TIME, SYMMETRY AND STRUCTURE: A STUDY IN THE FOUNDATIONS OF QUANTUM THEORY

by

Bryan W. Roberts

Submitted to the Graduate Faculty of

the Department of History & Philosophy of Science

in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2012

UNIVERSITY OF PITTSBURGH

HISTORY & PHILOSOPHY OF SCIENCE

This dissertation was presented

by

Bryan W. Roberts

It was defended on

May 30th, 2012

and approved by

John Earman, PhD, History & Philosophy of Science

John D. Norton, PhD, History & Philosophy of Science

Robert Batterman, PhD, Philosophy

Laura Ruetsche, PhD, Philosophy (University of Michigan)

Giovanni Valente, PhD, Philosophy

Dissertation Advisors: John Earman, PhD, History & Philosophy of Science,

John D. Norton, PhD, History & Philosophy of Science

TIME, SYMMETRY AND STRUCTURE:
A STUDY IN THE FOUNDATIONS OF
QUANTUM THEORY

Bryan W. Roberts, PhD

University of Pittsburgh, 2012

This dissertation is about the sense in which the laws of quantum theory distinguish between the past and the future. I begin with an account of what it means for quantum theory to make such a distinction, by providing a novel derivation of the meaning of "time reversal." I then show that if Galilei invariant quantum theory does distinguish a preferred direction in time, then this has consequences for the ontology of the theory. In particular, it requires matter to admit "internal" degrees of freedom, in that the position observable generates a maximal abelian algebra. I proceed to show that this is not a purely quantum phenomenon, but can be expressed in classical mechanics as well. I then illustrate three routes for generating quantum systems that distinguish a preferred temporal direction in this way.

Keywords: philosophy of physics, quantum theory, time, time reversal, T violation.

TABLE OF CONTENTS

LIST OF FIGURES

PREFACE

What follows are four papers on the nature of time reversal invariance in quantum theory. Each of them (Chapters 2, 3, 4 and 5) can be read independently. Chapter 3 involved formulating a precise expression of a theorem due to Josef Jauch, and Chapter 4 required working out a classical analogue of the theorem. These results are given in Appendix A, and also stand alone.

I owe a great debt of gratitude in this work to the exceptional supervision of John Earman and John D. Norton. Their feedback was wonderfully detailed, and their encouragement invaluable. I had the particular pleasure of spending the 2011-2012 academic year working in the Center for Philosophy of Science with John Norton. Those who have visited the Center during his tenure as director will know the pleasure of John wandering into the room grinning, working excitedly through a problem full of colorful diagrams, and then bouncing out with some offhand wise-crack. I wouldn't trade those experiences for the world, and I count myself a better philosopher because of them.

I owe more thanks to many other philosophers. Harvey Brown, Craig Callender, and David Malament have my special thanks for very kindly hosting me during my visits, and for many lovely and helpful discussions about this work. For their helpful comments at various stages I also thank my committee members Robert Batterman, Laura Ruetsche and Giovanni Valente, as well as Peter Distelzweig, Balázs Gyenis, James Ladyman, Christoph Lehner, Wayne Myrvold, Michael Tamir and David Wallace. Thanks also to my parents, for their wonderful support. Above all, I would like to thank my wife, Alma, for her unending support and encouragement during this journey. This would not have been possible without her.

I received generous financial aid during the writing of this dissertation from the University of Pittsburgh, the Andrew Mellon Foundation, and the Center for Philosophy of Science. The dissertation was greatly improved thanks to a Doctoral Dissertation Research Improvement Grant (#1058902) from the National Science Foundation, which allowed me to undertake a research stay at the University of California, Irvine and the University of California, San Diego during the Winter of 2011. This work also benefited from the generosity of the Wesley Salmon Foundation, which allowed me to present many of these results at the University of Oxford and the University of Bristol in the Spring of 2011.

Alice sighed wearily. ‘I think you might do something better with the time,’ she said, ‘than waste it in asking riddles that have no answers.’

‘If you knew Time as well as I do,’ said the Hatter, ‘you wouldn’t talk about wasting *it*. It’s *him*.’

‘I don’t know what you mean,’ said Alice.

‘Of course you don’t!’ the Hatter said, tossing his head contemptuously. ‘I dare say you never even spoke to Time!’

Lewis Carroll, *Alice’s Adventures in Wonderland*

1.0 INTRODUCTION

Philosophy and physics are not, historically speaking, two separated islands of thought. Newton's *Principles of Natural Philosophy* was a groundbreaking work of both physics and philosophy. Einstein's philosophical ideas were inextricably entwined with his physics. Today, the disciplines of physics and philosophy are often separated in distinct academic departments. However, there remain problems of common interest for which the border between the two is blurred. These problems are often deep, calling for methodologies from many different disciplines: physics, mathematics, and philosophy. This dissertation is an investigation into one such problem.

The problem I will be concerned with relates to the question of how and when to distinguish between the "past" and the "future." This question is sometimes referred to as a *problem of the direction of time*. It is a very old problem, tracing its origin at least to Aristotle's study of the "before" (proteron) and the "after" (husteron) in Book IV of the *Physics*. However, it became especially urgent when it was thrust brusquely into the attention of scientists at the end of the 19th century, in a form now referred to as the *reversibility problem*.

The reversibility problem was made famous in a correspondence between a chemist, Josef Loschmidt, and the physicist Ludwig Boltzmann[1]. It can be put in this way. Consider the class of phenomena referred to generally as "thermodynamic behavior," such as when water evaporates from a mug, or when a glass shatters. This behavior has a preferred direction in time: although a glass will often shatter, it will almost never *un*-shatter, as it appears to when we play a film of a shattering glass in reverse. Nevertheless, the fundamental laws governing the billions of particles that make up these systems do not seem to be so directed in

[1] A careful and detailed account of this debate has been given by Uffink (2007, §4.3).

time: if a fundamental process can occur, then it seems that it can occur in reverse. If that's right, then we have a problem. If a fundamental process involving basic particles produces some thermodynamic behavior like a shattering glass, then the "reverse processes" would produce the reverse behavior, like an "un-shattering" glass. This is not what we observe in nature.

Norton and Earman summarize the problem concisely:

> if a system whose microdynamics is governed by deterministic time reversible laws exhibits thermodynamic behaviour, then anti-thermodynamic behaviour can be produced by reversing the velocities of microconstituents. (Earman and Norton 1998, p.439).

The central questions of this dissertation are about the antecedent of this claim. Our best fundamental theory describing the "microdynamics" of matter is quantum theory. So, taking the "deterministic laws" to be those of quantum theory, my principal concerns will be the following.

- What does it mean for a deterministic law to be "time reversible"?
- When is the time reversibility of the fundamental laws guaranteed?
- When does the time reversibility of the fundamental laws fail?

The aim of this dissertation is to provide an answer to these questions. Despite an enormous corpus of philosophy and physics devoted to the reversibility problem, and to problems of the direction of time more generally, little has yet been done to answer these three questions, much less in the context of quantum theory. My hope is that this project will fill some of that gap.

1.1 The meaning of time reversal in quantum mechanics

The Russian physicist Vladimir Fock is rumored[2] to have remarked: "Physics is essentially a simple science. The main problem is to understand which symbol means what." The philosophy of physics often engages in a closely related problem, that of clarifying the meaning of

[2](See Khriplovich 2005, p.53)

a central term appearing in the foundation of a theory. One resilient example for both physicists and philosophers is the term, “time reversal.” Unlike a transformation like rotation or translation in space, there is no apparent physical act that would exchange the past and the future. Nevertheless, time reversal is often given a rough and intuitive meaning in the following way. Suppose we film a motion, and then play the film back in reverse. The result is a film displaying the “time reverse” of the original motion. In a given theory of physics, this intuitive idea is normally replaced with a precise and well-accepted transformation. But there remains a question of where this “precisified” transformation comes from, and what its relationship is to the original intuition.

Malament (2004) has shown one interesting way to answer this question, in the context of classical electromagnetism. His concern is a standard dogma, that the electric field $\mathbf{E}$ remains unchanged when it is time-reversed, while the magnetic field $\mathbf{B}$ reverses sign. Where do these conventions come from? Malament pointed out that they can be grounded by viewing electromagnetism as reversing the temporal orientation of a relativistic spacetime. Carrying this transformation through to the electromagnetic fields turns out to have exactly the effect on $\mathbf{E}$ and $\mathbf{B}$ that the common dogma suggests.

Although Malament's example has spawned considerable discussion among philosophers[3], nobody has yet pursued the question of how this kind of question can be answered in a theory like quantum mechanics. That is the problem that I aim to solve in Chapter 2.

The standard dogma in quantum theory is that time reversal is expressed by a transformation with a very unusual property, called *antilinearity*. It is also said to leave the position observable Q unchanged, while reversing the momentum observable P, as well as reversing the spin and angular momentum observables σ and J. Although these transformation rules were first identified over 80 years by Wigner (1931), a great deal of mythology and misunderstanding still exists around their origin.

Chapter 2 seeks to correct this confusion, by formulating and then dispelling three common myths about time reversal. In particular, I show a systematic way to build up the meaning of time reversal in quantum theory, through a sequence of three steps. I formulate these steps in a completely general way, which applies equally to ordinary quantum theory

[3]For example, see (Leeds 2006), (North 2008) and (Arntzenius and Greaves 2009).

and to relativistic quantum field theory. The result is a rigorous grounding of the meaning of time reversal in quantum theory.

1.2 When quantum theory is time reversal invariant

With the meaning of time reversal in quantum mechanics secure, a natural next step is to ask when quantum theory is time reversal *invariant*. To formulate the question somewhat fancifully: when does quantum theory distinguish between "the past" and "the future"? The assumption that quantum theory makes no such distinction (i.e., that it is time reversal invariant) is often taken for granted[4]. Its truth is crucial for Norton and Earman's expression of the reversibility problem to be of much interest. So, when is it it true?

As it turns out, there are many circumstances in which time reversal invariance can fail. However, Chapter 3 shows that there are surprisingly generic circumstances for which time reversal invariance is guaranteed.

The first condition that we consider is that quantum theory be *non-relativistic*. This is done by assuming that the theory is covariant under the translations in space, and under constant changes in speed (the so-called "Galilei Boosts.") These conditions, which find their origin in the *Dialogue* of Galileo Galilei (1967, Day 2), can expressed as follows. Suppose we set up and perform an experiment, then set up and perform the same experiment in a different location in space. *Spatial translation covariance* says that we will get the same results in both cases. Suppose similarly that we perform an experiment on a boat at rest, and then perform the same experiment on a boat moving with constant velocity across the ocean. *Galilei boost covariance* says that we will again get the results (Figure 1.1).

The second scenario that we consider is the lack of *internal degrees of freedom*. Internal degrees of freedom are properties of matter, first introduced in the 20th century when the electron's "intrinsic spin" was discovered. These properties are distinguished by their independence from position (and change in position) in space. Although a host of internal degrees of freedom are now posited by the standard model of particle physics, we will

[4] Wigner (1931, 1932) himself seems to have assumed this.

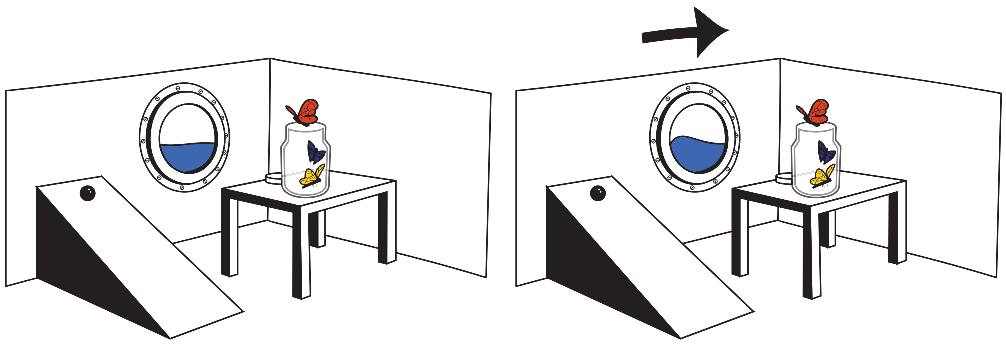


Figure 1.1: Galilei boost covariance says that if the same experiment is performed at different constant velocities, it will produce the same results.

consider what the world would be like in their absence.

Chapter 3 shows that the absence of internal degrees of freedom is deeply connected to time reversal invariance. In particular, we prove that if a Galilei covariant quantum system admits no internal degrees of freedom, then it must be time reversal invariant. To return to our fanciful expression, such a system cannot distinguish the past from the future. To put the same result a different way: in Galilei covariant quantum systems, one can *only* violate time reversal invariance through the presence of internal degrees of freedom. An ontology rich enough to include internal degrees of freedom thus plays an essential role in time asymmetry.

1.3 When classical theory is time reversal invariant

Although physicists typically study time reversal invariance and time reversal violation in the context of quantum theory, one may ask whether or not the discussion of Chapter 3 applies to other theories as well. Chapter 4 shows that a very similar discussion can indeed be carried out in classical mechanics.

Like quantum theory, classical mechanics is often assumed without reflection to be time reversal invariant, at least for systems in which energy is in some sense conserved. The first part of Chapter 4 shows that this common assumption cannot stand. Several classical systems are given, including some that conserve energy, in which time reversal invariance fails. The second part of Chapter 4 then argues that, in spite of these examples, there are at least two interesting ways to characterize the kinds of classical systems that *are* time reversal invariant. The result is an illustration of some general circumstances under which we do (and do not) a classical "arrow of time."

1.4 When quantum theory is *not* time reversal invariant

For the first six decades of the 20th century, fundamental physics was assumed nearly without reflection to be time reversal invariant. So it was with surprising (though characteristic) foresight that Paul Dirac wrote in 1949:

> A transformation... may involve a reflection of the coordinate system in the three spacial dimensions and it may involve a time reflection. ... I do not believe there is any need for physical laws to be invariant under these reflections, although all the exact laws of nature so far known do have this invariant. (Dirac 1949, p. 393)

Fifteen years later, Christenson, Cronin, Fitch and Turlay (1964) produced the first evidence of quantum systems in which time reversal invariance fails. Today, it is common wisdom that according to quantum field theory, "nature distinguishes between the past and future even on the microscopic level" (Bigi and Sanda 2009, p.457).

What enables these violations of time reversal invariance? Chapter 3 provided a partial answer: these systems must either fail to be Galilei invariant, or have internal degrees of freedom, or both. One might then describe the underpinnings of time reversal violation by simply producing a list of the experimental evidence. However, it would be more informative if there were some *general* way to describe these underpinnings.

Chapter 5 argues that there are three such general routes to violating time reversal invariance. In addition to giving three simple templates for describing time reversal violating systems, this chapter discusses a number of experimental research programs that test for the violation of time reversal invariance, and shows how each fits into one of the three general templates.

1.5 The mathematical structure of quantum theory

Quantum theory can be characterized quite generally as a theory about operators on Hilbert spaces, and the present work will make extensive use of these structures. A vector ψ (and more generally a density operator ρ) can represent a quantum state, and a linear self-adjoint operator A can represent an observable quantity that we can measure. The expected value of an observable A, given that we have prepared an initial state vector ψ, is given by the Born Rule, $E = \langle\psi, A\psi\rangle$. For a general state represented by a density operator ρ, that expectation value is given by the trace prescription, $E = \mathrm{Tr}(\rho A)$.

Following Fock's remark of Section 1.1, I would like to briefly discuss what these symbols mean in terms of statements about physical experiments. In order to bracket the difficulty that the so-called "measurement problem" poses for this task, I will adopt the strategy introduced by Mackey (1963), in connecting the above mathematical structures as directly as possible with experiment. I will only summarize this strategy in what follows; for more detailed remarks, see (Jauch 1964) or (Blank et al. 2008).

1.5.1 Grounding operators on Hilbert space

The basic starting point of this approach is the principle that experiments can be thought of as assigning truth values to propositions. For example, one might propose that there is a particle in some region of space Δ, accessible to a particle detector. A particle detection experiment will thus determine either a "true" or a "false" value for that proposition (Figure 1.2).

