories of this more general kind, so we may restrict our attention to T .

The problem that faces us, if [1.5] is to be preserved, is to find a good reason against the possibility of a theory of the kind T . A serious problem with T strikes us immediately: T makes the direction of time flow a *contingent matter*. For according to T , whether time flows in the direction $+t$ or $-t$ depends upon the contingent fact of whether the world has the property P or not. It goes against our normal intuitions about time flow that its direction could depend upon contingent features of the world in this way.¹⁴ For instance, suppose that the feature P was that

¹⁴This point has been brought out well in criticisms of the idea that 'the future direction' might be identified with some physical criterion, such as 'the direction

thermodynamic systems undergo entropy increase in the +t direction more frequently than in the -t direction. If one believes that it really is only a contingent fact that entropy increase in our own universe occurs mainly in the direction it does, then one can easily imagine that things might have been a little different, and major entropy increase might have occurred in the opposite direction. This would just require changes in certain boundary conditions on the physical state of the universe. But it is hardly plausible that this change in physical states could have the 'metaphysical' effect of reversing the real direction of the future, which is to say, reversing the direction of *real change*.¹⁵

This kind of intuition indicates that a contingent theory of the direction of time flow could not do justice to the concept of time flow, so that the theory T being considered could not really be a theory of time flow after all. I think this is so, but what is required is some further analysis, to make the basis of this intuition clear. There must be some deeper source for it in our concept of time flow: I will now attempt to describe this source.

of entropy increase in thermodynamic systems'. See particularly Grunbaum [1973].

¹⁵Of course, if by the 'future' is only meant something like 'the direction which we experience as 'the future' because of our psychological makeup', then it might be alright to identify it contingently. But we are here considering the idea of the future in its 'metaphysical' sense.

3.8 The direction of time flow cannot be contingent.

My discussion depends firstly upon the notion of *identities of moments across (nomologically possible) worlds*. The concept of cross-world identities is controversial, but the general assumptions about it that are needed here are already implicit in the problem at hand, and in the realistic treatment of physical theory that is the foundation of this inquiry.

In the first place, the notion of cross-world identity of the *directions of time* is already assumed in the problem we are considering. For the assumption that the theory T makes the temporal direction of the future $+t$ in world w , but $-t$ ($\neq +t$) in world w^R , already involves us in presuming that $+t$ is the 'same direction of time' in both worlds. It will quickly become obvious that the whole notion of ontology which we are presuming would not make sense without the idea of cross-world identities of moments, so this idea will be presumed as unproblematic.

The main reason we need to assume cross-world identities is so that we can make sense of certain kinds of counterfactuals. It seems that scientific theories must support these counterfactuals if they are to be realistically interpreted. The counterfactuals of prime concern here are of the form: *if the state of the world at the past time t had been s^* , (instead of s), then the present state of the world, at time t_1 , would be s_1^* (rather than s_1)*. Notice that, by its use of tenses (present and past), this counterfactual presupposes that time is dynamic, so I will call it a *dynamic counterfactual*. The first point is that

physical laws which entail time flow will support such dynamic counterfactuals. The theory T which we are considering will support dynamic counterfactuals like this.

I will now argue for two principles concerning cross-world identities of times. The first principle holds whether time is taken to be static or dynamic:

Principle 1. The temporal directions between moments are preserved across worlds.

This means that if the direction from t to t_1 is $+t$ in world w , then the direction from t to t_1 is $+t$ in any other world, w_1 . This is a logical truth.¹⁶

The second principle presupposes a dynamic view of time, since it is about the dynamic concept of 'the present moment':

Principle 2 When the present moment in world w is t , the present moment in any other world, w_1 , is also t .

This principle involves two major ideas. First is the idea that each world has a 'present moment': this is a necessary part of the dynamic concept of time, and is not under question here. The second idea is that relations of 'co-presentness' hold across

¹⁶The temporal directions $+t$ and $-t$ are properties of temporal vectors. A temporal vector t is definable as an ordered couple of moments, (t, t_1) . The direction *from* t to t_1 is $+t$ just in case the vector (t, t_1) has the property $+t$. Because the extensions of $+t$ and of $-t$ are world-independent, temporal directions are world-independent.

worlds. The whole point of Principle 1 is to postulate such relations: this will need considerable argument.

(i) First, it is implicit in the notion of the dynamic ontology that it makes sense to talk of *what would presently be the case if the actual world were w^* instead of w* . For this notion of co-presentness across worlds is required if we are to make sense of the counterfactuals of the form: *if the state of the world at the past time t had been s^* , (instead of s), then the present state of the world at t_1 would be s_1^* (rather than s_1)*, which are supported by dynamic theories, as observed above.

(ii) It might be objected that we do not need the notion of *what would presently be the case if the actual world were w^* instead of w* , which implies co-presentness across worlds. Instead the idea of *what would be the case at the moment t_1 in the history of w^** might suffice. But it does not suffice, simply because it does not suffice in the *actual world* (given that the actual world is a dynamic one.) That is to say, if we are trying to imagine that w^* is the actual world rather than w , then we must imagine that w^* has a certain *present moment*, just as the actual dynamic world does. Otherwise w^* simply is not a dynamic world. The present moment in w^* must have similar features to the real present moment, in particular it must undergo *change*. The only plausible way of representing these features in w^* seems to be to assume relations of co-presentness with the actual world.

(iii) If these relations of co-presentness across worlds are admitted, it still remains to be established that the relation should be as Principle 2 states, i.e. that when t is the present moment in w , that t , rather than some other moment t' , would be the present moment in w^* (if w^* rather than w was actual). I suggest that this simply be taken to be what the cross-world identity of moments *means* in the dynamic theory. *The relations of co-presentness of moments across worlds gives a substantial relation on which cross-world identity of moments can be based.* It is certainly the most natural assumption to make, and it is hard to imagine what might take its place.

(iv) Perhaps the strongest reasons for assuming relations of co-presentness across possible worlds emerge in the context of the kind of dynamic probabilistic ontology discussed in Chapter Seven. For in this ontology, it must be possible to represent the fact that, when the present time is t , the present state of the world leaves it undetermined whether the future state will be that of world w or of w^* . Hence it is not presently determined what world it is that *will* come about. Yet, whichever it is, its *present time is t*. Suppose that w is the world that really will come about. Then the counterfactual: *if the world to come about were w^* , its present moment would be t* would be true, and supported by physical law, implying the required relation of co-presentness across worlds. More generally, the 'branching futures' model of indeterminism developed in Thomason [1969] and McCall

[1976], and which has now become a popular device for representing an indeterministic ontology, presumes that there are relations of co-presentness across physically possible alternative future 'branches'. This shows that it is a natural assumption, that philosophers find very useful to make.

Having established a case for the conceptual naturalness of Principle 2, I will now show how it solves the problem with the theory T. T implies that there are possible worlds, w and w^R , in which time flows in the directions $+t$ and $-t$, respectively. But this contradicts Principles 1 and 2. For suppose that the present time in w is t . By Principle 2, the present time in w^R is also t . Now since time flows in the $+t$ direction in w , there is a time, t_1 , such that $t < t_1$, and t_1 will be the present moment in w (in the future). But when t_1 is the present moment in w , it is also the present moment in w^R , by Principle 2. Hence it is also true that t_1 will be the present moment in w^R (in the future). But this directly contradicts the hypothesis that time flows in the $-t$ direction in w^R . Hence a theory of the form T involves a contradiction, and we can conclude that [1.5] is justified after all.

I hope that this argument helps to bring out the deeper principles that underlie the intuition that the direction of time flow must be a necessary feature, not merely an accidental one. This thesis has little direct practical importance, for it is implausible that anyone will seriously maintain a reversible theory of time flow. The value of the discussion at this level is

rather to try to get a clear organisation of the conceptual framework.

3.9 Some basic theorems.

The following theorems are referred to in following chapters, and are essential for a basic grasp of the notion of reversibility.

[3.9] For any process P , $(P^R)^R = P$.

That is to say, double time-reversal is an identity operation. This follows from the definition of time reversal for processes, to be discussed in detail in the next chapter.

[3.10] For any theory T , $(T^R)^R = T$.

A process P is in T just in case P^R is in T^R (definition of T^R). Similarly, P^R is in T^R just in case $(P^R)^R$ is in $(T^R)^R$. Hence, since $(P^R)^R = P$, P is in T just in case P is in $(T^R)^R$. Thus $(T^R)^R = T$.

[3.11] For any theory T , if T entails T^R , then $T = T^R$.

For a reductio, assume that T entails T^R , but that $T \neq T^R$. Since T entails T^R , every T -process is a T^R -process. I.e. T^R contains T . Therefore, since $T \neq T^R$, there must be some process, P , which is in T^R but is not in T . Since P is in T^R , P^R must be in T (definition of T^R), and since T is in T^R , P^R must also be in T^R . Hence $(P^R)^R$,

which is just P , is in T (definition of T^R). But this contradicts the assumption that P is not in T .

[3.12] A theory T is reversible just in case, for every proposition L entailed by T , L^R is also entailed by T .

(i) Suppose T is reversible. Let L be any proposition entailed by T . This means that every T -process is an L -process, or that L contains T . Since L contains T , L^R contains T^R (by the definition of reversibility). But since T is reversible, $T=T^R$. Thus L^R contains T , i.e. T entails L^R . (ii) Suppose that for every proposition L entailed by T , L^R is also entailed by T . T is a proposition entailed by T , hence T^R is entailed by T . By [3.11], $T=T^R$, thus T is reversible. [3.11] can be used as an alternative definition of reversibility, and is often appealed to in the following chapters.¹⁷

[3.13] Reversible theories entail non-reversible consequences.

It is important to note that the reversibility of T does *not* mean that all consequences of T are also reversible. On the contrary,

¹⁷ Earman [1974] p.548 states: "...although laws may be reversible, they may be highly noninvariant under the interchange of *earlier* and *later*; that is, if C is a consequence of law L and C' is obtained from C by exchanging the roles of *earlier* and *later*, then C' may not be a consequence of L even though L is reversible". This appears to be a denial of the theorem above, and if so it is a mistake. It is a strange one, however, since Earman is well aware of this theorem (e.g. see his [1969, p.281]).

any (non-trivial) reversible theory T is bound to entail consequences which are not reversible. For instance, if T is any (non-trivial) theory, let P be a process which is not in T , and which is not identical to its reversal (i.e. $P \neq P^R$). Form a new proposition, T^* , as the class containing just T and P (and not containing P^R). T entails T^* , but $T^{*R} \neq T^*$ (since P is in T^* but is not in T^{*R}). Thus T^* represents an *irreversible* law of the theory T .

Although the last two theorems are elementary, a lack of appreciation of them is sometimes a source of confusion about the implications of reversibility.

3.10 Summary.

A clear analysis of the concepts of *time reversibility*, *directional symmetry of time*, *time reversal*, and *phenomenological directedness* has been arrived at. Here is a summary of some important features of these concepts:

1. *Theories* can be said to be time reversible or irreversible.
2. Relative to a theory, processes can also be said to be reversible or irreversible. A process of type P is *irreversible* in the context of T just in case P is a T -process while P^R is not.

3. *Time* can be said to be directionally symmetric or asymmetric. By this is always intended that time is *intrinsically* (nomologically) symmetric or asymmetric.
4. Time is symmetric in a T-world (i.e. a world for which T is the fundamental physical theory) just in case T is time reversible. Thus time reversibility and time symmetry are effectively equivalent.
5. There are *time reversal* operations for *theories*, *process-types*, and *statements*. These operations are closely linked. Important equivalences are: $T = T^R$ just in case $T = T^R$, and: P is in T just in case P^R is in T^R .
6. *Phenomenological processes* can be said to be phenomenologically directed or not. Whether a given process is phenomenologically directed is relative to one's environment or world.
7. Phenomenological directedness in the environment is not logically dependent on the reversibility or otherwise of the fundamental laws.

3.7 M. Bunge on reversibility.

Unfortunately there is not a complete standardization in the literature of the technical vocabulary, or the meanings to be attached to various terms, and a number of alternative treatments can be found. Most of these (e.g. Grunbaum [1973], Mehlberg [1980], Reichenbach [1956], Earman [67,69,74]) are essentially similar to what I have proposed here, at least in their

general distinction of categories. But at least one very different systematization of concepts, which cannot be accommodated with mine, has been proposed, by Bunge [1970]. I will conclude this chapter by criticizing Bunge's system in detail.

Bunge first complains of general confusion about the concept of time, and states:

One such confusion, perhaps the most harmful of all, is the conflation of three quite distinct ideas huddled under the umbrella of the so-called 'arrow of time': time asymmetry, non-invariance under time reversal, and irreversibility. Let us try to clear up this confusion even at the risk of error. [1970, p.122].

He is certainly right in this judgement. The term 'the arrow of time' has no precise meaning, and many physicists and philosophers slip in and out of various uses of it without any awareness of crucial distinctions. It is sometimes used ambiguously to mean both the *flow of time*, the *asymmetry of time*, and sometimes the *phenomenological directionality* of the environment, which is a considerably worse confusion than that which Bunge is complaining of here. Incredibly enough, Bunge himself slips into exactly this confusion in the first section of his paper.

Bunge attempts to give precise definitions of time asymmetry, non-invariance under time reversal, and irreversibility.

He starts by introducing the concept of a *local time function*, **T**. The purpose of **T** is to give the *temporal duration* from any

event, e , to any other event, e' . Such durations are relative to the choice of a *reference frame*, denoted k , and a *chronometric scale*, denoted s , and thus strictly T is a function which takes four arguments: $T(e,e',k,s)$. The value of $T(e,e',k,s)$ is a real number, t , representing the duration *from* e to e' , in the reference frame k and chronometric scale s .

For our purposes reference frame and chronometric scale are irrelevant, so I will assume they are fixed, and take T to simply map ordered pairs of events to real numbers. Thus: $T(e,e') = t$ means that the duration from e to e' is t .

The first point is that 'duration is an oriented interval', which just means that the duration *from* e to e' is the negative of the duration *from* e' to e . I.e. $T(e,e') = -T(e',e)$.

Pointing this out, Bunge then observes: "This is all there is to the asymmetry or anisotropy of time."

What Bunge is calling the *asymmetry* or *anisotropy* of time is what is normally meant by saying that there are *two temporal directions* - $+t$ and $-t$, or 'earlier' and 'later'. Having two directions is of course a basic topological property of time, equivalent to the normal (objectivist) understanding of time as a one-dimensional continuum of moments. (There are always two directions on a line.)

The trouble with Bunge's definition of *asymmetry* is that it is not what is normally meant by asymmetry at all. I know of no important writer who uses the term *time asymmetry* (or equivalently *anisotropy*) in Bunge's sense¹⁸. To express what

¹⁸However, see Reichenbach [1956, p.32] for some similarities with Bunge, and

Bunge calls the asymmetry of time we would normally just say that *there are two temporal directions*. When the question of *time symmetry* (or *anisotropy*) is discussed in any normal scientific or philosophical context, it is usually being asked whether the laws of physics render time asymmetric. I think that Bunge's definition of time asymmetry must be rejected as completely misleading.

After defining time asymmetry, he then makes the following strange statement:

In other words, the asymmetry of time ... is a fact, but the decision to count time forwards, i.e. in the direction of coming events, is arbitrary. Put in metaphorical terms: nature tells us that time 'flows', but not wither. Better: time has no arrow built into it. Arrows must be sought in whole processes not in one of the features of processes.

I hardly know of a better example of the confusion over temporal concepts that Bunge has just been complaining of than this passage. It is hardly possible to guess at what Bunge intends here, but one claim seems clear. This is that "nature tells us that time 'flows', but not wither." The implication is that time 'flows' *because* time is (in Bunge's sense) asymmetric. That is to say, time 'flows' because there are two temporal directions. This is bizarre indeed. Does the street in the front of my house also 'flow' because it has two directions?

Bunge's treatment of *time reversal* is better, although it lacks

Grunbaum [1973,p.218] for a criticism of Reichenbach on this point.

clarity at many points. He defines time reversal as the 'mathematical operation' of 'inversion of the sign of the time variable or coordinate'. This is of course the *syntactic* time reversal operator described in Section 3.3 above. In his discussion Bunge tries to make clear what this operation is intended to do. Basically, if P is the *description* of a process-type, then the time reversed description, P^R , should be a description of the reversed process-type. Bunge expresses this in the prescription *time reversal corresponds to process reversal* (p.126, his italics.) This not a very clear statement, but his intentions are clear enough if one already has a good grasp of the subject, and this section on time reversal is of some interest.

Bunge then defines the *invariance under time reversal*, or equivalently *T-invariance*, of law statements. If L is a law statement and its time reversal is L^R , then L is T-invariant just in case $L = L^R$.¹⁹ This of course is what I have called the *reversibility of L*. (It is what I called in Section 3.3 the 'physicists definition' of *reversibility*.)

Bunge calls it instead *time reversal invariance*, or *T-invariance*, but rejects the label *reversibility*. The former two terms are perhaps more common in the physics literature than the term *reversibility*, but nevertheless it is still common enough to call laws *reversible* rather than *T-invariant*. Certainly most other writers would understand L is time reversible to mean the

¹⁹Bunge uses a different nomenclature, denoting L as $L(t)$, its time reversal as: $L(-t)$. This is not an uncommon terminology, but it hides the real complexity of the syntactic time reversal operation, and I prefer to make the operator clear by denoting it as R .

same as L is T -invariant. Bunge however gives a completely different meaning to *reversibility*, and here I think he seriously misleads us a second time.

Bunge's definition of *reversibility*:

Reversibility is a property of certain processes, most microphysical. A reversible process is, strictly speaking, one in which both the system concerned and its surrounding can be restored to their original condition. (p.127).

'Strictly speaking' this hardly is: the definition as it stands is awfully vague, and unfortunately not much is done to elucidate it. For instance, does the *can* in 'can be restored' refer to human agency, or nomological possibility, or what? What does it mean to say that 'the *surroundings* can be restored to their original condition'? How much of the rest of the universe do 'the surroundings' include, and what exactly is the degree of 'restoration' that is required? Need the positions of cars in the street, the clouds in the sky, or the planets in the heavens be restored, before the 'reversal' of a given process is achieved?

Although Bunge does not clearly define his concept of 'reversibility', his intentions are revealed to a some extent in his discussion, and I will do my best to interpret the idea he is trying to get at.

What Bunge seems really to be thinking of is something that I find far more natural to call *restoration from change*. A *change* always destroys something – in general, it destroys the state of

some object or system and brings about a new state. For instance, breaking up a jig-saw puzzle into its separate pieces is a change to the state of the jig-saw. Smashing a cup is a change in the state of a cup. Burning a witch to death is a change in the state of the witch.

Some changes are regarded as 'reversible' in the sense that the original state of the object could be restored if one wished. We don't get too upset when someone dismantles a jig-saw puzzle, because we know that with a little patience we can restore it to its former state. Indeed this is the whole point of jig-saw puzzles. Other changes are more difficult to restore. A broken cup might be reassembled with glue, but it is difficult to do it convincingly. One always feels that the glued-together cup is not really the same as the original cup - restoration in this case is an illusion of restoration, not the real thing. Other changes again are accepted as quite impossible to restore - a burnt witch cannot be restored to her former living state at all. Death in general is taken to be an irreversible change in this sense, which is why it is considered fatal.

Notice that in discussing this kind of 'reversibility' it is natural to talk of the *restoration from changes*, rather than the *reversal of processes*. Consider, for instance, the process of a volume of water running from a reservoir down a creek-bed into a lower reservoir. In Bunge's sense, this process or change in the water is *reversible*, because the original state of the system can be restored: we can pump all the water back up to the first reservoir, and restore the original condition. But in fact this

restoration of the original state has nothing to do with the *reversal of the process* that brought about the change. The reversal of that process would involve water running back up the creek-bed, but that of course is (physically) impossible, and is not the process by which the original state is restored at all. Bunge's concept of reversal thus has nothing to do with *process reversal* in its usual sense.²⁰

More generally, Bunge's notion of reversibility is better treated as a relation between *states* than as a property of *processes*. Say that a system X has initial state s_1 and goes through a process which leaves it at some later time in state s_2 . This process is reversible (in Bunge's sense) just in case the system can be restored to the state s_1 . (See the last footnote). But of course whether this is so or not does not depend on the actual process (the path of states) which the system went through in getting from state s_1 to state s_2 . This process is irrelevant in deciding the reversibility or otherwise of the change of state of the system. Thus what is reversible is not strictly the process the system went through in getting to s_2 , but just the

²⁰According to Bunge we have to restore the 'surroundings' of the system to its original condition as well, but this is such a vague requirement that I will ignore it. However, if you are worried that, for instance, the motor that ran the pump is now lower on fuel than in the original situation, then just imagine getting some more fuel for it from outside the region that counts as the 'surroundings'. If you are worried that the clouds in the sky are in a state different from their original state, then you have taken such a strong view of 'restoring the surroundings' that it is doubtful whether any changes would be reversible - which conflicts with Bunge's clear supposition that some are.

change from s_1 to s_2 . Bunge should call the change from s_1 to s_2 reversible just in case the system in s_2 can be restored to the state s_1 .

What the motivation for developing this concept and calling it 'reversibility of processes' is I am not certain. However there are some textual hints that what Bunge is trying to capture with his notion of irreversibility is the notion that I have called *phenomenological directedness*. (In fact *phenomenological irreversibility* would be a reasonable alternative term, but I wish to preserve the term *irreversibility* with just its single key meaning.) If this is Bunge's intention, however, it fails badly.

Consider again the process of the water running down the creek-bed. Obviously this is a phenomenologically directed process, or if you prefer the term, phenomenologically irreversible. For the phenomenological reversal - water running up a creek - never occurs and is impossible to bring about.

But in Bunge's sense, this process is *reversible*, because the original state of the system can be restored - we can pump all the water back up to the first reservoir, and restore the original condition. Thus Bunge's concept of irreversibility does not capture phenomenological directedness of processes at all.

This is seen even more clearly with another example. Imagine that the actual universe is fully deterministic, and furthermore that it goes through a closed cycle of states. We are caught up in a deterministic Nietzschean eternal return. The precise state of the universe at any given moment is destined to recur after a finite period of time. We can imagine that this is so while

phenomenologically the world looks just as it does now. In such a universe, the state of every system would eventually be restored perfectly. Thus any change would eventually be restored *in fact*. Thus all changes are restorable - or in Bunge's sense, all processes are reversible. This includes, for instance, the process of the witch burning: this would be a reversible process in Bunge's terms, since the original state of the witch would one day be restored perfectly.²¹

Thus on Bunge's definition the *apparent* 'irreversibility' of a process is no guarantee of irreversibility. It is clear that Bunge's concept of reversibility does not capture *phenomenological directedness*, and I cannot see anything interesting that it does capture. It certainly does not capture any understanding of 'irreversibility' that is common in the literature.

I conclude then that Bunge's definitions of both *time asymmetry* and of *reversibility* are quite wrong. In his discussion of time reversal he makes some worthwhile points, but the bulk of his paper can only serve as an example of the dangers of a lack of conceptual clarity in this field.

²¹Notice also that since the laws governing such a universe could well be expected *not* to be T-invariant, Bunge's conclusion that "If a process is reversible then its laws are T-invariant" is wrong. This is a conclusion he stresses, (Section 4, statement 3(b)), but I can find no possible interpretation of 'irreversibility' to justify this.

CHAPTER FOUR

TIME REVERSAL OPERATIONS.

In previous chapters, *time reversibility* has been defined by:

A theory T is reversible just in case, for every T-process, P, the time-reversal P^R of P is also a T-process.

This will be called the *primary definition* of reversibility. Before this definition can be properly understood, or used in the practical analysis of any real theory, the concept of *process reversal* must be understood. This the main topic of the present chapter.

4.1 The physicist's definition and the syntactic reversal operator.

The key problem in applying the primary definition is interpreting the process reversal operator, R . It might be thought that this problem can be sidestepped by using the 'physicist's definition' of reversibility instead, which is:

A statement T is reversible just in case $T \equiv T^R$, where ' R ' denotes the syntactic time reversal operator.

This does not mention process reversal. Instead it employs the concept of the 'syntactic time reversal operator', R , but the syntactic operation is defined simply enough in most texts on the subject as just the replacement of the temporal variable, t , with its inverse, $-t$ (E.g. Davies [1974] p.23, Mehlberg [1980], p.205), which does not seem difficult to understand. So it appears that a direct practical analysis of reversibility is possible through the physicist's definition, without the need to understand process reversal.

To some extent it is true that the physicist's definition, rather than the primary definition, is what is employed in the practical analysis of reversibility. This is because in practice we always deal with theories by considering and manipulating *statements* of them, and the physicist's definition tells us how to manipulate the statement of a theory to find its time reversal. But this does not mean that we can ignore the understanding of process reversal. In fact the understanding of process reversal is

logically prior to the understanding of the syntactic reversal operator.

This is because the syntactic reversal operator is designed with a certain purpose in mind: roughly speaking, to duplicate process reversal in the statements used to describe processes. If a statement T captures a theory T , (which is to be thought of as a class of processes), then T^R must capture T^R (the class of reversed processes). This is the whole point of the syntactic operator, and if a syntactic operator does not have this effect, then it fails to be the syntactic time reversal operator. Thus, the syntactic reversal operator that is adopted must be *justified*, and its justification involves reasoning directly about the nature of process reversal. So we cannot escape consideration of process reversal.

A second point is that, although the syntactic operator is commonly described as simply involving 'the replacement of t with $-t$ ', in fact this is only a rule of thumb. It is not literally correct. To apply the rule correctly requires a certain degree of informal judgement, and this judgement depends upon intuitions about process reversal.

It is worth illustrating this with an example. Consider, for instance, the following law statement, L :

[L] If $s_1(t)$ then $s_2(t+\Delta t)$

This law states that if a system is in the state s_1 at t , then it will be in state s_2 at $t+\Delta t$.¹ A law of this kind will be called a

*deterministic transition law.*²

Consider what happens if we follow the rule 'replace temporal variables with their inverses' quite literally in the case of L . We get the following statement, L^* :

[L^*] If $s_1(-t)$ then $s_2(-t-\Delta t)$

Since t is implicitly universally quantified in this statement, this is just:

[L^*] If $s_1(t)$ then $s_2(t-\Delta t)$ ³

[L^*] is *not* the time reversal of L , however. L^* would be the time reversal of L only if *the time reversal of any L -process was an L^* -process*. But this is not so. To see why not, we must turn to consider directly what the time reversal of *processes* amounts to. This is examined in detail in following sections, but a rough explanation can be given here, based on the following important theorem which will be demonstrated in Section 4.4:

¹Conditional upon the system in question remaining isolated through the period.

²Physical theories are not usually directly represented in terms of deterministic or probabilistic transition laws, because there are far more elegant mathematical ways of representing dynamic laws.

³This states that the state s_1 must be *preceded* by the state s_2 at a moment Δt earlier.

[4.1] The time reversal of any process-type that involves a transition from a state s_1 to a state s_2 in a period Δt must be a process that involves a transition from a state s_2^R to a state s_1^R in a period Δt .

Here R is the time reversal operator for states, so that s_1^R and s_2^R are states which are the time reversals of the states s_1 and s_2 respectively. State reversal will be defined in the next section, but it is enough for the moment that (i) in general a given state is not identical to its time reversal, so that we may assume that: $s_1 \neq s_1^R$, and: $s_2 \neq s_2^R$, and (ii) as with process reversal, double application of time reversal for states returns us to the original state. That is, $(s^R)^R = s$, for any state s .

A process can now be found which is an L -process but not an L^* -process, as follows. Suppose first that s is some (arbitrary) state which is not the reversal of s_2 , i.e. $s^R \neq s_2$. Now consider a process P that involves a system running from the state s_1^R to the state s , in a period Δt . $P = [s_1^R(t) \rightarrow s(t+\Delta t)]$. P is an L -process. (Remember that L just puts a *constraint* on processes, viz: If s_1 at time t , then s_2 at time $t+\Delta t$. The process under consideration satisfies this constraint, and is hence an L -process.)

By [4.1], the reversal of this process must run from $s^R \neq s_2$ to $(s_1^R)^R = s_1$, in a period Δt . But this reversed process fails to be an L^* -process, since it breaks exactly the constraint that L^* imposes - namely, the constraint that *the state s_1 must be preceded by the state s_2 at a moment earlier by Δt* . Hence the

reversal of this L-process fails to be an L^* -process, and thus L^* cannot be the time reversal of L .

The rule of 'replacing t with $-t$ ' therefore cannot be taken literally, since taken literally it delivers quite the wrong result for the time reversal of a statement of the form of L . This will hardly surprise anyone acquainted with the subject, for in fact the rule is not applied literally in practice at all. Instead it is applied with considerable discretion, according to the physicist's intuitions about the desired result of the syntactic operation. In the present case, it is recognised intuitively that time reversal induces *state reversal*: so not only must the temporal variables t and Δt be replaced by $-t$ and $-\Delta t$, but the terms s_1 and s_2 denoting states must also be replaced by the terms s_1^R and s_2^R , denoting the time reversed states. The correct time reversal of L is in fact the following statement, L^R :

$$[L^R] \text{ If } s_1^R(t) \text{ then } s_2^R(t-\Delta t).^4$$

Behind the practical application of the syntactic reversal operator, therefore, lie direct intuitions about what *process reversal* amounts to. So let us to the concept of process reversal.

⁴In all other texts on the subject, the reversal of L is taken to be: *If* $s_2^R(t)$ *then* $s_1^R(t+\Delta t)$ (E.g. Davies [1974, ch.2], Sklarr [1974, p.365].) But this is incorrect in principle, as discussed in detail in the following chapter. However, a fully deterministic theory entails that L^R is equivalent to: *If* $s_2^R(t)$ *then* $s_1^R(t+\Delta t)$, as will also be shown.

4.2 Terminology: tokens and types of states and processes.

Before beginning the discussion it is important that the concepts of *process-types*, *token processes*, *state-types* and *token states* are clearly defined.

- (i) The notion of a *state-type* is fundamental. A state-type is a property that a system of physical objects might have. For instance, in classical physics a single particle might have a certain position and momentum. State-types are taken here to be completely specific *micro-states* (or logically atomic properties of systems), not macro-states or phenomenological states. The latter are taken as classes of micro-states. Whenever the discussion turns from micro-states to macro-states, this will be made clear, because many results concerning micro-states do not hold for macro-states.
- (ii) *Token states*. State-types will be denoted by the terms: s, s_1, s_2, \dots . These terms can take two kinds of arguments, denoting the *time* at which the state-type holds, and the *system* (or object) for which the state-type holds. Moments of time are denoted by t, t_1, t_2, \dots , and systems are denoted by X, Y, Z . (In fact these terms are used ambiguously as both *variables* ranging over moments and objects, and *constants* denoting specific moments and objects. Context will make clear which is intended. As is

common in physics, it is not worth the trouble to establish separate terms for variables and constants here.)

Thus we can have the following conventions:

s denotes a state-type;

$s(t)$ denotes that the state-type s holds at the moment t ;

$s(X)$ denotes that the state-type s holds of the object X ;

and

$s(X,t)$ denotes that the state-type s holds of object X at moment t .

The last, $s(X,t)$, is a *token state*, since it is the instantiation of a specific state at a specific moment. Where reference to the object is redundant (which is normally the case), the term for the object will be suppressed, and a token state will be denoted just by: $s(t)$.

It should also be recognised that in law statements, such as L of the previous section, there is an implicit universal quantifier over both moments and objects. I.e: *If $s_1(t)$ then $s_2(t+\Delta t)$* is strictly written: *(For all X,t)((If $s_1(X,t)$ then $s_2(X,t+\Delta t)$)*

Commonsense tells us where quantifiers are intended; to include them all explicitly would make the notation unnecessarily cumbersome.

- (iii) A *process-type* is a *sequence of state-types over a duration of time*.⁵ Process-types are denoted by: P, P_1, P_2, \dots

Process-types of infinite duration are allowed for, as are

⁵Duration is an oriented interval, as understood by Bunge [1972].

process-types of finite duration which do not have initial or final states. However, normally we will deal with process-types which have initial and final states, and these can be indicated by: $P=[s_1(t) \rightarrow s_2(\Delta t)]$, where s_1 is the initial state, s_2 the final state, and Δt is the duration of the process.

(iv). Just as *systems* can take on *state-types* at *specific moments*, to give token states, so systems can run through process-types in specific intervals of time to give *token processes*. The nomenclature used here is:

$P(t_1, t_2)$ denotes that process-type P occurs in the interval from t_1 to t_2 .

$P(X)$ denotes that object X goes through process-type P .

$P(X; t_1, t_2)$ denotes that X goes through P in the interval from t_1 to t_2 . This last is a fully specific token process.

Having made clear the logical relations between process-types, token processes, state-types and token states, I will frequently just talk of *processes* and *states*, and allow context to make clear whether types or tokens are intended.

4.3 The metrisation of time.

Process reversal is best understood through a consideration of how processes are *described*. To describe a token process, we have of course to refer to moments of time. Specifically, a sequence of states of an object has to be associated with a specific interval of time. This requires a language to describe time, and the language is provided by a *metrisation* of time.

By this is meant the mapping of moments of time onto the real numbers, so that the numerals then effectively provide names of moments, and consequently the mathematics of real numbers provides a language which can be used to talk about time. (A similar thing is done for space, mass, energy, and other physical quantities.)

In a very broad sense, any assignment of moments to numbers might be called a metrisation of time⁶. But not many such mappings are useful. More importantly, not many are *valid relative to the formalization of a given theory*.

This is easily seen. Metrisations are normally provided in practice by periodic physical processes, which we call clocks. These must 'run evenly' to be any good. The swing of a pendulum is a good approximation, and provides a metrisation (approximately) *valid relative to the formalization of our good scientific theories*. An irregular periodic process, such as the appearance of meteorites in a certain part of the sky, could also provide a metrisation of time, but it would not be valid relative to any known scientific theories.

For a more formal illustration, suppose that we adopt a certain, sensible, metrisation of time, f . f maps moments, t , onto real numbers, r , i.e. for any moment t there is a unique real number r such that: $f(t) = r$.

Suppose we describe a certain token process using this sensible metrisation as our means of referring to time. To use a

⁶Or more restrictively, any one-one mapping which reflects the order relation ('earlier than') among moments in an order relation ('less than' or 'greater than') among the corresponding numbers.

concrete example, suppose the process consists of a point particle X moving at a constant velocity of +3 units (in a one-dimensional space). We may express this feature of the process with an equation of motion:

$$[4.2] \quad \Psi(t_2) - \Psi(t_1) = 3(t_2 - t_1),$$

where $\Psi(t)$ is the function which gives the spatial co-ordinate of X at the moment t . Now of course, the variables t_1 and t_2 are being treated in this equation as *numerical* variables, ranging not over moments of time at all, but over numbers. They refer to moments only indirectly, through the mapping f which takes moments to numbers. If we wished to make this feature quite explicit, we could write instead:

$$[4.3] \quad \Psi(f(t_2)) - \Psi(f(t_1)) = 3(f(t_2) - f(t_1))$$

In this second version, t once again ranges over moments proper, and the mapping into numbers is exhibited explicitly by the function f .

Now let us suppose that the particle in question is not under any external forces, and that it *in fact obeys Newtonian mechanics*. In this case, f is an *adequate metrisation relative to the usual formalization of Newtonian mechanics*, roughly because it delivers equation [4.2], which is consistent with Newtonian mechanics. (Newtonian mechanics requires that a particle under no external forces has a constant velocity).

It might have turned out instead that [4.2] was *inconsistent* with Newtonian mechanics. In this case, the metrisation f would have been *an inadequate metrisation relative to the usual formalization of Newtonian mechanics*, for since (by supposition) the particle's motion actually does obey Newtonian mechanics, its equation of motion must be consistent with Newtonian theory.

In fact it is easy to construct such an inadequate metrisation from f . For instance, consider f^* , defined by: $f^*(t) = e^{f(t)}$. Using t_1^* and t_2^* as temporal variables in this new metrisation, the alternative equation of motion that results is:

$$[4.4] \quad \Psi^*(t_2^*) - \Psi^*(t_1^*) = 3(\ln(t_2^*) - \ln(t_1^*))$$

(since: $\Psi^*(t_1^*) = \Psi(\ln(t_1^*))$, and by [4.2]: $\Psi(\ln(t_2^*)) - \Psi(\ln(t_1^*)) = 3(\ln(t_2^*) - \ln(t_1^*))$). But of course, on this metrisation, the equation of motion, [4.4], for a particle under no external forces, represents exponential deceleration, which is incompatible with Newtonian mechanics. Hence f^* is *an inadequate metrisation relative to the usual formalization of Newtonian mechanics*.

One might object that, relative to the new metrisation f^* , Newtonian mechanics ought to be reformalized. If we wish to preserve the meaning of the original theory, then this is obviously true, but it misses the point being made. The point is that the *normal* formalization of a theory presupposes a certain metrisation, or class of metrisations. This is the class of *valid metrisations* (relative to the theory in question). When we

describe a process with the aim of examining whether it conforms to a theory, such as Newtonian mechanics, we must adopt a valid metrisation for the purposes of our description.

4.4 Symmetries.

While there are always narrow constraints on the class of valid metrisations relative to a given theory, there is normally also some freedom. This freedom results from *symmetries* of the theory. One of the most crucial symmetries, which has held for all seriously developed theories of physics, is the *translational symmetry of time*. This means essentially that it doesn't matter which specific moment of time is associated with the origin (zero) of the number line.⁷

Time translational symmetry is formally described as follows. Let us suppose that a specific function $f(t)$ achieves a valid metrisation, for definiteness let us say that of the Christian calendar, with a scale measured in days. f takes us from moments to (numerical) dates: its inverse, f^{-1} , takes us from dates back to moments. Thus $f^{-1}(0)$ is supposedly the exact moment of the birth of Christ, $f^{-1}(1)$ is the first moment of the

⁷In fact, the need for this symmetry can be seen to be the prime determinant of what the class of valid metrisations is presupposed to be when a theory is formalized. That is, the theory is formalized deliberately *so that* it is time-translation invariant: without this invariance, the metric placed on time normally becomes extremely inconvenient. It is presupposed here that all theories being dealt with are time-translation invariant. Without time translation invariance, time-reversal invariance becomes almost impossible.

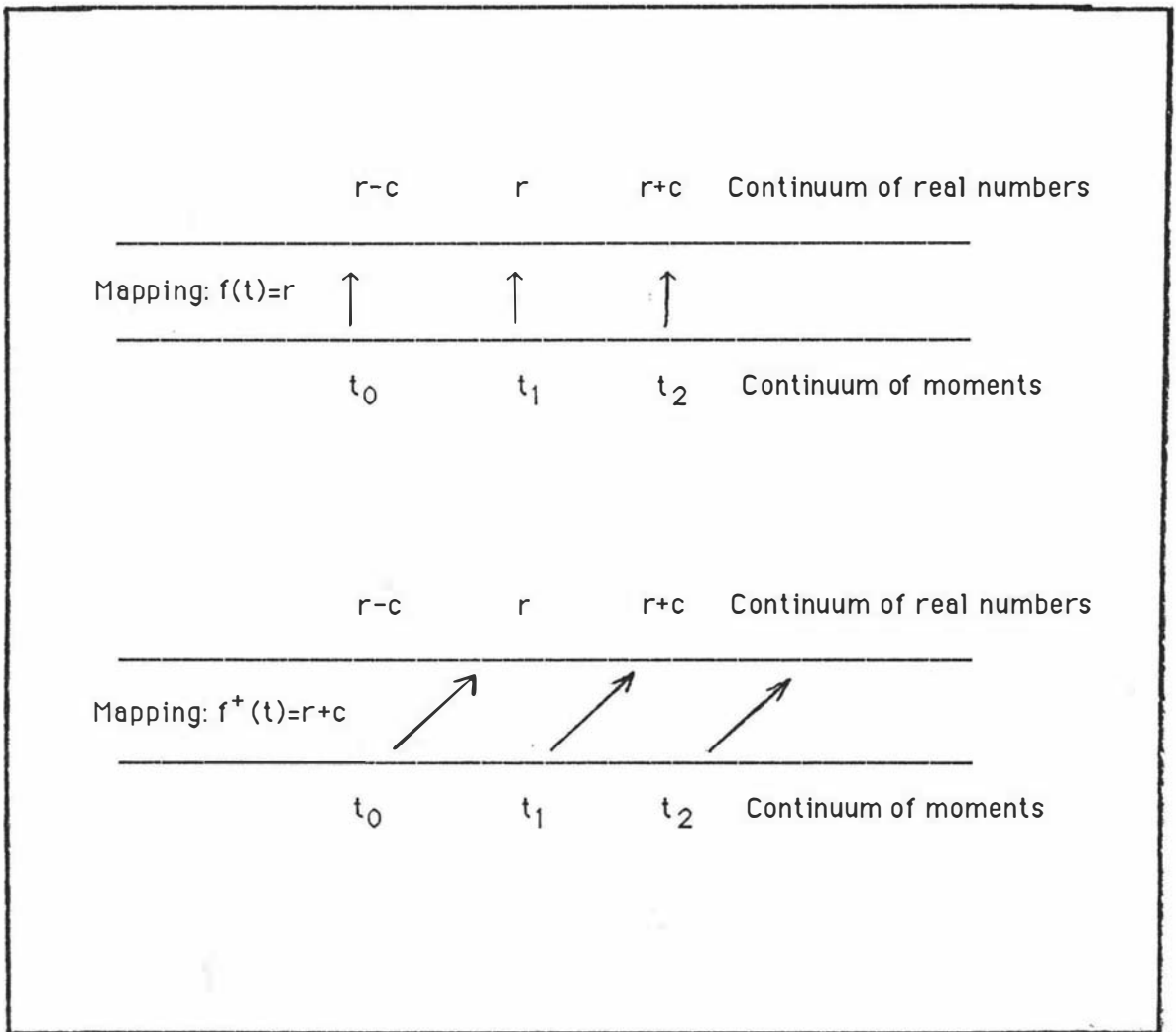


FIG. 4.1.
Time Translation Symmetry.

f^+ generates a valid metric if f does.

day after the birth, $f^{-1}(365)$ the first moment of the first year after the birth, and so on.

Now the *translational symmetry of time* means that any new function f^+ defined by: $f^+(t) = f(t)+c$, where c is a constant, also gives a valid metrisation. (See Fig. 4.1). f^+ simply makes the dates of all moments larger by the constant amount c .

Although the terms t, t_1, \dots , have so far been strictly treated as referring to *moments*, for many purposes in physics it is more convenient to consider them as referring directly to the *dates* attached to moments by the metrisation function. From now on I will freely regard these variables as referring to the dates, rather the moments themselves, whenever convenient: context will make it clear what is intended.

If a theory is *time-translation invariant*, then the transformation of the metrisation from f to f^+ is an *invariant transformation*. It may be regarded as a transformation of the 'frame of reference' for the description of processes that leaves the laws of physics invariant.⁸

Physicists commonly denote this transformation as: $t \rightarrow t+c$, meaning that the date t of each moment transforms to $t+c$. This is a useful shorthand for denoting transformations which I will often use.

⁸Invariant transformations can be equivalently viewed as (i) transformations of processes which leave them as lawlike processes, or (ii) transformations of the laws themselves, which leave them unchanged. The latter generally provides the mathematically simplest way of investigating invariances, and is the common approach in relativity theory for instance where physical laws are required to be invariant under the Lorentz transformations.

4.5 The time-reversal transformation.

Let us now turn to the transformation of special interest: *time reversal*. It too is a simple transformation: in the shorthand just noted, it is: $t \rightarrow -t$. (This is how it is usually denoted by physicists.) A little more fully, if f is a metrisation, then the time reversed metrisation is defined by: $f^R(t) = -f(t)$. (See Fig. 4.2).

Our interest is in the effect of this transformation on the *appearance* of process-types. A given process may, of course, *appear* different when described using different metrisations (or frames of reference). We already saw this when we considered the motion of point particle in the previous section. On the first metrisation, f , it appeared to be moving at a constant velocity of +3, while on the second metrisation, f^* , it appeared to be exponentially decelerating.

The question here is: given that a process appears to be of type P in the metrisation f , what type of process does it appear to be in the metrisation f^R ? Answering this will tell us the effect of the *time reversal operation* on process-types. By definition, a process that appears to be of type P on the metrisation f , appears to be of type P^R (the time reversal of P) on metrisation f^R . This is the very meaning of process reversal.⁹

⁹In these terms, *time reversibility* can be explained as follows. A theory T is time reversible just in case any process which *appears* to be a T -process on a metrisation f , also appears to be a T -process on the reversed metrisation, f^R . This gives the most picturesque way of visualizing the meaning of time

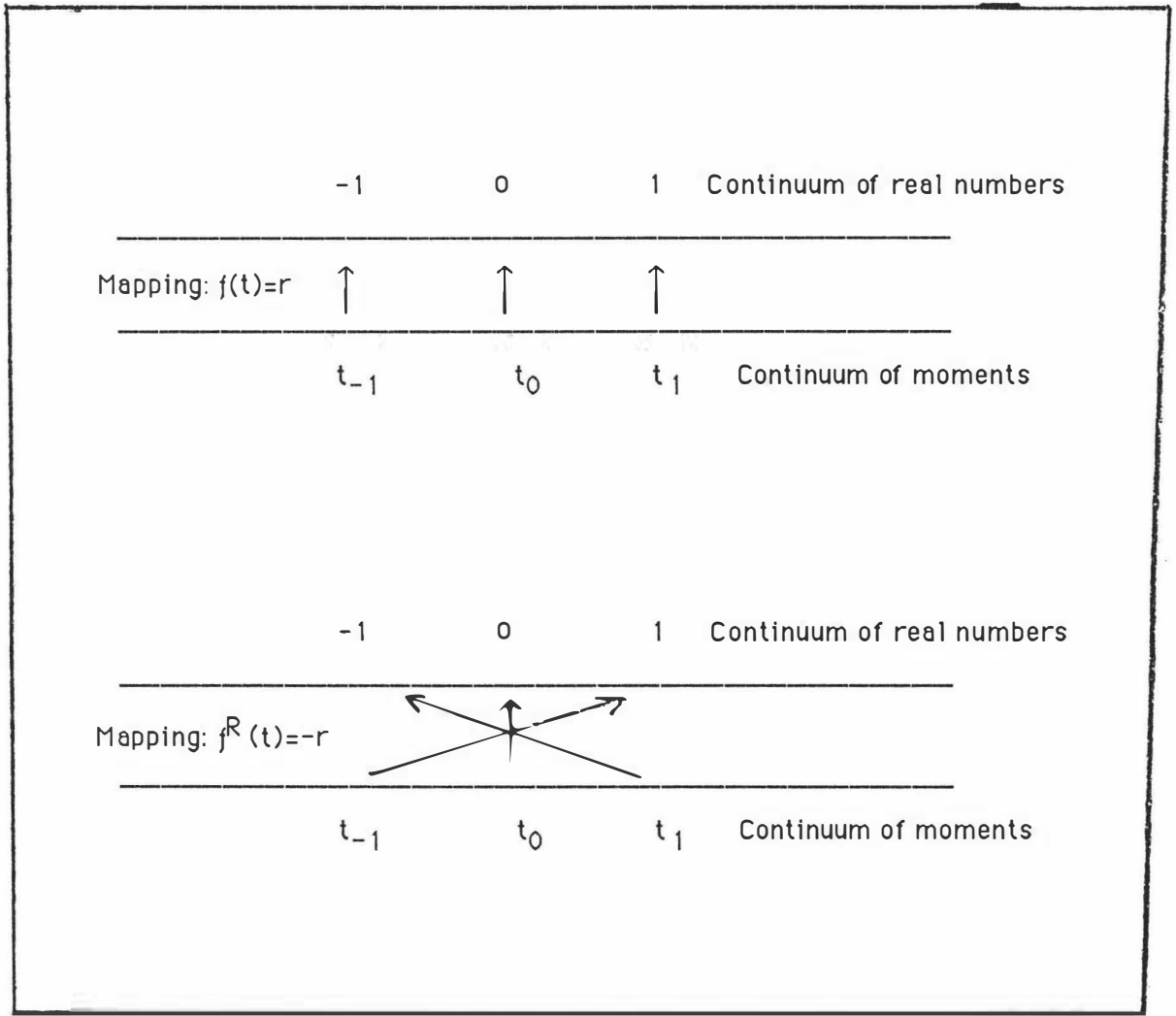


FIG. 4.2
Time Reversal Symmetry.

..

f^R generates a valid metric if f does.

The general effect of time reversal on processes is now easy enough to visualize. Obviously, if a process contains two states separated by a duration Δt , on the time reversed metrisation the states appear to be separated by the duration $-\Delta t$.¹⁰ This is just to say that the *temporal order of states is reversed by time reversal*. But this is not all there is to it: the *types of states* that occur in the process also appear to be 'time reversed'. (The time reversal of state-types will be defined in a moment). Hence the general rule for process reversal is this:

[4.5] The time reversal, P^R , of a process-type P consists of the reversed temporal sequence of the time reversed states of P . Symbolically, if $P=[s_1(t) \rightarrow s_2(t+\Delta t)]$, then $P^R=[s_2^R(t) \rightarrow s_1^R(t+\Delta t)]$

(Theorem [4.1] is an obvious corollary of this.)

4.6 Time reversal of states.

The chief question remaining is how time reversal affects state-types. What does a given state, s , appear to be like when the time axis is reversed?

The state s will be characterised in terms of a certain set of reversibility.

¹⁰Formally, $f(t) = -f^R(t)$ (definition of R), hence: $f(t_2) - f(t_1) = f^R(t_1) - f^R(t_2)$, for all moments t_1 and t_2 and all metrisations f .

parameters, such as, for example, the positions, masses, electric charges, velocities, and momenta, of a set of fundamental particles. What must be asked is: what effect does the transformation: $t \rightarrow -t$ have on the values of these parameters? This will tell us how the instantaneous state transforms.

It is recognised that some parameters (e.g. position, mass) are unaffected by time reversal, while others (e.g. velocity, momentum) are in fact reversed. Recognised transformations for classical mechanics are:

| | |
|--------------------|------------------------------------|
| $t \rightarrow -t$ | (time of course is being reversed) |
| $r \rightarrow r$ | (positions are invariant) |
| $m \rightarrow m$ | (masses are invariant) |
| $F \rightarrow F$ | (forces are invariant) |
| $v \rightarrow -v$ | (velocities reverse) |
| $a \rightarrow a$ | (accelerations are invariant) |
| $p \rightarrow -p$ | (momentums reverse) |
| $E \rightarrow E$ | (energies are invariant) |

While everyone undoubtedly feels they have a good intuitive understanding of why the different variables transform as they do¹¹, I know of no full explanation of the transformations on the variables. Indeed, so far as I know no one has even attempted a full explanation. Here and in Appendix 4.1 I try to offer one.

Part of the answer (the well-understood part) is that the group of variables are not *independent*, but conceptually related. For instance, velocity is *defined* in terms of time and position by:

¹¹E.g. Davies [1974, ch2], Sklarr [1974, ch12].

$v=dr/dt$. Thus, once we have decided that position is invariant under time reversal, we are forced by this definition to conclude that velocities reverse under time reversal. (Since the transformed velocity is given by: $v^R = dr/d(-t) = -v$.) Similarly, once it is decided that masses are invariant, it follows that momentum reverses. In fact the first three transformations, $t \rightarrow -t$, $r \rightarrow r$, and $m \rightarrow m$, obviously enough determine the transformations in all the other variables.¹²

It appears that a small set of variables (t , r , m , and F) is chosen as special: in some sense fundamental.¹³ I will call this the set of *fundamental variables*. The transformation: $t \rightarrow -t$ on time is considered to induce no transformation in the other *fundamental variables*, r and m , so that they are assigned the identity transformations, $r \rightarrow r$ and $m \rightarrow m$. This set of transformations then suffices to determine all transformations.

The (non-temporal) *fundamental variables* (mass and position) are clearly considered to be *independent of the metrisation of time*. That is, it has been decided that objects have masses and positions independent of how time is metricised. This seems to be the reason they are assigned the identity transformations.

¹²What is unexplained is why time reversal gives us $r \rightarrow r$, and $m \rightarrow m$ in the first place. For instance, why not take: $r \rightarrow r$ and $m \rightarrow -m$, which would then induce the transformations: $v \rightarrow -v$, $a \rightarrow a$, $p \rightarrow p$, $E \rightarrow -E$? This is an important theoretical question which has not been answered. A similar question is at the center of a long-standing puzzle about the correct definition of the state-reversal operator for quantum states. See Appendices 4.1 and 4.2.

¹³Forces must be taken as physically real if Newtonian mechanics is to be an empirical theory. I.e. $F=ma$ must be regarded as a postulate about forces, not as a mere *definition* of force.

(Obviously, this is what being 'independent of the metrisation of time means').

Other variables, such as velocity or momentum, arise by conceptual definition from the fundamental variables (including time). These will be called *secondary variables*. The transformations on the secondary variables follows from their definitions in terms of fundamental variables, and the transformations on the fundamental variables.

This indicates a more general procedure for finding the reversal transformations, namely,

- (i) specify a set of *fundamental variables*, v_1, v_2, \dots, v_n , plus time, t .
- (ii) Set their transformations as the identity transformations: $v_1 \rightarrow v_1, \dots, v_n \rightarrow v_n$, except for time which transforms according to: $t \rightarrow -t$,
- (iii) derive the transformations in all other secondary variables from their conceptual definitions in terms of v_1, \dots, v_n , and t .¹⁴

What the set of 'fundamental variables' is taken to be is critical. Its formal properties are simple enough. It must first be adequate for the definitions of all other variables. It should also have no redundancy, in the sense that no fundamental variable v_i should

¹⁴There are, of course, also various mathematical and logical terms: these are obviously invariant under time reversal, since the metrisation of time does not affect mathematical or logical objects or truths. There are also names of specific objects or systems (here the terms X, Y, Z) These are also invariant under time reversal, since particular objects retain their identities as objects on time reversal.

be definable in terms of the rest of the fundamental variables, or an inconsistency could arise.¹⁵

The informal property of the class of fundamental variables is that it represents the fundamental *ontology* of the theory. For example, in Newtonian mechanics, if the masses and positions of every particle are specified for every moment throughout a process, this uniquely identifies the process. There is no need to also specify the velocity (or momentum or energy) of the particles, since these are implicit in the specification of positions and masses. Specification of masses and positions at moments provides, on the most natural interpretation of the theory, the fundamental *ontology* of Newtonian mechanics.¹⁶

It is clear enough intuitively, therefore, how the time reversal transformations arise simply from the transformations in the fundamental variables that provide the fundamental ontology. Time reversal for *state-types* is determined by these transformations, since state-types can be fully defined in terms of the fundamental variables. (Otherwise the fundamental variables do not form an adequate class). Once we are provided with the time reversals of states, [4.5] implies that we have all the answers about time reversal for processes.

¹⁵For the conceptual definition of the redundant variable, v_i , plus the transformations on the other fundamental variables might entail that the transformation in v_i must be: $v_i \rightarrow -v_i$, while as a fundamental variable its transformation is already defined as: $v_i \rightarrow v_i$.

¹⁶What determines what the ontology is? See Appendix 4.1.

4.7 Return to the syntactic reversal operator.

Now that the nature of state and process reversal is understood clearly, we can consider the syntactic reversal operator again. A completely general rule for the syntactic reversal operation can now be defined. The syntactic reversal, L^R , of an expression L may be found in general by this procedure:

- (i). Rewrite L in *primitive form*, i.e. in a form which involves only terms for *fundamental variables* (and of course mathematical or logical terms, including proper names of objects.)
- (ii). Replace every *temporal term*, t , of L , with its inverse, $-t$, and leave all other terms alone. This generates L^R .

The substantial change from the physicists' normal rule of 'replace t with $-t$ ' is that the formulae must first be written in primitive form. Unfortunately writing an expression in primitive form may be difficult, and the following rule is much more practical:

- (i). Rewrite L in a form where all (non-mathematical) terms name either *fundamental or secondary variables*. As defined above, these variables all have well defined time reversal transformations. An expression in this form will be said to be in *secondary form*.

(ii). Replace all terms according to the time reversal transformations for the variables. I.e. if a variable ξ transforms according to: $\xi \rightarrow \chi$, then the *term* ξ is replaced by χ . (Mathematical and logical terms of course remain unchanged, their transformations being: $\xi \rightarrow \xi$). The expression that results is L^R .

It needs to be shown that these rules correctly capture the syntactic reversal operation. Here I will simply explain the strategy of the proof. It is by induction on the length of expressions. (i) If ξ is a simple term (e.g. t , r , m , or a logical term like $\&$), then it names a simple object, ξ . By definition, if the time reversal transformation for the object is: $\xi \rightarrow \chi$, then the term χ denotes the time reversal of ξ . I.e. $\xi^R = \chi$. (ii) It must then be shown in detail that, for the various ways of constructing composite terms, if ξ is a composite term with components ν and ε , replacing ν and ε with ν^R and ε^R respectively in ξ gives ξ^R .

The rules for finding syntactic reversals will shortly be illustrated with some examples, but first a summary of the time reversal transformation.

4.8 Time reversals of objects.

We really have a very wide concept of 'time reversal' now, which applies to any kind of object logically constructible from the fundamental ontology. If L names such an object L , then L^R names an object which is the time reversal of L . The main interest is in the case where L is a proposition, but it could be any number of other things: e.g. a moment, a position, a mass, a velocity, a token state, a state-type, a token process, a process-type, a propositional function, etc. An important point is that only in certain cases of L is its time reversal, L^R , *independent of the metric on time*. Only in these cases is there a useful concept of time reversal.

By way of illustration, I will consider the effect of time reversal on *moments, velocities, token states and state-types, token processes and process-types, propositions and laws*.

- (i) Moments. The time reversal transformation maps the class of moments back onto itself. Each moment t is mapped to a moment $-t$. But obviously this mapping is *relative to the choice of origin on the time-line*. Which is to say, the mapping is *relative to the choice of metric*. Let t_1 and t_2 be constants naming different particular moments. Suppose that the metric f_1 assigns t_1 as the origin, i.e. $f_1(t_1) = 0$, while f_2 assigns t_2 as the origin. In the first case, the time reversal transformation maps t_1 onto itself; in the second case, the time reversal transformation maps t_1 onto some other moment, t_1^* .

This shows in fact that there is no single time reversal transformation: rather there is a class of such transformations, one for each distinct choice of origin of the time line, or as I will say, one for each distinct metric.

Because the time reversal transformation on moments is relative to the metric in this way, it is not meaningful to talk of '*the* image of a moment, t , under time reversal'. For it has different images depending upon the (conventional) choice of origin metric.

- (ii) Velocities. By contrast, the image under time reversal of a *velocity* is independent of the choice of metric. A velocity v maps to $-v$ whatever the choice of metric. It is meaningful, therefore, to talk of *the* image under time reversal of velocities.

Similarly, it is meaningful to talk of *the* image under time reversal of position, mass, momentum, energy, and any other quantity which is independent of the metric on time.

- (iii) Token states. The time reversed image of a particular token state $s(X,t)$ is the particular token state $s^R(X,-t)$. Clearly this image depends upon the choice of origin of the metric. For instance, if the temporal origin is taken to be the moment of the birth of Christ, then the time reversed image of the token state which occurs at the first moment of 2,000 A.D. would be a token state

occurring approximately in the year 2,000 B.C. But if the origin is taken to be the first moment of 1990, then the time reversed image would be a token state occurring at the first moment of 1980.

There is, therefore, no meaningful concept of *the* image of token states under time reversal.

- (iv) State-types. By contrast, the image under time reversal of a state-type is independent of the choice of metric. This is why we can talk of *the* image of a state-type under time reversal.
- (v) Token processes and process-types. Token processes are sequences of token states: hence the image under time reversal of a token process is dependent on the choice of metric. Process-types are sequences of state-types: hence the image under time reversal of a process-type is independent of the choice of metric. Hence there is such a thing as *the* image of a process-type under time reversal; there is no such thing as *the* image of a token process under time reversal.
- (vi) Propositions and laws. Where a proposition is *time translation invariant*, its image under time reversal is independent of the metric on time. Otherwise its image is dependent on the metric on time - which is to say, it does not have an image under time reversal *per se*, but only relative to a choice of metric. Only in the former case, therefore, is it meaningful to ask whether $P = P^R$, since in the latter case P^R is not well defined. Hence in

general, as noted earlier, only where a law or theory is time translation invariant is it meaningful to ask whether it is time reversal invariant, i.e. reversible.

It is assumed throughout this discussion that the physical laws being talked about are time translation invariant. This is a slight simplification, for any theory must entail some laws which are not time translation invariant.¹⁷ This causes a problem with theorem [3.12], which states that:

[3.12] A theory T is time reversal invariant just in case, for any law L entailed by T , L^R is also entailed by T .

Since L^R is not well defined unless L is time translation invariant, there is a problem interpreting the meaning of this; but a minor modification solves it. The correct formulation is:

[3.12] A theory T is time reversal invariant just in case, *on any choice of metric f* , for any law L entailed by T , L^R is also entailed by T . ..

Normally it is not worth bothering with the complication of relativity of L^R with respect to the metric f , and this feature will be ignored unless there is some special reason to consider it.

¹⁷This is very easy to show. For instance, any general law of the form: *(For all t) P* , (which is time translation invariant), entails a law of the form: *(For all t)(If $t > 0$ then P)*, which is not time translation invariant.

4.9 The syntactic time reversal operation: some examples.

I will close this chapter by illustrating the general rules, formulated in Section 3.5, for finding syntactic time reversals, with some examples. The two final examples are particularly important, and will be discussed in more detail in the following chapter.

1. Consider the expression $s(X)$, which states that the object, X , is in the state s . We can find the syntactic reversal of this expression by applying the general rules formulated in Sec. 4.7.

(i) First $s(X)$ must be rendered in secondary form. Suppose for the sake of a definite example that s is the state of having a definite momentum, p_1 , and definite position, r_1 . I.e., for any X , $s(X) \equiv (p_1(X) \& r_1(X))$, so that $s(X)$ in secondary form is: $(p_1(X) \& r_1(X))$.

(ii) The transformations on momentum and position are: $p \rightarrow -p$ and $r \rightarrow r$, so replacement of terms gives us: $s^R(X) = (-p_1(X) \& r_1(X))$. Thus s^R names a state similar to s , except that the *momentum* involved has been reversed, which of course is what is expected.

2. Consider the earlier example of the *deterministic transition law*, L : *If $s_1(t)$ then $s_2(t+\Delta t)$.*

(i) First this must be written in secondary form. Let us suppose that s_1 and s_2 have already been included as secondary variables, so that the time reversal

transformations on them have been defined. These transformations may be simply represented as: $s_1 \rightarrow s_1^R$, and $s_2 \rightarrow s_2^R$. Hence L can be regarded as already being in secondary form.

(ii) Replacement of terms now gives: $L^R = \text{If } s_1^R(-t) \text{ then } s_2^R(-t-\Delta t)$. Since t is implicitly universally quantified, this is equivalent to: $\text{If } s_1^R(t) \text{ then } s_2^R(t-\Delta t)$.

This tells us that the time reversal of a law: $\text{If } s_1(t) \text{ then } s_2(t+\Delta t)$ is: $\text{If } s_1^R(t) \text{ then } s_2^R(t-\Delta t)$. Note that while the first law is a *future-directed* transition law, the second is a *past-directed* transition law.

Thus, *any reversible theory which entails a law of the form: If $s_1(t)$ then $s_2(t+\Delta t)$ must entail a corresponding law of the form: If $s_1^R(t)$ then $s_2^R(t-\Delta t)$ (by 3.12).*

This represents the proper criterion for reversibility of deterministic transition laws. It will be discussed further in the following chapter.

3. Let L be the probabilistic statement: $PROB(Q) = p$, where Q is some statement.

(i) The reversal transformation for *probability* has not yet been defined. It is discussed at length in the following chapter, but for now it may be accepted as intuitively obvious that metrisation of time does not affect the values of probabilities. They must keep the same value whatever the metrisation of time, since there

is nothing intrinsically dynamic about them, as there is about momentum for example. I.e. the transformation is: $PROB() = p \rightarrow (PROB() = p)$. (The argument of the term: $PROB()$ will of course suffer time reversal separately.)

We can take the transformation on Q to be: $Q \rightarrow Q^R$, and thus take the syntactic transformation to be: $Q \rightarrow Q^R$.

- (ii) Making the replacements, we therefore get that L^R is: $PROB(Q^R) = p$. It is obvious that this really is the time reversal of L , since $PROB(Q) = p$ is true on a given metrisation only if $PROB(Q^R) = p$ is true on the reversed metrisation. Thus we have a formal way of finding the time reversals of probabilistic statements.

4. Finally consider a statement, L , of the form:

$$[L] \quad PROB(s_2(t+\Delta t) | s_1(t)) = p.$$

This will be called a *probabilistic transition law* (compare it with the *deterministic transition law* considered above.)

- (i) By the definition of conditional probability, L is logically equivalent to:

$PROB(s_2(t+\Delta t) \& s_1(t)) / PROB(s_1(t)) = p$. This can be regarded as being in secondary form, the appropriate transformations just being: $t \rightarrow -t$, $\Delta t \rightarrow -\Delta t$, $s_1 \rightarrow s_1^R$, and: $(PROB() = p) \rightarrow (PROB() = p)$.

- (ii) Thus the time reversed statement, L^R , is:

$$[L^R] \text{PROB}(s_2^R(-t-\Delta t) \& s_1^R(-t)) / \text{PROB}(s_1^R(-t)) = p.$$

By the definition of conditional probability, this is equivalent to: $\text{PROB}(s_2^R(-t-\Delta t) | s_1^R(-t)) = p$. Being a law statement, there is as usual an implicit universal quantifier on the temporal variable, t , hence L^R is equivalent to:

$$[L^R] \text{PROB}(s_2^R(t-\Delta t) | s_1^R(t)) = p.$$

Note that this postulates a *past-directed probability*, in contrast to the original statement L , which postulates a future directed probability. Note also the L^R would have been derived correctly if we had directly made the substitutions in the conditional probability law, L , instead of translating to and from the formulation involving absolute probabilities.

This is a very important result: what it means is that the time reversal of any proposition of the form: $\text{PROB}(s_2(t+\Delta t) | s_1(t)) = p$ is a proposition of the form: $\text{PROB}(s_2^R(t-\Delta t) | s_1^R(t)) = p$. This immediately delivers the key criterion for the reversibility of probabilistic theories. It will be called the *Criterion for Probabilistic Reversibility*. It is the main topic of the following chapter.

CHAPTER FIVE

THE CRITERION FOR PROBABILISTIC REVERSIBILITY

A *probabilistic theory* is assumed here to be one that entails *probabilistic transition laws*, of the form:

$$\text{PROB}(s_2(t+\Delta t) | s_1(t)) = p, \quad \text{where: } 0 < p < 1.$$

As usual with dynamic laws of any kind, such a law applies only where there is no undue interference with the system concerned in the interval from t to $t+\Delta t$. Following the orthodox analysis of the reversibility of quantum theory it is assumed here that quantum theory entails such laws. The purpose of this chapter is to examine the conditions for the reversibility of theories that entail such probabilistic transition laws.

5.1 The CPR.

At the end of the previous section, a criterion for the reversibility of a theory that entails such laws was derived. It will be called the *Criterion for Probabilistic Reversibility*, abbreviated to *CPR*. By definition, a theory T satisfies the CPR just in case:

[CPR] For any law: $PROB(s_2(t+\Delta t) | s_1(t)) = p$ entailed by T , the law: $PROB(s_2^R(t-\Delta t) | s_1^R(t)) = p$ is also entailed by T .

It is crucial to note that the first-mentioned law in the CPR is *future-directed*, supplying a probabilistic connection from present to future, while the second is *past-directed*, supplying a probabilistic connection from present to past.

For the purposes of the main argument, it is only important that the CPR turns out to be a *necessary condition* for the reversibility of probabilistic theories. The main task of this chapter is to establish this in detail. It turns out in fact that the CPR is both a necessary and a sufficient condition for the reversibility of any theory which is logically equivalent to some class of deterministic or probabilistic transition laws, as the proof in Section 5.8 (a) shows.¹ But that it is a sufficient

¹Normal methods used so far in science for defining physical equations of motion can be treated as methods for defining classes of deterministic or probabilistic transition laws. Deterministic theories can be treated as special cases of

condition is not important for any arguments here.

It was effectively derived at the end of the previous chapter that the CPR is at least a necessary condition for reversibility. However, that derivation was rather abstract, and nothing was done to illustrate the meaning of the CPR, which is the purpose of this chapter. But before beginning on this, it will be useful to review the criterion for probabilistic reversibility that the orthodox analysis of the reversibility of quantum theory assumes.

5.2 The PPMR.

The orthodox analysis makes no mention of the CPR. An entirely different condition is taken as the criterion for the reversibility of a class of probabilistic laws. It is generally called the *principle of micro-reversibility*, or the *probabilistic principle of micro-reversibility*, abbreviated here to *PPMR*.² It is defined as follows:

[PPMR] A theory T satisfies the *probabilistic principle of micro-reversibility* (PPMR) just in case, for every law: $PROB(s_2(t+\Delta t) | s_1(t)) = p$ entailed by T, the law: $PROB(s_1^R(t+\Delta t) | s_2^R(t)) = p$ is also entailed by T.³

probabilistic theories, on the assumption that the transition probabilities between states are always 1 or 0. Thus the criterion for reversibility of probabilistic transition laws suffices as the general criterion for reversibility of all present theories.

²E.g. Lewis [1930], Davies [1974, ch.6].

³This is often abbreviated in physics texts as the principle that: $w = w_{rev}$, for

The orthodox analysis demonstrates that quantum theory satisfies this principle, and therefore concludes that the probabilistic part of quantum theory is time reversible⁴. My claim is that *the PPMR is irrelevant to time reversibility*. Therefore the orthodox analysis fails to establish the reversibility of quantum theory. I propose to demonstrate that:

- (a) The PPMR is neither a necessary nor a sufficient condition for reversibility. (It is simply irrelevant to reversibility).
- (b) The CPR is a necessary condition for reversibility.

In the following chapter I will demonstrate that :

- (c) Quantum theory fails the CPR.

If (b) and (c) are correct, it follows, of course, that quantum theory is irreversible.⁵

all w 's, where w 's represent transition probabilities between states. E.g. Davies [1974, p.157], Watanabe [1955c].

⁴This result relies upon having already shown that the deterministic part of quantum dynamics, i.e. the Schrodinger equation, is reversible.

⁵It is worth stressing that this result is *not* merely a conclusion that the 'measurement process' is irreversible. My argument makes no explicit mention of 'measurement' at all. It is a result solely about the *probabilistic* nature of quantum theory. There is a connection with 'measurement' only insofar as *probabilities come into play in 'measurements'*. Of course, one traditional view is that probabilities only come into play when 'measurements' are made. This would not affect my argument, but nevertheless it is almost certainly an

(a) and (b) are entirely general results about probabilistic theories, and not at all specifically about quantum theory. Quantum theory therefore requires no serious mention until the following chapter.

I will begin by illustrating the CPR, and why it is a necessary condition for reversibility, in a number of different ways.

5.3 CPR 1: A way of picturing the time reversal of probabilities.

This first illustration is very informal, and is intended just to provide a concrete picture of the CPR. Suppose that: $\text{PROB}(s_2(t+\Delta t)|s_1(t)) = p$ is a natural law, and that a particular system goes through the transition: $s_1(t) \rightarrow s_2(t+\Delta t)$. To account for the probability as physically objective, we need more than simply the *sequence*, $s_1(t) \rightarrow s_2(t+\Delta t)$, to represent the process, since this sequence does not indicate the existence of the probability involved in the transition. We might picture the probability as a temporally directed relation between the two states, $s_1(t)$ and $s_2(t+\Delta t)$. In a diagram of state plotted against time, we might draw an arrow or vector from s_1 at t to s_2 at $t+\Delta t$, labelled with a 'strength', p , to indicate the probability. This representation will be justified in Chapter Seven.

Represented in this way, the need for *past directed probabilities* to achieve time symmetry with the future directed probabilities becomes immediately apparent. For if the laws of inadequate view of quantum ontology. See Chapter Eight.

nature are reversible, then it should be equally valid to describe the given process in the reversed temporal metric. In the reversed metric, the *sequence* appears to be of the type: $s_2^R(t) \rightarrow s_1^R(t+\Delta t)$. The *probability* involved in the transition still has the value p , and points from $s_1^R(t+\Delta t)$ to $s_2^R(t)$, so it has reversed its temporal direction. Such a probability is *past-directed*, and if it exists intrinsically in nature, it must be founded on a *past-directed* probabilistic law $\text{PROB}(s_2^R(t)|s_1^R(t+\Delta t)) = p$. If quantum theory is reversible, it must therefore entail this past-directed probability law if it entails the corresponding future-directed probability law. This immediately gives the CPR as a necessary condition for reversibility.

5.4 CPR 2: Model-theoretic representation of probabilities.

The idea appealed to above of representing probabilities as temporally directed vectors is informal and its adequacy may be doubted. In Chapter Seven, an ontology for probabilities along such lines is explicitly set up, but a more orthodox representation of probabilities in model-theoretic terms will be shown here to deliver the same result, that the CPR is necessary for reversibility.

The model-theoretic approach represents probabilities as the *relative weights of classes of models*. Let the class of (nomologically possible) models (or worlds) in which a particular system under consideration has the state s at moment t be

denoted by: $\mu(s(t))$. The intersection of two such classes, $\mu(s_1)$ and $\mu(s_2)$, is: $\mu(s_1 \& s_2)$. The *relative weight* of two classes of models, μ_1 and μ_2 , will be denoted by: $\omega(\mu_1/\mu_2)$. Then the fact that: $\text{PROB}(s_2(t+\Delta t)|s_1(t)) = p$ is represented by the assignment of relative weight: $\omega(\mu(s_2(t+\Delta t) \& s_1(t))/\mu(s_1(t))) = p$.

This reflects the usual definition of conditional probability in terms of absolute probability as: $\text{PROB}(P|Q) = \text{PROB}(P \& Q)/\text{PROB}(Q)$.