Propositions representing experiments can stand in relations to each other just like ordinary propositions. For example, sometimes a proposition P_1 will be true whenever P_2 is; in this case we will write $P_1 \subset P_2$. Indeed, we will assume that all the propositions can be

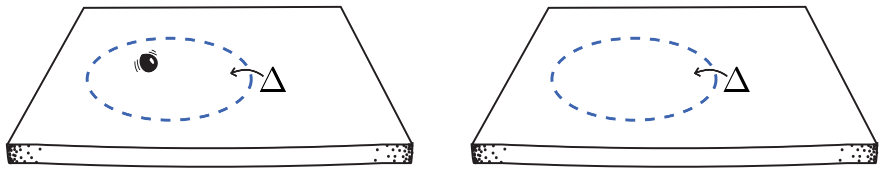


Figure 1.2: A particle detection experiment can assign a true or false value to the proposition, "There is a particle in the region Δ." It is true if the particle is in Δ (left), and false otherwise (right).

expressed as elements of a lattice $\mathcal{L}$, which is partially ordered by $\subset$. We will also introduce the expression $P_1 \cap P_2$ to denote that P_1 and P_2 are "simultaneously valid[5]"

Finally, we will assume that we can speak of the "negation" of propositions. This gets expressed as the claim that for any proposition $P \in \mathcal{L}$, there is a negated (or "orthocomplemented") proposition $P^\perp$ such that $(P^\perp)^\perp = P$, $P \cap P^\perp = \emptyset$, and $P_1 \subset P_2$ implies $P_2^\perp \subset P_1^\perp$. In particular, if P refers to the experiment of testing for a particle in some region Δ, its orthocomplement $P^\perp$ refers to the experiment of testing for a particle *outside* the region Δ.

The proposition lattices describing elementary quantum systems are well-known. As it turns out, the propositions of quantum theory forced by experiment to have a certain mathematical structure, which allows them to be expressed in terms of Hilbert space structures! In particular, the propositions of quantum theory must be expressible as *irreducible von Neumann lattices*[6]. An irreducible von Neumann lattice extends uniquely to a particular kind of algebra, called a *von Neumann algebra*[7]. These algebras have the following neat property[8]: they are all isomorphic to subalgebras of $\mathcal{B}(\mathcal{H})$, the algebra of bounded linear operators on a separable[9] Hilbert space. The propositions in the lattice correspond to projections onto a

[5]Mathematically, $P_1 \cap P_2$ is defined to be the "infimum" proposition such that for all $P \in \mathcal{L}$, $P \subset P_1 \cap P_2$ if and only if $P \subset P_1$ and $P \subset P_2$.

[6]These lattices share a property called *orthomodularity*, meaning that $A \subset B$ and $A^\perp \subset C$ implies that $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$. For a discussion of the classification of von Neumann algebras, see (Blank et al. 2008, §6). For a discussion of their experimental necessity, see (Rédei 1998).

[7]A *von Neumann algebra* is a C^*-algebra that is closed in the weak operator topology. A C^*-*algebra* is a Banach $*$-algebra, where the involution and the norm are related by $|A^*A| = |A|^2$.

[8](See Blank et al. 2008, §6.3)

[9]A Hilbert space is *separable* if it admits a countable basis.

subspace of $\mathcal{H}$. As a consequence, we may rest be assured: The propositions representing the experiments of quantum mechanics can be viewed as a sublattice of projections on a separable Hilbert space.

This grounds the use of a Hilbert space $\mathcal{H}$, and the choice of a von Neumann algebra of operators on $\mathcal{H}$, as an acceptable mathematical foundation for expressing the experiments of quantum theory.

1.5.2 Grounding self-adjoint operators

We have represented an experiment in quantum theory as a proposition with a true-or-false outcome. This allows us to think of "observable quantities," such as position or energy, as *sets* of such propositions. For example, if P_Δ represents the proposition, "the particle was detected in the region Δ of $\mathbb{R}^3$," we can represent the measurable quantity *space* in $\mathbb{R}^3$ as the set of all such propositions $\{P_\Delta \mid \Delta \subseteq \mathbb{R}^3\}$.

The so-called *spectral theorem* guarantees that a self-adjoint operator can always be thought of as an observable quantity in this sense. The reason is that, according to the spectral theorem, every self-adjoint operator A admits a unique decomposition into projection operators E_Δ (called a *projection-valued measure*). Namely,

$$A = \int_{\mathbb{R}} \lambda dE_\lambda,$$

which for finite dimensional Hilbert spaces can always be expressed $A = \sum_i \lambda_i E_i$. So, the self-adjoint operators can be represented as "weighted sums" of projection operators representing experiments, and can thus be thought of as representing measurable quantities[10].

1.5.3 Grounding probability

Quantum theory is at its core a probabilistic theory. In our picture this means that, instead making the prediction that a proposition P will be true or false (0 or 1), quantum theory assigns it an *expectation value* between 0 and 1. More generally, given a lattice of propositions

[10] They are not the only operators that can represented in this way; any *normal* operator A (meaning that $A^*A = AA^*$) has a unique spectral decomposition. This has led Penrose (2004, §22.5) to suggest that any normal operator should be counted as an observable in quantum theory.

$\mathcal{L}$, we will define a *lattice state* to be a probability measure $\mu : \mathcal{L} \to [0, 1]$, assigning each each proposition in the lattice an expectation value. Subject to some adequacy conditions[11], we can refer to such a measure μ as a *quantum probability measure*.

The connection between this notion of a probability measure, and the usual expression in terms of the Born rule and the trace, is the following. As we have seen, the von Neumann lattice of propositions $\mathcal{L}$ representing quantum experiments can always be thought of as a lattice of projections on a Hilbert space $\mathcal{H}$, and the von Neumann algebra $\mathfrak{A}$ that it generates as a subalgebra of $\mathcal{B}(\mathcal{H})$. *Gleason's theorem* then tells us that for every quantum probability measure $\mu : \mathcal{L} \to [0, 1]$ (over a Hilbert space of dimension ≥ 3), there exists a unique normal state[12] $\omega : \mathfrak{A} \to \mathbb{C}$ and a corresponding density operator ρ_ω such that, for each projection $P \in \mathcal{L}$, $\mu(P) = \omega(P) = \mathrm{Tr}(\rho_\omega P)$. And, when ω is a pure vector state[13], then there is in addition a Hilbert space vector ψ_ω such that $\mu(P) = \omega(P) = \langle \psi_\omega, P\psi_\omega \rangle$ for all projections P.

In this sense, the usual expressions of probabilities in terms of the trace prescription and the Born rule are the *unique* ways to characterize probabilities in quantum theory.

1.5.4 Going forward

Most of the discussion below will make use only of the familiar Hilbert space formalism for quantum theory. I hope the discussion above to have shown principally that this formalism is not arbitrary. It is in a certain sense forced on us by the nature of experimental practice. Although fancy objects like von Neumann algebras will make only infrequent appearances in the discussion that follows, I would like to emphasize that it is nevertheless this kind of thinking that grounds our discussion and makes it plausible.

[11] In particular, μ must be σ-additive on the families of mutually orthogonal projections.

[12] A *normal state* $\omega : \mathfrak{A} \to \mathbb{C}$ is a mapping that can be implemented by a density operator using the trace prescription, $\omega(A) = \mathrm{Tr}(\rho A)$ for all $A \in \mathfrak{A}$.

[13] A state ω is pure if $\omega = \alpha\omega_1 + (1 - \alpha)\omega_2$ with $0 < \alpha < 1$ only if $\omega_1 = \omega_2 = \omega$.

2.0 THREE MYTHS ABOUT TIME REVERSAL

2.1 Introduction

2.1.1 Prolegomena

Suppose we film a physical system in motion, and then play the film back in reverse. Will the resulting film display a motion that is possible, or impossible? This is a rough way of posing the question of the time reversal invariance: if reversed motion is always possible, the system is time reversal invariant. Otherwise, it is not.

Unfortunately, the practice of reversing films is not an appropriately practical or general way to understand the symmetries of time. Worse, the interpretation of the physical quantities being viewed in a reversed film is not always clear. Nevertheless, in the wealth of philosophical and physical discussions of time reversal in the foundations of quantum theory, the time reversal transformation is generally given a definite mathematical meaning. Where does this meaning come from?

There is a good deal of confusion about how to answer this question, as evidenced by a certain amount of mythology in the literature. Three myths are particularly prevalent.

Myth 1. *The preservation of transition probabilities ($|\langle T\psi, T\phi\rangle| = |\langle\psi,\phi\rangle|$) is a definitional feature of time reversal, with no further physical or mathematical justification.* Many presentations[1] begin with this assumption, referring to it as a definitional property of 'symmetry operators.' It was first introduced Wigner, who showed that any operator T satisfying $|\langle T\psi, T\phi\rangle| = |\langle\psi,\phi\rangle|$ is either unitary or antiunitary. A question remains as to *why* transition probabilities must be preserved. A common myth is that there is no good answer.

[1] An excellent one is (Ballentine 1998, §3.2).

Myth 2. *The antiunitary character of time reversal can only be established by fiat, or by appeal to particular transformation rules for 'position' and 'momentum.'* Textbook treatments[2] commonly begin with a representation of the canonical commutation relations $[Q, P]\psi = i\psi$, and then assume that time reversal has the effect, $Q \mapsto Q$, $P \mapsto -P$. They then observe that no unitary operator satisfies this condition. For if T were such a unitary operator, then it would follow that $i = TiT^* = T[Q, P]T^* = [TQT^*, TPT^*] = -[Q, P] = -i$, a contradiction. Therefore, if $|\langle T\psi, T\phi\rangle| = |\langle \psi, \phi\rangle|$, then T cannot be unitary, and thus must antiunitary by Wigner's theorem. This argument has an unfortunate limitation, in that many quantum systems do not admit a representation of the canonical commutation relations. The myth is that there is no other way to establish antiunitarity.

Myth 3. *The way that position and momentum transform under time reversal can only be justified by appeal to their classical analogues.* When asked to justify the transformations $Q \mapsto Q$ and $P \mapsto -P$, authors often appeal to the myth that this is only to comply with the transformation rules for the classical analogues $\mathbf{q}$ and $\mathbf{p}$. The myth is that this is the only way to establish the quantum transformation rules. This has left room for philosophers like Callender (2000), Albert (2000, §1)), and Maudlin (2007, §4.2) to argue that the standard definition of "time reversal" is not justified at all, or at least is not deserving of its name.

The present work will seek to dispel these myths. We proceed in three stages.

Stage 1 dissolves the first myth, by revisiting a theorem of Uhlhorn (1963). Uhlhorn considered a class of "orthogonality-preserving" transformations, and showed that under very general circumstances, these transformations always preserve transition probabilities. I argue that adopting this minimal characterization of time reversal provides a natural justification for the claim that time reversal preserves transition probabilities.

Stage 2 dispels the second myth, by providing a novel argument that T is antiunitary. This argument is completely independent of whether or not we have a representation of the canonical commutation relations, resting instead on the requirement that time reversal be a "non-trivial" concept in a certain sense.

Stage 3 goes after the third myth, by considering how to ground the meaning of time

[2] For example, see (Sachs 1987, §3.2).

reversal when position and momentum are in play. I offer philosophers of time a "flexible option" and a "rigid option." The flexible option allows one to countenance *any* antiunitary operator that is a "reversal" (precisely, an *involution*) as representing time reversal; I show that for any such operator, one can construct a representation of the canonical commutation relations in which Q (position) and P (momentum) transform in the standard way, $Q \mapsto Q$ and $P \mapsto -P$. Alternatively, the rigid option demands that time reversal satisfy certain further requirements with respect to the position degree of freedom; I show that in any given representation of the canonical commutation relations, these requirements uniquely fix the meaning of time reversal.

What is lost in clinging to the myths? In a sense, what is lost is a deeper understanding of the nature of time in quantum theory. Consider, by way of analogy, the canonical commutation relations $[Q, P]\psi = i\psi$. One might take this statement to be a primitive law of physics, and be done with it. But Dirac (2001), in investigating the foundations of this expression, suggested that it derives from the way position and momentum are understood in classical mechanics, as elements of a Poisson algebras satisfying $\{q, p\} = 1$. Dirac took quantum systems to derive from classical systems, by a 'quantization' homomorphism onto an irreducible Hilbert space representation. This homomorphism in particular is supposed to map q and p to self-adjoint operators Q and P such that $[Q, P] = i$. Unfortunately, there are many classical systems for which such a homomorphism does not exist (see (Woodhouse 1991, Ch.8-9), (Streater 2007, §12.7)). But more importantly: deriving the quantum from the classical is a foundationally backwards procedure. After all, it is quantum theory that governs the interactions of matter, with classical theory often providing a limiting approximation, and not the other way around.

Fortunately, a much more satisfying underpinning for the commutation relations was later found. In particular, the canonical commutation relations turn out to follow essentially from the homogeneity of space[3]. This is a more mathematically robust derivation, in that

[3] A detailed discussion of this derivation can be found in (Jauch 1968). But, here is a brief summary. Let $\Delta \mapsto E_\Delta$ be a set of projection operators, each bearing the interpretation, "the system is located in the region Δ." To the extent that physical space is *homogeneous*, it makes no difference to the predictions of quantum theory if these regions are transformed by a continuous spatial translation $\Delta \mapsto \Delta + a$. One may thus posit the existence of a strongly continuous one-parameter group of unitary operators U_a (implementing 'spatial translations') such that that $U_a E_\Delta U_a^* = E_{\Delta+a}$, where $\Delta + a = \{x : x - a \in \Delta\}$. By Stone's theorem, this group may be expressed $U_a = e^{iaP}$ for some self-adjoint operator P. Defining $Q = \int_{-\infty}^{\infty} \lambda dE_\lambda$, it is then

it applies even in circumstances in which quantization is not possible. It is also a more philosophically satisfying approach, in that it clarifies the significance of the commutation relations of quantum theory without appeal to classical mechanics.

In what follows, I hope to provide a small but similar insight into the interpretation and mathematical underpinnings of time reversal in quantum theory.

2.1.2 Notation

I begin by setting out my notational conventions; readers very familiar with this material may wish to skim it and head to the next section.

2.1.2.1 Hilbert space rays Let $\mathcal{H}$ be a separable Hilbert space. A *ray* of $\mathcal{H}$ is a set of vectors in $\mathcal{H}$ related by a constant of unit length. We will write vectors in lower-case, and rays in upper-case greek letters. Hence, $\Psi := \{\phi \mid \phi = c\psi, |c| = 1\}$ is a ray, consisting of unit multiples of the vector ψ. The set of rays of $\mathcal{H}$ forms a new Hilbert space $\mathfrak{H}$, with an inner product defined by $\langle\Psi, \Phi\rangle := |\langle\psi, \phi\rangle|^2, \psi \in \Psi, \phi \in \Phi$. The inner product is the same regardless of which ψ, ϕ in the respective rays are chosen. This new Hilbert space $\mathfrak{H}$ is called the *ray space* of $\mathcal{H}$. Two rays Ψ, Φ are said to be *orthogonal* if their inner product vanishes, $\langle\Psi, \Phi\rangle = 0$; this property is often expressed, $\Psi \perp \Phi$.

In quantum theory, vectors in the same ray give rise to the same probabilistic predictions about the outcomes of experiments. For this reason, it is the ray space $\mathfrak{H}$, rather than the original $\mathcal{H}$, that best captures our probabilistic predictions.

2.1.2.2 Hilbert space operators Measurable degrees of freedom will be represented by self-adjoint linear operators on Hilbert space. In a typical infinite-dimensional Hilbert space, many such operators do not act on the entire Hilbert space. So, if A is a linear operator on $\mathcal{H}$, we refer to the set $\mathcal{D}_A \subseteq \mathcal{H}$ on which A acts as the *domain* of A. The *spectrum* of a self-adjoint operator A is the set of real numbers λ such that $(A - \lambda I)$ does not have an

a simple matter to show that $e^{iaP}e^{ibQ} = e^{ia \cdot b}e^{ibQ}e^{iaP}$, from which it follows that $[Q, P]\psi = i\psi$ for all ψ in the common dense domain of Q and P.

inverse. When the spectrum of A is discrete, this is equivalent to "eigenvalue" definition of the spectrum, as the set of real numbers λ such that $A\psi = \lambda\psi$ for some vector ψ.

An operator $A : \mathcal{H} \to \mathcal{H}$ is *unitary* if it is both linear (i.e., $A(b\psi + c\phi) = bA\psi + cA\phi$) and $A^*A = AA^* = I$. It is *antiunitary* if it is both antilinear (i.e., $A(b\psi + c\phi) = b^*A\psi + c^*A\phi$) and $A^*A = AA^* = I$. Unitary operators preserve inner products, $\langle A\psi, A\phi\rangle = \langle\psi, \phi\rangle$. Antiunitary operators conjugate them, $\langle A\psi, A\phi\rangle = \langle\psi, \phi\rangle^*$.