The idea that there are physically real probabilities is simply the idea that these relative weights have an objective existence, that they are real physical variables.⁶

We can now see why the idea of intrinsic *past-directed probabilities* makes sense (as claimed in Chapter One), since such probabilities can be explicated in these terms. The past-directed probability: $\text{PROB}(s_2^R(t-\Delta t)|s_1^R(t)) = p$ is represented by the assignment of relative weight: $\omega(\mu(s_2^R(t-\Delta t) \& s_1^R(t))/\mu(s_1^R(t))) = p$. This too may or may not have an objective existence in nature.

Let us now consider what is required of the space of nomologically possible models for the laws of nature to be time reversible. The requirement of course is that the *time reversed*

⁶This may seem strange to the nominalist, since it is a postulate about *the structure of the space of nomologically possible worlds*, not simply a fact about the actual world. But, as pointed out in an earlier footnote, this is a feature of any 'laws of nature', realistically interpreted. *All* laws of nature concern the structure of the *space* of nomologically possible worlds (or processes): the postulate of intrinsic probabilities is not special in this respect. In the Chapter Seven a model of single-case physical probabilities as genuine physical entities will be presented.

space of nomologically possible models must be identical to the original space. Under time reversal, the class of models: $\mu(s_1(t))$ appears as the class: $\mu(s_1^R(-t))$, and: $\mu(s_2(t+\Delta t) \& s_1(t))$ appears as: $\mu(s_2^R(-t-\Delta t) \& s_1^R(-t))$. Hence the relative weight: $\omega(\mu(s_2(t+\Delta t) \& s_1(t)) / \mu(s_1(t))) = p$ representing the future-directed probability: $\text{PROB}(s_2(t+\Delta t) | s_1(t)) = p$ appears in the time-reversed space of models to be the relative weight: $\omega(\mu(s_2^R(-t-\Delta t) \& s_1^R(-t)) / \mu(s_1^R(-t))) = p$, which is the past-directed probability: $\text{PROB}(s_2(-t-\Delta t) | s_1(-t)) = p$. This only exists for this specific case where time = $-t$ if the general law: $\text{PROB}(s_2(t-\Delta t) | s_1(t)) = p$, (for all t) holds of the time-reversed space of worlds. Reversibility therefore requires that this law holds of the space of nomologically possible worlds if the original law holds, and we arrive again at the CPR.

5.5 CPR 3: A statistical picture.

Physical probabilities are reflected by statistical frequencies. For instance, suppose that: $\text{PROB}(s_2(t+\Delta t) | s_1(t)) = p$ is a natural law. If at time t a set of N systems are in the state: $s_1(t)$, then it is to be expected that at time $t+\Delta t$ the *frequency* of these systems found in the state: $s_2(t+\Delta t)$ will be $p \cdot N$.

The existence of the physical probabilities will make this relation of *frequencies* a lawlike feature of the world, subject of course to statistical deviations. Suppose then that this relation of frequencies is lawlike in a certain world, reflecting the intrinsic probability stated above. What feature does this

confer on the time reversed world?

In the time reversed world, the original frequency of $s_1(t)$ being followed by $s_2(t+\Delta t)$ appears as the frequency of $s_1^R(t)$ being preceded by $s_2^R(t-\Delta t)$ (disregarding the absolute value of t , since because of time translation invariance, it is only the sign of Δt that matters). This frequency relation is lawlike in the time reversed world, reflecting the time reversal of the probabilistic law: $\text{PROB}(s_2(t+\Delta t)|s_1(t)) = p$ that holds in the original world. Clearly, then, the time reversal of this law is the *past directed probability law*, $\text{PROB}(s_2^R(t-\Delta t)|s_1^R(t))$, since this is what the time reversed frequency reflects.

For the laws of the original world to be identical to those of the reversed world, this past-directed probability law must also hold in the original world, and we once again see that the CPR is a necessary condition for reversibility.

5.6 Previous recognition of the CPR.

The CPR seems an obvious condition for reversibility once it is recognised. After all, if a theory entails *future-directed* probabilities (or any sort of future-directed relation between events), then it is intuitively obvious that for time symmetry it must entail appropriate *past-directed* probabilities (an appropriate past-directed relation). Otherwise a particular direction of time is picked out nomologically as the direction in which probabilities are directed. It is surprising, therefore, that with the intensive work that has gone into the study of

reversibility, the CPR has not been widely recognised. In fact, although it has been very scantily recognised, it has not been entirely overlooked: there have been two important discussions of it that I am aware of. The first is by Watanabe [1955c,65,66,70] on the reversibility of quantum theory; the second is in work on the reversibility of thermodynamics.

Of the two contributions, Watanabe's is the most important. Watanabe did not formulate the CPR in the way I have done, but he showed decisively that the past-directed probabilities required by it fail to hold in quantum theory, and in later papers particularly, he stressed that this failure implies some kind of time asymmetry of quantum theory. Unfortunately he did not have a clear analysis of the meaning of time symmetry (or reversibility), and he did not stress the full importance of the CPR for reversibility. He came very close, however, and his series of papers on the subject are a very valuable contribution to the literature on the reversibility of quantum theory. It is a great pity that their importance has not been widely recognised.

The scientific community has failed to appreciate the significance of Watanabe's result, and perhaps this is partly because he did not show clearly that *the failure of the CPR in quantum theory means that quantum theory is irreversible*. For instance, he did not formulate the CPR as a necessary condition for reversibility. He did not discuss the existence of *intrinsic probabilities*, or discuss quantum ontology in any way, but instead expressed his result in the claim that *quantum physics is irretrodictable* ([1965, p.156]). And he continued to refer to

quantum theory as *reversible* because it satisfies the PPMR. The following statement is typical: "This basic asymmetry due to irretractability is compatible with reversibility or any other symmetry rules mentioned at the beginning of this section" [1965, p.157].

Also, the 'apparent' time asymmetry of the universe had already a long history as a scientific puzzle in the face of the reversibility of the fundamental equations, and physicists had come to terms with it, basically by explaining it as an effect of special boundary conditions for the universe. It must have been easy to attribute Watanabe's 'irretractability' of quantum theory to the same basic cause, and thereby explain it as the result of contingent features of the universe, rather than as a *structural* feature of quantum theory itself⁷. Watanabe did not argue lucidly enough that irretractability constitutes a structural irreversibility of the quantum theory, but he recognised the fact quite clearly, and there is no doubt that he should be credited with the discovery that quantum theory is irreversible.

The other connection in which the CPR has also been discussed is in thermodynamics. The principle was discussed by the Ehrenfests [a] and subsequently by many others (see particularly Mehlberg [1980] and Grünbaum [1973]). I will give a brief sketch of the situation.

Thermodynamics tells us that a system in a state of low entropy will very probably develop to a state of higher entropy,

⁷This is reflected in the work of a number of writers who followed Watanabe's lead in discussing the 'irretractability' of quantum theory, e.g. Aharonov *et alia* [1964], Bohm and Bub [1966], Cocke [1967], Penfield [1966].

but that a high-entropy system will probably not decrease its entropy. For a definite example, suppose that a specific theory predicts the following laws, where S_1 is a low-entropy state and S_2 a high-entropy state:⁸

$$\text{PROB}(S_2(t+\Delta t)|S_1(t)) = 0.999, \text{PROB}(S_1(t+\Delta t)|S_2(t)) = 0.001.$$

Thermodynamic states are normally indistinguishable from their time reversals, i.e. $S_i = S_i^R$. (For instance, the states might concern a set of particles in a box, S_1 might hold if the particles occupy a small volume, S_2 if they occupy a larger volume. Since these states depend only upon *position*, which is independent of the assignment of the temporal metric, they appear exactly the same in reverse). This appears to mean that the laws given above are sensitive to the direction of time. For consider, first, a normal thermodynamic process in which a system runs through the sequence of states: $S_1 \rightarrow S_2$, in a period Δt . This is perfectly legitimated by the probability laws above, having a transition probability of 0.999. However the time reversal of this process would be: $S_2^R \rightarrow S_1^R$, which is: $S_2 \rightarrow S_1$, in a period Δt . But this seems to be a thermodynamically *improbable* process, since the transition: $S_2 \rightarrow S_1$ has a probability of only 0.001.

This appears to show that the thermodynamic laws are *irreversible*, since a certain process that appears probable in one direction of time appears improbable if interpreted to occur in the reverse direction. However, this analysis is flawed, as the Ehrenfests [a] recognised.

⁸I have used capital S's here to indicate that the states are not *microstates* but *macrostates*. In some contexts the difference is crucial.

The flaw is illustrated by the following example. Suppose that the full set of transition probabilities is:

$$\text{PROB}(S_1(t+\Delta t)|S_1(t)) = 0.001$$

$$\text{PROB}(S_2(t+\Delta t)|S_1(t)) = 0.999$$

$$\text{PROB}(S_1(t+\Delta t)|S_2(t)) = 0.001$$

$$\text{PROB}(S_2(t+\Delta t)|S_2(t)) = 0.999.$$

A system governed by these laws will exhibit the following behaviour. It will spend most (99.9%) of its time in the high entropy state S_2 , occasionally making the transition to S_1 , in which it will spend 0.1% of its time. Now this behaviour is manifestly 'time symmetric', in the sense that it appears equally lawlike viewed in either direction of time. Let us examine the source of this time symmetry of the process in terms of the reversibility of the laws.

The process appears lawlike in the normal direction of time because it is being supposed that the system obeys, with normal statistical fluctuations, the transition probabilities above. E.g. in 99.9% of cases, when in S_1 at t , it is found in S_2 at $t+\Delta t$; in 0.1% of cases, when in S_2 at t , it is found in S_1 at $t+\Delta t$.

What is interesting is the reason that it appears lawlike in the reverse direction of time. Viewed in reverse, each occurrence of S_i remains an occurrence of S_i , and each transition of type: $S_i(t) \rightarrow S_j(t+\Delta t)$ appears as a transition of type: $S_j(t) \rightarrow S_i(t+\Delta t)$. What is important for the question of the lawlikeness of the time reversed process are *the frequencies of transitions of type: $S_j(t) \rightarrow S_i(t+\Delta t)$ relative to occurrences of S_j* . If the laws for the original process hold for the reversed process, these should

reflect the probabilities defined above; to be precise, the expected frequency of the transition $S_j(t) \rightarrow S_i(t+\Delta t)$ divided by the expected frequency of S_j should equal $PROB(S_i(t+\Delta t)|S_j(t))$, as given above. Let us consider as an example the case where $i=1$ and $j=2$.

The frequency of $S_2(t) \rightarrow S_1(t+\Delta t)$ in the reversed process is the frequency of $S_1(t) \rightarrow S_2(t+\Delta t)$ in the original process. The frequency of S_2 in the reversed process is equal to the frequency of S_2 in the original process. So we must calculate the relative frequencies of $S_1(t) \rightarrow S_2(t+\Delta t)$ and S_2 in the original process to find the relative frequency of $S_2(t) \rightarrow S_1(t+\Delta t)$ and S_2 in the reversed process.

The relative frequency of $S_1(t) \rightarrow S_2(t+\Delta t)$ and S_2 in the original process equals the *past-directed* probability: $PROB(S_1(t)|S_2(t+\Delta t))$ for the original process. For denoting the absolute frequencies in the original process of S_2 and $S_1(t) \rightarrow S_2(t+\Delta t)$ by: $FREQ(S_2)$ and $FREQ(S_1(t) \rightarrow S_2(t+\Delta t))$, respectively, the relative frequency is:

$$FREQ(S_1(t) \rightarrow S_2(t+\Delta t)) / FREQ(S_2),$$

and since: $FREQ(S_1(t) \rightarrow S_2(t+\Delta t))$ is (expected to be):

$$FREQ(S_2) \cdot PROB(S_1(t)|S_2(t+\Delta t)),$$

this equals:

$$FREQ(S_2) \cdot PROB(S_1(t)|S_2(t+\Delta t)) / FREQ(S_2) = PROB(S_1(t)|S_2(t+\Delta t)).$$

Past-directed probabilities are not generally entailed by future-directed probabilities: but in this case they are, because of the special nature of the system concerned. The system concerned is in a particular kind of *equilibrium*. This equilibrium entails that

the number of transitions *to* the state S_1 from any other state equals the number of transitions *from* the state S_1 to any other state⁹. It follows that:

$$\text{PROB}(S_1(t+\Delta t)|S_2(t)) = \text{PROB}(S_1(t-\Delta t)|S_2(t)).$$

In other words, for a given occurrence of the state S_2 , the probability that the system will develop *into* the state S_1 in the next interval equals the probability that it developed *from* S_1 in the last interval. This probability is 0.001 (by the laws above). Hence, the relative frequencies of $S_2(t) \rightarrow S_1(t+\Delta t)$ and S_2 in the reversed process is 0.001. This reflects the probability law for the reversed process:

$$\text{PROB}(S_1(t+\Delta t)|S_2(t)) = 0.001,$$

which is the same as the law for the original process. It may be verified in a similar way that all the probability laws for the original system hold for the time reversed system. The laws as they apply to this system are therefore time reversible.

A number of comments are in order.

(1) Firstly it is seen that the *failure* of the symmetry:

$$\text{PROB}(S_1(t+\Delta t)|S_2(t)) = \text{PROB}(S_2(t+\Delta t)|S_1(t)),$$

which is the PPMR symmetry for this theory, *has no implications for reversibility.*

(2) Secondly, the discussion above has illustrated once again

⁹This can be easily verified. It is not exactly true if we allow the system to have a *first* or *last* state, since this may upset the equality slightly, although only in a negligible way for a long process. However, if the laws as they stand represent the full theory, there can be no last state. Let us add a postulate that there can be no first state of a system either, and the theory is perfectly reversible.

that what is necessary for reversibility is the CPR - since for the law: $\text{PROB}(S_1(t+\Delta t)|S_2(t)) = p$ to hold of the time reversed system, it is necessary that the law: $\text{PROB}(S_1^R(t-\Delta t)|S_2^R(t)) = p$ hold for the original. (The fact that states are their own time reversals in this example blurs this message a little).

- (3) Thirdly, a word of caution about this type of example. All that has been shown is that a certain *system* which is governed by the probabilistic theory in question exhibits time-symmetric behaviour. But this by itself *does not show that the theory is reversible*. For the way that a *particular system* that is governed by a theory behaves is irrelevant to the reversibility of that theory. (A theory may be reversible, and yet *all* systems governed by it exhibit highly time-asymmetric behaviour.)

The example presented is saved only by the special nature of the theory, essentially by the fact that the *only* kind of process governed by the theory is an equilibrium process. It follows that the theory is reversible if this kind of process is suitably time symmetric, which is what was effectively demonstrated.¹⁰ This point is important, because many writers make the mistake of discussing the time-symmetry of specific *processes* governed by a theory, rather than the reversibility of the theory itself.

¹⁰Strictly, the theory considered is reversible only with the addition of the postulate mentioned in the last footnote. An irreversible theory can easily be obtained by adding alternative extra laws, for instance a law to the effect that *all processes begin in the state S_1* .

Lewis [1930], Aharanov [1964], and Cocke [1967] make this mistake. They all conclude that the probabilistic laws of quantum theory are reversible because *equilibrium processes* governed by these laws are 'time symmetric'. But this does not show that the laws are reversible: it would only show that a theory consisting of these laws in conjunction with a further postulate that *all quantum processes are equilibrium processes* would be time symmetric. The extra postulate manifestly fails of the real world, however, so this hypothetical 'reversible quantum theory' is not a viable one.

The general conclusion therefore is that for thermodynamics to be reversible, the failure of the 'PPMR' symmetry:

$$\text{PROB}(S_j(t+\Delta t)|S_i(t)) = \text{PROB}(S_i(t+\Delta t)|S_j(t))$$

can be ignored, and that what is important is the satisfaction of the 'CPR' symmetry:

$$\text{PROB}(S_j(t+\Delta t)|S_i(t)) = \text{PROB}(S_j(t-\Delta t)|S_i(t))$$

A number of commentators claim that the latter symmetry is satisfied by thermodynamics¹¹, but I am not convinced. It is true that the symmetry holds for classically-based thermodynamics if *all thermodynamic processes are expected to be equilibrium processes*. But this assumption must be considered highly speculative if it concerns the process that the universe as a whole is going through. I am unaware of any evidence that the actual universe is involved in a long-term

¹¹Mehlberg [1980] in particular, following the Ehrenfest's [a].

equilibrium process. The evidence seems to be quite the reverse, given the spectacularly *disequilibrium* state the universe has apparently been in for all of its known past, and the lack of detailed knowledge about its ultimate origins. Even if it is assumed that in the long run our universe is involved in an equilibrium process (from which the present disequilibrium would be an accidental fluctuation from the high-entropy equilibrium state), it is clear that the CPR symmetry cannot obtain for a theory of thermodynamics that has practical application to our universe in its present era. For to put it in Watanabe's terms, in our universe *retrodiction* is simply not valid in the same way as *prediction*, which means that the past-directed probabilities required for the thermodynamic version of the CPR fail empirically. This will be shown in the following chapter (or see Watanabe [1955c,65,66,70]).

What might be claimed is that thermodynamics should be formulated as a reversible theory, but that when it is *applied* to the actual universe, special features of the universe - viz., the peculiar boundary condition of extremely low entropy - must be taken into account. The result is that the past-directed probabilities entailed by the reversible thermodynamics fail of the actual universe, *but only as a contingent fact*.

It is not clear to me whether this is a viable treatment. It is certainly appropriate when applying a probabilistic theory to a particular object to take into account further factual knowledge about that object, and to modify probabilistic predictions accordingly, hence this approach to thermodynamics might be

valid. But reservations seem justified because of the special nature of the present case, where the object being considered is the entire universe. For what can our evidence for the nomological existence of past-directed thermodynamic probabilities be, if these probabilities dramatically fail to hold for the entire observable universe?

However this is not a question that needs to be considered any further here. In the first place, the aim of this discussion is only to illustrate the extent to which the CPR has already been essentially recognised in thermodynamics as a condition for reversibility. In the second place, the question arises only when considering a thermodynamics which is based upon a *reversible* micro-theory. A thermodynamics based upon an *irreversible* micro-theory is bound to be irreversible, and hence given that quantum theory is already irreversible by failing the CPR, a quantum-based thermodynamics will automatically be irreversible.

I will conclude this section by noting how an *irreversible* thermodynamics can be based upon a *reversible* micro-theory. Suppose for instance that the micro-theory for the universe is a classical, deterministic, fully reversible theory. And suppose that a thermodynamic theory is postulated which is *irreversible* in the following way. The thermodynamic theory deals with the transitions between macro-states, S_j , which are classes of micro-states of the reversible micro-theory. The thermodynamic theory postulates that various *future-directed* probabilities of the form: $\text{PROB}(S_j(t+\Delta t)|S_i(t))=p$ hold, but does not postulate any

past-directed probabilities. Hence it is irreversible because it fails the CPR.

The first question is: could this thermodynamic theory be obtained as a *logical consequence* of the reversible micro-theory? The *irreversibility* of the thermodynamics does *not* in itself rule this possibility out: an irreversible law can be a logical consequence of a reversible theory (theorem 3.13).¹² However, further practical considerations do rule this out. Call the fundamental micro-theory T , and the thermodynamic theory T^* . If T entails T^* , then, because T is reversible, T also entails $(T^*)^R$ (theorem 3.13). $(T^*)^R$ is the 'time-reversed' version of the thermodynamic theory, and entails *past-directed transition probabilities* in place of the *future-directed transition probabilities* entailed by T^* . Now since T entails $(T^*)^R$, it must be asked why the thermodynamic theory was not formulated as the conjunction: $T^* \& (T^*)^R$. This would give a reversible thermodynamics, which entails the past-directed probabilities of $(T^*)^R$ as well as the future-directed ones of T^* . It would be much stronger than T^* , and if observationally adequate, correspondingly more useful.

The reason that T^* rather than $T^* \& (T^*)^R$ is adopted will have to be because the postulate of the past-directed probabilities is empirically *false* of the real universe, unlike the postulate of the future-directed probabilities. Since it is assumed that T is observationally adequate, T cannot entail $(T^*)^R$. But if T does not entail $(T^*)^R$, it cannot entail T^* either (since if it entails T^*

¹²This is stressed by Earman in his [1967,69,74].

it entails $(T^*)^R$. Hence T does not entail T^* .

This is the expected result. In postulating the probabilities that it does, T^* undoubtedly goes beyond the content of T , which is being presumed to be a *deterministic* theory. T^* must involve a further postulate in which the probabilities make their appearance: this is usually known as the postulate of molecular chaos. It is here that the irreversibility of the thermodynamic theory is introduced.

The second important question is this: does the *irreversibility* of the thermodynamic theory indicate irreversibility in the *laws of nature*, despite the reversibility of the micro-theory? To date, the widely accepted answer to this question has been *no*. The usual view is that the postulate of molecular chaos (or whatever postulate it is that renders the thermodynamic theory irreversible) reflects a *contingent* fact about the world, rather than a natural law. But the basis for this view is unclear to me. We have a workable concept of the distinction between nomological and contingent facts for most purposes, but, as a number of writers have noted¹³, we do not have a clear enough theory of this distinction for settling the present question decisively.¹⁴

¹³E.g. Earman [1969], Grünbaum [19zzz].

¹⁴However I would suggest that the problem stems rather from a lack of clarity in our views about the *creation* of the universe, by which I mean the laws behind the selection of our particular universe as the actual one. For the 'molecular chaos' postulate (for a reversible, deterministic micro-theory) could be *nomological* only if it followed from laws governing the 'selection of the actual universe' from some larger class of possible universes.

5.7 Failure of the PPMR as a criterion for reversibility.

That the CPR is a necessary condition for reversibility has been sufficiently illustrated and emphasized. To give a fuller view of the subject, the failure of the PPMR as a criterion for reversibility will now be considered.

To begin with, it does not need to be further emphasized that the PPMR does not simply provide the *meaning* of reversibility for probabilistic laws, and does not count as the criterion for reversibility by definition, or *a priori*. The PPMR should be adopted as the criterion only if it can be shown by close analysis to guarantee reversibility.

No such analysis has been attempted. Hence there is no important positive argument *for* adopting the PPMR as the criterion of reversibility that needs to be overcome. Nevertheless, there are reasons why the community of physicists have adopted it, and it is worth considering these first. I think there are a number of features that conspired to the adoption of the PPMR:

- (i) The PPMR is an extremely important symmetry in its own right. Indeed, after time-translation symmetry, it is the most important temporal symmetry that *actually holds* of quantum theory. Physicists were bound to spend some effort exploring it, and they did so from early in the history of quantum theory.¹⁵

¹⁵See Lewis [1930].

- (ii) In contrast, the CPR is a temporal symmetry that blatantly fails to hold of quantum theory, and there has been no practical interest in it at all. (There is only the negative result that the CPR fails, no practically useful theorems are connected with the CPR.) Indeed, it is difficult even to get an intuitive grasp of the CPR, since it involves the concept of past-directed probabilities, and Watanabe is the only previous writer who has come near to formulating it for quantum theory (as noted above). Thus, the PPMR has received much attention, while the CPR has been ignored.
- (iii) Having established the PPMR, physicists found it very natural to call it 'time reversal invariance'. The PPMR has the *appearance* of being about time reversal for a number of reasons:
 - (iv) Firstly, it has a connection with the *concept* of time reversibility, since it concerns the *time reversals of states*.
 - (v) Secondly, establishing the PPMR follows on naturally from establishing the *reversibility of the (deterministic) Schrödinger time-dependent equation*. (This result is a pre-requisite for establishing the PPMR).
 - (vi) Thirdly, the PPMR seems to arise as a natural generalisation of a common formulation of the meaning of *reversibility of deterministic transition laws*. The formulation in question involves a way of *visualizing* the meaning of reversibility, and this method of visualization

makes the PPMR appear to be the natural criterion for reversibility of probabilistic transition laws. This method of visualization will be examined shortly.

These features make it understandable why the PPMR has been adopted as the criterion for reversibility. However, none of the points (i)-(vi) above has any weight as a *justification* for this conclusion, except the last, (vi). It is this last point that is undoubtedly the most telling, and I will now discuss it in detail. A serious flaw will be revealed in the way that reversibility is commonly visualized, not only for probabilistic systems, but, remarkably enough, for deterministic systems too.

5.8 A flaw in the interpretation of reversibility for deterministic laws.

Here is a typical interpretation of the meaning of reversibility (or time reversal invariance) for deterministic laws, by Sklarr [1974,p.367]:

..

"We start off with a system in a state s , allowing it to evolve, after time Δt to state s_1 . At the same time we start off another system, exactly like the first, except that its initial state is the "time-reversed state" of the *final* state of the original system. Call this new state s_1^R . If the laws of nature are time-reversal invariant, then at the end of the interval Δt we will find the second system in the state s^R ,

the time-reversed version of the *original* state of the first system."¹⁶

This is invariably the way we are asked to *visualize* the meaning of reversibility for deterministic laws. I will call it the orthodox way of visualizing reversibility. But there is a serious mistake in it. The correct way of visualizing reversibility for deterministic laws is this:

Suppose that a system started in the state s evolves to the state s_1 in time Δt . Then if the laws of nature are reversible, a system found in the state s^R *after evolving naturally for time Δt* must have been in the state s_1^R at the moment Δt earlier.

Sklarr's recipe for visualizing reversibility makes it appear that the time reversal of a *future-directed deterministic law*, L :

[L] If $s_1(t)$, then $s_2(t+\Delta t)$

is another *future-directed deterministic law*, L^* :

[L*] If $s_2^R(t)$, then $s_1^R(t+\Delta t)$.

But this is conceptually wrong: the time reversal of L is in fact

¹⁶Sklarr represents the time reversal operation on a state s by: $T(s)$. This is a common terminology with physicists, but here the superscripted R is used.

the *past-directed deterministic law* L^R :

$$[L^R] \quad \text{If } s_1^R(t), \text{ then } s_2^R(t-\Delta t).$$

(This will be shown shortly.) The mistaken adoption of the orthodox way of visualizing reversibility does not affect the result of the analysis of fully deterministic theories, because for a fully deterministic reversible theory, L^* and L^R are equivalent. (To be shown shortly). But the treatment of time reversal for *probabilistic* laws has been arrived at by generalising from the treatment for deterministic laws, and here the mistake leads to a real problem. For given the orthodox way of visualizing the reversibility of deterministic laws above, the natural generalisation to the case of probabilistic laws seems to be as follows:

"If the equations of motion contained a stochastic term, then the present characterization [the characterization of reversibility for *deterministic* transition laws] 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 s_f^R to s_i^R ." (Earman [1969,p.281]).¹⁷

¹⁷In fairness to Earman it must be noted that this is very much just a passing comment, and Earman nowhere investigates the idea of reversibility for probabilistic laws.

This implies that the time reversal of a *future-directed probabilistic transition law*, M :

$$[M] \text{ PROB}(s_2(t+\Delta t)|s_1(t)) = p$$

is another *future-directed probabilistic transition law*, M^* :

$$[M^*] \text{ PROB}(s_1^R(t+\Delta t)|s_2^R(t)) = p$$

which of course gives the PPMR as the criterion for reversibility of probabilistic laws. But as we have seen, the time reversal of M is really a *past-directed probabilistic law*, M^R :

$$[M^R] \text{ PROB}(s_2^R(t-\Delta t)|s_1^R(t)) = p.$$

Whereas M^* appears as the natural generalisation from L^* , M^R appears as the natural generalisation from L^R .¹⁸ If L^R had been correctly recognised as the time reversal of L , no doubt M^R would have been correctly recognised as the time reversal of M , and the CPR would have been adopted instead of the PPMR.

The conceptual mistake in the visualization of reversibility for *deterministic* laws is, therefore, at the root of the historical mistake in the treatment of reversibility for probabilistic laws.

To conclude, I will demonstrate the claims made above that (a)

¹⁸This is obvious if L is written in probabilistic form, as: $\text{PROB}(s_2(t+\Delta t)|s_1(t)) = 1$. Then L^* takes the form: $\text{PROB}(s_1^R(t+\Delta t)|s_2^R(t)) = 1$, which is just a special case of M^* , whereas L^R takes the form: $\text{PROB}(s_2^R(t-\Delta t)|s_1^R(t)) = 1$, which is just a special case of M^R .

L^R is the time reversal of L in general, while L^* is not, and (b) for fully deterministic reversible theories, L^R and L^* are materially equivalent, so that for such theories the conceptual mistake pointed out here does not have practical implications.

(a) A law L_1 is the time reversal of a law L just in case the following condition holds: for any process P , if P is an L -process, then P^R must be an L_1 -process. Consider the following law, L , and the following schematic depiction of a process, P :

[L] If $s_1(t)$ then $s_2(t+\Delta t)$.

[P] $s_2 \rightarrow s_1 \rightarrow s_2 \rightarrow s_0 \rightarrow s_2 \rightarrow s_0 \rightarrow s_2 \rightarrow s_1 \rightarrow s_2$

(Each arrow represents an interval of Δt). A quick examination shows that P is an L -process, since every occurrence of s_1 is followed by s_2 at a time Δt later. Now let us consider P^R , the time reversal of P . By definition of process reversal, P^R consists of the reversed sequence of reversed states of P :

[P^R] $s_2^R \rightarrow s_1^R \rightarrow s_2^R \rightarrow s_0^R \rightarrow s_2^R \rightarrow s_0^R \rightarrow s_2^R \rightarrow s_1^R \rightarrow s_2^R$

The law L^* : If $s_2^R(t)$, then $s_1^R(t+\Delta t)$ *fails* to hold of P^R (there are two instances in P^R of: $s_2^R \rightarrow s_0^R$, rather than: $s_2^R \rightarrow s_1^R$). Hence the law: If $s_2^R(t)$, then $s_1^R(t+\Delta t)$ *cannot be the time reversal of the law*: If $s_1(t)$, then $s_2(t+\Delta t)$. This proves that, in general, L^* is not the time reversal of L .

To prove that the law L^R : If $s_1^R(t)$, then $s_2^R(t-\Delta t)$ is the time

reversal of L , suppose for a *reductio* that it is not. In this case, there must be some process P which is an L -process, but the reversal, of which, P^R , is not an L^R -process. For the latter to hold, P^R must contain a sequence of states: $s_3^R \rightarrow s_1^R$, where $s_3^R \neq s_2^R$. Since $s_3^R \neq s_2^R$, $s_3 \neq s_2$. Since P^R contains: $s_3^R \rightarrow s_1^R$, by the definition of process reversal, P contains: $s_1 \rightarrow s_3$. Since $s_3 \neq s_2$, this part of the process P fails to obey the law L , and hence P is not an L -process. *Contradiction*. Hence the law L^R : *If $s_1^R(t)$, then $s_2^R(t-\Delta t)$* is the time reversal of the law L : *If $s_1(t)$, then $s_2(t+\Delta t)$* .

(b) Because the orthodox analysis takes L^* to be the time reversal of L , it adopts the following as the criterion for the reversibility of deterministic theories:

- [5.1] A deterministic theory T is reversible just in case for every law of the form L (i.e: *If $s_1(t)$, then $s_2(t+\Delta t)$*) that T entails, it also entails the law of the form L^* (i.e: *If $s_2^R(t)$ then $s_1^R(t+\Delta t)$*).

Since L^* is not the time reversal of L , we can see that this is wrong in principle; since L^R is the time reversal of L , the correct criterion is:

[5.2] A deterministic theory T is reversible just in case for every law of the form L (i.e: If $s_1(t)$, then $s_2(t+\Delta t)$) that T entails, it also entails the law of the form L^R (i.e: If $s_1^R(t)$ then $s_2^R(t-\Delta t)$).

The use of [5.1] instead of [5.2] is generally adequate in practice, however, because for theories which entail that all transitions between states are deterministic, [5.1] and [5.2] are materially equivalent. PROOF. Suppose that T entails that all transitions between states of systems are deterministic. (i) Suppose T satisfies [5.1], and that: *If $s_1(t)$, then $s_2(t+\Delta t)$* is a law of T , for some s_1 , s_2 , and Δt . The problem is to show that: *If $s_1^R(t)$, then $s_2^R(t-\Delta t)$* is also law of T . Suppose, for a *reductio*, that this is not a law of T . By the satisfaction of [5.1], *If $s_2^R(t)$, then $s_1^R(t+\Delta t)$* is a law of T . Hence: *If $s_1^R(t)$, then $s_2^R(t-\Delta t)$* can only fail to be a law of T if there is some other state, s_3 , such that: *If $s_3^R(t)$, then $s_1^R(t+\Delta t)$* is a law of T . (Otherwise s_2^R is the *unique* predecessor of s_1^R , and *If $s_1^R(t)$, then $s_2^R(t-\Delta t)$* must hold.) But then, by the satisfaction of [5.1], *If $s_1(t)$, then $s_3(t+\Delta t)$* is a law of T . But this contradicts the law: *If $s_1(t)$, then $s_2(t+\Delta t)$* , since s_2 and s_3 are distinct. *Reductio*. Hence *If $s_1^R(t)$, then $s_2^R(t-\Delta t)$* is a law of T .

(ii) Suppose T satisfies [5.2], and that: *If $s_1(t)$, then $s_2(t+\Delta t)$* is a law of T , for some s_1 , s_2 , and Δt . The problem is to show that: *If $s_2^R(t)$, then $s_1^R(t+\Delta t)$* is also law of T . By the satisfaction of [5.2], *If $s_1^R(t)$, then $s_2^R(t-\Delta t)$* is a law of T . Hence it is *nomologically possible* for s_2^R to evolve into $s_1^R(t+\Delta t)$. But

since all transitions are deterministic, this transition is possible only if it is nomologically necessary, i.e. if: *If $s_2^R(t)$, then $s_1^R(t+\Delta t)$* is a law of T. Hence *If $s_2^R(t)$, then $s_1^R(t+\Delta t)$* is a law of T.¹⁹

5.9 Failure of the PPMR as a criterion for reversibility.

What has been said so far in this chapter shows that there is no reason to expect that the PPMR is either a necessary or a sufficient condition for reversibility. I will conclude by showing explicitly that it is neither, and hence is quite irrelevant to reversibility. The proofs involve defining two simple hypothetical theories, T_1 and T_2 . T_1 is reversible, but fails the PPMR, showing that the PPMR is not a necessary condition for reversibility. T_2 is irreversible but satisfies the PPMR, showing that the PPMR is not a sufficient condition for reversibility.

T_1 : A reversible theory which fails the PPMR.

T_1 governs a single system, for which there are three possible kinds of states, s_1 , s_2 and s_3 , which are their own reversals, i.e. $s_i = s_i^R$. The laws of T_1 are the following probabilistic transition laws:

$$\begin{aligned} \text{PROB}(s_2(t+1)|s_1(t)) &= 1/2, & \text{PROB}(s_3(t+1)|s_1(t)) &= 1/2, \\ \text{PROB}(s_1(t+1)|s_2(t)) &= 1, & \text{PROB}(s_1(t+1)|s_3(t)) &= 1. \end{aligned}$$

(All other transitions have a probability of 0).

¹⁹ Earman [1986] has an equivalent proof.

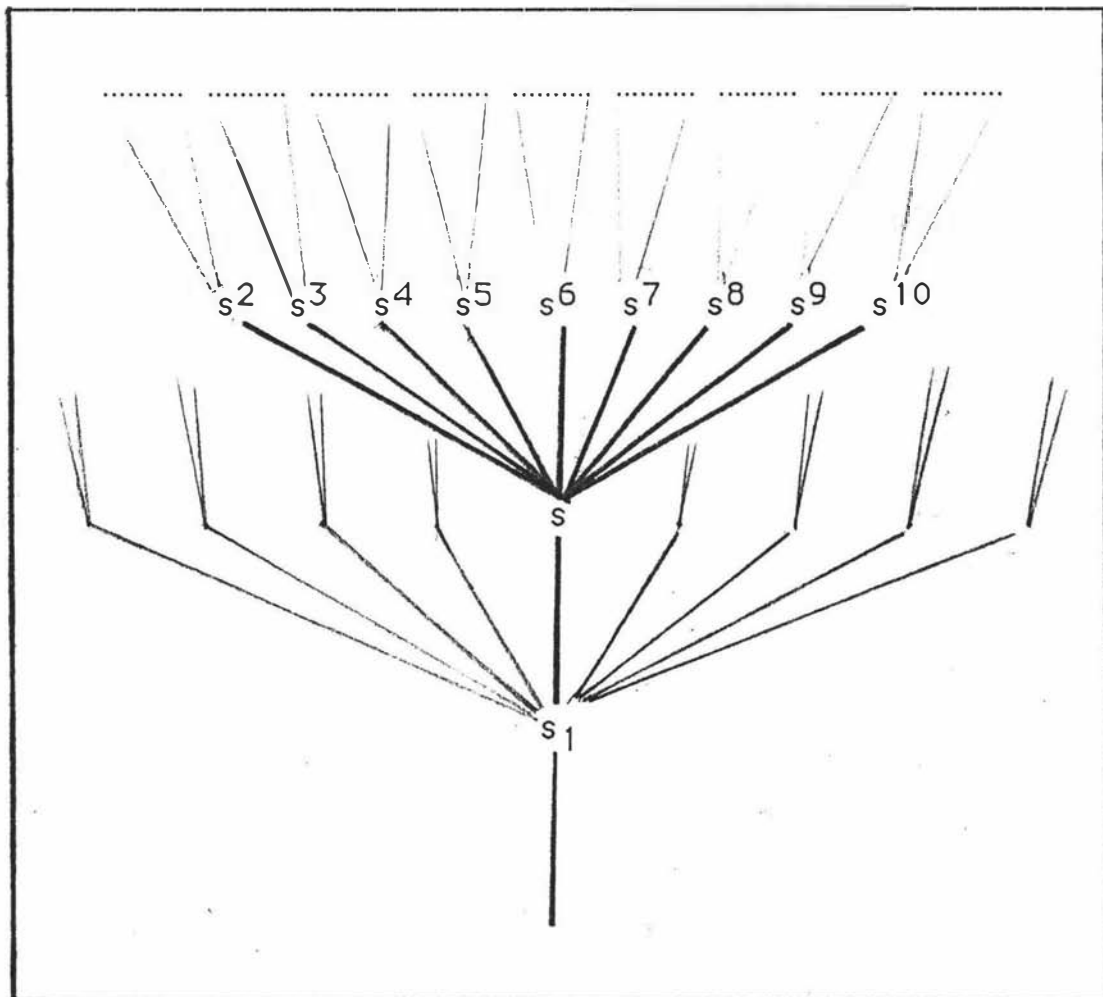


FIG. 5.1

The tree-structure of states for theory T_2 .

(i) T_1 fails the PPMR. For instance, $\text{PROB}(s_2(t+1)|s_1(t)) = 1/2$, but: $\text{PROB}(s_1^R(t+1)|s_2^R(t)) = \text{PROB}(s_1(t+1)|s_2(t)) = 1$.

(ii) T_1 is reversible. All T_1 processes are sequences composed of the following four kinds of transitions: $s_1 \rightarrow s_2$, $s_1 \rightarrow s_3$, $s_2 \rightarrow s_1$, $s_3 \rightarrow s_1$. Because: $s_i = s_i^R$, the first and third of these are the time reversals of each other, as are the second and fourth. Hence the time reversal of any T_1 -process is also a T_1 -process.²⁰

This example shows that the PPMR is not a necessary condition for reversibility.

T_2 : An irreversible theory which satisfies the PPMR.

T_2 is a little more involved than T_1 . It governs a single system which has an infinite number of possible states. These states are related to each other in a kind of tree-structure (see Fig. 5.1). Specifically, each state is *connected* to 10 other states: 9 of these are *higher states*, while one is a *lower state*. If s denotes a given state, then the single *lower state* to which s is connected will be denoted: s_1 , while the 9 *higher states* to which s is connected will be denoted: s^2, s^3, \dots, s^{10} . If a state s is connected to a state s' , then s' is connected to s . If s is higher than s' , then s' is lower than s , and vice versa. No state is both

²⁰A little more is really needed to show that T_1 satisfies the CPR, and hence to show conclusively that T_1 is reversible. I will leave this as an exercise for the reader, since it is obvious enough that T_1 is reversible.

higher and lower than another. Note that: $s = (s^i)_1$ (for any $1 < i < 11$).

Furthermore, each state is its own time reversal.

The laws of T_2 fall into two groups: *future-directed probabilistic laws*, and *past-directed probabilistic laws*. The future directed laws are very simple: there is an equal probability of $1/10$ of transition from any given state, s , to any connected state, s_1, s^2, \dots, s^{10} , in an interval of 1.

$$\text{PROB}(s_1(t+1)|s(t)) = 1/10 \text{ and } \text{PROB}(s^i(t+1)|s(t)) = 1/10.$$

It can be seen that this means that the system governed by T_2 evolves steadily into higher and higher states as time goes on, though with occasional fluctuations backwards into lower states. This is exactly like the evolution of a thermodynamic system relaxing through an infinite series of states of higher and higher entropy.

The *past-directed probability laws* are as follows. Given that a system is in a given state s at time t , the probability that the system was in the state s_1 at the earlier moment, $t-1$, equals $9/10$; the probability that it was in any one of the higher states, s^i , at the earlier moment equals $1/90$.²¹

$$\text{PROB}(s_1(t-1)|s(t)) = 9/10, \quad \text{PROB}(s^i(t-1)|s(t)) = 1/90.$$

It can be shown that this system of probabilities is mathematically consistent.²¹ These past-directed probabilities,

²¹The future-directed probabilities specified limit the consistent possibilities for the assignment of past-directed probabilities. Where the past directed

like the future-directed probabilities, are to be thought of as *objective*, not merely epistemic. Their *objectivity* may be thought of in this way. Imagine that the universe (all that physically exists) is a single system governed by the theory T_2 . Call the entire history of states that a system goes through a *world*. Thus only one *actual world* out of all the possible worlds has existence. Think of the choice of this actual world from the ensemble of possible worlds as being governed by T_2 , and by nothing more. That is, T_2 acts like a probabilistic selection device on the class of possible worlds. It confers a distinctive structure on the selected world, viz., that the selected world satisfies the statistical implications of the probabilities of T_2 . This seems to be essentially the *only* way to make sense of objective probabilities if we take a bloc universe view of existence, and thought of in this way, the past-directed probabilities make sense in exactly the same way as the future-directed probabilities.

The set of past-directed probabilities means that, whatever state a system is found in at a definite moment, it can be inferred to have reached that state by a process of moving steadily from lower to higher states for all of its history, with occasional probabilistic fluctuations. This is consistent with the future-directed probabilities, which imply that it continues to probabilities are assigned to make the value of: $\text{PROB}(s^i(t-1)|s(t)) = 1/10$ constant for all $2 \leq i \leq 10$, there are only two consistent possibilities: the probabilities specified above are one, and the alternative is that: $\text{PROB}(s_1(t-1)|s(t)) = 1/10$ and $\text{PROB}(s^i(t-1)|s(t)) = 1/10$.

move steadily into higher states in the future. Overall, the system acts like a thermodynamic system which relaxes for all time through an infinite series of higher and higher entropy states.

(i) T_2 satisfies the PPMR. T_2 satisfies the PPMR in virtue just of the *future-directed probabilities*, and the fact that each state is its own reversal, as can be easily checked. E.g. $\text{PROB}(s^i(t+1)|s(t)) = 1/10$. To satisfy the PPMR, this must equal the value of: $\text{PROB}(s^R(t+1)|(s^i)^R(t))$, which it does since: $s = s^R$, $(s^i)^R = s^i$, and: $s = (s^i)_1$, so that: $\text{PROB}(s^R(t+1)|(s^i)^R(t)) = \text{PROB}((s^i)_1(t+1)|(s^i)(t))$, which by the laws specified also equals $1/10$.

(j) T_2 is irreversible. This is obvious enough without the need for a formal proof. Any T_2 -process is expected to involve a system developing through higher and higher states for all of its history. The reversal of any such process will involve a system developing through a series of lower and lower states for all its history. These reversed processes clearly contradict the probabilities of T_2 , hence the space of *reversed* T_2 -processes is quite different from the space of T_2 -processes, and T_2 is irreversible. (The reader can easily verify formally that T_2 fails to satisfy the CPR, and since the CPR has been shown to be necessary for reversibility, this demonstrates the irreversibility of T_2 in a formal way.)

We must conclude therefore that the demonstration that quantum probabilities satisfy the PPMR is *irrelevant* to the reversibility

of quantum theory. To establish reversibility, what must be shown is that the CPR is satisfied. The CPR is not satisfied by quantum theory, as Watanabe's [1955c,65,66,70] all effectively show. This result will be explained in the following chapter.

CHAPTER SIX

THE IRREVERSIBILITY OF QUANTUM THEORY

Quantum theory as ordinarily formulated postulates future-directed probabilities, but does not overtly postulate past-directed probabilities. Hence *prima facie*, the CPR is not satisfied by quantum theory, and the theory is irreversible. It may be thought, however, that either quantum theory covertly entails the existence of past-directed probabilities which satisfy the CPR, even though it does not postulate them explicitly, or else that quantum theory could be satisfactorily extended to entail such past-directed probabilities. In the following section I will argue that *past-directed probabilities satisfying the CPR cannot be introduced into quantum theory without making the theory blatantly false of the actual world*. This argument is a simplified version of the more general treatment given by

Watanabe [1955c,65,66,70], but my emphasis is rather different from Watanabe's. In subsequent sections I will consider some possible objections to the argument.

6.1 The lack of nomological past-directed probabilities.

Firstly, it has been established independently that quantum theory satisfies the very important PPMR symmetry¹, i.e. for any quantum states ψ and φ :

$$[6.1] \text{PROB}(\varphi(t+\Delta t)|\psi(t)) = \text{PROB}(\psi^R(t+\Delta t)|\varphi^R(t)).$$

Suppose that quantum theory also satisfies the CPR, so that for any states ψ and φ :

$$[6.2] \text{PROB}(\varphi(t+\Delta t)|\psi(t)) = \text{PROB}(\varphi^R(t-\Delta t)|\psi^R(t)).$$

Substitution of ψ^R for φ and φ^R for ψ in [6.2] yields:

$$[6.3] \text{PROB}(\psi^R(t+\Delta t)|\varphi^R(t)) = \text{PROB}((\psi^R)^R(t-\Delta t)|(\varphi^R)^R(t)),$$

Since for all ψ , $(\psi^R)^R$ is physically identical to ψ (see Appendix 4.2), this entails:

¹See any textbook on quantum theory. The breaking of this symmetry by systems involving k^0 -mesons does not affect the conclusion of the argument since few systems are of this kind.

$$[6.4] \text{PROB}(\psi^R(t+\Delta t)|\phi^R(t)) = \text{PROB}(\psi(t-\Delta t)|\phi(t)),$$

[6.1] and [6.4] yeild:

$$[6.5] \text{PROB}(\phi(t+\Delta t)|\psi(t)) = \text{PROB}(\psi(t-\Delta t)|\phi(t)).$$

Thus we have that Quantum Theory + CPR entails [6.5]. It will now be shown that empirical phenomena blatantly contradict the retrodictive probabilities necessary for [6.5] to be satisfied, hence that Quantum Theory + CPR fails dramatically of the real world. This shows that no empirically satisfactory version of quantum theory could satisfy the CPR.

[6.5] means that it should be valid to retrodict the state ψ from the state ϕ with exactly the same probability that one can predict the state ϕ from the state ψ . But the empirical possibility of *controlling the initial states of systems independently of their final states* means that there is simply no possibility of nomological retrodiction of this kind. For a concrete example, suppose ψ to be the spin-up state on the x axis, and ϕ to be the spin-up state on the y axis, of an ²electron. Then quantum theory predicts that: $\text{PROB}(\phi(t+\Delta t)|\psi(t)) = 1/2$. This probability means that, in any sample of N systems chosen at time t in state ψ , it is to be expected that $N/2$ will be found in the state ϕ at the

² The fact that 'measurements' of spin are *preparations* of spin states means that a system following a measurement which has the eigenvalue corresponding to the eigenstate ϕ as its result is indeed in the state ϕ . It is well known that this is not so for most types of measurements.

later time $t+\Delta t$. If this expected frequency failed dramatically enough in a real experiment, then the idea that the predictive probability: $\text{PROB}(\varphi(t+\Delta t)|\psi(t)) = 1/2$ is nomological would have to be rejected. For instance, suppose $N = 1,000$, and suppose the experimental result that *no transitions from ψ to φ are observed in the sample of 1,000 transitions*. If the probability of this transition were really $1/2$, the chance of this happening by accident would be 2^{-1000} . This probability is so astronomically small that the envisaged result would provide a compelling disconfirmation of the probability law in question. To date however no such negative results have been observed. Instead, the future-directed laws of quantum theory have been strongly confirmed by experiment. This is why those laws are taken seriously.

Let us now turn to the retrodictive probability: $\text{PROB}(\psi(t-\Delta t)|\varphi(t))$ which forms the right-hand-side of equation [6.5]. To continue the example, [6.5] and the fact that: $\text{PROB}(\varphi(t+\Delta t)|\psi(t)) = 1/2$ entail that: $\text{PROB}(\psi(t-\Delta t)|\varphi(t)) = 1/2$. But consider the following simple experiment, which I will call Experiment 1. It has in effect been performed many times by experimental physicists.

Experiment 1. 1,000 systems are prepared in the state φ at the time $t-\Delta t$, and each of these is found in the state φ at the later time t (since: $\text{PROB}(\varphi|\varphi) = 1$.)

The retrodictive probability: $\text{PROB}(\psi(t-\Delta t)|\varphi(t))=1/2$ under consideration fails dramatically of this sample. This probability

should lead one to expect at time t that on average 500 systems presently in the state φ have evolved from the state ψ ; but in empirical fact none have evolved from ψ . More precisely, if: $\text{PROB}(\psi(t-\Delta t)|\varphi(t)) = 1/2$ were genuinely nomological, the probability of this result happening by accident is again the astronomically small figure of 2^{-1000} ; hence this experiment decisively disconfirms the existence of the retrodictive probability in question.

At first it may seem that this experiment is unfair, because of the fact that the systems being considered have been *deliberately manipulated* to contradict the probability: $\text{PROB}(\psi(t-\Delta t)|\varphi(t)) = 1/2$ (by the deliberate manipulation of the initial states). But in fact this is the whole point of the experiment: namely, to demonstrate that physical systems can be manipulated to behave in such a way that retrodictive probabilities such as: $\text{PROB}(\psi(t-\Delta t)|\varphi(t)) = 1/2$ fail! If this probability was genuinely nomological, and really governed the behaviour of the physical systems, then it would be impossible to manipulate the behaviour of those systems so that they contradicted it. This is certainly so with the future-directed probability: $\text{PROB}(\varphi(t+\Delta t)|\psi(t)) = 1/2$. Because this probability *is* nomological, it is impossible to manipulate the later states of systems so that it fails. What future-directed probabilities of this kind imply is the uncontrollability of the future behaviour of systems governed by it. There is no physical method of selecting, at time t , a sample of systems in the state ψ , which will evolve

so that the future-directed probability: $\text{PROB}(\varphi(t+\Delta t)|\psi(t)) = 1/2$ is seriously contravened. In contrast, Experiment 1 shows that there is a method of selecting, at time t , a sample of systems in the state φ , which have evolved in such a way that: $\text{PROB}(\psi(t-\Delta t)|\varphi(t))=1/2$ is contravened. Hence this probability cannot be nomological.

This shows that there are no retrodictive quantum probabilities in nature satisfying the CPR. In fact a much more general result is evident: *there can be no nomological probabilities of the form: $\text{PROB}(\psi(t-\Delta t)|\varphi(t))=p$ whatsoever, if there exists the physical possibility of (i) forcing systems into the state φ at time t , in such a way that: (ii) each system is (or alternatively, is not) in the state ψ at the earlier time $t-\Delta t$.*

The discussion of Experiment 1 shows how this possibility would allow for an experiment disconfirming the past-directed probability, since it allows directly for the production of a sample of systems which have undergone transitions such that the frequencies of past states relative to future states contradict any past-directed probabilities. Since for ordinary quantum states ψ and φ , both (i) and (ii) are possible, in general there are no past-directed quantum probabilities of the form: $\text{PROB}(\psi(t-\Delta t)|\varphi(t))=p$.

In following sections some possible objections to this basic argument will be considered, but it may be helpful to first consider a little more deeply the reason for the failure of the past-directed probabilities. The reason is clearly that nomological retrodiction does not work in the real world.³ But

what feature of the world entails its failure? The ultimate feature is in fact *the striking decrease of entropy of real systems in the past direction of time*. I will describe the basic reason for this, although the argument that follows is not fully precise.

Let: $FREQ(e)$ represent the actual frequency of occurrences of an event of type e in the history of the universe to date⁴. I will use the conventions that: $FREQ(\psi)$ denotes the frequency of distinct occurrences of the micro-state ψ , and that: $FREQ(\psi(t)\&\phi(t+\Delta t))$ denotes the frequency of transitions of systems from state ψ to state ϕ in an interval of Δt .

By the definition of conditional probability, the existence of a probability law: $PROB(\phi(t+\Delta t)|\psi(t))$ means that actual frequencies are *expected* to conform to:

$$[6.6] \text{ } PROB(\phi(t+\Delta t)|\psi(t)) = FREQ(\psi(t)\&\phi(t+\Delta t))/FREQ(\psi).$$

Similarly, if: $PROB(\psi(t-\Delta t)|\phi(t))$ were a probability law, then it would be expected that:

$$[6.7] \text{ } PROB(\psi(t-\Delta t)|\phi(t)) = FREQ(\psi(t-\Delta t)\&\phi(t))/FREQ(\phi).$$

Note that: $FREQ(\psi(t-\Delta t)\&\phi(t))$ is identical to: $FREQ(\psi(t)\&\phi(t+\Delta t))$,

³See also Grunbaum [1973].

⁴ Or if the universe is infinite, throughout a cosmically large local spatio-temporal region.

so we may rewrite [6.7] as:

$$[6.7] \text{PROB}(\psi(t-\Delta t)|\varphi(t)) = \text{FREQ}(\psi(t)\&\varphi(t+\Delta t))/\text{FREQ}(\varphi).$$

If [6.5] was true, it would therefore be *expected* that:

$$[6.8] \text{FREQ}(\psi(t)\&\varphi(t+\Delta t))/\text{FREQ}(\psi) = \text{FREQ}(\psi(t)\&\varphi(t+\Delta t))/\text{FREQ}(\varphi)$$

from which it follows that:

$$[6.9] \text{FREQ}(\psi) = \text{FREQ}(\varphi).$$

The implication of this is that all possible micro-states of a given type of system should be *expected* to occur with the same frequency. This condition will appear to be fulfilled only by systems in periods of thermodynamic equilibrium. Systems over periods of low entropy observably fail condition [6.9], because the low entropy state means that a certain tiny class of micro-states observably dominates in frequency over a far larger class. The fact that the universe has, throughout its known history, been evolving through a series of states of extraordinarily low entropy means that [6.9] can be seen to fail at a rather gross level of observation. Hence the low entropy state of the universe can be viewed as the prime evidence for the failure of [6.5].

There are of course three different ways in which [6.5] could fail: through the failure of appropriate past-directed probabilities, future-directed probabilities, or of both. It happens

that the past-directed probabilities fail. This is ultimately a reflection of the fact that while the increase of entropy in the future direction of time is consistent with the future-directed probabilities of quantum theory, the decrease of entropy in the past direction of time is inconsistent with past-directed probabilities satisfying [6.5].⁵

It is very interesting that two distinct sources of evidence for the lack of past-directed probabilities have now been established. (i) The first argument against past-directed nomological probabilities turned on the fact that it is possible to control the frequency of earlier states relative to later states, while in contrast it is not possible to control the frequency of later states relative to earlier states (these frequencies really being determined by future-directed probabilities). (ii) The second result was that what ultimately supplies the evidence for the lack of past-directed probabilities is the startling decrease of entropy of the universe in the past direction of time. It would seem that these two forms of evidence must be connected. The basic connection appears to be this: the possibility of control of frequencies of earlier states relative to later states is a reflection of the 'flow of information' from past to future. What is meant by the latter is that detailed information about states at earlier times can be preserved in the states of physical systems at later times (while temporal reverse of this does not hold). This generates the well-known asymmetry of knowledge

⁵See Watanabe [1955,65,66,70] for more detailed discussion.

about the past and future, for instance. It allows the requisite control of frequencies of earlier states relative to frequencies of later states basically because it provides for a physical method of choosing samples of systems for which the earlier states are known in detail at later times. (It can be seen that if this held for the future, i.e. if it could be presently known whether or not a system in present state ψ would develop to state ϕ in the future, then samples could be presently chosen which contradicted the future-directed probabilities.⁶) This 'flow of information' from past to future has in turn been traced to the increase of entropy in the future direction of time by a number of authors.⁷ Hence it appears that the increase of entropy in the universe is the fundamental reason for the failure of past-directed probabilities, and is what gives rise to possibility of control of frequencies of past states relative to future states. The full discussion of this subject is beyond the scope of the present work, however.

I will now turn to some possible objections to the first argument given against the existence of past-directed probabilities.

..

⁶There is the possibility that relative frequencies on a global level would still conform to the probabilities; but the empirical fact is that where such knowledge of the future is available, it is in our control to manipulate systems to upset any such global frequencies.

⁷E.g. Mehlberg [1980], Grunbaum [1973], Reichenbach [1956].

6.2 Objection 1: an accident?

It may at first appear that we can successfully combine the past-directed probability: $\text{PROB}(\varphi(t+\Delta t)|\psi(t)) = 1/2$ with the 'contradictory frequencies' of the proposed experiment by claiming that the frequencies in question occur only because our universe is, as a contingent matter, a special sort of universe. Specifically, it is a universe in which a special boundary condition holds (low entropy in the early universe), and this *contingent* feature can ultimately be regarded as *contingently* giving rise to the rogue frequencies which appear to contradict the past-directed probability.

Although the flaw in this objection is obvious, it is similar in some ways to a view which is popular with a great many writers on the subject, namely that the manifest temporal asymmetry of the world arises merely from special contingent boundary conditions on the universe, and has no implications for the reversibility of the fundamental theory.⁸ Therefore I will consider this objection carefully.

For an analogous case, suppose that we have a coin for which we postulate the following probabilistic laws. Let H and T represent the events of the coin coming up *heads* and *tails* respectively after a throw, and suppose that the coin is thrown once per unit of time. The following postulate states that the chance of throwing heads (H) after previously throwing tails (T)

⁸See particularly C.deBeauregard [1970,1977].

equals 1/2:

$$[6.10] \text{ PROB}(H(t+1)|T(t)) = 1/2.$$

Now suppose that in a long series (say 10^6) of observed throws, in only 1% of cases is H followed by T. Would we be prepared to maintain the truth of [6.10]?

Obviously not: we would credit [6.10] with almost no chance of being correct. But consider the following argument against this conclusion. *Despite appearances, the observed sample does not contradict [6.10]. For the deviation in the sample from the behaviour predicted from [6.10] arises only as a contingency: specifically, it arises because of the following contingent 'boundary condition' on the set of samples in question: that in the sample of all throws, in only 1% of cases is s_2 followed by s_1 .*

This argument of course is ridiculous: if this form of argument were allowed, it would make any probabilistic hypothesis equally compatible with any statistical observations. The argument comes down to saying that the deviation from the norm implied by [6.10] is just an *accident*, but of course we can only evaluate probabilistic or statistical hypotheses by assuming that the sample of evidence does not just reflect an incredible 'accident', but is a reasonably fair reflection of the probabilities involved. The argument in the opening paragraph is of exactly the same form as that just given, and equally mistaken.

It is perhaps also worth dispelling a variation on the argument

under dispute. The variation goes like this. *The probabilities of [6.10] are apparently contradicted by the sample of observations, but there is an explanation for the peculiarity of the sample. The explanation is that: (i) the sample obeys a boundary condition which implies the deviation from [6.10], and (ii) this boundary condition is not merely 'accidental', but nomological, i.e. its occurrence (or a good probability of its occurrence) is after all implied by, and explained by, certain natural laws.*

But (ii) is just to say that the probabilities of [6.10] are not nomological; for if there are natural laws that imply a blatant contradiction of the postulated probabilities, then the probabilities cannot themselves be nomological.

6.3 Objection 2: A biased sample?

There is only one way that a sample of observations which *prima facie* disconfirm a probabilistic hypothesis, can be reconciled with that hypothesis: this is if the sample is systematically biased in some way, not randomly chosen. Is the sample of observations appealed to in the hypothetical experiment of Section 6.1 'randomly chosen'? If it can be established that it is not, then the resulting argument might be rejected.

This seems at first a very promising objection. The sample that disconfirms the past-directed probability: $\text{PROB}(\psi(t-\Delta t)|\varphi(t))=1/2$ was deliberately produced so to do just that. The point, however, which was already made in Section 6.1, is that if this probability was nomological in the same way as the future-

directed probability of the form: $\text{PROB}(\varphi(t+\Delta t)|\psi(t))=1/2$ are nomological, then it would be quite impossible in practise to produce the sample in question.

The usual way in which a class of observations fails to be randomly selected is if it is specially selected from a larger class, to achieve a special bias. For instance, imagine that you were shown a film of a coin being thrown, in an apparently 'fair' way, ten times in succession, and landing heads each time. If you could think of no trick in the method of throwing which biased the coin, you would probably suspect that the person who made the film actually filmed a very long sequence of coin throws, that the sequence of ten heads occurred by chance somewhere in the much larger sequence, and that this special part of the entire sample was deliberately selected out to impress you. If so, it does not represent a random sample of throws of the coin, but a biased sample.

This is the most obvious method of biasing a sample, but it is clearly only possible to use it where the size of the full sample required to produce the 'special case' is of practical dimensions. For instance, it would be physically impossible to generate by the same method a film of 1,000 heads in a row, since somewhere in the general order of: 2^{1000} individual tosses of a fair coin would be needed before this special sequence could be expected to occur by chance. It is physically impossible that this astronomical number of tosses of a coin could be observed.

For the same reason, it is physically impossible that the

evidence that exists against past-directed probabilities could be counterfeited in any way. Experiment 1 shows that the improbability of nomological past directed probabilities is simply staggering.

6.4 Is Experiment 1 of the wrong type?

Another possible objection to Experiment 1 is that it is of the wrong *type* to test the probability in question. Let me first make the general idea clear with an example of an experiment similarly ill-designed to test the future-directed probability: $\text{PROB}(\varphi(t+\Delta t)|\psi(t))=1/2$. Imagine a sample of sample of 1,000 systems prepared in state ψ at t , and subjected to a measurement which does *not* have φ as an eigenstate. Suppose for instance that the measurement has ψ as an eigenstate. In this case, all transitions will be from initial state ψ to final state ψ ($\text{PROB}(\psi|\psi)=1$). Hence, in the 1,000 transitions of state, there will be *no* cases of transition to φ , despite the probability: $\text{PROB}(\varphi(t+\Delta t)|\psi(t))=1/2$. Of course this result does not contradict that probability, because that probability is further conditional upon a certain type of measurement being made, namely one which has φ as an eigenstate, and this condition is not fulfilled.

Perhaps then a similar objection may be raised against the idea that the 1,000 cases of the transition: $\varphi(t-\Delta t)\rightarrow\varphi(t)$ in Experiment 1 contradicts the probability: $\text{PROB}(\psi(t-\Delta t)|\varphi(t)) = 1/2$. For to consider this as a nomological probability of the same kind

as: $\text{PROB}(\varphi(t+\Delta t)|\psi(t))=1/2$, it has to be considered that there is a 'measurement' mediating the state $\varphi(t)$ and the state $\psi(t-\Delta t)$. This 'measurement' will have to be considered as occurring in the reverse direction of time. The analogous condition on the eigenstates of this 'past-directed' measurement will have to be that it has ψ as an eigenstate. Now what this condition means is rather difficult to say, since the notion of a 'past-directed measurement' is entirely unclear, but perhaps it is reasonable to hold that the experimental procedure must allow at least for the possibility that the state at $t-\Delta t$ is ψ . In the proposed experiment, it is reasonable to maintain that the states are all constrained to be φ , hence there is no such possibility, and the 'past-directed measurement' involved is of the wrong type to test the hypothesis that: $\text{PROB}(\psi(t-\Delta t)|\varphi(t)) = 1/2$.

I have two comments on this objection. (i) Firstly, it is apparent that the concept of a 'past-directed measurement' is undefined in quantum theory, and probably quite nonsensical. This simply stresses the failure of the existence of past-directed probabilities which are temporally symmetrical to future-directed probabilities, since the concept of past-directed measurements essential to bring such probabilities into play is apparently nonsensical.

It is nonsensical in particular because there is no definable concept of *the eigenstates of a past-directed measurement*. Future-directed measurements have well-defined eigenstates in the sense that the measurement process between initial state $t-\Delta t$ and final state t physically constrains the class

of possible final states of the measured system immediately *after* time t . But no measurement process between time $t-\Delta t$ and t constrains the state immediately *before* time $t-\Delta t$. This state is instead constrained by the evolution of the system before time $t-\Delta t$. This is a direct reflection of physical 'irretrodictability'.

(ii) Consider the following experiment. 2,000 systems are prepared in the state ψ at time $t-\Delta t$, and then subjected to a measurement which has ϕ as an eigenstate. It is to be expected that 1,000 systems make the transition to ϕ at t , while 1,000 do not. Let us suppose for simplicity that in fact exactly 1,000 systems make the transition to ϕ at t . Those which do not make this transition have no relevance to the probability: $\text{PROB}(\psi(t-\Delta t)|\phi(t))=1/2$, since the condition: $\phi(t)$ is not fulfilled. The sample relevant for testing this probability is hence just the set of 1,000 transitions of: $\psi(t-\Delta t)\rightarrow\phi(t)$. The experimentally produced sample of course strongly disconfirms the probability: $\text{PROB}(\psi(t-\Delta t)|\phi(t))=1/2$, since if this probability held, it would be expected that of the 1,000 systems in final state $\phi(t)$, only 500 began in the state $\psi(t-\Delta t)$, which departs spectacularly from the observed number. Moreover, if it is regarded that there is a 'past-directed measurement' mediating $\phi(t)$ and $\psi(t-\Delta t)$, then clearly this 'measurement' has ψ as an eigenstate in every case. This experiment therefore overcomes the present objection, even on the rather wild hypothesis that the concept of 'past directed measurement' is ultimately coherent.

6.5 Objection 3: anthropomorphic bias?

Consider the following argument. A sample is *biased towards meeting a condition, C* if the method used to choose the sample means that it is more likely than normal to meet the condition C. Any sample of thermodynamic systems used to demonstrate anything about the physical world must meet the condition that *it is selected in a universe containing intelligent (for our purposes, human) life*. The effect of this bias is considerable. Since life as we know it can only evolve and exist in an environment which is in a fairly extreme thermodynamic disequilibrium, that environment will exhibit all the consequences of thermodynamic disequilibrium, one of which is irretractability. Hence the irretractability illustrated by the sample of evidence in Experiment 1 has really been guaranteed by the simple fact of selection of the sample of evidence. The sample being viciously biased, the argument of Section 6.1 is invalid.

I think this objection can be quickly dismissed. Consider the following analogous case. A person conducts a survey amongst all her acquaintances to find out how many of them hate her enough to want to kill her. The number she arrives is quite low, let us say less than ten. A critic subsequently argues that this kind of result was inevitable, because if more than ten people hated her enough to want to kill her, she would have been dead before she could administer the survey. Therefore the very act of sampling

implies a certain result, hence a bias in the sample.

True as this is, it does not affect the accuracy of the result: the survey shows that less than ten people want to kill the paranoid statistician, and this is quite correct. The 'bias' does not invalidate the result.

Consider a second kind of survey, which is made to collect evidence on whether *life exists in the universe*. Since surveyors are themselves alive, it is inevitable, if the survey is carried out sensibly, that the evidence collected will be positive: it will be concluded that life exists. This conclusion remains true, despite the 'bias' represented by the fact that *the survey can only be carried out in a universe where life exists*.

The sampling of systems in the real universe which demonstrates irretrdictabilty seems to be exactly the same: the supposed 'bias' does not alter the fact that the sample gives the correct information about whether past-directed probabilities exist.

6.6 Objection 4: long-term equilibrium?

Irretrdictability arises ultimately from the thermodynamic disequilibrium of the universe. Could it be that the vast thermodynamic disequilibrium of the universe in its present era is merely an 'accidental' fluctuation from a normal state of equilibrium? If so, the present disequilibrium would be a highly unrepresentative state, and would not establish the lack of past-directed probabilities.⁹ In this situation, the 'anthropomorphic

bias' described above would become much more sensible. The situation would be that the universe is in a state of equilibrium for almost all periods of its history, so that any properly *representative* sample of processes would indicate the existence of past-directed probabilities in exact symmetry with future-directed probabilities. Periodically, however, vast disequilibriums would arise, by natural chance. A sampling of physical systems in such an era would indicate the failure of [6.9], hence of [6.5]. It may be assumed that only in such eras could the complex structures required for the existence of life develop, hence any actual sampling by living creatures such as ourselves would inevitably, but quite incorrectly, indicate the lack of past-directed probabilities. The apparent irreversibility of quantum theory would then quite genuinely be a mere artifact of the possibility of formulating the theory.

This seems to be the only serious objection to the argument of Section 6.1. Two different versions should be distinguished, one weak and one strong. The weak version has the major premise:

[6.12] It is *not known* whether the universe is merely in a chance fluctuation from equilibrium or not,

and the corresponding conclusion that:

⁹This would be a special case of a 'biased sample', since the sample of processes used to establish the lack of past-directed probabilities would in fact be unrepresentative.

[6.13] It is *not known* whether the ultimately correct version of quantum theory is reversible or not.

The strong version has the major premise:

[6.14] The universe is merely in a chance fluctuation from a normal state of equilibrium;

with the corresponding conclusion that:

[6.15] The ultimately correct version of quantum theory is reversible.

The lack of detailed knowledge about either the origin or the end of the present era of the universe means that this question cannot yet be firmly decided. But it can be stated that (a) there is as yet no observational or theoretical support for [6.14], and thus no reason at all to think it is correct.¹⁰ (b) [6.12] seems to be the

¹⁰There is also no guarantee that [6.15] follows from [6.14]. It is also worth noting the points made by Penrose [1986,p.41] against the quite unfounded idea entertained by Gold [1962], Cocke [1967], and others that, in a closed universe which undergoes cycles of expansion and subsequent gravitational contraction, entropy will *decrease* in the contraction period, in symmetry to the way that it is presently increasing in the present period of expansion. There is no reason to think that gravitational collapse implies entropy decrease, and there is especially no reason to think that it implies that the reversals of ordinary thermodynamic processes will occur. This would imply a blatant contradiction of future-directed quantum probabilities, and there is no reason to think that quantum theory will

kind of skeptical view which cannot be conclusively dismissed, but which there seems to be no pragmatic reasons for entertaining. According to our most general observational conception of the universe, the only plausible version of quantum theory is the irreversible version. However, until the process of the creation of our universe is understood, our understanding of the reversibility of the laws of nature remains ultimately incomplete.

6.7 Epistemic past-directed probabilities.

It has been argued that there are no past-directed quantum theoretic probabilities in temporal symmetry with the future-directed probabilities. Nevertheless, the future-directed probabilities may give rise, in special circumstances, to objective past-directed epistemic probabilities. Watanabe [1965,66,70], Aharonov *et alia* [1964], Cocke [1967], and Penfield [1966] have discussed this question, although there does not appear to be any systematic treatment of the subject. This topic must be central in a quantum theoretic account of knowledge. In particular, if our world is really a quantum theoretic world, and if all present knowledge of the past is ultimately grounded in the nomological implications that the present physical states of systems have for past states, then substantial knowledge of the past requires past-directed probabilities.¹¹

become blatantly false in the future.

There is also a fundamental problem about the status of future-directed probabilities, which has been noted but not discussed. This is that the future-directed probability laws of the form: $PROB(\varphi(t+\Delta t)|\psi(t))=p$ have an implicit second conditional, which is that a measurement for which φ is an eigenstate must be made. Unless such a measurement is made, there is no actual chance of $\varphi(t+\Delta t)$ occurring.

If quantum theory is to generate absolute probabilities of future events from present conditions (and not merely probabilities which are further conditional on events of measurement occurring), it must also recognise absolute probabilities of events of measurements occurring in the future.

Here I will just note the orthodox view of this problem. The orthodox view does not seem satisfactory at a fundamental level, but it is pragmatically adequate. The orthodox interpretation maintains a distinction between 'classical' concepts and quantum concepts, and holds that while the formal quantum theory is primarily about the quantum concepts, classical concepts are necessary for any understanding of the practical application of quantum theory¹². Measurement is a key 'classical' concept in this regard, and a corresponding 'classical' understanding of what measurements are is presumed. Although it is hardly transparent what 'classical' should be taken to mean here, it seems that our ordinary understanding of the processes of making measurements allows for the possibility of it being *physically determined at*

¹¹Including past-directed necessities, which are special cases of probabilities.

¹²See Jammer [1974], Murdoch [1987].

the present time that a measurement of a certain kind will be performed at a future time. If this is so, then there will be cases when all the conditions necessary for absolute probabilities of future events to arise are satisfied.

It is even more important that there are cases where it is physically determined at the present time that a measurement of a certain kind was performed at a past time. Indeed, orthodox quantum theory depends not only on this, but on the idea that the results of past measurements can be presently determined. The practical basis for this assumption is obvious enough, but the theoretical basis is not clear.

CHAPTER SEVEN

A DYNAMIC MODEL OF TIME

I turn now to the development of a formal model of time flow suited to a probabilistic world, in line with the intentions expressed in the Introduction. The model proposed here is a development of the dynamic model proposed by McCall [1976]. But after acknowledging a considerable debt to McCall, I must stress that I reject his theory on a number of fundamental points. The model proposed here will be shown to arise out of the adoption of natural solutions to certain fundamental problems with McCall's theory.¹

¹ Reichenbach [1953], Capek [1961], Whitrow [1961], Bondi [1952] have also suggested, like McCall, that the idea of indeterminism supports a dynamic view of time. McCall's theory owes something to their speculation, particularly to Reichenbach's. But although these other writers have articulated their intuitions

Section 7.1 begins with a very preliminary sketch of the model that will ultimately be defended. In Section 7.2 McCall's theory will be sketched, and the basic framework of concepts made clear. In following sections, important features of McCall's model will be criticised, and it will be shown how the adaptations necessary to solve the problems with his theory leads to the theory described in Section 7.1.