2.1.2.3 The canonical commutation relations An assignment of real numbers to Hilbert space operators $a \mapsto U_a$ is *strongly continuous* if the real-valued function $f(a) = |U_a\psi|$ is continuous for all $\psi \in \mathcal{H}$, where $|\psi|^2 = \langle\psi, \psi\rangle$. Let $a \mapsto U_a$, $b \mapsto V_b$ be two strongly continuous unitary representations of the additive group of reals, $(\mathbb{R}^n, +)$. The triple $(\mathcal{H}, a \mapsto U_a, b \mapsto V_b)$ is called a *unitary representation of the canonical commutation relations in Weyl form* if

$$U_aV_b = e^{a\cdot b}V_bU_a$$

for all $a, b \in \mathbb{R}^n$. The representation is *irreducible* if, whenever $\mathcal{H}' \subseteq \mathcal{H}$ is a subspace such that $(\mathcal{H}', a \mapsto U_a, b \mapsto V_b)$ is a unitary representation of the Weyl commutation relations, it follows that $\mathcal{H}' = 0$ or $\mathcal{H}' = \mathcal{H}$. Stone's theorem (Blank et al. 2008, Thm. 5.9.2) guarantees the existence of unique self-adjoint operators P and Q such that $U_a = e^{iaP}$ and $V_b = e^{ibQ}$. For all ψ in the common dense domain of Q and P, one can show that a continuous representation of the Weyl commutation relations implies that $[Q, P]\psi = i\psi$; for example, see (Jauch 1968, §12.2).

2.1.2.4 Dynamical systems A *dynamical system* in quantum theory is a pair $(\mathcal{H}, t \mapsto U_t)$, where $\mathcal{H}$ is a separable Hilbert space, and $t \mapsto U_t$ is a strongly continuous representation of $(\mathbb{R}, +)$ in terms of unitary operators on $\mathcal{H}$.

Note that this conception of the dynamics is fully compatible with both non-relativistic and relativistic quantum field theory. The only subtlety is that in the relativistic case, the group $t \mapsto U_t$ is defined relative to a given foliation of spacetime into spacelike hypersurfaces. Note also that, as above, Stone's theorem allows us to freely write $U_t = e^{-itH}$ for a unique self-adjoint operator H.

2.1.2.5 Invariance Let $(\mathcal{H}, t \mapsto e^{-itH})$ be a dynamical system. Then for any bijection $T : \mathcal{H} \to \mathcal{H}$, we say that this dynamical system is *T-reversal invariant* if it is the case that if $\psi(t) = e^{-itH}\psi$, then $T\psi(-t) = e^{-itH}T\psi$. An equivalent way to express T-reversal invariance is: if $\psi(t)$ is a solution to the Schrödinger equation, then so is $T\psi(-t)$.

This is just an application of the standard concept of what it means to be invariant under a transformation. Invariance under spatial rotations means that if a given dynamical trajectory is possible, then so is the rotated trajectory. Similarly, our statement of invariance under T-reversal means that if $\psi(t)$ is a possible trajectory of the unitary group e^{-itH}, then so is $T\psi(-t)$, the trajectory given by applying time reversal operation T is applied to every state, and by reversing the temporal order $(t \mapsto -t)$.

Readers new to this kind of transformation might imagine the operator T as analogous to the time reversal operator in classical Hamiltonian particle mechanics. Namely, it might have the effect of fixing positions and reversing momenta, $TQT^{-1} = Q$ and $TPT^{-1} = -P$. Then the transformation $\psi(t) \mapsto T\psi(-t)$ is analogous to reversing a film of the particle: the particle travels through the same positions $(Q \mapsto Q)$, but in the reverse temporal order $(t \mapsto -t)$, and with momenta in the opposite direction $(P \mapsto -P)$.

However, it should be emphasized that our discussion will seek to *derive* this kind of transformation rule, and will not just assume it. Our definition T-reversal requires nothing whatsoever about the nature of the operator T, other than that it is a bijection. In the remainder of the paper, we will aim to determine what is required of this operator in order to appropriately capture the meaning of "time reversal."

2.2 First Stage: Why T is Unitary or Antiunitary

Wigner's Theorem is the traditional starting point for the discussion of symmetry in quantum mechanics. One begins with the interpretive assumption that, if a transformation T represents a "symmetry transformation" like time reversal, rotation, translation, etc., then it preserves transition probabilities: $|\langle T\psi, T\phi\rangle| = |\langle\psi, \phi\rangle|$. Equivalently, one assumes that the associated transformation on ray space (denoted with a boldface $\mathbf{T}$) preserves ray inner products, $\langle \mathbf{T}\Psi, \mathbf{T}\Phi \rangle = \langle \Psi, \Phi \rangle$. Wigner's theorem shows that for any $\mathbf{T}$ satisfying this condition, there exists a Hilbert space operator T that implements it that is either unitary or antiunitary. So, applying Wigner's theorem to the interpretive assumption above guarantees that transformations like time reversal (and rotation, translation, etc.) can always be represented either by a unitary operator or by an antiunitary one.

But what justifies the interpretive assumption, that symmetries preserve transition probabilities? The question matters because, on the face of it, the assumption that symmetries preserve transition probabilities appears to 'sneak in' a certain kind of physical invariance. Of course, it is not the official definition of T-reversal invariance given at the end of the last section. But Wigner's assumption remains a substantial physical claim, not an a priori mathematical truth. If this is not the way the world actually works, then the standard representation of symmetries in terms of unitary or antiunitary operators would be wrong-headed.

Fortunately, there is a very compelling way to motivate the claim that time reversal preserves transition probabilities, although it is not very well-known[4]. It derives from the following theorem, due to Uhlhorn (1963).

Theorem 2.1 (Uhlhorn)**.** *Let* $\mathbf{T}$ *be any bijection on the ray space* $\mathfrak{H}$ *of a separable Hilbert space* $\mathcal{H}$ *with dimension greater than 2. Suppose that* $\Psi \perp \Phi$ *if and only if* $\mathbf{T}\Psi \perp \mathbf{T}\Phi$*. Then,*

$$\langle \mathbf{T}\Psi, \mathbf{T}\Phi \rangle = \langle \Psi, \Phi \rangle.$$

Moreover, there exists a unique (up to a constant) $T : \mathcal{H} \to \mathcal{H}$ *that implements* $\mathbf{T}$ *on* $\mathcal{H}$*, in the sense that* $\psi \in \Psi$ *iff* $T\psi \in \mathbf{T}\Psi$*, and which satisfies* $|\langle T\psi, T\phi \rangle| = |\langle \psi, \phi \rangle|$ *for all* $\psi, \phi \in \mathcal{H}$*.*

In short, in Hilbert spaces of dimension greater than 2, T automatically preserves transition probabilities whenever it is the case that $\Psi \perp \Phi$ iff $\mathbf{T}\Psi \perp \mathbf{T}\Phi$. So, how plausible is the latter condition when $\mathbf{T}$ is the operator representing time reversal?

It turns out to be very plausible indeed. Orthogonal rays can be thought of as representing mutually exclusive pure states. (For example, two eigenstates of a self-adjoint

[4]See (Varadarajan 2007, §4), especially Theorem 4.29, for a textbook treatment of this approach. I thank David Malament for drawing this to my attention, and Keith Hannabuss for directing me to the Uhlhorn reference.

operator are orthogonal if they have distinct eigenvalues[5].) So, Uhlhorn's condition means that two pure states are mutually exclusive if and only if the time-reversed states are mutually exclusive. To ground this intuition, consider any two mutually exclusive propositions, like "detecting a particle with spin up" and "detecting a particle with spin down" (Figure 2.1). The intuition underlying Uhlhorn's condition is that if we film a physical system that allows these two properties, then the two propositions remain mutually exclusive *no matter whether the film is playing forward or in reverse.* The facts about mutual exclusivity should be independent of anything to do with the facts about the direction of time.

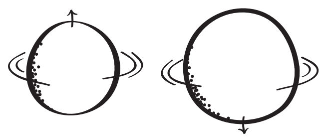


Figure 2.1: If two pure states are orthogonal, and hence mutually exclusive, then we assume this fact is independent of the direction of time.

With this assumption on board, Uhlhorn's theorem provides the following clear account of why time reversal really does preserve transition probabilities: it is because orthogonality between pure states is a matter that is independent of the direction of time.

It is worth highlighting that the requirement of dimensionality greater than 2 is a necessary condition for Uhlhorn's theorem[6]. This precludes the application of such a result to certain idealized Hilbert spaces of dimension 2. However, since most realistic quantum systems are thought to have a momentum (or energy-momentum) degree of freedom with a continuous spectrum, the applicability of the result remains extremely general.

For the remainder of this discussion, I will adopt the Uhlhorn perspective on time reversal, and thus take for granted that the time reversal operator T can be implemented by a unitary operator or an antiunitary one. In the next stage I will argue that, unless time reversal is

[5]This is because if A is a self-adjoint operator such that $A\psi_1 = \lambda_1$ and $A\psi_2 = \lambda_2\psi_2$, then since the eigenvalues of A are all real, $(\lambda_1-\lambda_2)\langle\psi_1,\psi_2\rangle = \langle\lambda_1\psi_1,\psi_2\rangle-\langle\psi_1,\lambda_2\psi_2\rangle = \langle A\psi_1,\psi_2\rangle-\langle\psi_1,A\psi_2\rangle = \langle\psi_1,A\psi_2\rangle-\langle\psi_1,A\psi_2\rangle = 0$. So, if $\lambda_1-\lambda_2\neq 0$, then $\langle\psi_1,\psi_2\rangle = 0$.

[6]A simple counterexample if $\mathcal{H}$ has dimension 2 is the following. Let Ψ, Φ be orthogonal rays, and let T be the mapping that exchanges Ψ and Φ, but is the identity on all other rays. Then T is orthogonality-preserving, but not angle-preserving. I thank John Norton for this example.

trivially uninteresting, this T must in fact be antiunitary.

2.3 Second Stage: Why T is Antiunitary

The time reversal operator is standardly taken to be antiunitary, and not unitary. Why? The most common answer, introduced in the first section, is that unitarity is impossible given the standard transformation rules for position and momentum. In the following proposition, we motivate this conclusion in a completely different way[7], by demanding the existence of at least one quantum system satisfying three adequacy conditions.

Proposition 2.1. *Let T be a unitary or antiunitary bijection on $\mathcal{H}$. Suppose there exists at least one densely-defined self-adjoint operator H on $\mathcal{H}$ such that, if $(\mathcal{H}, t \mapsto e^{-itH})$ is a dynamical system, then the following conditions hold.*

(i) (positive spectrum) $0 \leq \langle \psi, H\psi \rangle$ *for all* $\psi \in \mathcal{D}_H$.

(ii) (non-triviality) H is not the zero operator.

(iii) (invariance) The dynamical system $(H, t \mapsto e^{-itH})$ is T-reversal invariant.

Then T is antiunitary.

Proof. Let $\psi(t) = e^{-itH}\psi$, for any arbitrary ψ and for all t. Substituting $t \mapsto -t$ implies equally that $\psi(-t) = e^{itH}\psi$. Applying T to both sides, we thus have $T\psi(-t) = Te^{itH}\psi$. Moreover, Condition (iii) guarantees that $T\psi(-t) = e^{-itH}T\psi$. Equating these two, we have $Te^{itH}\psi = e^{-itH}T\psi$. But ψ was arbitrary, so it follows that $Te^{-itH}T^{-1} = e^{itH}$. Therefore,

$$e^{itH} = Te^{-itH}T^{-1} = e^{T(-itH)T^{-1}}, \tag{2.1}$$

where the second equality follows from the functional calculus. But e^{itH} is a strongly continuous unitary group, and so its self-adjoint generator is unique by Stone's theorem. Thus, in view of Equation (2.1), $itH = -TitHT^{-1}$.

[7] This proof makes precise a strategy that was originally suggested by Wigner (1931, §20). I thank David Malament for helpful suggestions that led to improvements in this formulation.

Suppose for reductio that T is unitary, and hence linear. Then $-itTHT^{-1} = itH$, and so $THT^{-1} = -H$. Moreover, since T is unitary, it preserves inner products. So,

$$\langle\psi, H\psi\rangle = \langle T\psi, TH\psi\rangle = -\langle T\psi, HT\psi\rangle$$

for all $\psi \in \mathcal{D}_H$. But H has a non-negative spectrum by (i), so both $\langle\psi, H\psi\rangle$ and $\langle T\psi, HT\psi\rangle$ are non-negative. Thus,

$$0 \leq \langle\psi, H\psi\rangle = -\langle T\psi, HT\psi\rangle \leq 0. \tag{2.2}$$

It follows that $\langle\psi, H\psi\rangle = 0$ for all $\psi \in \mathcal{D}_H$. Since H is densely defined, this is only possible if H is the zero operator, contradicting our assumption (ii). So, T is not unitary, and is therefore antiunitary. □

This proposition applies equally to both non-relativistic quantum mechanics and to relativistic quantum field theory. Let me discuss each of the premises in term.

Premise (i) expresses the widely-held belief that energy is non-negative. As far as we can tell, such systems are physically implausible[8].

Premise (ii) says that the expression of energy is non-trivial, in that it is not just the zero operator. Otherwise, unitary time evolution would be given by the identity operator, and our physical description would reduce to on in which "nothing ever happens." We would like to understand time reversal for systems in which things happen.

Premise (iii) guarantees that time reversal is an interesting enough concept to be an invariance of at least one quantum system. As it happens, we might come to believe that most realistic systems fail to be time reversal invariant. That would be perfectly compatible with the perspective advocated here. We demand only that there is at least one non-trivial

[8]Premise (i) can in fact be replaced with an alternative, (i*): the spectrum of H is bounded from below but not from above. This premise is perhaps more plausible in the context of quantum field theory, in which H is unbounded from above for any realistic system involving creation and annihilation operators, but is still bounded from below by a (possibly negative) real number r. This allows us to write $r \leq \langle\psi, H\psi\rangle$ and $r \leq \langle T\psi, HT\psi\rangle$. The proof of Proposition 2.1 proceeds exactly as before, except that Equation (2.2) becomes,

$$r \leq \langle\psi, H\psi\rangle = -\langle T\psi, HT\psi\rangle \leq -r.$$

This implies that H is bounded from above by $-r$, contradicting (i*). Therefore, T is not unitary. Assumption (ii) is not necessary on this variation of the proof. I thank David Wallace for pointing this variation out to me.

description of a quantum system that is time reversal invariant. (And it need not even be physically realizable!) Without an assumption like this, the concept of time reversal invariance would be uninteresting, in the sense that it would fail to apply to any non-trivial quantum system whatsoever.

Our proposition thus provides the following way to ground the antiunitary character of time reversal: if time reversal is interesting enough to be an invariance of at least one non-trivial, physically plausible system, then the time reversal operator must be antiunitary.

2.4 Third Stage: Position and Momentum

The above discussion provides a very general perspective on the nature of time reversal. It applies to both non-relativistic and relativistic quantum theory, and to any arbitrary algebra of observables. However, in such a general context, little more can be said about the time reversal operator T, beyond the result that it is antiunitary.

On the other hand, suppose one wishes to describe a particular physical system, such as a particle in space, or a free relativistic Bose field. Both descriptions require "momentum and position" degrees of freedom, given by two strongly continuous unitary groups $a \mapsto U_a$ and $b \mapsto V_b$, which satisfy the canonical commutation relations in Weyl form, $U_a V_b = e^{ia\cdot b} V_b U_a$. What can be said about the nature of time reversal for these specific systems?

The "standard" view, often given without argument, is that the time reversal operator is antiunitary, and has the effect $TQT^{-1} = Q$ and $TPT^{-1} = -P$. In terms of the Weyl operators $U_a = e^{iaP}$ and $V_b = e^{ibQ}$, these latter two relations are equivalent[9] to T having the effect $U_a \mapsto U_a$ and $V_b \mapsto V_{-b}$.

What justifies this view? Antiunitarity and the canonical commutation relations are not enough to guarantee these transformation rules will be satisfied[10]. I would like to argue that

[9]Write $TU_aT^{-1} = Te^{iaP}T^{-1} = e^{TiaPT^{-1}} = e^{-iaTPT^{-1}}$, where the last equality follows from antiunitarity. Suppose $TPT^{-1} = -P$. Then $TU_aT^{-1} = e^{-ia(-P)} = U_a$. Conversely, suppose $TU_aT^{-1} = U_a$. Then $e^{-iaTPT^{-1}} = e^{iaP}$. Then by the uniqueness clause of Stone's theorem, $-iaTPT^{-1} = e^{iaP}$. Therefore, $TPT^{-1} = -P$. (The obvious analogous argument can then be made for Q and V_b.)