7.1 Preliminary sketch of the model.

Envisage first of all a simple sort of probabilistic theory, which determines absolute (single-case) probabilities of future states of systems given their complete present micro-states (plus relevant present facts about their environs). Assume that this theory postulates a finite universe which at any moment is in a definite complete micro-state, and that this universe as a whole can be regarded as a system governed by the probabilistic laws. (This theory is, of course, non-relativistic, since it requires a physically real world-wide present state: the problem of placing relativistic constraints on the dynamic ontology described here will not be discussed until Chapter 9.)

The idea is to give a model of this probabilistic theory in which time is 'dynamic'. The broad idea is as follows. Firstly, there is presumed to be something that *presently exists*. This on the subject, they have not developed substantial theories, and I will not discuss their work here.

consists of (i) a concrete physical universe in its present momentary state, plus (ii) the physical probabilities of future states, as determined by the present state plus the probabilistic laws of nature, plus (iii) the physical possibilities of past states, these possibilities being determined by the present state plus the laws of nature. (No nomological past-directed probabilities are assumed, although there might well be objective epistemic past-directed probabilities, as described in the previous chapter.)

Note that (ii) requires a strongly realistic theory of physical probabilities. The logical framework for such a theory will be described in a later section. Probabilities will be explicated as genuine physical properties of physical entities. Physicists find it hard to imagine probabilities as physical things in this way, but this seems to be only because they are accustomed to thinking of simpler kinds of physical properties, such as mass or electric charge. Probabilities are of a much more complex logical type than such properties. This complexity makes them hard to imagine, but it does not prevent them from being physically real. At any rate, a formal model will later be proposed, and I trust it will prove logically coherent. **

This entity, which has components (i), (ii), and (iii) described above, is therefore regarded as the complete entity which has present existence. Following McCall's suggestive terminology, it will be called the *universe-tree*. The laws of nature will allow for many nomologically possible universe-trees.

Note that the present universe-tree is extended in time. For part (ii) of the universe-tree consists of probabilities of states

at times later than the present moment, while part (iii) consists of possibilities of states at times earlier than the present moment. These represent presently existing properties of times earlier and later than the present moment. Hence other moments participate in the object which presently exists. The fact that these probabilities and possibilities obtain in the present universe-tree simply means that there are present facts about earlier and later moments.

It must be recognised that the term 'present facts' means facts that have the present *mode of existence*. The central feature of the dynamic ontology is that it recognises three temporal modes of existence, namely past, present and future modes of existence. Let us now turn to these.

Since the ontology is supposed to be dynamic, we are required to recognise not only the existence of the present universe-tree, but also that there have been other universe-trees which existed in the past, and others again which will exist in the future. Since the past and future are being treated realistically, elements must be provided in the formal model which correspond to the entities which have past and future existence. These are past and future universe-trees: hence we are committed not only to the existence of a present universe-tree, but to the existence of a whole temporal sequence of universe-trees. (See Fig. 7.2)

There are three kinds of items in the temporal sequence: past universe-trees, the present universe-tree, and future universe-trees. It is crucial to recognise that these different kinds of

items have different modes of existence, viz. past, present and future existence, respectively. This is crucial to the nature of the existence of the temporal sequence: it means that the temporal sequence does not itself have present existence, but rather, has a mixed mode of existence, being partly past, partly present, partly future. This mixed mode is reflected formally by the logical type of the sequence. It is vital to recognise this mixed mode of existence if certain logical problems are to be avoided.

The model now provides, in an obvious way, for a semantics for past- and future-tensed propositions. If P is a proposition, then the past-tensed proposition: *it was the case that P* is true just in case P holds of a past universe-tree. Similarly, *it will be the case that P* is true just in case P holds of a future universe-tree.²

These tensed propositions, which are about past and future parts of the temporal sequence of universe-trees, must be contrasted with another kind of 'tensed' proposition, which I will call *untensed propositions*. Untensed propositions are about the present universe-tree. That is, they take truth-values with

²It will be seen that this preserves the idea that there is a 'unique past' and a 'unique future', in the sense that any past-tensed proposition: *it was the case that P* is either true or false, and similarly any future-tensed proposition: *it will be the case that P* is either true or false. But the truth of these propositions is not determined by the present universe-tree, at least not unless physical determinism holds. This construal of semantics is a serious departure from McCall's [1976] (and also Thomason's [1970]) theory, but I argue below that McCall is confused on this point.

respect to the single universe-trees, in contrast to tensed propositions, which take truth-values only with respect to temporal sequences of universe-trees. Untensed propositions are about facts in the present universe-tree, and this includes facts at moments other than the present moment. It is the temporal extension of the present universe-tree that provides for such propositions.³ It is natural to wonder why untensed propositions about times other than the present moment are required in addition to tensed propositions. This is really just the question of why universe-trees are temporally extended.⁴ The answer is that they must be temporally extended if there are to be present facts about times earlier and later than the present, and that on any satisfactory view of the world there must be such facts. This point will be discussed in detail in Section 7.9.

The basic idea, then, is that what exists presently is a universe-tree, which consists of a present instantaneous micro-state of the entire universe, plus the probabilities and possibilities which this micro-state generates via the laws of nature. What makes the model dynamic is that this entity *changes*. Change is represented through the relations of the present universe-tree with past and future universe-trees. In the abstract, these relations are that the past universe-trees *have*

³It will be seen that what McCall [1976] and Thomason [1970] regard as tensed propositions are in fact untensed propositions.

⁴Since the temporal extension is just a result of presently existing facts about times other than the present moment, and such facts immediately generate propositions about those times.

changed to the present universe-tree, and that the present universe-tree *will change* to future universe-trees. The key problem that concerns us is the physical basis for change. It will be remembered from the Introduction that we desire a physical theory of time flow. What this means is that the dynamic feature should not be added to the physical ontology through an arbitrary, 'metaphysical' postulate: instead what we require is that it corresponds to a respectable element of physical reality, postulated by the physical theory. What is it in the physical theory that provides for the interpretation of the dynamic feature, i.e. the relations generated by change?

My proposal is to identify change with the actualisation of probabilities. This interpretation depends upon the idea that the concept of actualisation of probability is implicit in the concept of physical probability appropriate to the physical theory in question. In another way of putting it, the notion of physical probabilities of future events cannot be understood unless the idea of the actualisation of these probabilities is already understood. The idea of actualisation of probabilities is that something must *happen* to render one possible probabilistic outcome actual, and the other possible outcomes non-actual. For instance, if there is a probability of $1/3$ that a system X has the state s_1 in one minutes time, and a probability of $2/3$ that it has the state s_2 in one minutes time, then it is implied that either s_1 will be actualised or s_2 will be actualised, in one minutes time.

It may seem that this procedure is circular, for it seems that

change is interpreted in terms of actualisation of probability, and then the idea of actualisation of probability refers back to the idea of change. This kind of circularity is inevitable, but not vicious. It simply reflects the fact that to add the idea of change to a conceptual scheme which does not already include it necessarily involves the addition of a new primitive concept. I am not therefore proposing to define change without introducing a primitive concept. My claim is only that the primitive concept is already well-established in physical theories: for it is implicit in the concept of actualisation of probability, which is an element of the concept of physical probability. (More on this in Section 7.10).

The physical probabilities are all future-directed, and this corresponds to the temporal direction of change. More precisely, the temporal direction of actualisation of probability is identified with the direction of change as follows: the state (of the world) U_1 *changes* to the state U_2 just in case the probabilities in U_1 are *actualised* in the outcome U_2 .

I will call the probabilities that support physical time flow in this way *dynamic physical probabilities*. The model developed in this chapter should show, at least, that such physical probabilities are possible.⁵ A further step is to argue that the probabilities that arise in quantum theory can be interpreted as dynamic physical probabilities. Two important facts supporting

⁵Note that I certainly do not maintain that all probabilities are dynamic physical probabilities. On the contrary, most probabilities referred to in everyday life seem to me to be merely epistemic probabilities.

this interpretation are: (i) The conclusion of the previous chapter that the fundamental quantum probabilities are future-directed. (ii) The fact that, in the orthodox interpretation of quantum theory, the full quantum state of a system provides maximal present information about the probabilities of future results of measurements.⁶

However, although I believe that this interpretation of quantum probabilities will appear sensible enough, the subject cannot be resolved before other foundational problems afflicting quantum theory are solved, so I will not attempt to offer conclusive arguments on the subject. See Chapter Eight for some further comments.

The general framework for the dynamic model should now be clear. In the following section the key concepts and features of the model will be discussed in more detail. These will be illustrated largely by examining the key similarities and differences between McCall's [1976] theory and my own.

--

⁶For without (ii) there is no possibility of interpreting the conditional probability law: $\text{PROB}(s_1|s_0) = p$ as generating an absolute, single-case probability: $\text{PROB}(s_1) = p$ from the satisfaction of the condition s_0 . I must thank Peter Milne for stressing this point to me. See later in this chapter, and also Milne [1986].

7.2 McCall's dynamic model.

Let us first of all return to the idea of what a theory or model of time is about. The fundamental point, as stressed in Chapter One, is that a theory of time is a theory of the nature of *existence*. This point is largely neglected in the scientific approach to time, where the concerns are primarily with metrical or topological features of time (or space-time). These are genuine concerns of course, but they are only a part of the story. It may appear to many scientists that the metrical or topological aspects of time are the real scientific concerns, whereas the theory of existence is properly a philosophical or metaphysical concern. This division between scientific and 'metaphysical' questions should be rejected, since the task of science is to construct good models of reality, and the theoretical framework for *existence* is a genuine part of this model, albeit one of its more abstract reaches.

There are two main historical rivals in the Western tradition for a theory of existence, which have come down from the time of Heraclitus and Parmenides: the *dynamic* view, and the *static* or *bloc universe* view. Since the dynamic view is only properly understood in relation to the bloc universe view, we will begin by considering the main features of the latter.

The bloc universe view is that there is but a single category of existence. This an 'eternal' category, which never changes in content, and includes all events that in normal language we would say have occurred in the completed history of the world. i.e. all

the events that are occurring, have occurred, and will occur. This class is picturesquely represented by a Minkowski diagram. In the bloc universe, there is no special moment which is *present* or *now*. The bloc universe is just a collection of events, which are spread out in time in a way which is ontologically no different to the way in which events are normally thought to be spread out in space. This view is appropriately said to *spatialise* time, since the temporal dimension is treated as ontologically the same as a spatial dimension. Any distinction between time and space is found only in the functional roles they play in the laws of nature.

In the bloc universe, objects are temporally extended just as they may be spatially extended. The physical body of a bloc universe person, for instance, is extended through his or her lifetime. A certain bloc universe body might stretch from 1920A.D. to 2000A.D., and would be 80 years long in the temporal dimension in the same way as it might be an axe-handle wide in a spatial dimension.

In the bloc universe there is no 'real change': there are only functional relations between different variables or properties. For instance, a physical body might 'change' its position with respect to time in a certain way. This just means that there is a certain kind of functional relationship between position and time for the four-dimensional wormlike object in question. In the same sense, snakes can be said to change their longitudes with respect to their latitudes, or birds to change their plumage with respect to their species.

Equally, there is no such thing as past existence or future existence in the bloc universe. There are just the temporal directions or relations, earlier than and later than. If our own world were a bloc universe, then it would be false to say at the time of writing that *the sinking of the Titanic is past*; all that could be truly said is something like: *the sinking of the Titanic occurs earlier than the moment of writing*. The latter statement does not imply the reality of temporal modes of existence, only of temporal relations between events.

This view of time gives an ontological framework that is very simple and seems logically convenient for many scientific purposes, and as noted in the Introduction, it is rather popular with scientifically-minded philosophers. But it is hardly the view assumed in our practical understanding of the world. Our practical view is that time does not simply exist like other things: rather it is the arena of existence. This is brought out by saying that objects *persist through time*. We normally assume that there is a world of objects which presently exists, and that the state of this world changes in time. The aim here is to show that a formal model which treats "existence as dynamic is possible.

The fundamental idea behind the model is well emphasised in McCall [1976]. He firstly illustrates the static view of time by sketching a number of different kinds of bloc universes (his Models A, B and C). The essential feature of these is very simple: they do not suffer *change*. He then proposes a dynamic model (his Model D), which he describes as follows:

The dynamic feature of Theory D, which differentiates it from the other [static] theories, consists in the following. The complete state-description of the universe, i.e. the universe-tree, is different at different times. (p.343).

By the 'complete state description' (the universe-tree), McCall means to indicate *the complete entity that presently exists*. That this is 'different at different times' means that what presently exists *changes*. McCall emphasises that this *change* is something that simply has no reality in the bloc universe, and hence that the notion of change is a new primitive notion which is simply not available in the set of concepts adequate to describe the bloc universe.

That the present universe-tree undergoes *change* necessarily gives it relations to other universe-trees. Presuming that time neither begins nor ends, then if U_1 is the present universe-tree, there must be another universe-tree, U_0 , from which it has changed, and there must be another universe-tree again, U_2 , into which it will change. I will represent these relations as: U_0 *has become* U_1 , and U_1 *will become* U_2 . Given these relations, and that U_1 is the present universe-tree, then U_0 is a *past universe-tree*, while U_2 is a *future universe-tree*. The reality of *change* generates 'has become' and 'will become' relations between the present universe-tree and other universe-trees, and these other universe-trees comprise the real past and future. These other universe-trees are *past* or *future* in the sense of having past or future *modes of existence*. The reality of change means that the present universe-tree must be located in a temporal sequence of

universe-trees.

Given this basic framework, there are three main parts left to fill in to generate a specific model. Firstly, a specific theory of the internal structure of the universe-trees is needed. Following this, a theory of what *change* amounts to, and further details of how it is to be formally modelled, is needed. Thirdly, the way in which the model provides for a semantics of temporal language must be sketched before the model has any practical application. I will quickly review McCall's position on these three points, and I will then turn to a number of problems with his theory.

(i) The nature of the universe-tree. McCall describes the universe-tree as having "the form of a tree, the space-time manifold which forms the trunk containing all past and present events (relative to the time in question), and the branches representing all physically possible courses of future events." (p.342). The most important points about McCall's universe-trees are as follows. First they are temporally extended - they contain elements at each moment of time. Secondly, a privileged moment within the universe-tree, the 'present moment', is defined. It is the latest moment before branching occurs. Thirdly, the branching is confined to the future, and represents the existence of physically real future *possibilities*, as befits a universe with a presently indeterminate future. The past in contrast does not branch, being presently determinate. Fourthly, and most importantly, the things that exist at all the different times in the universe-tree are fully *concrete* objects, events,

states, or whatever. This makes each McCall universe-tree a bloc universe, although of a somewhat unusual kind, since they have a complex topological structure of 'branches' of histories at times later than a certain moment.

(ii) Change. One such universe-tree comprises all that presently exists: this continually *changes*. The mechanism for change is simple enough. Change occurs through the process of one of the range of future possibilities becoming an actuality. Hence the notion of the *actualisation of possibilities* is taken as a primitive. This actualisation of possibilities corresponds to a progressive 'pruning' of branches of the universe-tree, giving rise to new universe-trees. This 'pruning' represents the disappearance of *future possibilities* from present reality, and at the same time it represents a progression of the *present moment* forwards in time (since the present moment is defined as the moment before the universe-tree begins to branch.)

It must be noted that this theory of change is quite substantial. What it determines in particular is a certain logical structure for the *temporal sequence of universe-trees* that change gives rise to. This can be seen by considering other mechanisms of change that might have been postulated instead. For instance an alternative theory might postulate that change in a given universe-tree gives rise not to a single new universe-tree, but to two new universe-trees (i.e. that two future possibilities rather than just one are always actualised). This would give rise to a 'branching' temporal sequence of universe-trees, rather than the usual linear temporal sequence.⁷ I am not

suggesting this branching model as a serious possibility, only pointing out that it would give another possible kind of model. This alternative possibility shows that McCall's rules governing change have substantial content.

McCall's rules for change determine that the temporal sequence of universe-trees is linear, and also that each later tree in the sequence is an object which literally forms a part of each earlier tree. This structure of temporal sequence of past and future universe-trees is depicted below.

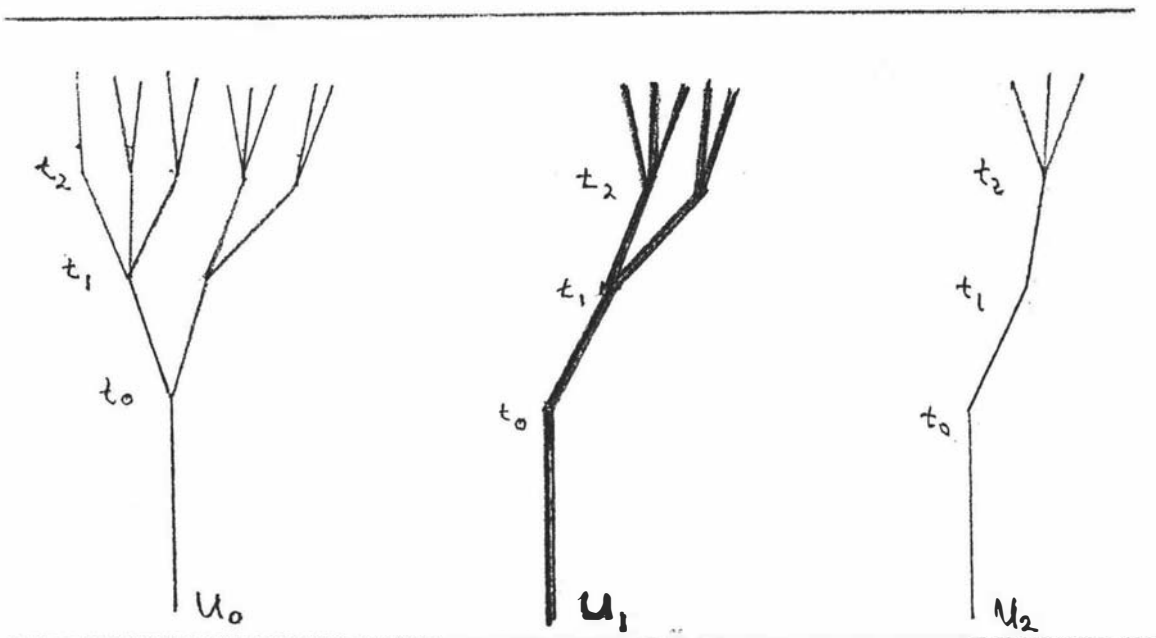


FIG 7.1

A temporal sequence of universe-trees.

⁷There would be no problems combining this idea of 'multiplying futures' with McCall's postulate of the determinateness of the past since each individual universe-tree would have a proliferation of futures in its temporal sequence, but a unique temporal sequence of past universe-trees.

For convenience I will henceforth use the terms U , U_0 , U_1 , U_2, \dots for universe-trees, and I will represent the type of temporal sequence depicted above as: $\dots U_0 \rightarrow U_1 \rightarrow U_2 \rightarrow \dots$. It is important to note that in the temporal sequence there is a special universe-tree picked out as the *present universe-tree*.⁸ The part of the temporal sequence earlier than the present universe-tree comprises the *past*, while the part of the temporal sequence later than the present universe-tree is the *future*. It is realism about change that requires realism about past and future.

(iii) Semantics. McCall proposes the following semantics for past- and future-tensed statements. Let t be the moment in the present universe-tree which is the present moment (defined to be the latest moment before branching occurs in the tree.) Then the future-tensed statement: *It will be the case that P* is true just in case P holds at some time in every branch of the universe-tree which is later than t . The branching of the universe-tree means that such future-tensed propositions can fail to take truth-values, for it can simultaneously fail to be the case that P holds in every later branch, and fail to be the case that not- P holds in every branch. Hence McCall's semantics allows future-tensed statements to fail the classical law of the excluded middle.⁹

The past-tensed statement: *it was the case that P* is taken to

⁸It will be pointed out later that McCall does not properly recognise the temporal sequence of universe-trees.

⁹An earlier semantics along these lines is provided by Thomason [1970]. See Fitzgerald [1985] and Yourgrau [1985] for some pertinent criticism.

be true just in case P holds at some time in the present universe-tree earlier than t. The past-tensed statements obey the law of the excluded middle.

This is the essence of McCall's theory. In following sections I will consider a number of key problems with it. These problems may seem rather diverse at first, but in fact they all have a single common source: a systematic confusion by McCall of the notions of past and future universe-trees with the notion of the earlier and later parts of the present universe-tree. Let us therefore begin by making this distinction clear.

7.3 Past/future in the temporal sequence versus earlier/later in the universe-tree.

The most notable feature of McCall's model is that time plays two different kinds of role in it. Firstly, it exists within the object that presently exists, the universe-tree, because the universe-tree is temporally extended. Secondly, it is represented in the temporal sequence of universe-trees. In the first role, time is a dimension much like space - McCall's present universe-tree is actually a certain kind of bloc universe (with a complex branching structure of space, or space-time). In the second role, time appears as *change*. Time thus has a kind of double aspect in this model, and one of the main problems is to explain how these two aspects can be thought of as combining into the single thing, 'time'. Unfortunately McCall does not discuss this problem,

because he doesn't recognise it: he is systematically confused about the distinction in question. His confusion on this point is fundamental, and leads his treatment astray on key points.

The temporal sequence of universe-trees represents the temporal modalities of existence, i.e. past, present and future. In contrast, the earlier and later parts of the present universe-tree represent *presently existing* facts about earlier and later times.¹⁰ The mistake McCall makes is to misinterpret the earlier and later parts of the present universe-tree as the past and future. For instance, he continually calls the 'trunk' of the universe-tree *the past*. But it is not the past, in the sense of being that which has *past existence*. What has past existence are *past universe-trees*. McCall similarly calls the later part of the universe-tree *the future*, but it is not the future. The future is the class of future universe-trees.

This confusion has dire consequences for most of the rest of his theory. The main faults in his treatment, which will be discussed in detail in following sections, are these: (i) McCall's semantics for tensed statements is wrong, because he takes tensed statements to be about the earlier and later parts of the present universe-tree, whereas they should be interpreted as

¹⁰Earlier and later are relations between moments of time; past, present and future are modes of existence. Relations of earlier and later are what constitute McTaggart's 'B-series', while attributes of pastness, presentness, or futurity characterize elements of his 'A-series'. The temporal sequence of universe-trees thus corresponds to the A-series, while the temporal sequence of facts that comprise the present universe-tree correspond to a B-series.

statements about past and future universe-trees. (ii) McCall fails to explain the reason for making the present universe-tree a temporally extended object. *Prima facie* this is an unexpected feature of the universe-tree, because since the present universe-tree is just what exists presently, it is naturally imagined that it should be a momentary object. (It seems at first sight that it should comprise just the world-wide state at the present instant). In fact, as will be argued in Section 7.9, the temporal extension of the universe-tree is a decisively successful feature of the model: but of course there is no value in having this feature unless the point of it is utilized. (iii) McCall's theory of the internal structure and content of universe-trees will be shown to be badly mistaken. The mistake stems from a misunderstanding of what the earlier and later parts of universe-trees represent. They do not represent the actual past and future, as McCall assumes, but only what is presently physically determined about the past and future. (iv) McCall fails to give a coherent account of the temporal modalities of existence, because he does not recognise how they must be represented in the model. This means that his theory is incomplete. (v) McCall also fails to give an adequate account of the nature of *change* which gives rise to the relations between universe-trees in the temporal sequence.

These points will be elaborated in detail in following sections. As a preliminary to the discussion, it will be necessary to clarify some further points about the notion of *temporal sequences of universe-trees*, since this notion plays a central role in the semantics of tensed propositions.

7.4 Temporal sequences of universe-trees.

The notion of a temporal sequence of universe-trees is implied by the notion of a presently existing universe-tree which suffers real change, as discussed above. Temporal sequences will be symbolised in the form: $(...U_0 \rightarrow U_1 \rightarrow U_2 \rightarrow \dots)$ ¹¹. This represents that U_1 is the universe-tree in the sequence which is *present*, while the sub-sequence of universe-trees: $(... \rightarrow U_0)$ preceding U_1 is *past*, and the sub-sequence: $(U_2 \rightarrow \dots)$ following U_1 is *future*.

The first important point to note is that a temporal sequence is composed of parts of three different ontological kinds: a past part, a present part, and a future part. This is very important in appreciating the kind of *existence* which a temporal sequence has. In particular, a temporal sequence does not simply have *present existence*. It has a mixed mode of existence, being partly past, partly present, partly future. In a formal model, this will be reflected by the fact that the logical type of the temporal sequence is a certain mixed type.¹²

It is important to realise that the temporal series has this

¹¹The model can easily be generalised so that the temporal sequence is continuous rather than discrete. I presume a discrete sequence only because it is easy to represent.

¹²The classical logician may well balk at this idea of three different types of existence: but to do so is just to reject the notion of the dynamic ontology out of hand, since it is founded on introducing these three temporal modalities or types of existence as primitive concepts.

mixed mode of existence, since the following kind of argument is sometimes put forward when this feature is overlooked. It is first of all observed that the dynamic model requires that the temporal sequence of U_i 's exists. This is assumed to mean that it exists *presently*, which can be so only if each U_i presently exists. But this immediately leads to incoherency in the model - for instance it means that U_3 above is both *future* (by virtue of being in the future of the presently existing U_2) and *present* (by virtue of existing as a part of the sequence). It is concluded that the dynamic model is incoherent.

However this argument is flawed, because although the sequence of U_i 's *exists* in a clear sense, it does not presently exist, but has a mixed mode of existence. When the mixture of modes is taken into account, no inconsistency arises.

The temporal sequence of universe-trees is generated by the *change* which occurs. The key problem which concerns us here is that this change in the present universe-tree also generates a kind of 'second-order' change in the temporal sequence of universe-trees. For clearly, as the present universe-tree changes, the temporal sequence of universe-trees changes as well, since the universe-tree which is *present* in the temporal sequence becomes a later tree. This second-order change is easy enough to depict. First-order change is depicted as simply: $\dots U_0 \rightarrow U_1 \rightarrow U_2 \rightarrow \dots$, i.e. first-order change is what the temporal sequence of universe-trees represents. The second-order change generated by this first-order change can be depicted:

$$\dots(\dots\mathbf{U}_0 \rightarrow U_1 \rightarrow U_2 \rightarrow \dots) \Rightarrow (\dots U_0 \rightarrow \mathbf{U}_1 \rightarrow U_2 \rightarrow \dots) \Rightarrow (\dots U_0 \rightarrow U_1 \rightarrow \mathbf{U}_2 \rightarrow \dots) \dots$$

Fig. 7.2

A temporal sequence of temporal sequences.

This depicts a temporal sequence of temporal sequences. To simplify the representation of temporal sequences, let us use T^0 to represent the sequence: $(\dots U_0 \rightarrow U_1 \rightarrow U_2 \rightarrow \dots)$, T^1 to represent: $(\dots U_0 \rightarrow U_1 \rightarrow U_2 \rightarrow \dots)$, and so on. Then Fig. 7.2 can be represented more concisely as:

$$\dots T^0 \Rightarrow T^1 \Rightarrow T^2 \dots$$

Here the bolding of T^1 (in Fig. 7.2, the shadowing of: $(\dots U_0 \rightarrow U_1 \rightarrow U_2 \rightarrow \dots)$) indicates that this temporal sequence is, in some sense of 'present', the *present temporal sequence*. Clearly if the notion of a temporal sequence of temporal sequences is to make sense, this notion of there being a *present temporal sequence* must be explicated: this is the main problem that will concern us. Before turning to that, note that this proliferation of temporal sequences continues indefinitely into higher and higher

orders. For we can define: $S^i = (... T^{i-1} \Rightarrow T^i \Rightarrow T^{i+1} ...)$, and construct the sequence: $...S^0 \Rightarrow S^1 \Rightarrow S^2...$, and so forth.

This infinite proliferation is not vicious in itself: the proliferation of higher-order logical objects from lower objects is normally inevitable in any logical scheme.¹³ But what badly needs to be explicated is the sense in which, in the second-order temporal sequences: $... T^0 \Rightarrow T^1 \Rightarrow T^2...$, the element T^1 is *present*. For it will be remembered that *being present* has so far just been taken as a modality of existence of individual universe-trees: clearly the temporal sequence T^1 cannot be present in this primary sense, since it is not a universe-tree. In fact, as we have already seen, since the temporal sequence T_1 is in fact an object which is part past, part present and part future, in the primary sense of those terms, it cannot possibly be said to be present in that sense. So how is the notion of the 'present temporal sequence' to be understood?

This problem is solved defining the appropriate second concept of present, call it *present*₂, from the first-order concepts of past, future and present. Present₂ may be defined as follows.

[7.1] The temporal sequence: $T^1 = (...U_0 \rightarrow U_1 \rightarrow U_2 \rightarrow ...)$ is *present*₂ just in case the partial sequence: $(... \rightarrow U_0)$ is past, the universe-tree U_1 is present, and the partial sequence: $(U_2 \rightarrow ...)$ is future.

¹³E.g. any sequence: $\langle 0_0, 0_1, 0_2, ... \rangle$ allows the definition of a higher order sequence-of-sequences: $\langle \langle 0_1, 0_2, ... \rangle, \langle 0_0, 0_2, ... \rangle, \langle 0_0, 0_1, ... \rangle, ... \rangle$, and so on indefinitely.

Appropriate second-order concepts of *past*₂ and *future*₂, can be similarly defined: the temporal sequence: $T^0 = (...U_{-1} \rightarrow U_0 \rightarrow U_1 \rightarrow ...)$ is *past*₂ (as it is in Fig. 7.2) just in case there is some $i > 0$ such that the partial sequence: $(... \rightarrow U_{i-1})$ is past, U_i is present, and the partial sequence: $(U_{i+1} \rightarrow ...)$ is future. Similarly, the temporal sequence: $T^2 = (...U_1 \rightarrow U_2 \rightarrow U_3 \rightarrow ...)$ is *future*₂ (as it is in Fig. 7.2) just in case there is some $i < 2$ such that the partial sequence: $(U_{i+1} \rightarrow ...)$ is future, U_i is present, and the partial sequence: $(... \rightarrow U_{i-1})$ is past.

These definitions show that the second-order concepts of past, present, and future, are reducible to the primitive first-order concepts. This is crucial, since it means that we do not require an endless proliferation of primitive concepts in the model: the first-order concepts of past, present and future suffice.

Clearly this result extends to all the higher-order concepts of change. The third-order concepts, *present*₃, *past*₃, and *future*₃ may be defined from *present*₂, *past*₂, and *future*₂ in exactly the same way as the latter have been defined from past, present and future. This shows how in general the n -th order concepts *present* _{n} , *past* _{n} , and *future* _{n} are ultimately reducible to the first-order concepts of past, present and future.

The proliferation of temporal sequences does not therefore entail a disastrous proliferation of primitive notions. In the following sections, I will refer not only to past, present or future universe-trees, but sometimes also to past, present or future

temporal sequences of universe-trees: the discussion here shows that this is legitimate.

7.5 Semantics of tensed propositions.

Tensed propositions will be symbolised using the tense operators P, N and F, defined as follows:

P(P) means that *it was the case that P*

N(P) means that *it is now the case that P*

F(P) means that *it will be the case that P*

The general framework of the dynamic model clearly means that the semantics for these tensed propositions must be as follows:

[7.2] Semantics of tensed propositions.

It was the case that P is true just in case there was at least one past universe-tree in which P is true.

It is now the case that P is true just in case P is true in the present universe-tree.

It will be the case that P is true just in case there will be at least one future universe-tree in which P is true.

It will also be useful to introduce a more general class of modal operators, T_t , where t is an index over times. The interpretation of these operators is:

[7.3] $T_t(P)$ means that *the proposition P was true, is true, or will be true, when t is the present time.*

The semantics for T_t is given by:

[7.4] $T_t(P)$ is true just in case P is true in the universe-tree in which t is the present moment.

This construal of semantics is very different from McCall's. For McCall takes the tensed propositions $P(P)$ and $F(P)$ to be propositions about *earlier and later parts of the present universe-tree*. But this is a misunderstanding of the model. McCall makes this mistake because he does not properly recognise the existence of past and future universe-trees in the first place. Only intermittently does he recognise the existence of *temporal sequences of universe-trees*. On page 343 he depicts part of such a sequence in the process of explaining how change is represented by his model. But in other places, particularly his treatment of the semantics of tensed statements, he fails to recognise the existence of past and future universe-trees, instead calling the earlier and later parts of the present universe-tree the past and future. This inconsistency in McCall's treatment is brought out by the following comments of Smart [1986]. Smart says:

McCall's picture suggests to me that there is a super-universe which is like a pack of ...playing cards, one above the other, cards higher in the pack

portraying a longer unbranched 'trunk' than those lower in the pack... In correspondence McCall has kindly commented on my interpretation of him, and has said that in one important respect it misses the spirit of what he intended to convey. He holds that the universe at a time t is not a slice of some super-universe, something analogous to a card in a deck. The universe at time t consists of just the universe at t , and the universe at t' ... does not exist at all. My worry is this: if the universe *now* is an entity, how can the universe at some other time be a non-entity? After all, McCall seems to be able to say things about it." [Smart, 1986 p.83]

Presuming that Smart has given an accurate summary of McCall's reply, his worry is entirely justified: how can McCall be a realist about the past and yet hold that the past "does not exist at all"? McCall has in fact given quite the wrong reply to Smart's original question, because he has failed to appreciate the need for the existence of temporal sequences in his model. The proper response to Smart's proposed 'super-universe' is to admit that it does indeed exist, although it is not quite as Smart imagines it to be. The 'super-universe' is the *present temporal sequence of universe-trees*: but as described in the previous section, this is an object which genuinely changes. It is not an atemporal bloc universe, as Smart presumes. The recognition of such higher-order objects which suffer higher-order versions of change, as outlined in the previous section, is obviously essential to understanding the dynamic model.

Having concluded that McCall's construal of the semantics for tensed statements is wrong, let us return to our own account. It

is seen from [7.2] or [7.4] that tensed propositions $P(P)$, $N(P)$, $F(P)$, and $T_t(P)$ can be logically identified with *classes of temporal sequences of universe-trees*. For instance, $P(P)$ can be regarded as the class of temporal sequences in which P is true in some past universe-tree. $T_t(P)$ can be regarded as the class of temporal sequences in which P is true in the universe-tree which has t as its the present moment.

Two important notions are presumed in this semantic theory. First, the notion of the *present moment in a universe-tree*. Secondly, the notion of *a proposition P being true in a universe-tree*. These will be considered in the following sections.

7.5 The present moment.

The concept of *the present moment in a universe-tree* has been referred to repeatedly above. The idea is that each universe-tree in the temporal sequence represents what the state of the world was, is, or will be, *at a moment*. The moment in question has the special status *within* that universe-tree of being *the present moment*. I will begin by reviewing some elementary formal features of this concept, and I will then turn to the specific interpretation of it that McCall's theory offers us.

In the first place, each universe-tree must have exactly one present moment:

[7.6] For each universe-tree there is exactly one moment which is *the present moment* in that universe-tree.

There will generally be many possible universe-trees which have the same present moment, hence there must exist a many-to-one function from universe-trees to moments which picks out the present moment in each universe-tree. I will call this the *present-moment function*, and symbolise it by Π .

[7.7] $\Pi(U) = t$ just in case t is the present moment in U .

Moments of time have a primary ordering represented by the *earlier than* relation. This ordering (the 'B-series' ordering) is independent of the flow of time, since it characterises time in the bloc universe as well as in the dynamic universe.¹⁴ The universe-trees in the temporal sequence are also ordered by the relations of *change* (such as: U_0 became U_1 , and U_1 will become U_2). Given that there is a *present moment* defined for each universe-tree, the ordering of the temporal sequence generates a second ordering of moments of time in an obvious way. This second ordering is normally required to correspond to the 'earlier than' ordering, in the sense of the following condition:

¹⁴ The *earlier than* relation orders the moments of time as they exist *within* the temporally-extended universe-trees.

[7.8] The temporal series of universe-trees must be ordered in such a way that U_1 is in the past of U_2 only if the present moment of U_1 is earlier than the present moment of U_2 .

This should be regarded as a logical condition on the interpretation of *the present moment*. One of its implications is that, as change occurs, the present moment becomes later and later.

Another obvious formal condition is that any temporal sequence of universe-trees be complete in the sense of including one (and only one) universe-tree for each moment of time:

[7.9] For every moment t , there is exactly one past, present, or future universe-tree in which t is the present moment.

The primary concepts of past, present and future apply not to moments of time, but to universe-trees in temporal sequences: the further concepts of *past moments* and *future moments* can be defined as follows:

[7.10] The present moment of each past (future) universe-tree is a *past (future) moment*. Or more formally: If U is past (future) and $\Pi(U) = t$, then t is a past (future) moment.

Notice that while the *earlier than* relations among moments are

eternal facts, whether a moment has the property of *being a future moment* or *being a past moment* or not can change. For as time passes, future moments cease to be future, and moments that were not past become past. Note also that by [7.8] and [7.10], every moment later than the present moment is a future moment, and every moment earlier than the present moment is a past moment.