[10]Example: if $[Q,P]\psi = i\psi$, then $[Q+P,P]\psi = i\psi$. But although both pairs (Q,P) and $(Q+P,P)$ satisfy the canonical commutation relations, an antiunitary T cannot transform preserve both Q and $Q+P$ while

the answer depends on one's level of commitment about what it means to "reverse time." In the next two subsections, I will illustrate two such levels of commitment available to philosophers of time.

2.4.1 The flexible option

The first option is flexible about what is required of an operator representing "time reversal." Suppose a philosopher of time is satisfied to call any antiunitary mapping $\mathbf{T}$ "time reversal," as long as it is an *involution*. An involution $T : \mathcal{H} \to \mathcal{H}$ is a mapping that if applied twice, is the identity on ray space: $T^2 = cI$ for some complex unit c. This is is the mathematical expression of what it means to be a "reversal" in quantum theory: applying a reversal twice brings us back to where we started, up to an arbitrary non-physical phase (Figure 2.2).

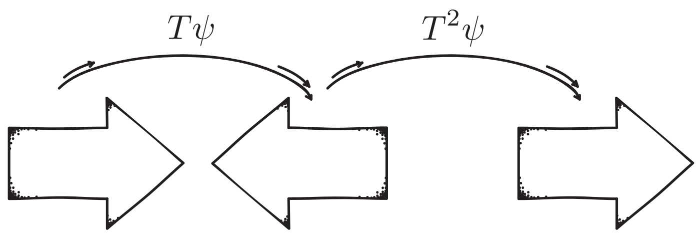


Figure 2.2: An involution is a transformation that is the identity operator (up to an arbitrary phase) when applied twice, $T^2\psi = c\psi$.

There turn out to be many antiunitary involutions, one for each orthonormal basis (Messiah 1999, §15.5). However, for each such antiunitary involution T, there turns out to exist a representation in which T behaves like the standard time reversal operator. Namely, there exists a representation in which T has the effect of transforming $U_a \mapsto U_a$ and $V_b \mapsto V_{-b}$. This is established by the following.

Proposition 2.2. *Let $\mathcal{H}$ be a separable Hilbert space of infinite dimension. Suppose there exists a bijection $T : \mathcal{H} \to \mathcal{H}$ satisfing the following two conditions.*

(i) (antiunitarity) $\langle T\psi, T\phi\rangle = \langle\psi, \phi\rangle^*$ *for all* $\psi, \phi \in \mathcal{H}$.

reversing P.

(ii) (involution) $T^2 = cI$ *for some non-zero* $c \in \mathbb{C}$.

Then there exists a unitary representation $(\mathcal{H}, a \mapsto U'_a, b \mapsto V'_b)$ *of the Weyl commutation relations on* $\mathcal{H}$ *such that* $TU'_aT^{-1} = U'_a$ *and* $TV'_bT^{-1} = V'_{-b}$.

Proof. We use the Schrödinger representation $(L^2(\mathbb{R}), a \mapsto U_a, b \mapsto V_b)$ to construct the desired representation. Let K be the conjugation operator defined by $K\psi(x) = \psi^*(x)$, for all $\psi(x) \in L^2(\mathbb{R})$. Recall that K is an antiunitary operator, and that $K^2 = I$, $KU_aK = U_a$, and $KV_bK = V_{-b}$. Define $A := TK$, and note that it satisfies $AK = TK^2 = T$. This A is unitary, because it is the product of two antiunitary operators. By the involution property, it also satisfies $cI = T^2 = AKAK$, and thus $KAK = cA^*$. Since A is unitary, this c is also a unit, $c^*c = 1$.

The fact that A is unitary entails that there is a well-defined (though not unique) 'square root of A' operator $\tilde{A}$, which has the properties that $\tilde{A}^2 = A$ and $K\tilde{A}K = c\tilde{A}^*$. To check this, let us write A in its spectral resolution,

$$A = \int_0^{2\pi} e^{i\lambda} dE_\lambda$$

where E_λ denotes the projection of A corresponding to the interval $(0, \lambda)$. Applying the functional calculus, define

$$\tilde{A} := \int_0^{2\pi} e^{i\lambda/2} dE_\lambda.$$

Viewing $\tilde{A} = f(A)$ as a function of A, this definition implies that $f(e^{i\lambda})^2 = (e^{i\lambda/2})^2 = e^{i\lambda}$, and it follows that $\tilde{A}^2 = f(A)^2 = A$. Moreover, since $KAK = cA^*$, it follows immediately that $K\tilde{A}K = c\tilde{A}^*$.

Now let $(L^2(\mathbb{R}), a \mapsto U'_a, b \mapsto V'_b)$ be a new representation, in which we define

$$U'_a := \tilde{A}U_a\tilde{A}^*, \quad V'_b := \tilde{A}V_b\tilde{A}^*.$$

Since $\tilde{A}$ is unitary, this is an irreducible unitary representation of the Weyl commutation relations. But $\tilde{A}^2 = A$, so $A\tilde{A}^* = \tilde{A}$. Therefore, applying $T = AK$ to both sides of U'_a and

V_b', we have that

$$\begin{aligned}TU_a'T^{-1} &= (AK)\tilde{A}U_a\tilde{A}^*(KA^*)\\ &= Ac\tilde{A}^*(KU_aK)\tilde{A}c^*A^*\\ &= c^*c\tilde{A}U_a\tilde{A}^* = U_a'\end{aligned}$$

and

$$\begin{aligned}TV_b'T^{-1} &= (AK)\tilde{A}V_b\tilde{A}^*(KA^*)\\ &= Ac\tilde{A}^*(KV_bK)\tilde{A}c^*A^*\\ &= c^*c\tilde{A}V_{-b}\tilde{A}^* = V_{-b}',\end{aligned}$$

which is the desired transformation. $\square$

Write $U_a' = e^{iaP}$ and $V_b' = e^{ibQ}$, where P and Q are self-adjoint operators. Proposition 2.2 can then be seen to show the existence of a representation in which time reversal has the "standard" effect, $Q \mapsto Q$ and $P \mapsto -P$.

Here is the upshot of this proposition. Suppose we are flexible in what we're willing to call "time reversal", and assume only that the time reversal operator is an antiunitary involution. There is a sense in which this is sufficient to characterize the concept of time reversal. Namely, we can choose *any* antiunitary T whatsoever to be our time reversal operator, and Proposition 2.2 will provide a representation of the canonical commutation relations in which that T behaves just like the standard one: $T : Q \mapsto Q$ and $T : P \mapsto -P$. For finitely many degrees of freedom, all representations are empirically (i.e. unitarily) equivalent (Blank et al. 2008, Theorem 8.2.4). So, in this context, the flexible option is all that is needed to understand time reversal (up to unitary equivalence).

Nevertheless, one may still wish to ask a more informative question. Namely, suppose that we fix a representation of the canonical commutation relations. Is there then some way to uniquely determine the action of the time reversal operator on Q and P? Such a uniqueness result is indeed available, but it requires more commitments about the nature of time reversal.

2.4.2 The rigid option.

Suppose we demand a more "rigid" level of commitment about the nature of time reversal, in the form of two further demands about T. I will refer to these demands as the "homogeneity" and "spatial independence" conditions.

Homogeneity is the demand that, whatever time reversal is like, it does not pick out any preferred region of space. To give this requirement mathematical expression, we draw on the interpretation of $U_a = e^{iaP}$ as the group of spatial translations, mentioned in the introduction to this paper. That interpretation is motivated as follows. Let $\Delta \mapsto E_\Delta$ is the spectral measure of the "position" operator Q, which generates V_b in the sense that $V_b = e^{ibQ}$. Then each projection E_Δ can be interpreted as the proposition, "the system is located in the spatial region Δ." But the canonical commutation relations in Weyl form imply that

$$U_a E_\Delta U_a^* = E_{\Delta - a}$$

(Jauch 1968, §12.2). That is, U_a transforms a system located in the spatial region Δ to the one located in the 'translated' region $\Delta - a$ (Figure). The homogeneity condition $[T, U_a] = 0$ thus amounts to the claim that spatial translation followed by time reversal is equivalent to time reversal followed by spatial translation. In other words, time reversal does not treat any particular spatial region differently than any other. Insofar as T is supposed to implement the reversal of *time* and not *space*, this is a sensible enough requirement.

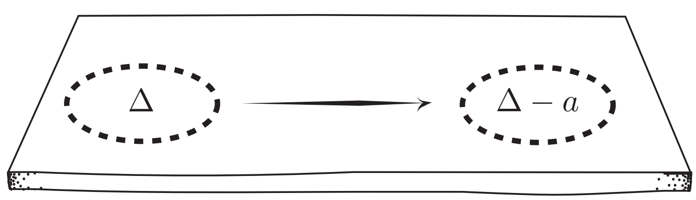


Figure 2.3: Time reversal transforms a system in the spatial region Δ in the same way that it does a system in the translated region $\Delta - a$.

The spatial independence demand is that, however the time reversal operator transforms Q, it must map it to an operator that is statistically independent of Q. The mathematical expression of statistical equivalence is that TQT^{-1} and Q commute. This captures the idea that time reversal does not introduce any detectable changes in the measurement of position. For example, it precludes the possibility that time reversal involve *time evolution*, generated by a Hamiltonian $H = (1/2m)P^2 + v(Q)$, since $e^{itH}Qe^{-itH}$ does not commute with Q. Of course, one might just demand the stronger assumption that "time reversal does not affect location in space," and write $TQT^{-1} = Q$ in place of spatial indendence. I have no problem with this assumption. However, its strength is simply not needed; the weaker concept of statistical independence $[TQT^{-1}, Q] = 0$ turns out to be enough to uniquely determine the form of T. This is established by our final proposition.

Proposition 2.3. *Let $(\mathcal{H}, a \mapsto U_a, b \mapsto V_b)$ be a strongly continuous irreducible unitary representation of the Weyl commutation relations. Suppose there exists a bijection $T : \mathcal{H} \to \mathcal{H}$ satisfying the following conditions.*

(i) (antinitarity) $\langle T\psi, T\phi\rangle = \langle\psi, \phi\rangle^*$ for all $\psi, \phi \in \mathcal{H}$

(ii) (involution) $T^2 = c$ for some non-zero $c \in \mathbb{C}$.

(iii) (homogeneity) $[T, U_a] = 0$

(iv) (spatial independence) $[TQT^{-1}, Q] = 0$,

where Q is the self-adjoint generator of V_b. Then $TV_bT^{-1} = V_{-b}$.

Proof. Since TQT^{-1} commutes with Q, it commutes with $e^{ibQ} = V_b$. Thus, the quantity $(TQT^1 - Q)$ commutes with V_b as well. But this same quantity also commutes with U_a. To see this, we first note that the commutation relations imply $U_aQU_a^* = Q + aI$ (Jauch 1968, §12.2). Then, recalling that $TU_a = U_aT$ by (iii), we have,

$$
\begin{aligned}
U_a(TQT^{-1} - Q) &= TU_aQT^{-1} - U_aQ \\
&= T(QU_a + aU_a)T^{-1} - (QU_a + aU_a) \\
&= TQT^{-1}U_a + aU_a - QU_a - aU_a \\
&= (TQT^{-1} - Q)U_a.
\end{aligned}
$$

But the representation is irreducible, so Schur's lemma implies that the only operators commuting with both U_a and V_b are constant multiples of the identity. This allows us to write,

$$TQT^{-1} - Q = kI,$$

for some $k \in \mathbb{C}$. But the involution condition (ii) implies that $T^2 = cI$ and $T^{-2} = c^{-1}I$ for some non-zero constant c. These facts together imply that

$$\begin{aligned} Q &= cQc^{-1} = T^2QT^{-2} \\ &= T(Q + kI)T^{-1} = TQT^{-1} + k^*I \\ &= Q + (k + k^*)I, \end{aligned}$$

where we have applied the antiunitarity of T in writing $TkT^{-1} = k^*I$. So, $k + k^* = 0$. But since $Q^* = Q$ and $T^* = T^{-1}$, we have that $(TQT^{-1})^* = TQ^*T^{-1} = TQT^{-1}$. So, TQT^{-1} is self-adjoint, and its spectrum is real, so $k = k^*$. Thus, $0 = k + k^* = k + k$, which implies that $k = 0$. Therefore, $TQT^{-1} = Q$, and

$$TV_bT^{-1} = e^{TibQT^{-1}} = e^{-ibTQT^{-1}} = e^{-ibQ} = V_{-b}.$$

□

If one is of the disposition to demand this more rigid characterization of time reversal, then this proposition allows us to take any given representation of the canonical commutation relations, and uniquely determine the way that T transforms U_a and V_b.

Note that, in the context of relativistic quantum field theory, there is some subtlety in the interpretation of the projections E_Δ as capturing propositions about 'space.' For example, Malament (1996) and Halvorson and Clifton (2002) have shown that under very weak conditions, such projections will fail to satisfy a natural condition of localizability, namely that if Δ_1 and Δ_2 are disjoint open regions in the same spacelike hypersurface, then

$$E_{\Delta_1}E_{\Delta_2} = E_{\Delta_2}E_{\Delta_1} = \mathbf{0}.$$

However, these challenges to localizability do not themselves challenge the postulate of a more general spatial measure $\Delta \mapsto E_\Delta$ appearing in Proposition 2.3, since we have avoided the requirement that this measure be localizable.

In summary, when we restrict attention to a particular kind of quantum system $(\mathcal{H}, a \mapsto U_a, b \mapsto V_b)$ satisfying the Weyl commutation relations, there is good reason to think time reversal can be treated in the standard way, as having the effect $U_a \mapsto U_a$ and $V_b \mapsto V_{-b}$. However, whether or not this is the *only* way to treat time reversal depends on one's interpretive commitments. The flexible interpreter of time reversal can take T to be any antiunitary operator, and Proposition 2.2 will provide a representation in which it behaves like the standard time reversal operator. On this view, although the standard transformation rules are always available, they are no more a necessity than is any particular representation of the commutation relations. On the other hand, the more rigid interpreter of time reversal can take Proposition 2.3 to *guarantee* these transformation rules, as a matter of necessity, when taken together with the stronger interpretive commitments of of the homogeneity and spatial independence.

2.5 Conclusion

Apart from dissolving some common mythology, the picutre I would like to advocate is one in which the meaning of time reversal is built up in stages. The first stage of commitment demands that the direction of time does not determine whether or not two states are mutually exclusive. This implies that time reversal preserves transition probabilities, and so is either unitary or antiunitary. The second stage demands that time reversal be an invariance of at least one non-trivial, physically plausible system. This guarantees that time reversal is antiunitary. Contrary to critics of the standard view, I see no good way to get reject this property. Resilient philosophers try to reject the assumptions introduced here that guarantee antiunitarity. But given Uhlhorn's theorem, together with Proposition 1, it seems to me that the price of this resilience is an implausibly peculiar characterization of time reversal.

The third stage of commitment involves interpreting systems described by the canonical commutation relations. This stage appears affords some interpretive freedom for philosophers of time. From the requirement that the time reversal operator be *any* antiunitary involution, it follows that the standard transformation rules are always available as a convention, since

Proposition 2.2 provides a representation in which these transformation rules hold. If the philosopher of time brings further requirements to the nature of time reversal, then these transformation rules can be viewed as more than a convention. Namely, if time reversal does not select a preferred region in space, and if it respects spatial independence, then the standard transformation rules are the only ones available.

I expect that some philosophers of time may disagree as to which level of commitment is the correct one. My hope is that, regardless of which level of commitment one choses, the approach advocated here may help clarify the mathematical and physical consequences that the interpreter of time reversal is committed to.

3.0 DOES QUANTUM TIME HAVE A PREFERRED DIRECTION?

3.1 Introduction

There is an old question about the extent to which spatial position and motion in time are sufficient to characterize the nature of matter. Robert Boyle hesitantly aligned himself with one response to this question, writing,

> whereas those other philosophers give only a general and superficial account of the phaenomena of nature... both the Cartesians and the Atomists explicate the same phaenomena by little bodies variously figured and moved (Boyle 1772, p.355).

In this paper I would like to discuss an analogous view that was ubiquitous in the early days of quantum theory, that spatial position and motion are sufficient to characterize the behavior of an elementary particle.

The real heyday of this ontology was before the electron's spin was discovered. The hydrogen atom was at the time characterized entirely by placement of electrons in "orbit" around the nucleus. Indeed, Heisenberg later reported being "psychologically" unprepared for Kronig's proposal that the electron had an internal spin not reducible to changes in position, recalling, "I just said, 'That is a very funny idea and very interesting,' but in some way I pushed it away" (AIP 1963).

I would like to point out one consequence of the ontology preferred by Heisenberg, which is perhaps unexpected. Namely, there is a sense in which, if a Galilei invariant system has no internal degrees of freedom, then motion cannot develop in a preferred direction in time. If such a system can develop in time at all, then the time reverse of that development can also occur.

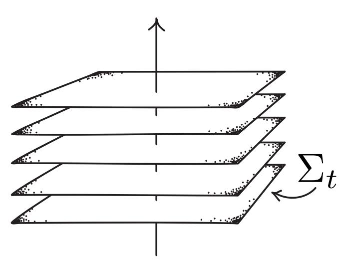
In physical language, the conclusion is that such systems *time reversal invariant.* The precise ontology of "no internal degrees of freedom"that I have in mind is one in which the position observable forms a *complete set of commuting observables.* The claim that I will argue for, then, is that whenever position forms a complete set of commuting observables, ordinary (Galilei invariant) quantum theory is time reversal invariant.

In what follows I will state and prove a precise expression of this claim. Section 2 introduces the basic structures and notation for the discussion at an elementary level; this section may be skimmed or skipped. Section 3 sets out the definition of Galilei invariance, and defines what it means to have no internal degrees of freedom. In Section 4, I discuss the meaning of time reversal invariance, and then state and prove a theorem that I argue captures the claim above. Section 5 discusses how the machinery of the theorem in more informal terms.

3.2 Basic Structures

3.2.1 Space and time

We will characterize space and time in terms of the smooth 4-dimensional manifold $\mathbb{R}^4$. To individuate space from time, we slice the block into a family of parallel hypersurfaces $\{\Sigma_t : t \in \mathbb{R}\}$, in such a way that each surface Σ_t is a 3-dimensional Euclidean space representing some moment in time t.



The 3-dimensional spaces Σ_t are made up of regions. More precisely, a *spatial region* $\Delta \subseteq \Sigma_t$ will be any open set or countable union or intersection of open sets – these are sometimes called the *Borel* sets. They can be assigned three Cartesian coordinate axes, allowing us to label a point p in a spatial region as $p = (x, y, z)$. For convenience of exposition, let us restrict attention to a single one of these axes. That is, let Δ represent a (Borel) set of the real numbers $\mathbb{R}$.

Here is how quantum theory appropriates these structures. Pure states (which capture the probabilities associated with basic experimental setups in quantum theory) can be represented by rays in a Hilbert space $\mathcal{H}$. This Hilbert space is a vector space, among other things[1], which we take to have a countably infinite basis set. But it is also a considerable abstraction from the spacetime structure set out above. A connection must be made between the two. This can be done in two steps: first, we connect $\mathcal{H}$ to space; then we connect $\mathcal{H}$ to time.

We begin the first step by recognizing that, like a vector space, the Hilbert space $\mathcal{H}$ contains subspaces. A *projection operator* projects each vector in $\mathcal{H}$ onto a subspace of $\mathcal{H}$. The connection to the Euclidean spatial surface Σ_t is then made as follows: take each spatial region $\Delta \subseteq \Sigma_t$ and associate it with a projection operator E_Δ of $\mathcal{H}$. Since projection operators have eigenvalues 1 and 0, they have interpretive significance: we follow Mackey (1963) in taking them to represent the true-or-false outcomes of physical experiments. In particular, we take each *spatial projection* E_Δ to represent the proposition, "there is a particle in the spatial region Δ" (Figure 3.1). A state for which this proposition is true is an eigenvector of E_Δ with eigenvalue 1. A state for which it is false is an eigenstate with eigenvalue 0. A truth value can be assigned to this proposition by way of a particle detection experiment[2].

This completes the first step: quantum mechanics can talk about spatial position. In the second step, we need to talk about time.

For this, we recall that by slicing our spacetime into spatial surfaces Σ_t indexed by $t \in \mathbb{R}$, we introduced a time axis. This axis admits a natural notion of "time development" or *translation* forward or backward in time, represented by the group of real numbers under

[1] A *Hilbert space* $\mathcal{H}$ is a vector space over the complex number field $\mathbb{C}$, equipped with a definite inner product $\langle \cdot, \cdot \rangle : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$, with respect to which it is Cauchy complete. A Hilbert space with a countable basis set is called *separable*.

[2] I have said nothing about how to interpret a "detection experiment," and philosophers of quantum mechanics are also welcome to do so however they like. The question of which interpretation of quantum theory appropriately captures measurement will be completely bracketed in the present discussion, so long as the basic structures of orthodox quantum mechanics are retained.

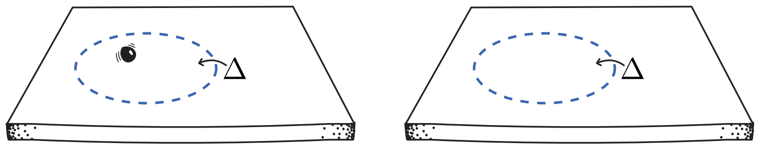


Figure 3.1: The projection operator E_Δ has eigenvalue 1 when an experimental detection occurs in the spatial region Δ (left), and eigenvalue 0 when it does not (right).

addition, $(\mathbb{R}, +)$. Each $t \in \mathbb{R}$ in this group can be interpreted as a time development for a duration t.

The connection to Hilbert space $\mathcal{H}$ is now made by assigning each t to an operator $\mathcal{U}_t : \mathcal{H} \to \mathcal{H}$. We do this in a way that is strongly continuous[3], in order to preserve the assumption that systems evolve continuously in time. We also do it in a way that preserves the way that temporal durations can be added and subtracted, by requiring that $\mathcal{U}_t\mathcal{U}_{t'} = \mathcal{U}_{t+t'}$. The assignment $t \mapsto \mathcal{U}_t$ is then called a *strongly continuous representation* of the group of time translations on $\mathcal{H}$. We interpret it to represent the translation of a pure state $\psi \in \mathcal{H}$ forward or backward in time by a duration t (Figure 3.2).

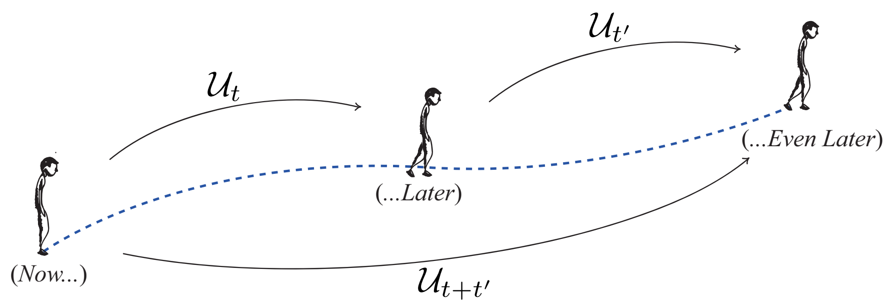


Figure 3.2: The operators $\mathcal{U}_t$ represent time translation by a duration t, and are isomorphic to the group $(\mathbb{R}, +)$.

Finally, motivated by the constraints of Wigner's theorem, we take each $\mathcal{U}_t$ in the

[3] A one-parameter group of operators is U_t *strongly continuous* if the function $f(t) = \langle \psi, U_t\psi \rangle$ is continuous for all $\psi \in \mathcal{H}$.

representation to be a *unitary operator*, meaning that it is a linear bijection such that $\mathcal{U}^*\mathcal{U} = \mathcal{U}\mathcal{U}^* = I$. A generalization of Wigner's theorem due to Uhlhorn (1963) shows that if a transformation U preserves *orthogonality*, in that $\langle\psi, \phi\rangle = 0$ if and only if $\langle U\psi, U\phi\rangle$, then U is unitary or antiunitary, the latter being excluded when $\mathcal{U}_t$ is a continuous group[4]. Orthogonality captures what it means for two states to represent two mutually exclusive measurement outcomes. (For example, a spin-up and a spin-down eigenstate are orthogonal; their inner product is zero.) Thus, the assumption that $\mathcal{U}_t$ is unitary can be viewed as capturing the requirement that facts about orthogonality do not change over time. For a discussion of this motivation in greater detail, see the Chapter 2 Section 2.2.

In summary, let $\mathcal{H}$ be a Hilbert space with a countably infinite basis set, whose rays represent the pure states of a quantum system; let $\Delta \mapsto E_\Delta$ be a projection valued measure, from regions of space to the lattice of projections on $\mathcal{H}$; and let $t \mapsto U_t$ be a strongly continuous unitary representation of the time translation group $(\mathbb{R}, +)$. The triple $(\mathcal{H}, \Delta \mapsto E_\Delta, t \mapsto U_t)$ contains the basic elements of a quantum description of a particle in space and time. It will be the object of our analysis.

3.2.2 Initial position and velocity observables

The projection valued measure $\Delta \mapsto E_\Delta$ on regions of space uniquely defines[5] a self-adjoint operator Q, which I will refer to as the position observable associated with E_Δ, or simply the *position observable*. Those familiar with the formalism will recognize this as the object standing in the canonical commutation relation,

$$[Q, P]\psi := (QP - PQ)\psi = i\psi, \tag{3.1}$$

for all ψ in the common dense domain of Q and P. However, note that the triple $(\mathcal{H}, \Delta \mapsto E_\Delta, t \mapsto \mathcal{U}_t)$ does not *presume* the existence of a self-adjoint operator P satisfying the canonical commutation relation. Nevertheless, we will now see that this triple does allow us to *construct* an operator $\dot{Q}$, which we may interpret as "velocity," using the given notions of

[4]If $\mathcal{U}_t$ is continuous or *Borel* Hilbert space symmetries, then it is a unitary group (Varadarajan 2007, p.288-9).

[5]In particular, $Q := \int_{\mathbb{R}} \lambda dE_\lambda$, where E_λ is the projection associated with the set $(-\infty, \lambda)$, and the $\int$ is the Lebesgue-Stieltjes integral. See (Jauch 1968, esp. §4.3) for an introduction.

position and time translation. In the next section, we will then introduce assumptions that allow us to *prove* that there exists a non-zero real number μ such that Q and $\mu\dot{Q}$ satisfy the canonical commutation relation (3.1).

To construct a velocity operator, we make use of our representation of time translation $t \mapsto \mathcal{U}_t$. In the Heisenberg picture, $\mathcal{U}_t$ determines how each operator, such as the position observable Q defined above, changes over time. In particular, at an arbitrary time t, position changes over time as

$$Q(t) = \mathcal{U}_t Q \mathcal{U}_t^*,$$

This allows us to refer to the operator $Q = Q(0)$ as the *initial position* observable (although we will often drop the "initial" part). It also allows us to consider the rate of change of this position observable with respect to time, that is, the velocity $\frac{d}{dt}Q(t)$. We can thus similarly refer to the operator $\dot{Q} := \frac{d}{dt}Q(t)|_{t=0}$ as the *initial velocity* operator.

To get a fix on the particular form of $\dot{Q}$, recall that $\mathcal{U}_t$ can always be written $\mathcal{U}_t = e^{itH}$, for a unique self-adjoint operator H called the *Hamiltonian*[6]. So, since $Q = Q(0)$ is independent of time, we have that

$$\begin{aligned}
\dot{Q} : &= \frac{d}{dt}Q(t)|_{t=0} = \frac{d}{dt}\left(\mathcal{U}_t Q \mathcal{U}_t^*\right)|_{t=0} \\
&= \left((\frac{d}{dt}e^{itH})(Qe^{-itH}) + (e^{itH})Q(\frac{d}{dt}e^{-itH}) \right)|_{t=0} \\
&= i(He^{itH}Q - e^{itH}QH)e^{-itH}|_{t=0} \\
&= i[H, Q].
\end{aligned}$$

where the second line makes use of the chain rule, the third a formal derivative, and the final one evaluates at $t = 0$. We may check that this $\dot{Q}$ is Hermitian by observing that $i^* = -i$ and $[H, Q]^*\psi = -[H, Q]\psi$, and hence that $\dot{Q}^*\psi = \dot{Q}\psi$, for all ψ in the common domain of $\dot{Q}$ and $\dot{Q}^*$. Since the domains of $\dot{Q}$ and $\dot{Q}^*$ coincide, it is also self-adjoint[7]. In this sense, we can think of the initial velocity $\dot{Q}$ as an "observable."

[6] This fact follows from Stone's theorem (Blank et al. 2008, Thm. 5.9.2).

[7] An operator A is *Hermitian* if $A\psi = A^*\psi$ for all ψ in the domain of A. It is *self-adjoint* if in addition A and A^* have the same domain. The former need not imply the latter when A is unbounded. In this case of $\dot{Q} = i[H, Q]$, both H and Q are self-adjoint. Hence, the domains of H and H^* coincide, as to those of Q and Q^*, and so the domains of $\dot{Q}$ and $\dot{Q}^*$ coincide as well.

3.3 Galilei Invariance

Ordinary, low-relative-velocity quantum mechanics is Galilei invariant. For us, the experimental significance of this will be the following.

Suppose two particle physics experiments are performed in different laboratories, and that the only difference between them is that they were set up in different spatial locations. Suppose one assumes that these two experiments will record the same result; this expresses the fact that particle physics is invariant under *spatial translations*. Similarly, suppose the two experiments are performed at different constant velocities, but are otherwise identical. For example, one experiment might take place on a boat traveling with uniform speed, while the other takes place on shore. These two experiments will again record the same results. This captures the claim that particle physics is invariant under *velocity boosts*. In particular, in the present case of ordinary quantum mechanics, these will be the Galilei boosts. The two assumptions of invariance under spatial translations and Galilei boosts will be referred to collectively as *Galilei invariance*.

We need to translate these assumptions into constraints on an elementary quantum system $(\mathcal{H}, \Delta \mapsto E_\Delta, t \mapsto \mathcal{U}_t)$. Such a system is associated with an initial position Q and an initial velocity $\dot{Q}$, as we saw in the previous section. The definition of translations and Galilei boosts will be formulated in terms of these operators.

To begin, we take spatial translation by a length $a \in \mathbb{R}$ to have the effect of translating the initial position $Q \mapsto Q + aI$, while leaving velocity fixed $\dot{Q} \mapsto \dot{Q}$. One might note that this definition of translations acts on self-adjoint operators, such as Q, rather than actual measurable values of spatial position. The latter are represented by the *spectrum* of Q. However, there is no harm in this definition of translations. For, as one can easily check[8], the mapping $Q \mapsto Q + aI$ can be a derived from a translation of the spectrum by $E_\Delta \mapsto E_{\Delta - a}$. Thus, our definition really does amount to a spatial translation in the required sense. For the same reason, we can take Galilei boosts by a velocity $b \in \mathbb{R}$ to have the effect

[8] To verify, consider how $E_\Delta \mapsto E_{\Delta-a}$ effects the position observable Q, defined as $Q := \int_{\mathbb{R}} \lambda dE_\lambda$. By the functional calculus, $\int_{\mathbb{R}} f(\lambda) dE_\lambda = f(Q)$ for any Borel function f. So, $E_\Delta \mapsto E_{\Delta-a}$ has the effect of mapping $Q \mapsto \int_{\mathbb{R}} \lambda dE_{\lambda-a} = \int_{\mathbb{R}} (\lambda + a) dE_\lambda = Q + aI$, where the first equality substitutes $\lambda + a$ for λ, and the second follows from the functional calculus.

of "boosting" the initial velocity observable $\dot{Q} \mapsto \dot{Q} + bI$, while fixing the initial position $Q \mapsto Q$.

What does it mean for a system to be *invariant* under translations and Galilei boosts? *We will not* require that these transformations leave the dynamics $t \mapsto \mathcal{U}_t$ (or the Hamiltonian H) unchanged. Rather, we require only that these be transformations these transformations be unitary, motivated by the constraints of Wigner's Theorem discussed in the previous section. This leads to the following.

Definition 3.1 (Galilei invariance)**.** The structure $(\mathcal{H}, \Delta \mapsto E_\Delta, t \mapsto \mathcal{U}_t)$ is *Galilei invariant* if there exist two strongly continuous one-parameter unitary groups S_a (translations) and R_b (boosts), which form a representation of the additive real vectors $(\mathbb{R} \times \mathbb{R}, +)$, and such that

$$S_a Q S_a^{-1} = Q + aI \qquad\qquad S_a \dot{Q} S_a^{-1} = \dot{Q}$$
$$R_b Q R_b^{-1} = Q \qquad\qquad R_b \dot{Q} R_b^{-1} = \dot{Q} + bI$$

for all $a, b \in \mathbb{R}$.

Another way to put Definition 3.1 is to say that for any $a, b \in \mathbb{R}$, the pairs of operators $(Q, \dot{Q})$, $(Q + aI, \dot{Q})$ and $(Q, \dot{Q} + bI)$ are all *unitarily equivalent*, meaning that each is related to the other by a unitary transformation.