What has been said so far indicates some rather elementary formal properties of *the present moment*. Any adequate interpretation must satisfy these requirements. Let us now consider whether the interpretation which McCall's theory offers is adequate.

McCall's interpretation of *the present moment* is a part of his larger theory of the nature of the universe-tree, which was sketched in Section 7.3. Briefly, McCall's universe-trees have a 'trunk' of *earlier* events or facts, an instantaneous world-wide state at the *present moment* within the tree, and a branching structure of *later* events. The present moment is defined as the latest moment in the universe-tree before branching begins.

This interpretation satisfies the formal requirements represented by [7.6] - [7.10] above. The structural feature of 'being the latest moment before branching' is objective enough, and it objectively picks out a unique moment in each universe-tree. Furthermore, McCall's rules for change means that his present moment changes in the appropriate way as the universe-tree changes, i.e. it becomes later as we move through into the

future. But although it satisfies these formal requirements, it fails to satisfy another requirement which is vital for an adequate interpretation. I will call this the requirement of the *ontological specialness of the present*.

The idea of ontological specialness is roughly that the present moment, or more accurately, what exists at the present moment in the universe-tree, must be 'real' in a way that what exists at other moments is not real. The point is that the earlier and later parts of McCall's universe-trees have been made *too real*, and this robs the present part of his universe-tree of the uniqueness it should have.

What exists in the earlier and later parts of McCall's trees are events which are as fully concrete as those that exist in the present part.¹⁵ This makes each universe-tree a full-blooded bloc universe (with a single spatial 'branch' of events up to a certain moment, and a number of spatially distinct branches at times later than that moment). This overdose of reality means that the present moment within the universe-tree lacks any special ontological status. Consider what the *present moment within the universe-tree* is supposed to be. It is supposed to be something quite ontologically special: in an intuitive sense, it should pick out the most 'real' part of the universe-tree. The nature of this ontological specialness is brought out by considering the idea of 'telling the time'. Within a given universe-

¹⁵The problem is worse, not better, for the later branched part of the trees, where there are numerous *incompatible* events which are all supposed to be fully real.

tree, there should be a single moment which is *what the time presently is* when that universe-tree has present existence. When inhabitants of a universe-tree are looking at an accurate clock, for instance, the clock ideally should tell them what this unique *present time* is. But this is not so in McCall's universe-trees, because in his trees there are fully real people at all sorts of different temporal locations, whose clocks give them all sorts of different perceptions of the time. For instance, in McCall's model, a fully real Napoleon exists at the first moment of 1800 in the *present universe-tree*: this Napoleon will be quite mistaken about what the present time is.

McCall's definition of the present moment within his universe-tree has some good features: it is objective, and it changes in the appropriate way as change occurs to the universe-tree. But these *formal* features do not in themselves make this interpretation of the present moment an adequate one. We can see this by noting that the moment picked out by a second structural feature, *being the moment two hours earlier than the first branch*, is equally objective, and also changes in the appropriate way. But clearly it cannot count as the present moment as well.

This problem with McCall's present moment points to a fundamental inadequacy in his conception of the universe-tree. The problem is a very useful one to focus on, for it must be faced by any conception of the universe-tree: *how is the universe-tree to be conceived so as to make the present moment ontologically special enough?* I will now suggest a modification of McCall's

theory which solves this problem. The modification gives a completely different theory of the universe-tree, and it seems to be the only plausible kind of theory there can be (it will later be generalised when I present my own theory of the *probabilistic universe-tree*).

7.6 Modification of McCall's theory: the possibilistic universe-tree.

The modified theory is very simple, given by the following two postulates: (i) The present part of the universe-tree is the concrete instantaneous world-wide state at the present moment of time (just as in McCall's theory). (ii) The earlier and later parts of the universe-tree consist of the *possibilities* of past and future events that are entailed by the present part of the universe-tree and the laws of nature.

The first postulate means that the present part of the universe-tree is a fully real, concrete entity (as in McCall's model.) This will provide for such physical things as people and clocks existing at the present moment, and for these people or clocks having a unique 'present time'. This gives the present moment the requisite 'concreteness' or 'reality'.

The second postulate means that the earlier and later parts of the universe-tree do not contain concrete events, but merely 'possibilities' of events. These rather ghostly 'possibilities' allow the present part of the universe-tree to remain ontologically special, because they do not compete with its 'reality'. For instance, if the present moment in the present universe-tree is t ,

then there can only be people *actually* telling the time at the moment t : at moments other than t there can presently be only possibilities of such events as people telling the time.

This sounds nice enough, but it is hardly satisfactory unless a precise theory of what the *possibilities* being talked about here can be given. There are two useful approaches to understanding these possibilities. The first is to consider what the purpose of having them in the model is: by considering the kind of role they are required to play we will see what kind of things they represent. This leads on to the second approach, which is to identify the *logical type* of things these possibilities are.

A few words on the subject of logical type are in order. Understanding the logical type of an object is very important: it should be compared with understanding the physical dimensions of a physical quantity, which is so necessary for the precise understanding of scientific concepts. For instance, in classical physics the concept of *energy* was only clearly understood when it was realised that energy is a quantity with the dimensions: $(\text{mass} \cdot \text{distance}^2 \cdot \text{time}^{-2})$. The problem of identifying the dimensions of a physical quantity in this way is in fact *exactly* the problem of identifying its logical type, as is immediately apparent when the ontology of classical physics is formally analysed. Therefore, if *possibilities* are considered to be physically real, as they are here, then identifying their logical type should be seen as identifying the kind of *physical entities* that they are. This is the key step to viewing possibilities as

physically real (as it was to viewing energy, momentum, etc, as physically real). Once it is achieved, the claim that possibilities can be understood as physically real 'entities' becomes respectable. It will be no surprise to the logician that possibilities (and as we will later see, *probabilities*) turn out to be rather high-order entities. Many philosophers refuse to contemplate the reality of any but first-order entities, so there will be resistance in many philosophical quarters to the very idea of physically real higher-order entities such as possibilities (or probabilities).¹⁶ But it seems to me that a great lesson of modern science is that if you can make a coherent mathematical or formal model of something, then that thing is possible and imaginable, regardless of any 'philosophical' prejudices against it. Hence I offer no apology for treating higher-order entities as physically real.

The discussion of semantics in the next section will show clearly what the role of possibilities in the present model is. This will prepare the ground for the formal analysis of the logical

¹⁶I am taking particular objects to be first-order entities. Properties such as 'having mass such-and-such', 'having electric charge such-and such', are of the same order as *classes* of first-order entities, hence second-order. Realism about these therefore requires admitting second-order entities as *primitive in nature*. The nominalist balks at this (for reasons that are a mystery to me), but does not offer, to my knowledge, any successful alternative way of understanding physical theories realistically. (Even Quine [1981, p.182] disavows a nominalistic view of physics for this reason). It will be seen that realism about possibilities and probabilities goes a step further still, involving recognising certain *third-order* entities as primitive in nature.

type of possibilities, which will be presented in a later section again. For the moment the reader is asked to accept these *possibilities* as primitive entities. These possibilities are to be thought of as *possibilities of propositions obtaining at times*. They will be symbolically represented in the form: $\diamond(P,t)$, which has the following interpretation:

[7.12] $\diamond(P,t)$ means that it is physically possible that the proposition P holds at time t .

The idea that such possibilities exist within the universe-tree really just means that such possibilities are determined *relative to individual universe-trees*. That is to say, relative to a universe-tree U , the proposition that: $\diamond(P,t)$ takes a truth-value. Hence they are unlike the tensed propositions, $P(P)$, $T_t(P)$, etc, which we have seen have truth-values only relative to temporal sequences of universe-trees. This will be examined in detail in the next section, where we return to the topic of the semantics of propositions about universe-trees. To conclude this section I will point out some implications of the present model.

The physical possibilities can also give rise to physical *necessities*. For if both: $\diamond(P,t)$ and: $\text{not-}\diamond(\text{not-}P,t)$ hold¹⁷, then it is not only physically possible, but physically *necessary*, that P holds at t . This necessity will be symbolised: (P,t) , and is interdefinable with possibility:

¹⁷The negation a proposition P will be symbolised: $\text{not-}P$. Standard truth-functional semantics for logical connectives (not- , and , or) is assumed.

[7.13] (P,t) just in case: $(\diamond(P,t)$ and $\text{not-}\diamond(\text{not-}P,t))$.

and:

[7.14] $\diamond(P,t)$ just in case: $\text{not-}(\text{not-}P,t)$.

Since possibility and necessity are interdefinable, either might be taken as primitive; possibility has been assumed as the primitive because this is generally more convenient.

The possibilities contained in the universe-tree are just those fixed by the instantaneous state at the present moment plus the laws of nature. Hence the *present part* of any universe-tree determines, via the laws of nature, the whole of the universe-tree. Equally, specifying the class of nomologically possible universe-trees is to specify the laws of nature.

Different types of natural laws will be reflected in different kinds of universe-trees. For instance, (i) where the laws of nature are deterministic towards the future, universe-trees will contain no 'branching' of possibilities at times later than the present moment. That is to say, for any t later than the present moment, and any P , it will either be the case that: $(\diamond(P,t)$ and $\text{not-}\diamond(\text{not-}P,t))$, or it will be the case that: $(\diamond(\text{not-}P,t)$ and $\text{not-}\diamond(P,t))$. (More concisely, either: (P,t) , or else: $(\text{not-}P,t)$).¹⁸

¹⁸Note that in the present model, 'branching' is not literally a branching of the space or space-time manifold of the universe-tree itself, as in McCall's model. The 'branching of possibilities' means nothing but the co-existence of: $\diamond(P,t)$ and:

(ii) Where the laws of nature are deterministic towards the past, there will similarly be no branching of possibilities in the earlier part of universe-trees. (iii) Indeterminism towards the future will require branching of possibilities in the later parts of at least some universe-trees. Indeterminism towards the past will require branching of possibilities in the earlier parts of at least some universe-trees.

We see therefore that the 'branching' of the universe-tree is something determined solely by formal aspects of the laws of nature, and this alerts us to another key mistake of McCall's. McCall specified that while there can be a branching of future possibilities, there can be no branching of past possibilities, and he justifies this claim as follows:

If possible futures are admitted as a part of complete state-descriptions, what reason, other than an arbitrary one, can be given for the exclusion of possible pasts? Answer: a metaphysical reason reflecting the common belief that the past is unique. [1976, p.349].

But given the distinction between the past/future of a universe-tree and earlier/later parts of a universe-tree, this is mistaken. The past (like the future) is 'unique' independently of the content of the *present* universe-tree. Its 'uniqueness' results just because the past is a definite temporal sequence of past universe-trees: this has nothing to do with the content of the earlier part of the present universe-tree. McCall's postulate about the earlier part of the present universe-tree guarantees only that

◊(not-P,t).

the present instantaneous state determines all past states – hence it is really an arbitrary postulate that the laws of nature are deterministic towards the past. It is entirely misplaced.

Having made it clear how different the present model of the universe-tree is from McCall's, let us turn to the semantics it supports.

7.7 The semantics for untensed propositions.

The semantics for tensed propositions, given in [7.2] and [7.4], clearly presupposes that there is a more fundamental class of propositions, namely the P's in: P(P), F(P), N(P), and $T_t(P)$. These will be called *untensed propositions*. Whereas the tensed propositions take truth-values with respect to temporal sequences, the more fundamental untensed propositions take truth-values with respect to individual universe-trees. Untensed propositions are clearly intended to be about facts that hold in individual universe-trees. The semantics of untensed propositions is fundamental, providing the foundation for the semantics of tensed propositions.

Untensed propositions are of two distinct kinds, corresponding to the two kinds of facts that hold in universe-trees.

(i) First are what will be called *concrete propositions*. These are about the facts which obtain in the present instantaneous part of the universe-tree (i.e. the 'concrete' part of the universe-tree). Some suggestive examples from natural language are: *it is*

raining, Monkey is drunk, Mr. Palmer is Prime Minister. Examples from physics might be: *system X has an electric charge of -1, or: the universe contains more than 10^9 electrons.* These are presently true or false according to whether or not the facts to which they refer obtain at the present moment in the present universe-tree. Notice that these propositions can generally change their truth-values as the universe-tree changes. For instance, Mr. Palmer was at first not the Prime Minister, and then he became the Prime Minister.

(ii) The second kind of untensed propositions will be called *time-indexed propositions*. These are generally about the facts which obtain in the earlier or later parts of the universe-tree. In the present model, these facts are just the *possibilities* described in the previous section, hence the time-indexed propositions are of the basic form: $\diamond(P,t)$. Note that these propositions also change their truth-values. For instance, it may have been possible last year that Mr. Palmer would not be the Prime Minister in January, 1990; this possibility is no longer actual.¹⁹

It is assumed that there is a class of *atomic* concrete propositions, which will be denoted: P, P_1, P_2, \dots : these give rise to the wider class of concrete propositions through the standard logical operations (truth-functional connectives, quantification).

¹⁹In a wider sense, it is *logically possible* (and undoubtedly, nomologically possible) that Mr. Palmer is not Prime Minister in 1990, since there are logically possible universe-trees (nomologically possible universe-trees) in which he is not Prime Minister in 1990. But these universe-trees are not physically compatible with the existence of the present actual universe-tree.

There is also a class of *atomic* time-indexed propositions, which are of the form: $\diamond(P,t)$, where P is any concrete proposition, and a wider class again is generated, from the joint class of these propositions and the full class of concrete propositions, by standard logical operations. This gives the full class of non-tensed propositions.

Each non-tensed proposition takes a truth-value in each universe-tree, hence a non-tensed proposition can be represented by the class of universe-trees in which it is true. I will not describe the framework for a formal semantics of this kind, since it may be easily inferred from well-known systems of formal semantics. I will also not discuss the concrete propositions, since these are also very familiar. What are interesting are the time-indexed propositions, $\diamond(P,t)$. The interpretation of these needs to be made clear.

Clearly: $\diamond(P,t)$ is intended to mean that *it is physically possible at present that the proposition P could be true when the present time is t* . When t is a past or future moment, $\diamond(P,t)$ therefore has implications about the past or future. I will try to bring out clearly what these implications are.

For brevity I will henceforth refer to temporal sequences of universe-trees as *histories*. The term *history* suggests an atemporal item (a bloc universe history), but of course, in our use of the term, a *history* is a temporal sequence with a *present* picked out, and consequently a past and a future as well. It is, so to speak, a 'bloc universe history' viewed from a particular

temporal perspective.

Two important classes of histories need to be defined: (i) Z is the full class of *nomologically possible histories*. The class Z is determined by the laws of nature. (ii) $Z^{P,t}$ is the subclass of histories in Z in which P is true at t (i.e., in which P is true in the universe-tree in which t is the present moment.)²⁰ $Z^{P,t}$ thus represents a particular *type of history*, namely, the type of history in which P *actually happens* at time t .

It is correct to read the proposition: $\diamond(P,t)$ as meaning that *it is presently possible*²¹ that the history of the universe is such that P actually happens at the time t . This can in turn be taken as attributing a certain property to the history-type: $Z^{P,t}$. Namely, the property that *it is presently possible that the history of the universe is of type $Z^{P,t}$* . This will be symbolised as: $POSS(Z^{P,t})$.

$POSS(\cdot)$ is a rather high-order property: a property of types of histories. Reanalysing: $\diamond(P,t)$ as: $POSS(Z^{P,t})$ makes clear the sense in which this proposition, although it is true or false of *individual universe-trees*, is at the same time *about the history of the universe-tree*, hence about the past or future.

It may seem that it is rather arbitrary to analyse: $\diamond(P,t)$ in this way. Why not, for instance, leave it as: $\diamond(P,t)$, and interpret:

²⁰For an example, suppose P is true in U_1 , but false in U_1^* . Let: $T^0 = (\dots U_0 \rightarrow U_1 \rightarrow U_2 \rightarrow \dots)$, and $T^{0*} = (\dots U_0 \rightarrow U_1^* \rightarrow U_2 \rightarrow \dots)$, and suppose that T^0 and T^{0*} are both in Z . Then $Z^{P,1}$ contains T^0 but does not contain T^{0*} .

²¹In the sense of *it is nomologically consistent with the present state of the universe*.

$\diamond(\cdot)$ simply as a primitive function which maps proposition-time couples, $((P,t)$'s), to propositions ($\diamond(P,t)$'s)? The point of reanalysing it as: $\text{POSS}(Z^P,t)$ is simply to render a certain feature of it transparent. The analysis as: $\text{POSS}(Z^P,t)$ is no more formally correct than the analysis as: $\diamond(P,t)$, but it is a more convenient representation for the purpose of perceiving what the proposition is about.²² It should be noted that whether we analyse the statement as $\diamond(P,t)$ or as $\text{POSS}(Z^P,t)$, we are committed to objects of the same order of logical complexity.²³ $\text{POSS}(\cdot)$ is a property of a class of histories, hence of the order of a class of classes of histories. $\diamond(\cdot)$ is a function from proposition-time couples to propositions; propositions are classes of universe-trees, hence of the same logical order as *histories*, hence $\diamond(\cdot)$ is of the logical order of a mapping from classes of histories to classes of histories, hence also of the order of a class of classes of histories.²⁴

The reason that: $\text{POSS}(Z^P,t)$ gives a more transparent analysis than: $\diamond(P,t)$ seems to be that we have no intuitive understanding

²²Compare with the analysis of kinetic energy as (i) $1/2 \cdot m \cdot v^2$, (ii) $p^2/2m$. These are formally equivalent, but they make different relations transparent. When we want to depict the dependence of energy on velocity and mass, we use (i); when we want to depict the dependence of energy on momentum and mass, we use (ii).

²³Compare this with the fact that: $1/2 \cdot m \cdot v^2$ and: $p^2/2m$ have the same physical dimensions.

²⁴Of course, the proposition $\diamond(P,t)$ considered as a logical entity is just of the much lower order of a class of universe-trees, like all the non-tensed propositions. What we are considering when we analyse the proposition $\diamond(P,t)$ is how it is *constructed* from parts. P. Tichý makes the idea of constructions clear in his [1988].

of what the function $\diamond(\cdot)$ represents, while we do have an intuitive understanding of what the property $\text{POSS}(\cdot)$ represents – namely, *the possibility that the history of the universe is of the type such-and-such*. The analysis: $\text{POSS}(Z^P, t)$ allows us to identify something as being '*the possibility*', namely, the 'possibility property', $\text{POSS}(\cdot)$. It is a possibility *of* something, namely, a type of history. This fits our natural language model reasonably closely. We cannot easily gain such an intuitive understanding of the function: $\diamond(\cdot)$: this is a function from proposition-times to propositions, but how on earth is it to be imagined?

In the following section a further reason for analysing possibility as $\text{POSS}(Z^P, t)$ is given: this is that it allows the simple idea of *possibility* to be fitted naturally into a far more complex system of concepts, the concepts of *probabilities*. (It will be seen why the plural is used). Viewing possibility in the context of probabilities will give us a far more systematic grasp of the subject.

It is seen, then, that possibilities are about past and future in the sense that they represent *properties of types of histories*. The curious thing is that these properties are actual in the present. I.e. types of histories *presently* have properties. These properties are just what the dynamic laws of nature imply. This is because those laws postulate 'connections' across time between the present instantaneous state and past and future states.

This approach to representing the physical ontology gives a unified and systematic view, for it means that laws of nature are represented through facts (of a rather high order) which obtain in the *present*. (There are some more comments on this in the final section.)

To conclude, note the relation between McCall's (or Thomason's [1970]) construal of past and future tensed propositions, and our time-indexed propositions. McCall construes: *P will be true in the future* as we would construe: *There is some future time t such that (P,t)*. This is not a tensed proposition but a time-indexed one: it is rendered more formally as: $(\exists t)(t > \Pi \ \& \ (P,t))$ (where Π is the function defined in [7.7] which maps universe-trees to their present moments, $>$ is the relation defined on moments with the universe-tree of *being future with respect to*, and \Box is necessity as defined in [7.13].) Similarly, on McCall's construal, *P was true in the past* is our time indexed proposition: *There is some past time t such that (P,t)*. (Formally: $(\exists t)(t < \Pi \ \& \ (P,t))$). Hence, what McCall calls past and future tensed propositions are really propositions about what is *presently necessary about the past and future*. McCall and Thomason [1970] fail to adequately justify their treatment of past and future tenses (see Fitzgerald [1985], Yourgrau [1985]), and given our present understanding of the subject, it seems that their treatment is quite unjustifiable.

It should also be noted that McCall's claim that the indeterministic model supports a non-classical logic, which fails the law of the excluded middle, is wrong. This failure was

supposedly based on the fact that, where a proposition P is neither determined to be true at a (future) time t , nor determined to be false at t , the proposition that P will be true at t has no truth-value. But this false, whether P will be true at t is construed as: $T_t(P)$ or (following McCall) as: (P,t) . For: $T_t(P)$ is true or false depending upon whether P is true of the universe-tree at t in the temporal sequence, and the fact that it is undetermined by the present universe-tree is irrelevant to this, while: (P,t) is simply false under this supposition, since it is not presently determined that P will true at t . The peculiarity is of course that both: $\text{not-}(P,t)$ and: $\text{not-}(\text{not-}P,t)$ can presently be true (this is really the feature that McCall's view that the 'indeterminist' logic is non-classical turns upon), but this does not imply the failure of the excluded middle.

7.8 The probabilistic universe-tree.

We have now arrived at a model very similar to that promised in Section 7.1: all that remains is to add appropriate *probabilities* to the ontology, over and above the simple *possibilities* that we already have. The *probabilistic universe-tree* consists of (i) a concrete physical universe in its present momentary state, plus (ii) the physical probabilities of future states, as determined by the present state plus the probabilistic laws of nature, plus (iii) the physical possibilities of past states, as determined by the present state plus the laws of nature.

The addition of probabilities adds a great deal more structure to the universe-tree. Propositions about probabilities will normally be symbolised: $PROB(P,t)=p$, which is read: *there is a probability of p that P will be true in the universe-tree at t* . The model now obviously supports non-tensed propositions of three basic kinds, viz. concrete propositions, P , and two kinds of basic time-indexed propositions: $\diamond(P,t)$ and: $PROB(P,t)=p$.

The problem that faces us is the interpretation of the *probabilities*. They will be analysed here as *properties of types of histories*, i.e. as being of exactly the same logical type as *possibilities*. It should be recognised that this analysis only determines the logical type of probabilities, i.e. the logical place they have in the ontological scheme. It does not determine the *content* of the properties that constitute probabilities, in the sense of specifying rules for assigning probabilities in practical situations. The latter problem has proved very difficult, and it is not required that it be solved here. A few comments will be made on it at the end of this section.

The analysis we already have of possibilities indicates what our analysis of probabilities will be, because possibilities are just a special case of probabilities, being definable in terms of probability thus:

$$[7.15] \quad \diamond(P,t) =_{df.} \text{PROB}(P,t) \neq 0.$$

$PROB(P,t)=p$ is naturally read as saying that *there is a*

probability p that the history of the universe is of a certain type, namely the type in which P is true at time t . This history-type is just that represented by: Z^P, t , and the probability can be represented as assigning a certain property to this history-type. Clearly every degree of probability, p , gives rise to a distinct such 'probability property'. (For the propositions: $\text{PROB}(P, t) = p$ and: $\text{PROB}(P, t) = q$, where $p \neq q$, are incompatible just because they assign two incompatible properties to the history-type Z^P, t .) The 'probability property' corresponding to the degree of probability p will be represented as: PROB_p (where p is a real index ranging from 0 to 1.) This analysis therefore allows us to rewrite: $\text{PROB}(P, t) = p$ as: $\text{PROB}_p(Z^P, t)$.

$$[7.16] \text{PROB}(P, t) = p \text{ =df. } \text{PROB}_p(Z^P, t).$$

The similarity with the analysis of possibilities as: $\text{POSS}(Z^P, t)$ is obvious. Indeed, [7.15] allows us to identify the possibility property, $\text{POSS}(\cdot)$, as a disjunction of probability properties: this disjunction is appropriately symbolised: $\text{PROB}_{\neq 0}$. $\text{PROB}_{\neq 0}$ is defined by:

$$[7.17] \text{PROB}_{\neq 0}(Z^P, t) \text{ =df. } (\exists p)(p \neq 0 \ \& \ \text{PROB}(P, t) = p).$$

It follows that:

$$[7.18] \text{POSS}(Z^P, t) \text{ =df. } \text{PROB}_{\neq 0}(Z^P, t).$$

The set of probabilistic facts obviously represents a far more complex structure than the set of possibilistic facts, since the latter is so simply modeled in it. But they are facts of the same logical type. On the analysis which represents them as: $\text{PROB}_p(Z^P, t)$, probabilistic facts consist of assignments of *probability properties*, from the continuous range of the PROB_p 's, to *types of histories*. The class of probability properties has a very rich structure, because the set of PROB_p 's form a continuum supporting the rich structure of interrelations represented by the axioms of probability theory. The structure which the set of probabilistic facts confers on the logical space is correspondingly rich. By comparison, the structure conferred by the much simpler set of facts about *possibilities* is very poor.²⁵

To conclude, I will make some comments on the interpretation of physical probability. Probabilities have been explicated as physically real (high-order) properties of *types of histories*. This only tells us their logical type, not what kinds of properties they are. That is, it does not determine the interpretation of the PROB_p 's. (Analogously, the fact that energy has the physical dimensions: $(\text{mass} \cdot \text{distance}^2 \cdot \text{time}^{-2})$ does not determine what energy is.)

Explicating what probabilities are requires spelling out the

²⁵The probabilities can be imagined as *possibilities* to which 'degrees' (between 0 and 1) have been assigned. It is from these 'degrees' that the richness of the structure of probabilities arises.

implications or connections they have with other areas.²⁶ There are two different types of implication: (i) implications internal to the theoretical ontology, (ii) implications for experience. Let us consider these in turn.

(i) In the first place probabilities have implications for other probabilities, e.g. if: $\text{PROB}(P,t) = p$, then: $\text{PROB}(\text{not-}P,t)=1-p$. It is necessary to understand these implications to some extent if one is to understand the nature of probabilities, but alone they provide only a certain *formal* structure for the set of probabilistic propositions or facts. There are also implications for other areas of the theoretical ontology; for instance, if it is presently true that: $\text{PROB}(P,t)=1$, this implies that: $T_t(P)$ (i.e. that the proposition P is true in the universe-tree in the present temporal sequence of universe-trees which has t as its present moment.) More generally, the laws of nature directly relate probabilities to propositions about the states of systems, by the laws of the form: $Q \Rightarrow \text{PROB}(P,t)=p$. Probabilities therefore play an integral part in the theoretical ontology, reflected by the role they play in structuring the nomologically possible universe-trees.

But this does not give an adequate understanding of how to *empirically judge facts about probabilities*. The laws: $Q \Rightarrow \text{PROB}(P,t)=p$ formulated in physical theory may provide a means of judging probabilities after the laws have been confirmed, but

²⁶A rationalist might claim that probabilities are grasped directly, by some sort of primitive rational intuition; but this would seem to mean that there is no way of *explicating* them at all.

how are the laws to be confirmed in the first place? To judge the accuracy of our hypotheses about probabilities obviously requires some grasp of the implications of probabilistic statements for *experience*. Hence we must accept that probabilistic statements have implications of the second kind, i.e. implications for experience.

(ii) Such implications for experience are widely recognised, and for practical purposes, well understood. For instance, if a theory predicts a single-case objective probability of 10^{-100} of an event E occurring, and in a solitary experiment E occurs, then the prediction of the theory would be dismissed as incorrect by any reasonable scientist. More generally, belief in propositions about probabilities have implications for the *degrees of expectation* we should rationally have. It is clear that probabilities are taken to have a powerful range of implications for observations that we make, and scientists generally agree objectively on these implications. The epistemology of probability may be a difficult subject, but nevertheless there are well-known procedures for confirming and disconfirming probabilistic statements. If this were not so, quantum theory would not have been confirmed. There is no reason to think that the concept of probability is any less respectable than any other concept in science.

7.9 The need for temporally extended universe-trees.

The temporal extension of the universe-tree is perhaps the most distinctive feature of the model; but why does the universe-tree need to be temporally extended? This is an important question which has been noted a few times in the discussion, but has not yet been properly discussed.

The temporal extension of the universe-tree at first seems a peculiar feature. Would it not be far more natural to have the universe-tree consisting of just the world-wide instantaneous state at the present moment, and nothing more? I will show here that this would be inadequate, and that the temporal extension of the universe-tree is necessary. The reason that it is necessary is to allow for the existence of *present facts about past and future times*. Such facts are required by the 'commonsense' view of the world as much as by the theoretical scientific view,²⁷ but I will only consider the situation for the scientific view here.

Dynamic physical laws entail, at the very least, conditional propositions about the possibilities of earlier and later states given present states. For instance a theory might entail that *given the total state of the universe is U at time t_0 , it is physically possible that the proposition P is true at t* . And it might also entail that: *given the total state of the universe is U at time t_0 , it is not physically possible that the proposition P^**

²⁷For instance, it is a present fact in the actual world that *New Zealand and Germany were at war in 1940*.

is true at t. Clearly a theory must entail that some processes are impossible, as well as that some are possible, if it is to have any notable content.

Suppose that: *given the total state of the universe is U at time t_0 , it is physically possible that the proposition P is true at t_1* is indeed a physical law; that the present time is t_0 , and t is a future time; and that the present state of the universe is indeed U. Then there will *presently* exist a real physical possibility that P will be true at t .

This possibility, being physically real, must be represented in the model of the physical ontology somehow. One way would be to represent it directly as a primitive fact: $\diamond(P,t)$, as done in the theory of the possibilistic universe-tree. In this case we automatically have a *temporally extended universe-tree*, since it contains a substantial fact about a moment later than the present. However, it might be thought the existence of the possibility might be represented in the model in some other way, which does not generate a temporally extended universe-tree. But it is easy to show that this cannot be done.

The feature we have noted means that natural laws make it a physical fact that in the universe-trees in a certain class (the class in which: $\diamond(P,t)$ is true) it is true that *it is possible that P holds when the time is t*. In the remaining class (represented by $\text{not-}\diamond(P,t)$), it is a physical fact that *it is not possible that P holds when the time is t*. Hence the laws provide an ontological basis for a partition of the space of nomologically universe-trees that corresponds to the proposition: $\diamond(P,t)$. It doesn't matter how

this partition is represented (by 'primitive facts' about probabilities, or in some round-about way), the partition is physically real, hence the propositions of the form: $\diamond(P,t)$ are supported by the physical ontology. And the existence of these propositions is all that is all that the 'temporal extension of the universe-tree' involves.

7.10 Some concluding comments.

(i) The representation of the laws of nature. In the dynamic theories that I have argued for, the laws of nature have existence *in the universe-trees*. They are represented by the internal structure of the universe-trees; this structure determines the possibilities and probabilities of past and future events. Strictly speaking, (a) the full set of the laws of nature are represented by the class of nomologically possible universe-trees; but it is also interesting that (b) the laws of nature *that can possibly come into play in the future of a given universe-tree* are fully represented in that universe-tree, (by the part of it which is later than the present moment). Hence, on this representation of physical worlds, the laws of nature that have any real application to a world at a moment are represented as *a part of what presently exists*. Laws of nature become facts that have present existence. It must be noted, however, that there is something about 'reality' that remains outside the set of present facts: this is the reality of change, or equivalently, of the past and the

future. Change, by its very nature, *cannot be represented as existing in the present*. However, certain features of the mechanism of change can be represented in the future: these features are what the temporal extension of the universe-tree represent. To represent change itself, however, the model must include more than just the *present* as real: it must also recognise a past and a future.

(ii) Change is introduced as a primitive concept. It must be recognised that the notion of *change* which the dynamic theory employs is introduced as a new *primitive* concept, which has no place in the static theory of existence.

The concept of *change* can be explicated in two equivalent ways, using two different sets of primitive concepts not available in the bloc universe: (a) the relations *became* and *will become*, or (b) the properties of *being past* and *being future*.²⁸ Taking universe-trees to be the entities that suffer change, these are interdefinable as follows:

U is past just in case U *became* the present universe.

U is future just in case the present universe *will become* U

(*U is present* just in case U is the present universe).

²⁸The concept of 'present existence' is also required, but it is required in the bloc universe too. What has 'present existence' in the bloc universe is just the whole bloc universe.

U_i became U_j just in case *U_i* has past existence, and when *U_i* had present existence, *U_j* had future existence.

U_i will become U_j just in case *U_j* has future existence, and when *U_j* will have present existence, *U_i* will have past existence.

The *primitiveness* of the new concept of change must be stressed. The critics of the dynamic ontology often seem to demand, for instance, that the concepts of *past* and *future* should be explicated in terms already available in the bloc universe theory (which are essentially the bloc universe concept of *existence*, and the temporal relations *earlier than* and *later than*.²⁹) The fact that this explication is impossible is then taken as a *reductio* of the dynamic theory. But it is really only a healthy sign that the dynamic theory is genuinely different from the static theory.

²⁹For instance, it might be supposed that *X has future existence* is just equivalent to: *X exists at a later moment of time*. But this must be rejected. See Section 1.25 and 1.26.

CHAPTER EIGHT

THE ROLE OF PROBABILITIES IN QUANTUM THEORY

Two main presumptions about the role of probabilities in quantum theory have been made in preceding chapters. The first was in Chapter Six, where the main argument relied on the premise that quantum theory postulates objective probabilities in nature. The second was in Chapter Seven where a model was given which represented probabilities as transition probabilities between definite states of physical systems (or of the entire universe), with the implication that quantum probabilities are of this kind. The first presumption is on much firmer ground than the second, but the unsolved problems of interpreting quantum theory pose problems about both. I will first argue that an interpretation which treats quantum theory as the *fundamental theory* must make provision for objective probabilities sufficient for the argument of Chapter Six. I will then consider the second

presumption, and the kind of view of quantum theory that it relies upon. I should stress that the model constructed in Chapter Seven is not an attempt to say anything substantial about interpretation of quantum theory, but is intended primarily to show that a coherent model of time flow is possible, and that probabilities can be naturally accommodated in such a model. This is not the only way for probabilities to be modeled, nor is there much evidence at present that quantum probabilities should be interpreted in this way. But I will argue that it is not an implausible view. First, however, the objective status of probability in quantum theory.

8.1 The objectivity of quantum probabilities.

The reason for thinking that quantum probabilities are objective is simply that the theory seems to be *fundamental*, and probabilistic laws are central to the theory. These laws are generally given by the rule: $\text{PROB}(|\psi\rangle | |\varphi\rangle) = |\langle\psi|\varphi\rangle|^2$, and without them quantum theory would have little empirical content. But exactly what such laws mean is controversial. They are often taken to represent transition probabilities between states, but this needs considerable elaboration. Firstly, the role of 'measurement' must be recognised in the transition from $|\varphi\rangle$ to $|\psi\rangle$. The probability is conditional on a measurement being made which has $|\psi\rangle$ as an eigenvector. Secondly, it is not generally true that the measured system makes a transition to the state $|\psi\rangle$. It

is only true when the measurement constitutes a preparation, as for instance when a polarizing filter is used to prepare particles with a certain polarized state. Thirdly, although it is common to refer to the state vector, $|\psi\rangle$, as the quantum state, in fact it is unclear what the state vector characterises. These problems derive from fundamental foundational problems of quantum theory, and make the understanding of the role of probabilities in the theory difficult. Indeed it is obvious that a full understanding of what quantum probabilities are requires a full solution to the interpretational problems that presently plague the theory, for these problems are largely centered on the process through which probabilities come into play. This is evident in the way in which different views on the interpretation of the theory give rise to different roles in the theory for probabilities. For instance, to consider three very different kinds of interpretation: (i) The orthodox positivistic view tends to treat the theory as a formalism for making predictions about the results of measurements or observations. Here the rule: $\text{PROB}(|\psi\rangle || \phi\rangle) = |\langle\psi|\phi\rangle|^2$ may be taken as somewhat indirectly expressing a probabilistic relation between *observation events*. (ii) On the fully realistic interpretation proposed by Maxwell [1988], wave packet collapse is a generic process in nature for which the dynamics are intrinsically probabilistic. The rule: $\text{PROB}(|\psi\rangle || \phi\rangle) = |\langle\psi|\phi\rangle|^2$ in this case expresses objective transition probabilities between physical states of quantum systems. (Considerable elaboration of the orthodox version of the theory is involved in describing the conditions for wave packet collapse however.) (iii)

The Everett-Wheeler 'many-worlds' interpretation is on the face of it a deterministic interpretation of the theory. The idea is roughly that a measurement gives rise to a proliferation of universes, one for each possible result of the measurement. The role of probabilities is a tricky point in this theory, but I will soon argue that they must be preserved as probabilities about the expected results of measurements from the point of view of physical observers within the universe.

I will first argue that each of these interpretations, insofar as they do justice to quantum theory, provides probabilities sufficient for the argument of Chapter Six. I will then consider the possibility of deterministic interpretations of the theory. (i) This interpretation provides probabilities relating observations to each other. The argument of Chapter Six assumes probabilities of transitions between *states*, but it is readily seen that it could be reformulated in terms of probabilities of the corresponding observations instead. I.e. we could understand: $\text{PROB}(\varphi(t+\Delta t)|\psi(t)) = p$ to mean that the probability of *observing* φ at $t+\Delta t$ given the observation of ψ at t equals p . The argument proceeds with little change.