3.4 A minimalist ontology

Our task is now to determine the consequences of requiring that a Galilei invariant quantum system admit no "internal" degrees of freedom. Such systems are not described in terms of any "spin" or "charge" parameters. More precisely, we will be restricting our attention to quantum systems in which the position observable Q forms a *complete set of commuting observables*. In this section, I would like to discuss the precise meaning of this requirement.

Here is a fanciful way to put it[9]: our ontology will be one in which particles "have no hair." If particles have hair, then we can distinguish them using properties of that hair.

[9] I owe this expression to John Earman.

For example, one particle might have a patch of hair oriented in the $+x$ direction. Another might have a patch oriented in the $-x$ direction. But if we take the hair away, then the only way to distinguish the particles at any given moment in time is through their spatial position.

One makes this idea precise by requiring all observables that are "independent" of initial spatial position Q to be expressible as functions of Q. To see what it means to be a "function" of Q, notice that composing the position observable with itself gives rise to a new self-adjoint operator, which maps a vector ψ in the domain of Q to $Q^2\psi = (Q \circ Q)\psi$. This operator is different than the position observable. Nevertheless, it is in an obvious sense a "derived entirely" from it. The same holds of any polynomial in Q, such as $Q^2 + Q + 41$. In fact, we need not even restrict ourselves to polynomials; any Borel function[10] of Q will suffice, and the set of all Borel function forms a commutative algebra, $\mathfrak{A}_Q$. We will refer to it as the *algebra generated by* Q. In an important sense, this algebra captures the class of operators that are"derived entirely from" the position observable[11].

Now, consider an observable S that commutes with the position operator Q. One may interpret S to be "independent" of (or "measurable simultaneously" with) Q. The set of all such observables that are not functions of each other is called a *complete set of commuting observables*. One typically uses such a set to characterize the independent degrees of freedom of a quantum system[12]. The quantum systems of interest to us, quite simply, are those for which Q is the only observable in that set. We formulate this requirement as follows

Definition 3.2 (Q is a CSCO)**.** A self-adjoint operator Q forms a *complete set of commuting observables* if for every (closed[13]) linear operator A, if $AQ = QA$, then A is in the algebra $\mathfrak{A}_Q$ of functions of Q.

This requirement characterizes the ontology of "no internal degrees of freedom" that is

[10] Let $f : \mathbb{R} \to \mathbb{C}$ be a Borel function that is defined almost everywhere on the spectral measure $\Delta \mapsto E_\Delta$. Then f defines a function of $Q = \int_{\mathbb{R}} \lambda dE_\lambda$, given by $f(Q) := \int_{\mathbb{R}} f(\lambda) dE_\lambda$; see (Blank et al. 2008, §5.5).

[11] One may also recognize it as the *von Neumann algebra generated by* Q, equal to the weak closure of $\{Q\}$, and to the double-commutant $\{Q\}''$.

[12] Earman (2008, §5) points out that the existence of such a set can be viewed as a *sine qua non* in the description of a quantum system.

[13] Closure is a technical requirement guaranteeing A will be sufficiently well-behaved. An operator A is *closed* if, whenever a sequence ψ_i in the domain of A is such that $\psi_i \to \psi$ and $A\psi_i \to \phi$, then it follows that ψ is in the domain of A, and $A\psi = \phi$. Closed operators are continuous on their domain, but need not be bounded.

of interest to us.

3.5 Time reversal invariance

We now have a handle on a what it means for an ordinary, Galilei invariant quantum system to have no internal degrees of freedom: it means that the position observable Q forms a complete set of commuting observables. What I would now like to point out is that there is a precise sense in which these systems *fail to distinguish the past from the future.*

3.5.1 The definition of time reversal

To begin, let me briefly introduce the standard definition of the *time reversal transformation.* Roughly, this is the transformation that takes one between a future-directed quantum system and a past-directed one.

One can ground intuitions about this transformation by thinking about what happens to a film when it is run in reverse. For example, reversing a film of a ball rolling down an inclined plane leads to a depiction in which the directions of velocities are reversed, while all the same positions occur, although they occur in the reverse order (Figure 3.3).

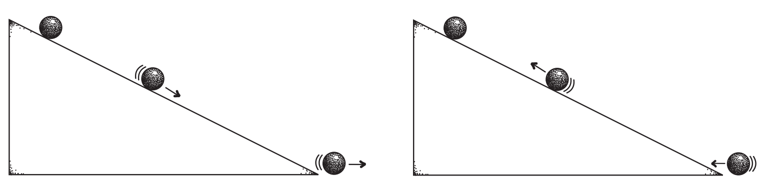


Figure 3.3: Time reverse of a ball rolling down an inclined plane.

As with our previous definitions, we capture time reversal as a transformation of an elementary quantum system $(\mathcal{H}, \Delta \mapsto E_\Delta, t \mapsto \mathcal{U}_t)$. As I have argued elsewhere[14], the

[14]See the previous chapter. Note in particular that, in the course of proving the T-theorem, I show that Q and $\mu\dot{Q}$ form a complete set of commuting observables. So, the argument of the previous chapter allows one to conclude that time reversal has the effect $Q \mapsto Q$ and $\mu\dot{Q} \mapsto -\mu\dot{Q}$.

standard definition of time reversal can be given a systematic philosophical underpinning. Rather than repeat that argument here, let me simply summarize the standard definition of the time reversal transformation[15]. It is the transformation that takes each trajectory $\psi(t)$ to $T\psi(-t)$, where T is a time reversal operator in the following sense.

Definition 3.3. A *time reversal operator* for a quantum system $(\mathcal{H}, \Delta \mapsto E_\Delta, t \mapsto \mathcal{U}_t)$ is a bijection $T : \mathcal{H} \to \mathcal{H}$ such that the following hold.

i. T is *faithful* in that it preserves positions and reverses velocities; $TQT^{-1}\psi = Q\psi$ and $T\dot{Q}T^{-1}\phi = -\dot{Q}\phi$, for all $\psi \in \mathcal{D}_Q$, $\phi \in \mathcal{D}_{\dot{Q}}$.
ii. $T : \mathcal{H} \to \mathcal{H}$ is *antiunitary* in that T is antilinear and $T^*T = TT^* = I$.
iii. T is an *involution* in that $T^2 = cI$ for some $c \in \mathbb{C}_{unit}$. In other words, T is a 'reversal,' which when applied twice brings us back to where we started (up to an arbitrary phase factor).

Notice that we do not yet have any assurance that there is a unique such operator. That uniqueness is indeed desirable, and it will be established in the theorem below.

3.5.2 Time reversal invariance

If we have a time reversal operator T, then we can say what it means for a quantum system $(\mathcal{H}, \Delta \mapsto E_\Delta, t \mapsto \mathcal{U}_t)$ to be T-reversal invariant. In short, it means that the dynamics for that system are reversible under the transformation of a time reversal operator T. More precisely:

Definition 3.4. A quantum system $(\mathcal{H}, \Delta \mapsto E_\Delta, t \mapsto \mathcal{U}_t)$ is *T-reversal invariant* for some bijection T just in case T satisfies the requirements (i)-(iii) of a time reversal operator, together with the requirement that,

iv. $T\mathcal{U}_tT^{-1} = \mathcal{U}_{-t}$,

for all $t \in \mathbb{R}$.

[15]Note that these properties are not satisfied by the "non-standard" definitions of time reversal advocated by Albert (2000, p.11), Callender (2000, §V), and Maudlin (2007, §4.2). Advocates of the non-standard view may substitute their preferred name for what I am calling "time reversal"; for example, Callender calls it "Wigner time reversal."

If there is a *unique* T satisfying (i), (ii) and (iii), and if that T also satisfies (iv), then there is no ambiguity in simply saying that our system is *time reversal invariant*. This is the precise characterization of what it means for a system to "fail to distinguish between past and future."

It will be useful to observe a few equivalent statements of time reversal invariance, in view of the argument of the next section. Namely, if T is a time reversal operator, then one can show[16] that the following statements are equivalent.

- T satisfies (iv): $T\mathcal{U}_t T^{-1} = \mathcal{U}_{-t}$.
- If $\psi(t) = \mathcal{U}_t\psi$ (where $\mathcal{U}_t = e^{itH}$ and $\psi \in \mathcal{H}$) is a solution to the Schrödinger equation $i(d/dt)\psi(t) = H\psi(t)$, then $T\psi(-t)$ is also a solution to the Schrödinger equation with the same Hamiltonian H.
- $[T, H]\psi = 0$, for all ψ in the domain of H (and where $\mathcal{U}_t = e^{itH}$).

For a time reversal operator T, all of these statements equally capture the meaning of T-reversal invariance.

3.5.3 A T-Theorem

Here is the central fact that I would like to point out about all this. It can be stated as a theorem, which follows straightforwardly from a lemma due to Jauch. That lemma shows

[16] The equivalence of the second and third points was pointed out by Earman (2002, p.248). The equivalence of the first and the third is established as follows. Write $T\mathcal{U}_t T^{-1} = Te^{itH}T^{-1} = e^{-itTHT^{-1}}$, where the last equality follows from the functional calculus. (Here is another way to see this, for the so-called "analytic" vectors for H. For these vectors, e^{itH} can be written in as an infinite sum (Blank et al. 2008, Proposition 5.6.1). That is,

$$\begin{aligned} Te^{itH}T^{-1} &= T(I + (itH) + (1/2!)(itH)^2 + \cdots)T^{-1} \\ &= I + T(itH)T^{-1} + (1/2!)T(itH)^2T^{-1} + \cdots \\ &= I + T(itH)T^{-1} + (1/2!)(T(itH)T^{-1})^2 + \cdots \\ &= e^{T(itH)T^{-1}}, \end{aligned}$$

where the penultimate equality follows from the fact that $T^{-1}T = I$. We assumed T is antiunitary, and all antiunitary operators are antilinear, meaning that they conjugate complex numbers. So, $Te^{itH}T^{-1} = e^{T(itH)T^{-1}} = e^{-itTHT^{-1}}$.) Thus, if the third point holds and $THT^{-1} = H$, then $Te^{itH}T^{-1} = e^{-itH}$ and we have the first point. Conversely, if the first point holds and $Te^{itH}T^{-1} = e^{-itH}$, then $e^{-itH} = Te^{itH}T^{-1} = e^{-itTHT^{-1}}$. But since both H and THT^{-1} are self-adjoint, and Stone's theorem guarantees $\mathcal{U}_{-t}$ has a *unique* self-adjoint generator, it follows that $THT^{-1} = H$, and we have the third point.

that our requirements provide a very strong restriction on the form of the Hamiltonian H.

Lemma 3.1 (Jauch)**.** *Suppose $(\mathcal{H}, \Delta \mapsto E_\Delta, t \mapsto \mathcal{U}_t)$ is Galilei invariant, and that the self-adjoint operator Q associated with $\Delta \mapsto E_\Delta$ forms a complete set of commuting observables. Then $(\mathcal{H}, e^{ibQ}, e^{ia\mu\dot{Q}})$ is an irreducible unitary representation of the canonical commutation relations in Weyl form*[17]*, and $\mathcal{U}_t = e^{itH}$, where*

$$H\psi = \frac{\mu}{2}\dot{Q}^2\psi + v(Q)\psi$$

for some non-zero real number μ, some Borel function v, and for all ψ in the domain of H.

Proof. See Jauch (1964, 1968). My reconstruction of the proof is given in the Appendix. □

Proposition 3.1. *Suppose $(\mathcal{H}, \Delta \mapsto E_\Delta, t \mapsto \mathcal{U}_t)$ is Galilei invariant, and that the self-adjoint operator Q associated with $\Delta \mapsto E_\Delta$ forms a complete set of commuting observables. Then there exists a unique bijection $T : \mathcal{H} \to \mathcal{H}$ such that,*

(i) (faithfulness) $TQT^{-1} = Q$ and $T\dot{Q}T^{-1} = -\dot{Q}$;

(ii) (antiunitarity) T is antiunitary;

(iii) (involution) $T^2 = cI$ for some $c \in \mathbb{C}_{unit}$.

Moreover, this T satisfies

(iv) (T-reversal invariance) $T\mathcal{U}_tT^{-1} = \mathcal{U}_{-t}$.

Proof. Let $(\mathcal{H}, e^{ia\mu\dot{Q}}, e^{ibQ})$ be the representation guaranteed by Jauch's lemma. We begin by constructing a distinct (Schrödinger) representation, and define a "conjugation operator" K_Q with respect to Q in that representation. This K_Q can then be used to construct an operator T satisfying conditions (i)-(iv).

Let $\mathcal{H}_Q$ be the Hilbert space of $L^2(\mathbb{R})$ functions $\psi(x)$ for which Q is the multiplication operator, $Q\psi(x) = x\psi(x)$. Let $P\psi(x) := i(d/dx)\psi(x)$, so that $(\mathcal{H}_Q, e^{iaP}, e^{ibQ})$ is the Schrödinger representation. Define the operator *conjugation operator* $K_Q : \mathcal{H}_Q \to \mathcal{H}_Q$ to be the operator such that $K_Q\psi(x) = \psi^*(x)$. We follow Messiah (1999, §XV.5) in noting three

[17] A triple $(\mathcal{H}, e^{iaA}, e^{ibB})$ is a unitary representation of the canonical commutation relations in Weyl form if A and B are self-adjoint and $e^{iaA}e^{ibB} = e^{iab}i^{ibB}e^{iaA}$. It is irreducible if it admits only trivial subrepresentations. These relations imply the more familiar canonical commutation relations $[A, B]\psi = i\psi$ for all $\psi \in \mathcal{D}_{A,B}$.

relevant properties of K_Q. First, it is antiunitary, since $K_Q^* = K_Q^{-1}$ and K is clearly antilinear. Second, it is an involution, since $K_Q^2\psi(x) = \psi^{**}(x) = \psi(x)$. Finally, since Q is pure real and P is pure imaginary in this representation, K_Q has the property that $K_Q Q K_Q = Q$ and $K_Q P K_Q = -P$.

This K_Q is not our desired time reversal operator, but will be used to construct it. The Stone-von Neumann theorem (Blank et al. 2008, Theorem 8.2.4) guarantees that there is a unitary bijection from $\mathcal{H}_Q$ to our original Hilbert space $\mathcal{H}$ (which we will call $W : \mathcal{H}_Q \to \mathcal{H}$) such that $WQW^* = Q$ and $WPW^* = \mu\dot{Q}$. We now define our time reversal operator to be the image of K_Q under this mapping:

$$T := WK_QW^*.$$

One may easily verify[18] that this T inherits properties (i), (ii) and (iii) from K_Q. To show that this T is unique up to a constant, suppose that both T and $\tilde{T}$ satisfy conditions (i)-(iii). So, they are both antilinear involutions satisfying

$$\begin{aligned} TQT^{-1} &= Q & \qquad \tilde{T}Q\tilde{T}^{-1} &= Q \\ T\dot{Q}T^{-1} &= -\dot{Q} & \qquad \tilde{T}\dot{Q}\tilde{T}^{-1} &= -\dot{Q}. \end{aligned}$$

Then $T\tilde{T}$ is a linear operator that commutes with both Q and $\mu\dot{Q}$, since

$$\begin{aligned} (T\tilde{T})Q(\tilde{T}^{-1}T^{-1}) &= TQT^{-1} = Q \\ (T\tilde{T})\mu\dot{Q}(\tilde{T}^{-1}T^{-1}) &= T(-\mu\dot{Q})T^{-1} = \mu\dot{Q}. \end{aligned}$$

[18] T is antiunitary because it is the composition of two unitaries and an antiunitary. It is also an involution:

$$T^2 = (WK_QW^*)(WK_QW^*) = WK_Q^2W^* = I.$$

And, it has the desired effect on Q and $\mu\dot{Q}$:

$$\begin{aligned} TQT^{-1} &= (WK_QW^*)(WQW^*)(WK_QW^*) \\ &= W(K_QQK_Q)W^* = WQW^* = Q \\ T\mu\dot{Q}T^{-1} &= (WK_QW^*)(WPW^*)(WK_QW^*) \\ &= W(K_QPK_Q)W^* = -(WPW^*) = -\mu\dot{Q}. \end{aligned}$$

.

But the representation $(Q, \mu\dot{Q})$ provided by Jauch's lemma is irreducible. By Schur's lemma, this implies that the only linear operators commuting with both Q and $\mu\dot{Q}$ are constant multiples of the identity. So, for some $k \in \mathbb{C}$,

$$kI = T\tilde{T} = T\tilde{T}^{-1},$$

where we have used the fact that $\tilde{T}$ is an involution in the second equality. Therefore, $T = k\tilde{T}$ as claimed.

Finally, we show that this T satisfies time reversal invariance (iv). Since $H = (\mu/2)\dot{Q}^2 + v(Q)$ by Jauch's lemma. So, from the fact that $\mu\dot{Q}$ and H are both self-adjoint, it follows that $v(Q)$ is self-adjoint as well. So, applying T to both sides of H, we get

$$THT^{-1} = \frac{\mu}{2}(T\dot{Q}T^{-1})^2 + Tv(Q)T^{-1} = \frac{\mu}{2}\dot{Q}^2 + Tv(Q)T^{-1}, \tag{3.2}$$

where by the functional calculus, $v(Q) = \int_{\mathbb{R}} v(\lambda)dE_\lambda$. But $v(Q)$ is known to be self-adjoint. The $v(\lambda)$ are thus real, and

$$Tv(Q)T^{-1} = \int_{\mathbb{R}} v(\lambda)dTE_\lambda T^{-1} = \int_{\mathbb{R}} v(\lambda)dE_\lambda = v(Q).$$

From this, together with (3.2), it follows that $THT^{-1} = H$, which establishes time reversal invariance. □

The force of the result is that in Galilei invariant quantum theory, if a system has no internal degrees of freedom, then time reversal invariance is guaranteed. Some might find the equivalent contrapositive statement even more interesting. Namely, if we have reason to believe that time reversal invariance fails in a Galilei invariant system, then that system *must* admit internal degrees of freedom.

3.6 Discussion of the T-Theorem

At first glance, the T Theorem may seem mysterious. So, let me make a few remarks about how the conditions of the T-theorem eliminate the possibility of specific systems that violate time reversal invariance.

3.6.1 Arbitrary vector potentials

The main step of the T-theorem is Jauch's lemma. This lemma shows that Galilei invariance and a lack of internal degrees of freedom restrict the Hamiltonian to have the form, $H = \frac{\mu}{2}\dot{Q}^2\psi + v(Q)$. However, the restriction is even more severe, in that it also eliminates the possibility that $\dot{Q}$ contain any non-trivial vector potentials.

Let us move to 3 dimensions, and write position and velocity as $\mathbf{Q} = (Q_1, Q_2, Q_3)$ and $\dot{\mathbf{Q}} = (\dot{Q}_1, \dot{Q}_2, \dot{Q}_3)$. Let $\mathbf{P} = (P_1, P_2, P_3)$ be the generator of the spatial translation group S_a (where we now interpret the parameter a to be a vector in $\mathbb{R}^3$). In electromagnetism, the velocity $\dot{\mathbf{Q}}$ can be expressed as the sum of a linear momentum $\mathbf{P}$ and a "vector potential" term $\mathbf{a}(\mathbf{Q})$:

$$\mu\dot{\mathbf{Q}} = \mathbf{P} + \mathbf{a}(\mathbf{Q}).$$

But the Hamiltonian of Jauch's lemma does not allow non-trivial vector potentials. More precisely, the assumptions of Galilei invariance and no internal degrees of freedom imply that $\mathbf{a}(\mathbf{Q})$ must be a constant (and, therefore, gauge-equivalent to zero).

To see why, note first that $\mathbf{a}(\mathbf{Q})$ commutes with $\mathbf{Q}$, since it is a function of $\mathbf{Q}$. But $\mathbf{a}(\mathbf{Q}) = \mu\dot{\mathbf{Q}} - \mathbf{P}$, and the right-hand side commutes with $S_a = e^{ia\cdot\mathbf{P}}$ by the definition of Galilei invariance. So, $\mathbf{a}(\mathbf{Q})$ commutes with both $\mathbf{Q}$ and $\mathbf{P}$. Since the representation is irreducible, this is only possible if $\mathbf{a}(\mathbf{Q})$ is a constant multiple of the identity, by Schur's lemma.

This is crucial for the T-Theorem. For, if a non-trivial vector potential were allowed, then one could immediately violate time reversal invariance by setting the function $\mathbf{a}(\mathbf{Q}) = Q$. Since T reverses P but not Q, it does not commute with $(P - Q)^2$. Therefore, T does not commute with H. As we noted in the previous section, this is equivalent to the failure of time reversal invariance by our definition (iv) $T\mathcal{U}_tT^{-1} = \mathcal{U}_{-t}$. Thus, the ontology under consideration makes time reversal invariance possible in part by eliminating the possibility of non-trivial vector potentials.

3.6.2 Systems that are not square in $\dot{Q}$

There are more systems that violate time reversal invariance, which Galilei invariance rules out in other ways. For example, consider a system with the Hamiltonian,

$$H = \dot{Q}.$$

Although it is somewhat unphysical, this system fails to be time reversal invariant as well. Namely, since a time reversal operator T satisfies $T\dot{Q}T^{-1} = -\dot{Q}$, it follows immediately that T does not commute with H, and so time reversal invariance fails.

This kind of system is also ruled out by the conditions of the T-theorem. To verify, recall our definition $\dot{Q} := i[H, Q]$ from the first section. Substituting $H = \dot{Q}$ then allows one to write,

$$\dot{Q} = i[\dot{Q}, Q]. \tag{3.3}$$

Given this constraint, there cannot exist a group R_b that "boosts velocity while fixing position," in that $R_b\dot{Q}R_b^{-1} = \dot{Q} + b$ and $R_bQR_b^{-1} = Q$. This can be seen immediately by surrounding both sides of this equation with R_b and R_b^{-1}, and noticing that this fixes the operator on the right hand side but not on the left[19]. The existence of such a group R_b would thus contradict Equation (3.3).

3.6.3 Arbitrary degrees of freedom

The lack of internal degrees of freedom also eliminates certain systems that violate time reversal invariance. Consider a quantum system with a degree of freedom σ that is independent of Q and not a function of it, so that our complete set of commuting observables is given by the set $\{Q, \sigma\}$. Suppose σ changes sign under time reversal, $T\sigma T^{-1} = -\sigma$ (as, for example, the "spin" observable does). This is possible independently of how T transforms Q, because σ is by assumption not a function of Q.

[19] The explicit calculation: $R_b i[\dot{Q}, Q]R_b^{-1} = i[R_b\dot{Q}R_b^{-1}, Q] = i[\dot{Q} + bI, Q] = i[\dot{Q}, Q]$, so the right hand side is fixed. On the other hand, $R_b\dot{Q}R_b^{-1} = \dot{Q} + bI$, so the left hand side is not.

This operator σ can now enter into a Hamiltonian in a way that violates time reversal invariance. Consider the Hamiltonian,

$$H = \frac{\mu}{2}\dot{Q}^2 + \sigma.$$

Then time reversal invariance again fails, because $THT^{-1} = (\mu/2)\dot{Q}^2 - \sigma \neq H$. Requiring that Q be a complete set of commuting observables eliminates the possibility of this kind of failure.

These remarks suggest there ought to be a more general T-theorem available. After all, time reversal invariance will follow whenever (1) all the observables appearing in the Hamiltonian either reverse or stay fixed under time reversal, and (2) the Hamiltonian is an even function of all the observables that reverse. (As a simple example, the Hamiltonian $H = \dot{Q}^4$ describes a time reversal invariant system, but is not allowed by the conditions of our theorem.) From this perspective, the conditions of Galilei invariance and no internal degrees of freedom are far stronger than necessary. However, whether there is an interesting generalization of the T-theorem along these lines remains an open question.

3.7 Conclusion

Those whose basic ontological commitments include only position have this fact to contend with: if the ontology of Galilean quantum theory includes no internal degrees of freedom, then time reversal invariance is guaranteed. As it turns out, this kind of result is not unique to quantum theory: one can prove a related theorem in classical Hamiltonian mechanics, which I discuss in a Chapter 4, Section 4.3.2. This suggests a fairly robust sense in which a minimal ontology prohibits time asymmetric phenomena.

It should be noted that this result is perfectly compatible with the famous experimental violation of time reversal invariance performed by Christenson, Cronin, Fitch and Turlay (1964), in the decay of neutral kaons brought on by the weak interaction. In fact, both of the major premises of our theorem are negated in case of neutral kaons: that model is assumed to be Lorentz invariant (negating our Galilei invariance assumption), and to admit the internal degree of freedom known as “strangeness” (negating our assumption of no internal degrees of freedom). The experimental setup is therefore well outside the scope of our theorem.

An important open question is thus whether time reversal invariance can be established on the assumption of no internal degrees of freedom, but with Lorentz invariance instead of Galilei invariance. This is a question for future research. For now, the claim that I would like to commit to is simply that, at least in Galilei invariant quantum theory, it is no accident that time asymmetric systems admit internal degrees of freedom. Such properties are absolutely essential to the phenomenon.

4.0 WHEN WE DO (AND DO NOT) HAVE A CLASSICAL ARROW OF TIME

4.1 Introduction

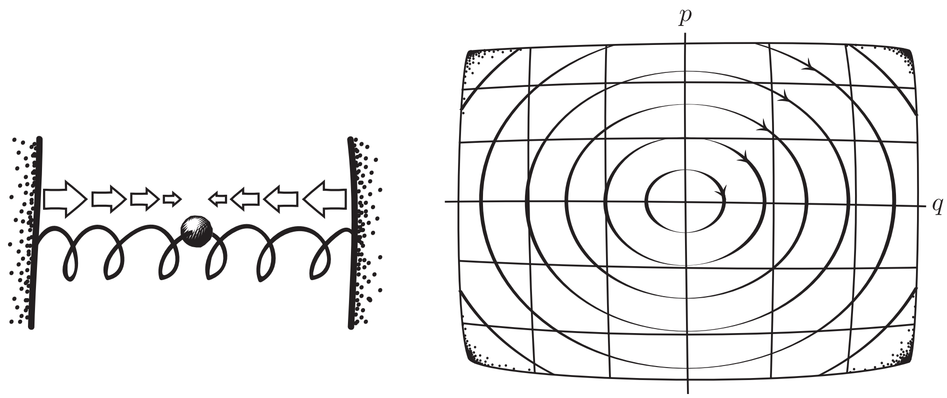
Contrary to popular belief, there are a number of ways that time reversal invariance can fail in classical mechanics. In this chapter, I review several common claims about time reversal invariance in classical mechanics, and show how they are incorrect without further qualification. I then propose two positive qualifications within the Hamiltonian formulation of mechanics, and show that they are sufficient for time reversal invariance. First, I point out that time reversal invariance follows whenever velocity is proportional to momentum. Second, I show that time reversal invariance can be seen to hold for a broad class of "ordinary" classical systems, where "ordinary" is qualified by the presence of Galilei covariance.

4.1.1 Example: the harmonic oscillator

To get an appropriate fix on the meaning of time reversal, we'll need to discuss both the Newtonian "force" formulation as well as the Hamiltonian formulation of classical particle mechanics. To keep things simple, let's begin with the example of a bob on a spring.

In the Newtonian formulation, the furniture of the world consists of force fields and point particles in space. For the simple harmonic oscillator, there is a force field that grows in proportion to the distance from a central point, and there is a massive particle located somewhere in that force field, as shown in Figure 1(a). The motion of the system is governed by Newton's law, which sets the acceleration of the particle proportional to the force.

On the Hamiltonian formulation, we don't really need to say what the furniture of the



(a) (b)

Figure 4.1: (a) Force field and (b) phase diagram for a harmonic oscillator.

world is, except that it can be characterized by a manifold of states ("phase space"), each of which can be assigned an energy value by a smooth function that we call the Hamiltonian. For the harmonic oscillator, those states can be written as position-momentum pairs $(q, p) \in \mathbb{R}^2$, and the energy values increase with the square of the distance and momentum, $h = ap^2 + bq^2$. The motion of the system is determined by Hamilton's equations, which for the oscillator say that the change in q and p goes along an ellipse of constant energy, such as those in Figure 1(b).

Notice that in this diagram, *the clockwise direction of the arrows matters.* We take a positive p to represent "rightward pointing momentum," and a negative p to represent "leftward-pointing momentum." So, as the state of the bob winds clockwise along the top half of an ellipse, the bob moves to the right in space with rightward-pointing momentum. As it winds along the bottom half, the bob moves to the left in space with leftward-pointing momentum. But, on this interpretation of the coordinate axes, it makes no sense for the state of the bob to move around the ellipse in the counterclockwise direction. That would imply that the bob's momentum goes in the opposite direction of its motion, which contradicts a background assumption about the harmonic oscillator, that $\dot{q} = mp$ (with $m > 0$).

This is an important distinction: I say that such a motion would "contradict a background assumption" about velocity and momentum in this system. This is *not* to say that it would be “impossible according to the laws of motion.” Indeed, the statement that $\dot{q} = mp$ is logically independent of whether or not Hamilton’s equations are satisfied. Thus, violating that statement is not a violation of the laws of motion. It is simply a meaningless way to describe a system like a bob on a spring. This is perhaps an obvious point, but crucial to bear in mind as we turn to the meaning of time reversal.

4.1.2 The meaning of time reversal

Suppose we film our harmonic oscillator bobbing back and forth, and then play the film in reverse. The result would be a new “reversed” motion of a bob on a spring. This transformation is roughly what will be meant by the *time reversal transformation*.

How should this transformation be described mathematically? In the Newtonian formulation, it is simply the reversal of the order of events in a trajectory $x(t)$. That is, if $x(t)$ is the curve describing the position of the bob over time, then the time-reversed trajectory is given by $x(-t)$. In this formulation, time reversal has no effect on the initial state $x(0) \in \mathbb{R}$. Reversing the order of events in a trajectory enough.

In the Hamiltonian formulation, reversing the order of events in a trajectory $(q(t), p(t))$ is not enough. In terms of the phase space depicted in Figure 1(b), an order-reversal by itself would just reverse the direction of the arrows, from the clockwise to the counterclockwise orientation. As we noted at the end of the last subsection, this would say that as the bob moves to the right, it has leftward-pointing momentum. This is *not* an example of the failure of time reversal invariance, because it is not a violation of the laws of motion. It is simply a meaningless statement about the system, given that we take it to satisfy $p = m\dot{q}$. This transformation is therefore not a plausible candidate for time reversal.

A better candidate for time reversal is obtained by observing that time reversal in the Hamiltonian formulation has two parts. First, time reversal requires a transformation of phase space that reverses momenta and preserves position: $T(q, p) = (q, -p)$. Second, it requires reversing the order of events in each trajectory. These two parts of the transformation are displayed Figure 4.2. The result is a transformation that satisfies our background assumption that $p = m\dot{q}$.

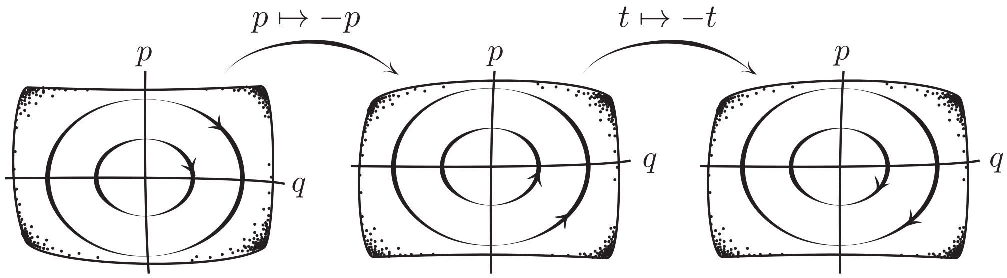


Figure 4.2: Time reversal of the harmonic oscillator, in two steps. The first ($p \mapsto -p$) is a mirror flip about the q axis; the second ($t \mapsto -t$) reverses the order of events.

The operator T appearing in the first part of the transformation is referred to as the *time reversal operator.* One can speak quite generally about time reversal operators, in a way that applies to both the Hamiltonian and the Newtonian formulations. In general, the time reversal operator is a bijection on a theory's space of states, whatever that space may be. In the Hamiltonian description of the harmonic oscillator, it is an operator on phase space, $T : \mathbb{R}^2 \to \mathbb{R}^2$. In the Newtonian description, it is a transformation of *physical* space, $T : \mathbb{R} \to \mathbb{R}$. The latter is easy to miss, because it is simply the identity transformation $T(x) = x$. The former is more conspicuous, because it is in general not the identity. However, both are examples of a time reversal operator. We will take *time reversal* to refer to both the application of the time reversal operator on a given theory's state space, together with the reversal of the order of events in trajectories.

4.1.3 Time reversal invariance

Our discussion is about the circumstances under which a system is time reversal *invariant.* A system is time reversal invariant if time reversal takes each trajectory satisfying the laws of motion to another trajectory, which *also* satisfies the laws of motion.