(ii) Maxwell's realistic interpretation is perfectly suited to the argument of Chapter Six.

(iii) The many-worlds interpretation is problematic for the argument of Chapter Six because it seems to render quantum theory *deterministic*. However, insofar as it fails to provide an interpretation of the quantum probability laws it fails to be an

adequate interpretation, and insofar as it provides for such probability laws it legitimates the argument of Chapter Six.¹ The interpretation is described roughly by saying that upon a measurement being made, the universe 'branches' into a number of separate parts, one for each possible result of the measurement. Since this branching is deterministic, it is difficult to see how to make sense of the idea of relative probabilities of results of measurements.² But this is only because it is difficult to see how the notion of *observation* itself is to be accommodated in the theory. The evidence for quantum probabilities is provided by relative frequencies of *observations of specific results of measurements*: how are such observations to be accommodated with the fact that *all measurement results are actual*? If they are not accommodated, then there is no place to fit probabilities into the theory, but the failure to make sense of observation means the interpretation is inadequate. The usual line, however, is to maintain that there are observers existing in specific branches from whose points of views measurements have specific results. But this must allow for the interpretation of *probabilities of results of measurements* as probabilities for observers to experience themselves as being in branches where certain results obtain. Hence the theory remains a probabilistic one. With the appropriate interpretation, the future-directed

¹It is also very doubtful that the many worlds-interpretation is an adequate interpretation for the reasons noted by Earman [1986, p.224-5].

²For instance, how could the probabilities of two possible results of a measurement be $1/3$ and $2/3$ respectively, if both results happen?

probabilities of quantum theory must be assigned their normal values, and the argument of Chapter Six against the existence of past-directed probabilities will clearly proceed.

These three examples of interpretations of quantum theory illustrate how insensitive the argument of Chapter Six is to questions of interpretation. This is obviously because those three interpretations all render quantum theory a *probabilistic* theory. But this is so on any interpretation of the theory, even a genuinely deterministic interpretation: quantum theory, like thermodynamics, simply is a probabilistic theory because it entails probabilistic laws. The possibility of a 'deterministic interpretation' does not alter this fact, because a deterministic interpretation does not render *quantum theory* deterministic: what it does is to render the *fundamental theory* deterministic. This simply implies that quantum theory is not the fundamental theory (because if it was, the fundamental theory would be probabilistic, not deterministic). Hence what a deterministic interpretation changes is not the result that quantum theory is irreversible, but the meaning of this result. If quantum theory is not a fundamental theory, then the fact that it is irreversible is of limited consequence for the symmetry of physical time.

To see this clearly, consider the possibility of a deterministic hidden-variable interpretation. Consider a quantum system in a state: $|\phi\rangle = \sum a_i |\psi_i\rangle$ on which a measurement A with eigenvectors $|\psi_i\rangle$ and corresponding eigenvalues λ_i is performed³. Orthodox

³For simplicity it is presumed the measurement has discrete and non-degenerate eigenvalues.

quantum theory tells us that the probability of any result, λ_i , is equal to $|a_i|^2$, which I will represent as: $\text{PROB}(\psi_i|\phi) = |a_i|^2$. These probabilities certainly exist in the weak sense that they give rise to correct statistical predictions. But what has been doubted is whether these probabilities are to be interpreted as *objective in nature*. The alternative is that there is some underlying deterministic reality to each quantum state which would determine the result of any measurement A before it was performed. This would mean that the state vector $|\phi\rangle$ is an incomplete representation of the fundamental micro-state, in a way similar to that in which thermodynamic states are incomplete.

The implication of such a result for my argument in Chapter Six, that quantum theory is irreversible, is not obvious. In the first place, the new deterministic theory must go considerably beyond the content of what is presently regarded as quantum theory, and may well depart from the present theory in certain predictions⁴. To this extent, the deterministic theory is a quite distinct theory, and not simply an *interpretation of the quantum*

⁴For this point, the failure of the proposal of Daneri, Loinger and Prosperi [1962], as noted by Earman [1986,p.223], is important. For such a proposal might remove probabilities from quantum theory by theoretically *simplifying* it, roughly by dispensing with the 'projection postulate' which generates the probabilistic laws governing wave packet collapse, in favour of purely deterministic evolution at all times. The trouble with this, as Earman notes, is that the purely deterministic Schrödinger evolution just cannot produce *measurements*.

formalism. The new deterministic theory would provide a reinterpretation of what the present quantum formalism is about, in the same kind of way that sub-atomic particle theories provide for a reinterpretation of what atomic chemistry is about. It is most plausible to regard this new interpretation of the present quantum formalism as the 'deterministic interpretation of quantum theory'. Let us call this theory QT-D, and the complete deterministic micro-theory DT. Since QT-D employs essentially the same formalism as quantum theory, it postulates the same set of future-directed probabilities: $\text{PROB}(\psi_j|\phi)$. Hence it remains a probabilistic theory, and must still be judged reversible or not according to whether it satisfies the CPR.⁵

The only difference may be that the new fundamental theory, DT, may provide reasons for introducing past-directed probabilities satisfying the CPR into the QT-D. But there would have to be some very special reasons for this to be the case, and at any rate, *the version of QT-D that applies to the actual universe* would remain irreversible in the same way as the version of classically-based thermodynamics that applies to the actual universe is irreversible (see Chapter Five).

The new interpretation therefore will not alter the fact that quantum theory is irreversible: what it would alter would be the *meaning* of this irreversibility. Very simply, QT-D would not be a fundamental theory, hence its irreversibility would not directly reflect on the symmetry of physical time. DT would be the

⁵Remember that thermodynamics, as an inherently probabilistic theory, must also be judged reversible or not according to the same criterion.

fundamental theory instead, and the symmetry of physical time would depend upon whether DT were reversible. Without knowing precise details of DT it is, of course, impossible to say whether it is reversible or not.

The problem posed by the possibility of a deterministic interpretation is therefore really just the problem that quantum theory might not be the fundamental theory. This possibility is admitted from the start, and does nothing to undermine the analysis of Chapter Six.

It is also interesting to consider the likelihood of such a deterministic interpretation, because it is thought by many present writers that Bell's theorem provides strong arguments against this possibility. This is a conclusion I would welcome, but in fact I am skeptical of it. My reasons are skepticism about the possibility of evidence for the postulate of *counterfactual definiteness*, without which Bell's theorem has no implications against determinism (see Earman [1986, Ch.XI] for an excellent discussion and further references.) The present success of quantum theory is consistent with the possibility of a more fundamental deterministic theory⁶. But this is the kind of possibility that must simply be lived with.

⁶There are reasons for believing quantum theory (no doubt in a more complete form than is presently known) to be fundamental, but these reasons just have to do with the very general success of the theory. If, on the contrary, it turns out that quantum theory is not fundamental, then it is not possible to tell from purely quantum theoretical results whether the fundamental theory is deterministic or probabilistic.

I will now turn to a more speculative topic: the representation of quantum probabilities as transition probabilities between states.

8.2 The nature of quantum probabilities.

The model of probabilities provided in Chapter Seven represents probabilism in nature obtaining by means of simple *transition probabilities* between states of physical systems. Can quantum probabilities be construed in this way? Maxwell [1988] provides the framework for a realistic interpretation of quantum theory which treats quantum probabilities in just this way.

The first main point about Maxwell's interpretation is its treatment of the 'wave-packet collapse' which occurs upon measurement. The probability: $\text{PROB}(\phi|\psi) = p$ comes into play, according to orthodox quantum theory, as the 'collapse' of the quantum state from ψ to ϕ upon an event of *measurement*. The orthodox theory is most inadequate in its treatment of the concept of measurement, which it leaves as an unanalyzed primitive. Critical of this aspect of the orthodox theory, Maxwell desires an interpretation in which *measurement* is not an extra primitive: instead the operation of measuring devices is to be analysable in terms of purely physical micro-processes. The main reason for this is rather convincing: measurements involve interactions between physical systems which are being measured and physical measuring devices which measure, and this interaction is fundamentally at the micro-physical level, and

should be explicable in terms of micro-physical laws. I will call this the thesis of *complete micro-physical analysability of measurements*. Of course there are other possibilities (for instance some form of *mentalism*, e.g. Wigner [1963]), but these all seem to have troublesome peculiarities of one kind or another. I think it is fair to say that these various alternatives have arisen largely in response to the difficulties of establishing an interpretation which delivers the *complete micro-physical analysability of measurements*, and not because they have much intrinsic plausibility. At any rate, the *complete micro-physical analysability of measurements* seems a desirable goal of an interpretation, so let us consider its implications.

The reason that measurement is troublesome in the first place is that it involves the strange phenomenon of *wave function collapse*. On the desired interpretation, wave function collapse (or whatever represents it) must become a generic phenomena in nature, governed by fundamental physical laws. This implies what I will call *spontaneous collapse of the wave function*. By this is meant that the states of quantum systems must sometimes 'collapse' to component states purely as a result of their intrinsic physical evolution, without the collapse being provoked by interaction with an external system. The reason for *spontaneous collapse* is very simple. A system measured by a measuring device undergoes wave function collapse as the result of the interaction with the device. Hence the combined measured-plus-measuring system undergoes collapse as well (see Earman

[1986,p.219-222)]⁷. This collapse is without external provocation (for anything that counts as external provocation may be included in the measuring device: otherwise the postulate of micro-physical analysability of measurement fails). It is therefore an unprovoked 'wave function collapse' which is governed by micro-physical laws. This is what is being called *spontaneous collapse of the wave function*.

Perhaps a word about the dynamics of such a collapse is in order. Maxwell treats the dynamics for quantum systems as falling into two distinct parts, the deterministic Schrodinger evolution and the probabilistic wave function collapse. His achievement is in trying to show that there might be specific micro-physical laws governing the latter, which is only described very roughly by orthodox quantum theory.⁸ This separation of two distinct dynamic principles has a long history (beginning with von Neumann's distinction of Type 1 and Type 2 processes), but is frequently objected to. One reason for complaint is that the principles governing probabilistic evolution are not precise enough to represent a dynamic theory, but Maxwell's theory is an attempt to solve exactly this problem, so this charge cannot be brought against him. A second objection is that it is implausible

⁷On Maxwell's interpretation, the combined measured-plus -measuring-system may lack certain properties peculiar to a *quantum* system (e.g. lack of phase relations among its parts). But it is a quantum system in the wider sense of being a physical system governed by quantum theory.

⁸See also Penrose [1986] suggests that wave function collapse is a gravitational phenomena.

that physical systems are governed by two distinct dynamic principles. This is convincing, for the separation of dynamics into two separate types must really be just a convenient way of representing what is a single dynamic principle. But the only practical way of formulating what this single principle is may be by separately formulating two distinct dynamic principles (dynamic and probabilistic respectively) and concocting a suitable mathematical combination of them.⁹ I will not consider this problem any further here.

Supposing generic spontaneous collapse of the wave function in nature as Maxwell does solves two important problems. First it allows for conditional probabilities which are not further conditional on measurements being made. This means that conditional probabilities in quantum theory can be understood as in Chapter Seven, without the need to worry about the further complications of secondary conditionals about measurements mentioned in Chapter Six. Secondly it allows a mechanism for the quantum state of the whole universe to undergo collapse, which cannot be assumed in the orthodox version of the theory. The interpretation of quantum theory that results consequently fits the dynamic model presented in Chapter Seven very well: the universe as a whole can be attributed a definite quantum state (relative to a specific frame of reference if necessary), and it undergoes probabilistic change to new states.¹⁰ Of course it is

⁹Others (see Pearle [1976]) have considered directly modifying the Schrodinger equation to obtain effects that might serve for 'wave function collapse', but with little positive result.

necessary to interpret the nature of the quantum state before a specific theory results, and Maxwell makes a definite proposal in this regard also, but the general fit of Maxwell's *type of interpretation* with the model of Chapter Seven is independent of these details.

This type of interpretation is the simplest realistic interpretation I know of which seems to have a real chance of being adequate. It is a controversial proposal which needs development in a variety of ways (particularly in the formulation of the dynamics of wave function collapse), and it may turn out to be wrong. However it fits nicely with the dynamic ontology of Chapter Seven and renders it plausible that quantum theory requires such a dynamic ontology, and hence that a good scientific theory requires time flow. Of course, without a conclusive solution to various foundational problems of quantum theory the question of the correct model for quantum theory cannot be conclusively answered.

¹⁰The separate principles for deterministic and probabilistic evolution is a little troublesome. One possibility is that all *change* (movement between universe-trees) be taken to be probabilistic, and the deterministic Schrödinger evolution be regarded simply as a part of the system for determining the probabilities that exist in the universe-tree. At any rate, the Schrödinger equation is already fully represented by the class of possible universe-trees, and there seems to be no need to interpret it in a dynamic way. (This fits well with orthodox scepticism about the reality of the deterministic evolution due to the fact that it is not directly observable - which for instance allows the Heisenberg and Schrodinger pictures of that evolution to appear equally valid). Another possibility is to take Schrödinger evolution as producing continuous change of the universe-tree.

CHAPTER NINE

PROBLEMS WITH RELATIVITY THEORY

Relativity theory poses a famous difficulty for the dynamic ontology proposed in Chapter Seven. The difficulty arises from the denial, in the orthodox interpretation of relativity theory, of any physically real relations of simultaneity between spatially separated events. Consider two space-like separated events, E_1 and E_2 . The dynamic ontology requires that at some time E_1 is *present*. But when E_1 is present, then according to relativity theory there is no *physical* fact about whether E_2 is present or not. In the context of the ontology presented in Chapter Seven, this would mean that the notion of the *present universe-tree* is not a physical notion, since there is no physical reality to the *co-presentness of spatially separated events*. For this reason, physical time flow is widely considered to be incompatible with

relativity theory, and this has convinced many to embrace a static rather than dynamic view of time.¹ This is the most serious scientific problem for a dynamic view of time, and a number of possible responses to it will be considered here.²

9.1 The 'metaphysical' postulate of simultaneity relations.

A possible response by the defender of dynamic time is to maintain that relations of simultaneity among distant events obtain 'metaphysically', despite relativity theory.³ This is possible because the physical content of relativity theory implies only that *there are no physical observations that provide evidence for relations of simultaneity*, and it is concluded that there are no such relations only because they would be physically superfluous.⁴ That it is possible to maintain the existence of an

¹See for instance Gödel [1949], Quine [1960, p.172], Costa de Beauregard [1963], Grünbaum [1963,p.329], Williams [1951], Putnam [1967], Fitzgerald [1969].

²McCall's [1976] theory of time flow requires world-wide simultaneity, and he devotes a section to the discussion of relativity theory. However, he does not discuss the problem posed by the physical lack of simultaneity relations: he appears simply to assume that simultaneity relations exist 'metaphysically'. McCall tries to establish some connections of his theory with quantum theory, and at least sees his assumption of indeterminacy in nature as legitimated by the latter, but in fact his theory of flow is clearly a metaphysical theory, and not a physical theory of the kind that I desire, hence the elaboration of his theory on the content of physical theory may be of little concern to him.

³ E.g. Prior [1972] p.322-3.

⁴"The introduction of a 'luminiferous ether' will prove to be superfluous in as

absolute frame consistently with relativity theory is illustrated by the consistency of the pre-relativistic theory of the 'luminiferous ether' in which moving objects contract and expand, and so forth, to generate all the relativistic phenomena. The reason the 'ether' hypothesis has been abandoned is not because it is inconsistent with the phenomena, but because a simpler form of explanation has been found in which the ether is redundant. Because of its physical redundancy, the hypothesis of an ether (equally, of relations of simultaneity) is regarded as 'metaphysical' in the positivistic sense of that term.

The 'metaphysical' postulate of real simultaneity may be acceptable to writers such as Prior who are content with a metaphysical theory of time flow, but it is against the spirit of the physical theory of time flow. I am concerned to find a place for time flow in the physical ontology, and one does not have to be a positivist to find the *ad hoc* postulate of relations of simultaneity as 'metaphysically' real but physically undetectable highly unsatisfactory. So let us consider other possibilities.

--

much as the view here to be developed will not require an 'absolutely stationary space' provided with special properties...", Einstein [1905].

9.2 Is the denial of simultaneity relations compatible with quantum theory?

There are a number of problems in combining relativity theory with quantum theory which throw doubt on the interpretation of the former, and might mean that an interpretation is needed in which physical simultaneity relations are postulated after all.

(i) Maxwell [1984] and [1988] argues for a quantum theory in which there are physical relations of simultaneity. Moreover, Maxwell's theory implies certain deviations from the predictions of orthodox quantum theory, and would be physically preferable to the latter if these were verified. And although Maxwell's theory implies the existence of physical simultaneity, this does not conflict with the observational implications of relativity theory, because it remains physically impossible to observe what the *actual* relations of simultaneity are.⁵

At first this sounds like the same kind of 'metaphysical' postulation of unverifiable simultaneity relations that was considered in the previous section, but in fact the situation is very different. The reason is that empirical evidence for the existence of simultaneity relations does not have to involve the direct observation of *actual* relations, or even the possibility of such observations. Maxwell would argue, that since his theory

⁵It is generally accepted that quantum theory entails the impossibility of faster-than-light transmission of information, given the impossibility of faster-than-light velocities of material particles, because of such proofs as Jordan [1983], Ghirardi *et alia* [1980].

entails simultaneity relations, evidence supporting his version of quantum theory provides evidence for the existence of simultaneity relations. This is a case of an empirical theory which entails the existence of facts of a certain *kind* (simultaneity relations), while at the same time entailing that it is physically impossible (by relativistic covariance) to observe what the *actual* facts (of simultaneity) are. There are many such situations in science where it is required that facts of a certain kind must hold, though it is unknowable even in principle what the specific facts are.⁶ Hence Maxwell's position of maintaining an absolute frame of reference, while denying that it can be known experimentally what this frame is, cannot be immediately rejected on methodological grounds.

This illustrates an interesting possibility - of physical evidence for the existence of simultaneity relations consistent with the observational implications of relativity theory - but there is some doubt about this aspect of his theory. For it seems likely that the unverifiability of *actual* simultaneity relations will eventually be reflected in unverifiability of the aspect of his interpretation that gives rise to them. However, since his interpretation is highly speculative anyway, I will not discuss this further here, but merely conclude that his arguments show

⁶E.g. a theory of processes occurring in the center of the sun might have the result that (i) at a moment t there is a definite state of a given region in the center of the sun, but (ii) it is impossible even in principle for this state to be known because there is no possible physical mechanism for the necessary information to be recorded.

that there may be some possibility that quantum theory requires the assumption of simultaneity relations consistently with the observational implications of relativity theory after all⁷.

(ii) More generally, any interpretation of quantum theory which treats wave packet collapse as a realistic phenomenon has a problem with the relativistic denial of simultaneity, since it appears that the collapse must be instantaneous across spatially extended wave packets. This is particularly so if 'collapse' is used to provide a mechanism for the non-local quantum correlations (which have been empirically verified by Aspect *et alia* [1982], Shimony [1978]). This is a deep problem which seems far from being solved.

(iii) There are deeper problems still with the unification of quantum theory and general relativity. Until some real progress is achieved in this area, the scope and meaning of relativity theory in the quantum realm remains unclear.

In summary, relativistic quantum theory is poorly understood, and it is unclear whether simultaneity relations are needed or not in the context of quantum theory. There is certainly no conclusive argument as yet to reject the existence of physical simultaneity relations necessary for the universe-trees of the dynamic model. However, there is perhaps somewhat less evidence for the existence of such relations than against them,

⁷Note that there is only consistency with the observational implications of relativity, but not with the relativistic principle that all physical laws are valid in all inertial frames, for in Maxwell's theory, wave packet collapse is not frame invariant.

and it would clearly be an advantage if a dynamic model could be provided which did not require such relations. This possibility will be considered next.

9.3 Are simultaneity relations necessary for a dynamic model?

That time flow requires world-wide simultaneity has been accepted by almost all writers⁸: a notable exception is Čápek [1966]. Čápek develops a remarkable view that relativity theory actually requires a dynamic conception of time. He accepts that relativity *unifies* time and space, but rejects the orthodox interpretation that this requires the *spatialisation of time*. Instead "the relativistic unification of space with time is far more appropriately represented as a *dynamisation of space* rather than a spatialisation of time." (p.515). And of the static conception of time he says: "Despite its superficial plausibility and widespread popularity, hardly any other view is more incompatible both with the spirit and the letter of the relativity theory." (p.519).

His argument that relativity theory requires a dynamic view of time is very interesting. He firstly denies any reality to simultaneity relations among distant events, and stresses that there is no world-wide 'now' or present, such as is required by the universe-trees of Chapter Seven. But he does not see such

⁸Including McCall [1976, sec.3]. McCall just assumes the existence of simultaneity relations, and displays no concern for their *physical* reality.

simultaneity relations as necessary for dynamic time. Instead, he bases a theory of dynamic time on a merely local conception of 'now', the 'here-now'. The reason is that relativity theory supports an objective distinction between the past and future parts of the world-lines of specific objects. The idea, as far as I can make it out, is that there is time flow for each specific object in the world considered individually, although not for the world as a whole, since in this respect the world cannot be considered as a whole.

Two parts of his argument are usefully separated. First is what may be called his *relativistic theory of time flow*, which is roughly that time flow is real for each object considered separately. Second is his argument that relativity theory *demand*s this interpretation of time. The latter argument seems strained. What is interesting is former, which implies the *possibility* of a relativistic theory of time flow. Unfortunately Capek does not develop his proposal for relativistic time flow very far: he certainly does not say enough to indicate how an explicit model of relativistic version of time flow might be developed, and the idea of 'dynamic time' he is working with remains vague. What follows is an attempt to develop the implications of his idea.⁹

If time flows for each object individually, then presumably *from the point of view of each object in the world* there is a different set of past, present, and future events. This might be

⁹I do not imply that Capek himself would approve of this development of his ideas.

accommodated by having a distinct universe-tree for each object. What do these universe-trees contain?

Let us consider first the hypothesis that *the state of the object at the present moment is the only present real event (for the universe-tree of that object)*. This means that the present part of a universe-tree for an object consists entirely of the state of that object at that moment. What do the *earlier* and *later* parts of the universe-tree contain? Since they contain facts (including probabilities) about the *present parts of past and future universe-trees*, they contain just facts about past and future states of the object in question. Therefore they contain nothing about any events *except events in which the object in question participates*. This of course may include interactions with other objects, but does not include any events which may occur to objects at different places. Let us assume therefore that the universe-tree for an object X at a moment t contains (i) the event occurring at the *place* of X at the moment t, (which is the present part of the universe-tree) (ii) probabilities of events occurring at the place of X at past and future moments (which gives the earlier and later parts of the universe-tree for X).

There are at least two serious problems with this. Firstly, *the probabilities of past and future events involving X* depend not merely upon the present state of X, but also upon other objects in the vicinity of X which may interact with X. Therefore the state of the larger universe must be taken into account in determining

the earlier and later parts of the universe-tree. How is this to be done? Secondly, the world is composed of many different objects, but in the model being considered there is no single universe-tree for 'all reality': there is only a collection of different universe-trees, one for each object. This seems to imply that to each object there corresponds a distinct reality. This contradicts the normal understanding that there is a single unified reality in which everything participates.

The first problem shows that a universe-tree for X must contain more than just the events involving X . A larger class of events is required to determine the probabilities of future events. This class need not be a class of *present events*, in the sense of events occurring at the present moment: but it must at least be a class of *presently determinate events*. What might this class be?

There are three obvious requirements on this class. First it must be independent of reference-frame if it is to be physically real given the constraints of relativity theory.¹⁰ Second it cannot include events which clearly lie in the future. Third it must suffice to determine probabilities of events indefinitely far in the future. The smallest, and I think the only plausible, class fulfilling these requirements is the *class of events comprising the surface of the backward light-cone from the 'here-now' of the object in question* (see Fig. 9.1). I will call this the class of *PD events* (*presently determinate events*), or the *PD class*. It must be attributed a special ontological status in the model. This

¹⁰This rules out any world-wide class of 'simultaneous' events, for instance.

status is weaker than that of being *present*, since the PD events are in fact clearly *past* events.¹¹ But it is stronger than being merely *past*, since some past events may be undetermined by the present universe-tree, while the PD events are all determinate. The physical basis for interpreting this class of events as 'presently determinate' from the point of view of the object concerned is not particularly clear, but it does not contradict relativity theory, so let us consider the implications of the idea.

The introduction of the PD class into the universe-tree generates a full set of probabilities for all past and future events in the universe, and thus it solves the first problem. But the second problem, of the lack of 'unification of reality', remains. Let us consider this problem in the context of a simple universe which contains only two objects, X and Y.

From the 'point of view' of X there is a *present universe-tree*, say U_X . This universe-tree represents the 'reality' that presently exists *for X*, or *from X's point of view*. This reality includes some facts about Y, but it is a strange set of facts. In particular, it does not determine what *the present universe-tree for Y is*. This reflects the fact that from X's point of view, there is in fact no such thing as *the present universe-tree for Y*, since the lack of simultaneity relations means that there is no such thing as *the present state of Y* (only the presently determinate state of Y). Yet being committed to the existence of Y, we are committed

¹¹ Shown by the case of a universe-tree for a photon: the previous trajectory of the photon is included in the PD class, and of course this trajectory must be considered to be in the absolute *past* of the present here-now of the photon.

to the existence of a *present universe-tree for Y*.

What this means therefore is that (i) there is a present universe-tree for X and a present universe-tree for Y, but (ii) there is no simultaneous reality to these universe-trees. This at first sounds contradictory, and it is at least very mysterious, but I am not certain it is incoherent. What it implies is that *there is no single all-encompassing point of view from which the dynamic relativistic universe can be regarded*. Instead there are only a number of *distinct points of view*, one for each individual object that exists within the relativistic universe. We might say that any point of view from which the universe appears *dynamic* is necessarily the 'point of view' of an object within the universe. There is no 'external' (or *objective*, or *frame-independent*) point of view of the universe which represents its dynamic nature.

This is a radical 'subjectivisation' of reality: there is no single reality, only 'reality as it appears from a point of view'. This subjective view of reality jars with our normal idea of what 'reality' means. The normal idea is that there is at least one single objective point of view of reality which represents all.¹² But must this be so? Is it possible to accept that the *dynamic relativistic universe* is metaphysically very different from a bloc universe in that there simply is no single *objective point of*

¹²Such a point of view seems to be provided, for instance, by a complete Minkowski picture of the universe in some specific frame of reference. But of course this requires that the universe is taken to be a bloc universe, and hence that there is no time flow.

view from which the universe can be described completely? This is obviously a speculative idea, which will set the logical intuitions of most people protesting, and of which I am entirely unsure myself. Therefore I will not try argue for it any further here, except to note the following points. (i) It seems to be the only way to treat time as dynamic given the relativistic denial of simultaneity relations. This of course might be an indication that dynamic time is inconsistent with relativity, but: (ii) the existence of distinct *consciousnesses* in the world seems to provide a model for both the idea of a 'points of view of reality', and of dynamic time. The actual existence of many different consciousness within the relativistic universe thus seems to provide a model for a *dynamic* relativistic universe, and hence to show that such a thing is not only logically coherent, but actually real.

9.4 Summary.

In summary, relativity theory poses a serious problem for the dynamic theory of time, and it is as yet unclear what the resolution will be. There appear to be two possibilities favourable to the dynamic view of time: (i) that despite the observational validity of relativity theory, the orthodox interpretation that this implies the absence of simultaneity relations is incorrect, and such relations are physically real after all. (ii) A more radical solution, suggested by Capek [1963], that

a dynamic view of relativistic time is possible. Perhaps I might add that: (iii) If a more detailed analysis shows that relativity theory is indeed inconsistent with time flow, the truth of relativity theory could be rejected, on the grounds that the evidence for time flow is far more convincing than the evidence for relativity theory.

APPENDIX 1.1

THE SPATIALISATION OF TIME THROUGH THE USE OF
SPATIAL DIAGRAMMS

"Discussion of Zeno's paradoxes, as of much else, is aided by graphing time against distance. Note then that such graphs are quite literally a treatment of time as spacelike."

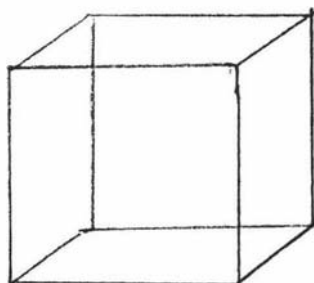
Quine [1960,p.372].

Physicists constantly use graphs or diagrams to represent relationships between different variables. The diagramisation of *time* is particularly popular, and particularly useful, as Quine notes. From the beginning of our scientific educations we encounter diagrams representing the behaviour of systems with respect to time. It becomes second nature to the physicist to draw a line on a page and call it the 'time line', and to plot such things as the distances, velocities, energies and so forth against it. The Minkowski diagram is perhaps the most celebrated of such visual aids, and is used extensively in thinking about problems in relativity theory. In fact because it is so widely used, the Minkowski diagram has become a popular informal representation of the *ontology* of relativistic physics. But I will argue that to take it as a representation of the *ontology* is a mistake, for

although the spatial representation of time in diagrams is useful for problem-solving purposes, when it comes to thinking about *the nature of time itself* the representation of time by space is dangerous. The reason, I will argue, is that it incorrectly leads to a transference of the ontological properties of space to time. This transference seems to come about through a presumption that the spatial dimension can be used to give an *analogue representation of time*. Because there is nothing about space that is an *analogue for the flow of time*, it is easily concluded that time flow cannot be represented in the scientific model of time, and is non-existent.

Let me first indicate what I mean by an *analogue representation*, or *analogue diagram*. An analogue diagram is one in which the elements of the diagram have real properties of the same kind as the thing being represented. For instance, a miniature scale model of a ship counts as an analogue diagram in respect of certain spatial or topological relations among parts of the ship. A colored copy of a painting counts as an analogue representation of the painting in respect of the arrangement of colours. Also, the representation of spatial relations by a two-dimensional spatial diagram (think of a map or floorplan) can be an analogue diagram in respect of space, as long as the number of spatial dimensions in the object depicted are available in the diagram, which would normally be two-dimensional (a piece of paper.) Otherwise, as in many physical diagrams, two or more dimensions of the object must be compressed into one dimension

of the diagram, as for instance when we depict a cube thus:



Here there is a failure of the spatial analogue between object and representation, in the obvious sense that there is no proper analogue in the picture of the real distances between all the edges of the cube. However, in this case there is still a pretty close analogue in many other respects.

The point I want to make is that any kind of spatial diagram to represent the flow of time must automatically fail to be an analogue diagram because there is no analogue of time flow in the diagram. This becomes vividly apparent when a real analogue representation of time is considered. The best example of this is a film of a temporal process being shown in real time. The stretch of real time in which this representation (the film) is viewed is of course not the time of the original process, but only refers to it through diagrammatic conventions. But since the time of the representation is *real time*, it obviously provides a perfect analogue representation of the stretch of time that is being represented.

We feel immediately how different this analogue representation of time is to a spatial diagram of time. The extra

feature of the analogue representation (the film) that differentiates it from the spatial diagram seems to be that what appears on the movie screen, in real time, *changes*. Let us say that Frame₁ of the movie depicts State₁ of some physical system, Frame₂ depicts State₂, and so forth. Then the *temporal relation* between State₁ and State₂ is depicted *analogously* in the movie by the temporal relation between Frame₁ and Frame₂ as the movie runs through in real time. The film is fundamentally *dynamic*, exactly because it is an analogue depiction of time, in contrast to the static spatial diagram.

The point is not, of course, that representations have to be analogue representations to be any good. Modern physics in a sense is the triumph of using purely mathematical language to represent physical phenomena, where there is no analogue between the representation (sentences and equations) and the things depicted. I also believe that a perfectly adequate mathematical/linguistic representation of time flow can be given, so there is no necessity to offer analogue diagrams of time (see Chapter 7). All I am suggesting is that the use of static spatial diagrams to represent time has misled people about the nature of time, for roughly this reason. The spatial dimension used to represent time is a partial analogue of time, in that there is an analogue in it for certain features of time: linear ordering, continuity. But there is no analogue at all for the *dynamic* feature of time, the flow or movement of time. This lack is seized upon, and rather than being perceived as counting against

the truth of the diagram, the fault is transferred instead to time itself: it is inferred that there is no real dynamic feature of time, since no such feature is (nor can be) exhibited in the spatial diagram.

This may seem a simplistic kind of mistake to make, but I think it can be seen in operation in many of the criticisms of time flow. For instance, there is a widespread scepticism about the idea of the temporal modalities of existence, (past, present, and future existence) required by the dynamic view of time. This seems to be inspired by the fact that in any static diagram of time, (where time is plotted against the state of the world at that time), we can easily point to any two moments *simultaneously*, and see that the diagram represents them as real *together*. Where, in the diagram, are the *modes* of reality or existence that are required by the dynamic view?

But this problem dissolves in the dynamic representation, where there is always just one part of the depicted process that we can point to at any time as *presently real*, namely the one depicted on the movie screen *at the present moment*. This makes us far less ready to concede that nothing in the depiction of time can represent time flow, because in this analogue depiction it is our direct experience that something does represent time flow, namely the real flow of the film itself as it is shown in time. Or again, if someone said that there could be no representation of real *change* in the representation of time (or temporal processes), we could easily reply by asking them whether the

scene depicted in the running movie was really changing or not.

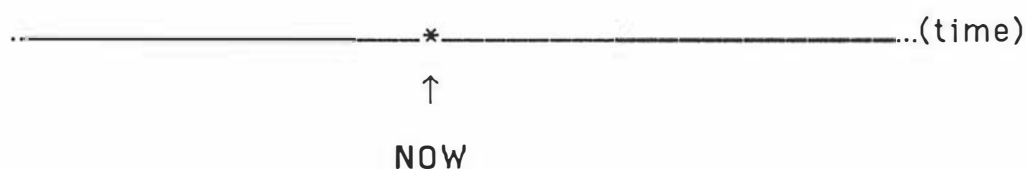
Or again, consider the following argument against time flow, which has some resemblance to McTaggart's famous argument of [1908]. *Time flow requires that all actual events are at first future, then present, and later past. But any event is past with respect to some later event, present with respect to a simultaneous event, and future with respect to some earlier event. Therefore every event is past, present and future. But past, present and future are incompatible qualities. Therefore the dynamic view of time, which treats past present and future as modes of existence, leads to a contradiction, and is incoherent.*

The trouble with this argument is that events are not past, present or future merely *with respect to other events*: they are past, present or future *per se*. (Their pastness, presentness or futurity of course *changes*.) This is easily perceived when the dynamic representation of time is considered, because here the pastness, presentness or futurity of the event depicted corresponds to the pastness, presentness, or futurity of the object that depicts, i.e. the image on the movie screen. Suppose that Frame₁ and Frame₂ of the film have been shown, and are past, while Frame₃ is presently displayed. Clearly the showing of Frame₂ is *absolutely past*: there is no temptation to say that it is *future with respect to Frame₁*, and to conclude that it is therefore future. It can only be said that it *was later than Frame₁*.

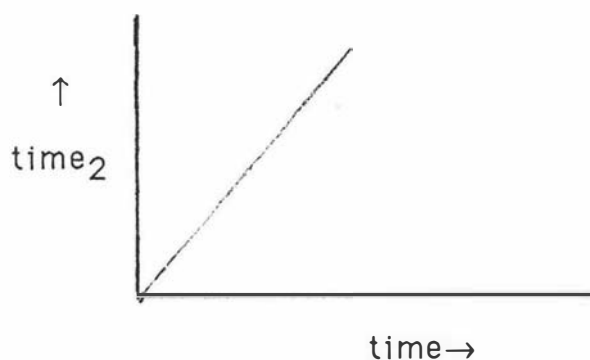
In the static diagram this point is easily overlooked, because

the diagrammatic objects that represent events at different moments all exist simultaneously, and this makes it natural to look for qualities of pastness, presentness and futurity as merely *relative qualities*, since there are no absolute qualities depicted which correspond to them.

For a final example, consider the common argument that any representation of real *change* requires an infinity of temporal dimensions. Suppose to begin with that we depict the sequence of temporal states of the world, and indicate which state is *present* or *now*, as in the following diagram, where each point on the 'time line' represents a state of the world:



This does not yet represent the dynamic nature of time, because there is no representation of the fact that the present *changes*. To represent change, another time dimension may be introduced, as depicted in the following diagram:



Here the dimension $time_2$ has been added so that the *change* of the now can be represented - it is seen that as $time_2$ gets later and later, the now gets later and later (along the dimension *time*).

But even apart from the peculiarity of having two time dimensions, we are no closer to representing *change*. For the 'change' of the now on the dimension *time* relies on change of a second 'now' in the second dimension $time_2$. But no representation of change of the now of $time_2$ has been given: hence we need to postulate a third time dimension, $time_3$, to represent this. Thus we get an infinite regress of dimensions of time, without ever getting any closer to anything truly dynamic in our representation.

This problem dissolves when the analogue representation of time is considered. Consider the film, depicting a process, which plays in real time. Here *change* in the depicted process is represented by *real change in the representation*- i.e. the changing of the image on the screen. Real change is genuinely depicted without a need for an infinity of time-dimensions, and

without paradox. Of course one might object that using time itself in the depiction of time flow is circular, since it gives no *analysis* of what time flow is. This is true, but the point being made here is simply that a depiction of time flow is possible. In Chapter Seven a logical model of time flow is given, which comprises an analytical representation of time flow.

If we carried movie projectors around with us, and whenever someone demanded to see a representation of time flow we showed them a movie running in real time, a great many popular objections to the cogency of time flow would evaporate.

APPENDIX 4.1

THE DEPENDENCE OF REVERSIBILITY ON INTERPRETATION.

In Section 4.6 it is argued that the time reversal transformations for a theory T are well defined once a definite class of *fundamental variables* have been chosen. It is assumed that t (time) is always a fundamental variable, and if the other fundamental variables are v_1, \dots, v_n , then the time reversal transformations are: $t \rightarrow -t, v_1 \rightarrow v_1, \dots, v_n \rightarrow v_n$.

This raises the important problem of how the class of fundamental variables is determined. What are the rules for choosing the class of fundamental variables? Some rules are needed to ensure mathematical completeness and consistency: basically, the class of fundamental variables must be adequate to define all other variables, and must have no redundancies. But this requirement does not suffice to determine a unique class of fundamental variables for a given theory. For instance, it seems normal and natural in classical mechanics to take time (t), force (F), position (r) and mass (m) as fundamental variables. (Specifying the values of these variables for a particle over period completely defines the instantaneous state of the particle at each moment). But an odd new variable, such as: $\beta = m + dr/dt$ (mass plus velocity) could be defined, and the alternative class of $t, F, r,$ and β would be just as *mathematically* adequate for the definition of the theory (since mass is extractable from β by the

equation: $m = \beta - dr/dt$). How are we to decide between two alternative classes of variables which are both mathematically adequate?

Since the class of fundamental variables determines the ontology, this is really the question of how the *interpretation* is to be decided upon. There seems to be no answer to this question. So long as two alternative interpretations are adequate in certain respects, there seems to be nothing else (except our own convenience) that we can use to decide between them, and so in the end the choice of interpretation appears to be a conventional matter.

This raises a very serious problem, alluded to a number of times by Earman (see his [1974], p.24.) Suppose that there are two different sets of variables, V and W , which are both mathematically adequate to serve as sets of fundamental variables for the definition of a certain theory. Let V be a class of variables: t, v_1, \dots, v_n , and in terms of these variables let the theory be formulated as: $T_V(t, v_1, \dots, v_n)$. Let W be a class of variables: t, w_1, \dots, w_m , and in terms of these variables let the theory be formulated as: $T_W(t, w_1, \dots, w_m)$. The adequacy of the two interpretations means of course that: $T_V(t, v_1, \dots, v_n) = T_W(t, w_1, \dots, w_m)$.

Now on the assumption that V is the appropriate class of fundamental variables, the time reversal transformations are: $t \rightarrow -t, v_1 \rightarrow v_1, \dots, v_n \rightarrow v_n$, and the time reversal of the theory appears as: $T_V(-t, v_1, \dots, v_n)$. On the other hand, if W is taken as the

appropriate class of fundamental variables, the time reversal transformations are: $t \rightarrow -t$, $w_1 \rightarrow w_1$, ..., $w_n \rightarrow w_n$, and the time reversal of the theory appears as: $T_W(-t, w_1, \dots, w_n)$. But what if it turns out that $T_V(t, v_1, \dots, v_n) = T_V(-t, v_1, \dots, v_n)$, so that $T_V(t, v_1, \dots, v_n)$ is formally *reversible*, while: $T_W(t, w_1, \dots, w_n) \neq T_W(-t, w_1, \dots, w_n)$, so that $T_W(t, w_1, \dots, w_n)$ is formally *irreversible*? The reversibility of the theory T would then appear to depend upon the interpretation taken: if the choice between the two interpretations (V and W) is merely conventional, then reversibility would be a merely conventional matter. Reversibility would effectively become merely a feature of the *formalism* chosen to express the theory, rather than an objective feature of the ontology (there would be no unique ontology). It need hardly be emphasized what a serious problem this would be for the notion of reversibility.

A simple example will show what a real problem this is. Consider the Newtonian law: $F = m \cdot d^2r / (dt)^2$ as a simple theory. In this formulation, the fundamental variables are naturally enough taken to be t , F , r and m . The time reversal transformations are thus: $t \rightarrow -t$, $F \rightarrow F$, $r \rightarrow r$, $m \rightarrow m$, and the theory is reversible, since its time reversal is: $F = m \cdot d^2r / (d(-t))^2$, which is identical to the original theory, $F = m \cdot d^2r / (dt)^2$. However, this law could be formulated in an alternative set of fundamental variables: consider the set t , F , r and β , where β is the new variable defined above by: $\beta = m + dr/dt$. Since $m = \beta - dr/dt$, the theory is now formulated as: $F = (\beta - dr/dt) \cdot d^2r / (dt)^2$. The time reversal transformations for the new set of fundamental

variables are: $t \rightarrow -t$, $F \rightarrow F$, $r \rightarrow r$, $\beta \rightarrow \beta$, thus the reversal of the theory is: $F = (\beta - dr/d(-t)) \cdot d^2r/(d(-t))^2$, which is: $F = (\beta + dr/d(t)) \cdot d^2r/(dt)^2$, and clearly quite different to the original theory. Hence, on the adoption of this alternative set of fundamental variables, the theory is *irreversible*.

The only possible solution to this problem is that, for some reason, the second choice of fundamental variables is inadequate, and can be ruled out. If this is not so - if it is mere convention that chooses between the two sets of fundamental variables, or two interpretations - then the whole notion of reversibility is on the rocks.

What could be wrong with the second interpretation? The first thought is that the variable β is not defined *independently* of other fundamental variables, particularly time. It will be remembered from Section 4.6 that the feature about fundamental variables that persuaded us to adopt their time reversal transformation as the identity transformation was that they are *independent of time*. Because of this, they cannot depend upon the metrisation of time, and hence altering the metrisation (time reversal) must leave them unchanged.

But why is it claimed that β is not defined independently of time? It cannot simply be that it is introduced by the formulae (conceptual definition): $\beta = m + dr/dt$, because this could be equivalently taken as a conceptual definition of *mass* in terms of β , r and t , rather than of β in terms of m , r and t . That is to say, someone committed to holding that β was really fundamental and

m was not would run the same argument against m.

But I think there is still something in this notion that β is not defined independently of time. The point seems to be that β is not defined in *practical* terms independently of time, while mass is. Take a particular particle at a particular moment. In a Newtonian world, it has a definite and objective mass, which is independent of the metrisation of time. But does it have a definite value of β independently of the metrisation of time? No, for the only way we have of *measuring* the value of β is by measuring its components m and dr/dt , and the latter depends upon the metrisation of time (including the direction of the metrisation). Thus β cannot be taken as fundamental, if being 'fundamental' implies taking a value independently of the metrisation of time.¹

It would be difficult to spell this argument out in a completely formal or general way, and I will not try to do so here. Instead I will assume that it contains the central truth of the matter, and I will try to encapsulate this truth in the following additional condition for the adequacy of a set of fundamental variables:

¹Alternatively, one might take β as 'fundamental', but hold that since its values reverse on time reversal, the correct time reversal transformation is: $\beta \rightarrow -\beta$, rather than $\beta \rightarrow \beta$. But I will continue to assume that the reversal transformations of fundamental variables must be the identity transformation. The key point is that we must look to the epistemology of β , i.e. the procedure for measuring it, to realise that: $\beta \rightarrow -\beta$ is the appropriate transformation.

Translatability Condition. If $V = \{v_1, \dots, v_n\}$ and $W = \{w_1, \dots, w_m\}$ are both adequate sets of fundamental variables (for a given theory), then for each i , there is a function F_i such that it is logically necessary that : $v_i = F_i(w_1, \dots, w_m)$. (Where the functions F_i are expressible, we can formulate definitions of the v_i 's in terms of $\{w_1, \dots, w_m\}$.)

This is a very strong condition: it means that *time* need play no part in the conceptual definitions relating v_i 's to w_j 's. This would rule out the possibility that both t, F, r, m and t, F, r, β are adequate sets of fundamental variables, since β cannot be defined in terms of just F, r and m . (Equivalently m cannot be defined in terms just of F, r and β).²

In this form the condition may be too strong, but it has considerable plausibility. For consider an object which takes a momentary state *at a moment* t , expressed, in the system of W -variables, by a set of values assigned to these variables. If the system of W -variables is equivalent to the V -variables, then this set of values alone must suffice to determine the description of this state in terms of V -variables. But this is so only if

²It is interesting to note that this postulate rules out t, F, r and p (momentum) as an adequate set of fundamental variables, given that t, F, r, m is an adequate set, since t occurs in the definition: $p = m \cdot dr/dt$. And this is fortunate, since (i) t, F, r and p do not suffice for the definition of t, F, r, m , since although generally $m = p / (dr/dt)$, where $dr/dt = 0$, m is left undefined by this equation, (ii) taking t, F, r, p as the fundamental variables of classical mechanics would make it an *irreversible* theory.

translations of the form required by the above condition exist.

However, although I think the proposed condition is defensible, it is too difficult to go into properly here, and I will not try to elaborate on it further. Instead I will assume that it is along the right general lines, and consider its implications.³

The problem of dependence of reversibility on interpretation is solved by this condition. Suppose that t, v_1, \dots, v_n and t, w_1, \dots, w_m meet this condition, and are adequate to define a theory: $T = T_V(t, v_1, \dots, v_n) = T_W(t, w_1, \dots, w_m)$. It is easily shown that if: $T_V(t, v_1, \dots, v_n) = T_V(-t, v_1, \dots, v_n)$ (i.e. T is reversible in the form: $T_V(t, v_1, \dots, v_n)$), then: $T_W(t, w_1, \dots, w_m) = T_W(-t, w_1, \dots, w_m)$ (T is reversible in the form: $T_W(t, w_1, \dots, w_m)$.) PROOF: Suppose:

$$(1) T_V(t, v_1, \dots, v_n) = T_W(t, w_1, \dots, w_m).$$

$$(2) T_V(t, v_1, \dots, v_n) = T_V(-t, v_1, \dots, v_n), \text{ and:}$$

$$(3) v_i = F_i(w_1, \dots, w_m), \text{ for each } i.$$

Then:

$$(4) T_V(t, v_1, \dots, v_n) = T_V(t, F_1(w_1, \dots, w_m), \dots, F_n(w_1, \dots, w_m))$$

(3, substitution)

$$(5) T_W(t, w_1, \dots, w_m) = T_V(t, F_1(w_1, \dots, w_m), \dots, F_n(w_1, \dots, w_m))$$

(1 and 4)

$$(6) T_W(-t, w_1, \dots, w_m) = T_V(-t, F_1(w_1, \dots, w_m), \dots, F_n(w_1, \dots, w_m))$$

(5, subst.)

³I realise the proposed condition is speculative, but it will be seen that some condition along these general lines is necessary if the problem of relativity of reversibility to interpretation is to be solved. I think my formulation of this condition therefore illustrates what the problem is.

$$(7) T_V(-t, F_1(w_1, \dots, w_m), \dots, F_n(w_1, \dots, w_m)) = T_V(-t, v_1, \dots, v_n) \quad (3, \text{subst.})$$

$$(8) T_W(t, w_1, \dots, w_m) = T_W(-t, w_1, \dots, w_m) \quad (1, 2, 6, 7).$$

Thus the new Translatability Condition on interpretations guarantees the desired result that *if a given theory is reversible in one adequate interpretation it must be reversible in all adequate interpretations*. I will call this the *independence of reversibility from interpretation*.⁴

So much for the successful side of the Translatability Condition - let us now consider some problems that remain. The main problem is that the Translatability Condition only tells us whether possible interpretations are *compatible*; i.e. it effectively partitions the class of possible interpretations into equivalence classes. But there is still no indication as to which class of interpretations is *correct*. For instance, the Translatability Condition means that the Classical interpretations based on the classes of fundamental variables: t, F, r, m and t, F, r, β are incompatible, i.e. they give rise to different theories. But which is correct? This is a vital question, since if the first is adopted, classical mechanics is reversible, whereas if the second

⁴The translatability condition is really a special case of the following more general condition: *Where at least one v_j has a definition: $F_j(w_1, \dots, w_m)$ (i.e. a definition in which t plays no part), then, for each i : $F_i(t, w_1, \dots, w_m) = F_i(-t, w_1, \dots, w_m)$.* This is a more general condition that appears to guarantee independence of reversibility from interpretation, although I will not prove this here.

is adopted, it is irreversible.

I have no answer to offer to this problem. Physicists usually have clear and strong intuitions about interpretations - all would choose t, F, r, m rather than t, F, r, β as the fundamental variables for classical mechanics, for instance. For the moment it seems that we can only hope that these intuitions have an objective basis. If they do not, then not only our concept of reversibility, but probably our whole concept of physical ontology is under serious threat.

My comments in this Appendix are clearly not complete, but I hope they at least provide some illumination of the depth of the problem. I feel, along with Earman but precious few others, that the questions raised here about the time reversal transformations are fundamental, and need to be given proper answers before a firm foundation for the subject is provided.⁵ In the meantime we have no choice but to trust our intuitions.

The following appendix has a close connection with this one, discussing the time reversal transformations for quantum mechanics.

⁵See Oddie [1986, Ch.7] for an extended discussion of what is essentially the same problem of relativity of truthlikeness, structure, confirmation, etc, with respect to interpretation.

APPENDIX 4.2

TIME REVERSAL FOR QUANTUM STATES

Although the time reversal transformation for quantum states is well agreed upon, it has a disturbing peculiarity, which has never been very satisfactorily explained. Very briefly, the obvious transformation to effect time reversal of the quantum state vector, $|\psi\rangle$, is the linear, unitary operator which transforms: $t \rightarrow -t$, and operators: $r \rightarrow r$, $p \rightarrow -p$. I will denote this operator by R^* . (Or in line with the treatment of reversal in Chapter Four, if: $\Psi(t)$ gives the state function for a system as a function of time, then: $\Psi^{R^*}(t) = \Psi(-t)$.) But adopting R^* as the time reversal operator has a problem. Quantum theory postulates that the state vector obeys the Schrodinger equation: $H|\psi\rangle = i\partial|\psi\rangle/\partial t$. But it is easily shown that such a state transformed by R^* instead obeys: $H|\psi^{R^*}\rangle = -i\partial|\psi^{R^*}\rangle/\partial t$. (See any standard text). Thus the adoption of R^* as the time reversal operator for the quantum state vector means the Schrödinger equation fails for time reversed state vectors.

The solution is to modify the operator R^* to an operator: $R = KR^*$, where K is an operator which carries out complex conjugation. The Schrödinger equation holds for time reversed state vectors under the adoption of R as the reversal operator.

What is problematic is why the transformation R is chosen instead of R^* . It may appear at first that the operator R has just been chosen conventionally as the time reversal operator with the explicit aim of rendering the Schrödinger equation reversible. This would mean that the reversibility of the Schrödinger equation is not objective, but merely depends upon a conventional interpretation of the time reversal operator. If this is so it is hardly a scientific fact that the Schrödinger equation is reversible.¹

Davies [1974] passes off the switch from R^* to R by cryptically observing that "Nevertheless, a solution of the Schrödinger equation is not itself observable, and the symmetry is restored by simply reversing the sign of i ..." (p.156). He then observes that R is anti-unitary, so that $|\langle \psi^R | \phi^R \rangle| = |\langle \psi | \phi \rangle|$, thus " R leaves the physical content of quantum mechanics unchanged". I will try to bring out what underlies these comments.

Firstly, let us think of the time reversal of $|\psi\rangle$ as the description of $|\psi\rangle$ as it would appear in a reversed time metric. What properties must this time reversed version of $|\psi\rangle$ have? Intuitively (and a little crudely put): (i) it should have the same *energy* as $|\psi\rangle$ itself, (ii) it should go through the reversed sequence of 'positions', (iii) it should have the inverse of the momentum of $|\psi\rangle$. It turns out that the object represented by $|\psi^R\rangle$

¹Into the bargain there is the mathematical peculiarity that R is non-linear. Watanabe [1955c] makes an interesting attempt to reformulate the quantum state function so that the time reversal operator becomes linear, although I do not think there is any hope for his proposal to succeed.

has exactly these properties (or rather, the appropriate expected values for the measurements of them). This is what is implied by the identity: $|\langle \psi^R | \phi^R \rangle| = |\langle \psi | \phi \rangle|$.

What this fundamentally means is that *the state function, $|\psi\rangle$, is not being taken as the fundamental physical object* – for if it were, we would have to interpret directly what its time reversal would be, and we could only arrive at the answer that it is given by: $|\psi^{R*}\rangle$. Instead, what is being taken as physically real are the values of observable quantities such as energy, position, and momentum, or more exactly, the *probabilities* (or probability amplitudes) for the system to be found to have various values of these quantities.

This clearly results from the *interpretation* that we are taking of quantum theory. The interpretation implicit in the treatment of time reversal means that the state vector $|\psi\rangle$ is not being treated as real in itself, but is taken as a mathematical device representing some other reality. The controversy over the interpretation of quantum theory means that there is no complete consensus about what the 'reality' that $|\psi\rangle$ represents is. But there is enough consensus for physicists to agree on the nature of the time reversal operator R . This agreement seems to imply that what is taken as real are the values of the *probability amplitudes* for the outcomes of certain measurements, and this makes it obvious why the state function $|\psi\rangle$ is not taken as directly real itself. This of, course, squares perfectly with the orthodox views of quantum theory.

This position reinforces the view taken in the previous

appendix: reversibility is dependent on interpretation. The interpretation that renders the Schrödinger equation reversible (giving R as the reversal operator) is *not* chosen arbitrarily, as first seemed, but is chosen to reflect the desired interpretation, roughly that values (or expected values) of observables such as energy, position, momentum, etc, are primarily real, and the state function itself only an indirect representation of them.

But although this is reasonably convincing, problems remain to be answered before we can be satisfied with this treatment. In particular, why is it thought that, in the time reversed frame of reference, energy and position are invariant while momentum reverses? The obvious reason is the *analogy with classical mechanics*, which provides so much of our insight into many aspects of quantum mechanics. But is the analogy a good one here? A serious *disanalogy* between classical and quantum theory on this point is that, whereas in classical mechanics the concept of the time reversal of a *measurement* makes good sense, in quantum mechanics there is no readily apparent sense to it. To make a classical measurement of *mass*, for instance, we might observe an object displacing the pointer of a spring-balance by a certain amount. This event would constitute the same observation of mass whether we considered it 'forwards' or 'backwards' in time. (I.e. we could run a film of this measurement event or process in either direction, and we would make the same inference about the mass of the object from what we saw.) I think a little thought about the reversibility of classical physics will show that this holds for any classical measurement

whatsoever. (Many processes will be exceedingly hard to interpret when considered in reverse, but in principle, the idea of *measurement* makes sense in either direction of time in classical physics.)

But along with most others, I cannot easily imagine what 'measurement in reverse' *means* in quantum physics. This brings into doubt whether the time reversed state vector can be considered to have well-defined properties of energy, momentum, and so forth. If not, then it is incorrect to characterise the reversal operator as R.

This is a problem which cannot be solved without a definitive interpretation of quantum theory. The lack of such an interpretation has understandably dissuaded some from the discussion of time reversal in quantum theory.² However rather than abandoning the discussion, I continue it under the assumption that the orthodox interpretation of the reversal operator is correct.

²E.g. Earman [1974] p.280. See also Penrose [1979] p.583-586 for a related problem about the interpretation of the time reversal of a quantum process that includes a measurement. Note that the *irreversibility* that I claim for quantum theory does *not* depend upon any theory of the irreversibility of measurement, but only upon the contention that quantum theory is irredeemably *probabilistic*. Of course, probabilities are traditionally thought to arise *only* from measurements, but apart from being implausible, that makes no difference to my claim.

BIBLIOGRAPHY

- Aharanov, Y. Bergmann, P. and Lebowitz, J. 1964 "Time Symmetry in Quantum Process of Measurement", *Phys. Rev.*, **134** (6B), (June 1964). B1410-B1418.
- Arronson, B.S. 1970 "Time, time stance, and existence" In Frazer et alia, 1972.
- Aspect, A. Dalibard, J. and Roger, G. 1982 *Phys.Rev.Letters* **49** 1804.
- Atkins, P.W. 1986 "Time and dispersal: the second law" In Flood, R., and Lockwood, M. (eds) 1986.
- deBeauregard, O.C. 1977 "Two Lectures on the direction of time", *Synthese* **35** 129-154.
- 1970 "No paradox in the theory of time anisotropy" In Frazer et alia, 1972.
 - 1966 "Time in relativity theory: Arguments for a philosophy of being" In Frazer.(ed) 1966.
- Benjamin, A.C. 1966 "Ideas of time in the history of philosophy" In Frazer (ed) 1966.
- Belinfante, F. 1973 *A Survey of Hidden Variable Theories*. Pergamon Press.
- Beltrametti, E. and Cassinalli, G. 1981 *The Logic of Quantum Mechanics*. Addison-Wesley.

- Bergson, Henri. 1913 "Duration and Intuition", in Smart, 1959.
(Extract from *An Introduction to Metaphysics*, 1913,
trans. T.E.Hulme, Macmillan and Co.)
- Black, M. 1958-9. "The 'direction' of time." *Analysis*, 19, 54-63.
- Bohm, D. 1980 *Wholeness and the Implicate Order* Routledge and Kegan.
- Bohm, D. and Bub, J. 1966 "A proposed solution of the measurement problem in quantum mechanics by a hidden variable theory." *Rev.Mod.Phys.* 38 (3) 453-469.
- Bondi, H. 1952 "Relativity and indeterminacy", *Nature* 169, 660.
- Bradley, F.H. 1930 "The unreality of space and time", in Smart, 1964. (Extract from *Appearance and Reality*, 1930).
- Brandon, S.G.F. 1970 "The deification of time" In Frazer et alia, 1972.
- von Bretzel, Philip. 1977 "Concerning a probabilistic theory of adequate causation for the causal theory of time." *Synthese*, 35, 173-190.
- Broad, C.D. 1938, *Examination of McTaggart's Philosophy* Vol.2, Part 1. Cambridge.
- 1938a "Ostensible Temporality", extract from Broad [1938], in Smart, 1964.
 - 1959 "Reply to my critics", in *The Philosophy of C.D.Broad*, ed. P.A.Schilpp.
- Bunge, Mario. 1970 "Time asymmetry, time reversal, and irreversibility" In Frazer et alia, 1972.
- Cápek, M. 1961 *The Philosophical Impact of Contemporary Physics*, Van Nostrand.

- 1970 "The fiction of instants" In Frazer et alia, 1972.
 - 1966 "Time in relativity theory: arguments for a philosophy of becoming" In Frazer (ed) 1966.
- Cassirer, E. 1970 "On the reality of becoming" In Frazer et alia, 1972.
- Chapman, T. 1982 *Time: A philosophical Analysis*. D.Reidel.
- Christiansen, Ferrel. 1976 "The source of the river of time", *Ratio* 18-19, 131-144.
- Clauser, J.F. and Shimony, A. 1978 *Rep.Prog.Phys.*41. 1881.
- Cocke, W.J. 1967 "Statistical time symmetry and two-time boundary conditions in physics and cosmology" , *Phys. Rev.* , 160 (5), 1165-1170.
- Cooper, L.N. 1973 "Wavefunction and observer in the quantum theory", in J.Mehra (ed.) *The Physicist's Conception of Nature*. D.Reidel .
- Cunningham, E. 1914 *The Principle of Relativity*. Cambridge.
- Davies, P.C.W. 1974 *The Physics of Time Asymmetry*. Surrey University Press.
- 1986 "Time asymmetry and quantum mechanics" In Flood, R., and Lockwood, M. (eds) 1986.
- Denbigh, K.G. 1970 "In defence of the direction of time", In Frazer et alia, 1972.
- 1981 *Three Concepts of Time* Springer-Verlag.
- Dobbs, H.A.C. 1969 "The 'Present' in physics", *Brit.J.Phil.Sc.* 19, 317-324.
- 1970 "The dimensions of the sensible present", In Frazer et alia, 1972.

- Dresden, M. 1986 "Partical physics and the interpretation of quantum theory." In Roth (ed.) [1986].
- Earman, John. 1967 "Irreversibility and temporal asymmetry", *J.Phil.*, LXIV (18), (Sept.1967) 543-549.
- 1969 "The anisotropy of time", *Aust.J.Phil.* 47(3), 273-295.
 - 1974 "An attempt to add a little direction to 'the problem of the direction of time' ", *Phil. Science*, 41, 15-47.
 - 1986 *A Primer on Determinism*. D.Reidel.
- Eddington, A.S. 1928 *The Nature of the Physical World*. Collins.
- Ehrenfest, P. and T. [a] "Begriffliche grundlagen der statistischen auffassung in der mechanik" *Encyklopädie der mathematischen Wissenschaften*, IV,2,11.
- Elliott, J.P., and Dawber, P.C. 1979 *Symmetry in Physics. (Vol 1 and 2.)* Oxford.
- Findlay, J.N. 1941 "Time: a treatment of some puzzles", in Smart, 1964.
- Fizgerald, P. 1969 "The truth about tomorrow's sea fight." *J.Phil.* 66 307-329.
- 1985 "The crash in market futures" *J.Phil* 82 (10) 560-61.
- Flood, R., and Lockwood,M. (eds) 1986 *The Nature of Time* Basil Blackwell.
- van Frassen, Bas C. 1980 "A temporal frame work for conditionals and chance" in W.L.Harper, et alia, *Ifs* D.Reidel.
- Frazer,J.T., Haber,F.C. and Muller,G.H. (eds) 1972 *The Study of Time* Springer -Verlag.

- (ed) 1966 *The Voices of Time* George Braziller.
- Ghirardi, G.C., Rimini, A. and Weber, T. 1980 "A general argument against superluminal transmission through the quantum mechanical measuring process". *Lettre al Nuovo Cimento*, 27 293-298.
- Gödel, K. 1949. "A remark about the relationship between relativity theory and idealistic philosophy." *Albert Einstein: Philosopher-Scientist*, (Ed. Schlipp, P.), Evanston 1949.
- Gold, T. 1962 "The Arrow of Time", *Am.J.Phys.* 30, 403-410.
- Goodman, Nelson. 1951 "Time and language, and the passage of time.", in Smart, 1964 (From The structure of appearance, Harvard University Press)
- Grünbaum, Adolf. 1964 "Time, irreversible processes, and the physical status of becoming", in Smart 1964. (From The philosophy of Rudolf Carnap, ed. P.A.Schilpp, Open Court.)
- 1969 "Are physical events themselves transiently past, present and future? A reply to H.A.C.Dobbs", *B.J.Phil.Sc.* 20, 145-162.
- 1970 "The meaning of time" in *Essays in Honour of Carl G. Hempel* (ed. Rescher), 1970, Reidel.
- 1973 *Philosophical Problems of Space and Time* D.Reidel.
- Hamblin, C.L. 1972 "Instants and intervals" In Frazer et alia, 1972.
- Horwich, Paul. 1987 *Asymmetries in Time*, MIT Press.

- Husserl, E. 1928 *The Phenomenology of Time-Consciousness*, trans. James S.Churchill. Indiana University Press.
- Jammer, M. 1974 *The Philosophy of Quantum Mechanics*. John Wiley & Sons.
- Jordan, T.F. 1983 "Quantum correlations do not transmit signals." *Phys.Letters*,94A 263.
- Kripps, H. 1989 "The objectivity of quantum probabilities." *Aust.J.Phil.* 67 (4) 423-431.
- Kroes, P. 1986 *Time, its Structure and Role in Physical Theories*.
- Lakatos, I. 1981 *The Methodology of Scientific Research Programmes: Philosophical Papers Volume I*. (ed. Worrall and Currie) Cambridge University Press.
- Landsberg, P.T. 1970 "Time in statistical physics and special relativity" In Frazer et alia, 1972.
- 1982 *The Enigma of Time* Adam Hilger.
 - 1982a "Thermodynamics, cosmology, and the physical constants", in Landsberg, P.T. (ed) 1982.
- Leggett, A.J. 1982 "The 'arrow of time' and quantum mechanics", in Landsberg, P.T. (ed) 1982.
- Lewis, Gilbert N. 1930 "The symmetry of time in physics", in Landsberg, P.T. (ed) 1982.
- Ludwig, G. 1985 *Foundations of Quantum Mechanics I and II*. (Trans. C.A.Hein). Springer-Verlag.
- Matthews, G. 1978. "Time's arrow and the structure of spacetime." *Phil.Sc.* 46 82-97.

- Maxwell, N. 1988 "Quantum propensiton theory: a testable resolution of the wave/particle dilemma." *Brit.J.Phil.Sc.* 39, 1-50.
- 1985 "Are probabilism and special relativity incompatible?" *Phil.Sc.* 52, 23-43.
- Mays, W. 1970 "Whitehead and the philosophy of time" In Frazer et alia, 1972.
- McCall, Storrs. 1976 "Objective time flow", *Phil.Sc.*, 43. 337-362.
- McTaggart, J.E. 1908 "The unreality of time", *Mind*, new series, 17, 457-474.
- Mehlberg, H. 1980 *Time, Causality, and the Quantum Theory*. Vols. 1 and 2. Reidel.
- Mellor, D.H. 1981 *Real Time*. Cambridge.
- Michon, J.A. 1972 "Processing of temporal information and the cognitive theory of time experience" In Frazer et alia, 1972.
- Milne, P. 1986 "Can there be a realist single-case interpretation of probability?" *Erkenntnis* 25 129-132.
- Mirman, R. 1975 "The direction of time", *Found.Phys.* 5 (3), 491-511
- Morrison, P. 1965 "Time's arrow and external perturbations", in Landsberg, P.T. (ed) 1982.
- Mundle, W.C.K. 1959 "Broad's views about time", in *The Philosophy of C.D.Broad*, ed. P.A.Schilpp.
- Ne'eman, Y. 1982 "Time reversal asymmetry at the fundamental level - and its reflection on the problem of the arrow of

- time" in Landsberg, P.T. (ed) 1982.
- Newton, I. 1962 *The Mathematical Principles of Natural Philosophy*. Trans. A.Motte, revised F.Cajori. University of California Press.
- North, J.D. 1970 "The time coordinate in Einstein's restricted theory of relativity" In Frazer et alia, 1972.
- Oddie, G. 1986 *Likeness to Truth*. D.Reidel.
- Park, D. 1970 "The myth of passage" In Frazer et alia, 1972.
- 1986 "Time in quantum mechanics", in Roth (ed) [1986].
- Pears, D.F. 1959 "Time, truth and inference", in A.Flew (ed) *Essays in Conceptual Analysis* Macmillan. 228-52
- Penfield, Robert. 1966 "More on the arrow of time", *Am.J.Phys.* **34**, 422-426.
- Penrose, 1979 "Singularities and time-asymmetry", in Hawking and Israel (ed.) *General Relativity*. Cambridge.
- 1986 "Big bangs, black holes, and 'time's arrow'". In Flood, R., and Lockwood, M. (eds) 1986.
- Piaget, Jean. 1966 "Time perception in children" In Frazer (ed) 1966.
- Popper, K.R. 1956 "The arrow of time", *Nature*, **177** (4507) (March 17), 538.
- Price, Hew. 1984 "Affecting the past", *Synthese*, **61**, 299-323.
- Prior, A.N. 1968 *Time and Tense*, Oxford.
- 1959 "Thank goodness thats over", *Philososophy* **34**, 12-17.
 - 1970 "The notion of the present" In Frazer et alia, 1972.
 - 1958 "Time after time" *Mind* **67**, 244-46

- Putnam, H. 1967 "Time and physical geometry" *The Journal of Philosophy* 64 240-247.
- Quine, W.V. 1936 "Truth by convention", in *Philosophical Essays for A.N. Whitehead* (ed. O.H.Lee), Longmans, Green and Company.
- 1951 "Two dogmas of empiricism", *Phil.Rev.*, 60.
 - 1960 "Time", in Smart, 1964. (From *Word and Object*, John Wiley and sons.)
 - 1981 *Theories and Things*. Belknap Press.
- Reichenbach, H. 1947. *Elements of Symbolic Logic*. Macmillan.
- 1953 "Les fondements logiques de la mécanique des quanta", *Annales de l'Institut Henri Poincaré* 13.
 - 1956 *The Direction of Time*. (ed. M. Reichenbach) University of California Press.
- Roth, L. and Inomata, A. 1986 *Fundamental Questions in Quantum Mechanics*, Gordon and Breach, 1986.
- Russell, B. 1915 "On the experience of time." *The Monist* 25.
- Sachs, Mendel. 1985 "On Einstein's later view of the twin paradox" *Found.Phys* 15 (9) 977-980.
- Schiller, R. 1962. "Deterministic interpretations of the quantum theory." In Wartofsky (ed) [1962]. *Boston Studies in the*
- Schilpp, P.A. 1959 *The Philosophy of C.D.Broad*. Tudor.
- Schlesinger, George N. 1980 *Aspects of Time* Hackett.
- Schrodinger, E. 1950 "Irreversibility", *Proc.Irish Royal Acad*, 53(A) 189-195.
- Schuster, M.M. 1986 "Is the flow of time subjective?" *Review of Metaphysics* 39 (June 1986) 695-174.

- Seddon, Keith. 1987 *Time; a philosophical treatment* Croom Helm.
- Shallis, Michael. 1986 "Time and cosmology", In Flood, R., and Lockwood, M. (eds) 1986.
- Shoemaker, Sydney. 1969 "Time without change", *J.Phil.* **66** (12), June 19, 1969, 363-381.
- Silberstein, L. *The Theory of Relativity*. Macmillan.
- Smart, J.J.C. (ed) 1954 "The temporal asymmetry of the world", *Analysis*, **3**, 79-83.
- 1958 "Critical study: 'The Direction of Time', by Hans Reichenbach" *Phil.Quart.* **8**, 72-77.
 - 1963 *Philosophy and Scientific Realism*, Routledge and Keegan Paul.
 - 1964 *Problems of Space and Time*, Collier-Macmillan.
 - 1987 *Essays Metaphysical and Moral*, Basil Blackwell.
 - 1987a "Time and becoming" in Smart, 1987.
- Sklarr, L. 1974 *Space, Time, and Spacetime*. University of California Press.
- 1977 "What might be right about the causal theory of time" *Synthese* **35**, 173-190.
 - 1981 "Up and down, left and right, past and future" *Nous*, **XV**: 1 (March, 1981), 111-129.
- Sudbery, A. 1984 "The observation of decay." *Annals of Physics*, **157**, 512-536.
- Swinburn, R (ed). 1983 *Space, Time and Causality*. D.Reidel.
- Taylor, J.G. 1970 "Time in particle physics" In Frazer et alia, 1972.

- Taylor, R. 1955 "Spatial and temporal analogies and the concept of identity", in Smart 1964.
- Thomason, R.H. 1970 "Indeterminist time and truth-value gaps" *Theoria* 36 264-81
- Tichý, P. 1980 "The transiency of truth." *Theoria* XLVI 162-182.
- 1988 *The Foundations of Frege's Logic*. Walter de Gruyter.
- Wartofsky, M. 1962. *Proceedings of the Boston Colloquium for the Philosophy of Science, 1961/62*. Reidel.
- Watanabe, S. 1955 "Symmetry of physical laws. Part 1. Symmetry in space-time and balance theorems.", *Rev. Mod. Phys* 27, 26-39.
- 1955a "Symmetry of physical laws. Part 2. Q-number theory of space-time inversions and charge conjugation.", *Rev. Mod. Phys.* 27, 40-76.
 - 1955a "Symmetry of physical laws. Part 3. Prediction and Retrodiction", *Rev. Mod. Phys* 27, 179-186.
 - 1965 "Conditional probability in physics." *Suppl. Prog. Theor. Phys (Kyoto) Extra Number (Commemorative Issue for the 30th Anniversary of the Meson Theory by Dr. H.Yukawa)*, 135-167.
 - 1966 "Time and the probabilistic view of the world" in J.T.Fraser (ed) *The Voices of Time* (1966) 527-564.
 - 1970 "Creative time" In Frazer et alia, 1972.
- Webb, C.W. 1977 "Could space be time-like?" *J.Phil.* 74, 463-474.
- Weyl 1949 *Philosophy of Mathematics and Natural Science*. Princeton University Press.

- Whitrow, G.J. 1961 *The Natural Philosophy of Time* Clarendon, 1980.
- 1970 "Reflections on the history of the concept of time", In Frazer et alia, 1972.
- Whyte, L.L. "One-way processes in physics and biology" *Brit.J.Phil.Sc.* 6.
- Wigner, Eugene P. 1963 "The problem of measurement", *Am.J.Phys.* 31, 6-15.
- Yourgrau, Palle. 1985 "On the logic of indeterminist time", *J.Phil.* 82 (10), Oct.1985.
- Zenzen, M. 1977 "Popper, Grünbaum and *de facto* irreversibility" *Brit.J.Phil.Sc.* 28, 313-57
- Zimmerman, E.J. "Time and the quantum theory" In Frazer (ed) 1966.
- Zwart, P.J. 1972 "The flow of time" *Synthese* 24 133-57