One can quickly see that the harmonic oscillator is time reversal invariant in this sense. In terms of Figure 4.2, each elliptical trajectory is transformed to another elliptical trajectory, and indeed to the very same ellipse. (Time reversal invariance does not always require that each trajectory be mapped to itself, but this happens to be the case for the harmonic oscillator.) More formally, one can verify time reversal invariance by observing the effect that time reversal has on Hamilton's equations or on Newton's laws. We will show this formal fact explicitly in the next subsection.

The general definition of time reversal invariance in classical mechanics is just like the harmonic oscillator. It can be stated as follows.

Definition 4.1 (time reversal invariance)**.** Let $\gamma(t) : \mathbb{R} \to \mathcal{M}$ be a curve through some manifold of states $\mathcal{M}$ that characterizes a dynamical trajectory. Let $T : \mathcal{M} \to \mathcal{M}$ be the time reversal operator with respect to $\mathcal{M}$. A theory of curves on $\mathcal{M}$ is called *T-reversal invariant* (or simply *time reversal invariant*) if, whenever $\gamma(t)$ is a possible trajectory according to the theory, then so is $T\gamma(-t)$.

We say "the time reversal operator with respect to $\mathcal{M}$" because, at this level of generality, one cannot say much more than that about T. Its meaning depends on physical facts about the degrees of freedom that the space of states $\mathcal{M}$ represents. If $\mathcal{M} = \mathbb{R}^3$ represents the location of a particle in space, T is the identity operator. Given Newton's laws, time reversal invariance then means, "$x(t)$ solves Newton's equation only if $Tx(-t) = x(-t)$ does." On the other hand, if $\mathcal{M} = \mathbb{R}^6$ represents the position and linear momentum of a particle, T is not the identity. Given Hamilton's equations with $p = m\dot{q}$, time reversal invariance then means, "$(q(t), p(t))$ solves Hamilton's equations only if $T(q(-t), p(-t)) = (q(-t), -p(-t))$ does." In general, specifying the precise meaning of the time reversal operator $T : \mathcal{M} \to \mathcal{M}$ requires specifying facts about what the space of states $\mathcal{M}$ represents in the physical world.

4.1.4 Two useful facts

Let me conclude this section by stating two useful facts, which will facilitate the identification of time reversal invariance in the remainder of our discussion.

Lemma 4.1. *The statement that $F(x, t) = F(x, -t)$ is equivalent to the statement that $x(-t)$ satisfies Newton's equation whenever $x(t)$ does.*

Proof. ($\Rightarrow$): Suppose $F(x, -t) = F(x, t)$. Let $x(t)$ satisfy Newton's equation with this force. Since this equation holds for all times $t \in \mathbb{R}$, we can substitute $t \mapsto -t$ to get $\frac{d^2}{d(-t)^2}x(-t) = F(x,-t)$. Given our assumption, this implies $m\frac{d^2}{dt^2}x(-t) = F(x,t)$, which says that the time reversed trajectory $x(-t)$ satisfies Newton's equation.

($\Leftarrow$): Suppose that $x(t)$ is a solution, then so is $x(-t)$, for some force $F(x,t)$. Then we have that both $m\frac{d^2}{dt^2}x(-t) = F(x,t)$ and $m\frac{d^2}{dt^2}x(t) = F(x,t)$. Substituting $t \mapsto -t$ into the former gives $m\frac{d^2}{dt^2}x(t) = F(x,-t)$, and substituting that into the latter we have $F(x,t) = F(x,-t)$. □

Lemma 4.2. *The statement that $h(q,p) = h(q,-p) + k$ (for some $k \in \mathbb{R}$) is equivalent to the statement that $(q(-t), -p(-t))$ satisfies Hamilton's equations whenever $(q(t), p(t))$ does.*

Proof. ($\Rightarrow$): Suppose $h(q,p) = h(q,-p) + k$. Let $(q(t), p(t))$ satisfy Hamilton's equations with $h(q,p)$. Since these equations hold for all t, we can substitute $t \mapsto -t$ to get $\frac{dq(-t)}{d(-t)} = \frac{\partial h(q,p)}{\partial p}$ and $\frac{dp(-t)}{d(-t)} = -\frac{\partial h(q,p)}{\partial q}$. The former implies $\frac{dq(-t)}{dt} = \frac{\partial h(q,-p)}{\partial(-p)}$, and the latter implies $\frac{d(-p(-t))}{dt} = -\frac{\partial h(q,-p)}{\partial q}$, by simply pushing negative signs around and by our hypothesis that $h(q,p) = h(q,-p) + k$. But this just says that the time-reversed trajectory $(q(-t), -p(-t))$ satisfies Hamilton's equations.

($\Leftarrow$): Suppose that (a) $(q(t), p(t))$ and (b) $q(-t), -p(-t)$ are both solutions. Substituting $t \mapsto -t$ into Hamilton's equations with (a) gives $\frac{d}{d(-t)}q(-t) = \frac{\partial h(q,p)}{\partial p}$ and $\frac{d}{d(-t)}p(-t) = -\frac{\partial h(q,p)}{\partial q}$. Hamilton's equations with (b) give $\frac{d}{dt}q(-t) = \frac{\partial h(q,-p)}{\partial(-p)}$ and $\frac{d}{dt}(-p(-t)) = -\frac{\partial h(q,-p)}{\partial q}$. Combining, we find that

$$\frac{\partial h(q,p)}{\partial p} = \frac{\partial h(q,-p)}{\partial p}, \qquad \frac{\partial h(q,p)}{\partial q} = \frac{\partial h(q,-p)}{\partial q}.$$

This implies that $h(q,p) = h(q,-p) + f(q,p)$, for some function f such that $\partial f/\partial q = \partial f/\partial q = 0$. But the only such function is a constant function, so $h(q,p) = h(q,-p) + k$ (for some $k \in \mathbb{R}$). □

There is nothing novel about these well-known facts. I state their proof here only for convenience, as we will be make significant use of both in the next section.

4.2 What does *not* underpin classical TRI

Overzealous textbook authors have been known to make the following sweeping claim.

Claim 1. *Classical mechanics is time reversal invariant.*

Philosophers have often fallen for this ruse as well. For example, Frigg (2008) writes that time reversal invariance (TRI) cannot fail in the Hamiltonian formulation of classical mechanics (which he calls HM).

> HM is TRI in this sense. This can be seen by time-reversing the Hamiltonian equations: carry out the transformations $t \to \tau$ [where $\tau = -t$] and $(q,p) \to R(q,p)$ and after some elementary algbraic manipulations you find $dq_i/d\tau = \partial H/\partial p_i$ and $dp_i/d\tau = -\partial H/\partial p_i, i = 1, \ldots, m$. Hence the equations have the same form in either direction of time. (Frigg 2008, p.181)

Frigg's conclusion, like Claim 1, is strictly incorrect. A simple counterexample is a classical system with a so-called "dissipative" force[1]. For example, Newton's laws (and Hamilton's equations) allow trajectories in which a block slides along a smooth surface, subject to the force of friction, until eventually coming to a stop. However, the time-reversed trajectory of a block that spontaneously begins accelerating from rest is not a possible solution. These systems are described by Hamiltonians for which $h(q,p) \neq h(q,-p) + k$. As we observed in Lemma 4.2, this is sufficient for the failure of time reversal invariance. The significance of such examples for time reversal has been emphasized by Hutchison (1993).

More charitably, Frigg and other authors sympathetic to Claim 1 must make a tacit assumption about the scope of classical mechanics. For example, many would avoid considering dissipative forces in the description of elementary classical systems, by requiring (for example) that $dh/dt = 0$ in the Hamiltonian formulation. This may be what Frigg had in mind: assuming Hamiltonian mechanics is about Hamiltonians that are "conservatice" in the sense that $dh/dt = 0$, no dissipative forces are allowed. Similarly, Callender (1995) responds to Hutchison by arguing that systems with dissipative forces like friction are not "interesting" examples of classical systems, at least from a foundational perspective. The apparent

[1]To put an even finer point on the problem with Frigg's statement: there are many Hamiltonians with the property that $h(q,p) \neq h(q,-p)$. As we noted at the end of the last section, this is sufficient for the failure of time reversal invariance.

"force" of friction only arises out of an incomplete description of the block on the surface. If the more elementary interactions between the block and the surface were accounted for, then the force describing the system would take a very different form. Time reversal invariance would stand a chance of being regained.

However, we are not out of the woods yet. When authors more explicitly state the assumptions underlying what they take to be classical mechanics, one often finds the following claim.

Claim 2. *Classical mechanical systems that are "conservative" are also time reversal invariant.*

For example, Callender (1995, p.334) writes that, on the assumption that there are no non-conservative forces, "it is easy to verify that classical mechanics is TRI." The correctness of that claim, however, hinges on the precise definition of the term 'conservative.' There are two points that I would like to make about this. First, on the usual definition of a conservative system, as one that "conserves energy" in some sense, Claim 2 is simply false. I provide counterexamples in Sections 4.2.1 and 4.2.2. In the former, I show that the usual "no free work" definition of a conservative system is not enough for time reversal invariance. In the latter I show that the more charitable reading of Frigg's claim, that the Hamiltonian formulation is time reversal invariant when $dh/dt = 0$, also fails. Second, there is a stronger definition of "conservative" that requires the force $F(x,t)$ (or the Hamiltonian $h(q,p)$) to take a certain functional form. This *is* sufficient for time reversal invariance, and so perhaps this is the definition that Callender and Frigg have in mind. However, the physical motivation for that stronger requirement has not yet been made clear. I will discuss this point in Section 4.2.3.

4.2.1 Conservative but not TRI, part I

Here is a typical textbook definition of a conservative system in the Newtonian force formulation. This definition makes use of the quantity of work $W_{12} = \oint F \cdot dx$ required to transport a system between two points 1 and 2 along a path through configuration space.

If the force field is such that the work W_{12} is the same for any physically possible path between points 1 and 2, then the force (and the system) is said to be *conservative*. (Goldstein et al. 2002, p.3).

This definition is equivalent to the statement that the work around a closed loop in configuration space is zero. A conservative system is thus one in which there is no "free work": if a procedure ends in exactly the same state that it started in, then no total work has been done.

This definition of a conservative system is not sufficient to guarantee time reversal invariance. Here is a simple example to illustrate[2]. Take a particle in three spatial dimensions, with position $x = (x_1, x_2, x_3)$. As a shorthand, we will write $\dot{x} := dx/dt$, and thus denote the particle's velocity by $\dot{x} = (\dot{x}_1, \dot{x}_2, \dot{x}_3)$. Suppose the particle is subject to a force field defined by,

$$F = x \times \dot{x}.$$

That is, the force on the particle is orthogonal to both its position and velocity vectors.

This system is "conservative" on the definition above. The reason is that the cross product $(x \times \dot{x})$ is orthogonal to x, and hence to dx. So, the line integral characterizing work W_P along any path P is given by

$$W_P := \oint_P F \cdot dx = \oint_P (x \times \dot{x}) \cdot dx = 0.$$

The system is thus system "maximally lazy": no work is ever done, along any path whatsoever. It is therefore trivially conservative.

Nevertheless, the system fails to be time reversal invariant. The motion of the system is strange, typically displaying a "spiraling" behavior that occurs in a preferred orientation. Namely, the particle accelerates in the direction orthogonal to x and $\dot{x}$ that is given by the right-hand-rule[3] (Figure 4.3). Because this preferred orientation is not preserved under time reversal, the system fails to be time reversal invariant.

To verify this formally, we simply observe that $F(x, -t) = x \times (-\dot{x}) = -F(x, t)$. So, $F(x, -t) \neq F(x, t)$, and time reversal invariance fails by Lemma 4.1. Thus, being conservative in the sense of "no free work" is not sufficient for time reversal invariance.

[2] I thank Wayne Myrvold for drawing this kind of example to my attention.

[3] I have provided an animation of a particle undergoing this motion, available at `http://www.youtube.com/watch?v=-3hv3-YVA-E`. Thanks to Peter Distelzweig for showing me how to visualize this motion using vPython.

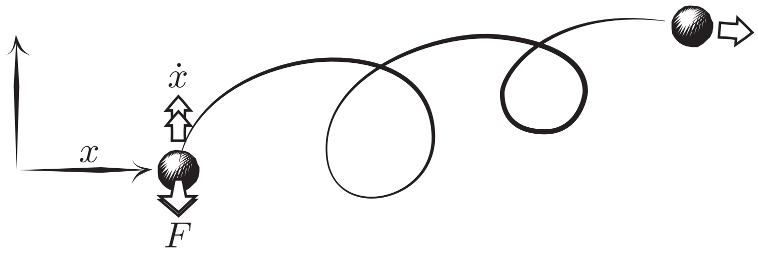


Figure 4.3: A particle experiencing the force $F = x \times \dot{x}$ with positive initial position and velocity.

4.2.2 Conservative but not TRI, part II

There is another natural definition of a "conservative" system in the context of Hamiltonian mechanics. Namely, since we interpret the Hamiltonian h to represent a system's total energy, "conservative" can naturally be taken to mean that h is a conserved quantity, $dh/dt = 0$. This perhaps is what Frigg (2008) hoped might imply that Hamiltonian mechanics is time reversal invariant.

However, the implication fails. There are many conservative systems of this kind that violate time reversal invariance. A simple example is a particle described by the Hamiltonian $h = p$. Since $dh/dt = \partial h/\partial t = 0$, this system is conservative in the required sense[4]. However, since $h(q, -p) \neq h(q, p) + k$, Lemma 4.2 implies that the system is not time reversal invariant.

A somewhat more interesting example is the system described by the Hamiltonian $h = \frac{m}{2}\dot{q}^2$, where $m\dot{q} = (p - q)$. This system can be interpreted as representing a free particle, in that the energy of the system is given entirely by a "kinetic energy" term. Like the previous system, it is conservative in that $dh/dt = \partial h/\partial t = 0$. But this system is also fails to be time reversal invariant, because $h(q, -p) = h(q, p) + \frac{m}{2}qp$, and hence $h(q, -p) \neq h(q, p)$.

There are thus various ways in which a system that is conservative in the sense of "conserving energy" can nevertheless violate time reversal invariant. If one wishes to guarantee

[4] Here I make use of the fact that $dh/dt = \partial h/\partial t$. This is because, by the chain rule, $\frac{d}{dt}h(q, p) = \frac{\partial h(q,p)}{\partial t} + \frac{\partial h(q,p)}{\partial q}\frac{dq}{dt} + \frac{\partial h(q,p)}{\partial p}\frac{dp}{dt}$. But, substituting Hamilton's equations $\frac{dq}{dt} = \frac{\partial h(q,p)}{\partial p}$ and $\frac{dp}{dt} = -\frac{\partial h(q,p)}{\partial q}$, we see that the latter two terms sum to 0, and so we get $dh/dt = \partial h/\partial t$.

time reversal invariance, a stronger condition is needed.

4.2.3 'Strong' Conservative implies TRI

In the context of Newtonian force mechanics, Arnold (1989, p.22) defines a conservative system to be one in which all forces have a particular functional form:

$$\mathbf{F}(\mathbf{x}, t) = \nabla V(\mathbf{x}), \tag{4.1}$$

for some scalar field $V(\mathbf{x})$, which (crucially) depends only on position. We might refer to this as "strong" conservativeness. On this definition, Newton's equation is manifestly time reversal invariant, because the right hand side of Equation (4.1) has no t-dependence, and thus $\mathbf{F}(\mathbf{x}, t) = \mathbf{F}(\mathbf{x}, -t)$.

This is certainly one way to guarantee time reversal invariance. But what reason do we have to believe that classical forces must take this functional form? It is certainly the case the *some* forces can be written in this way. But asserting this is no better than asserting the obvious fact that some systems are time reversal invariant. Such statements about particular contingent facts are unhelpful, if the goal is to understand the general sense in which classical mechanics is time reversal invariant.

The problem has an analogue in the Hamiltonian formulation. If the Hamiltonian h has its "common" form $h = (m/2)p^2 + v(q)$, then $h(q, p) = h(q, -p)$, and we are guaranteed time reversal invariance. But what reason do we have to think that the Hamiltonian must have this functional form? If we are to go beyond the banal fact that many classical systems just happen to take this form, we must minimally seek a *reason* why classical Hamiltonians (or classical force fields) have the functional form required by "strong" conservativeness.

There are reasons to think that ordinary systems in classical particle mechanics will take a restricted form, and indeed a form that is time reversal invariant. In the next section, I will point out two such reasons.

4.3 What *does* underpin classical TRI

An account of the sense in which classical mechanics is time reversal invariant should do things. First, it should state a general premise or set of premises, which may plausibly be taken to hold of some important subset of classical mechanical systems. Second, it should show that these premises are sufficient to establish time reversal invariance. In this section, I would like to point out two such accounts available in the Hamiltonian formulation of classical mechanics.

4.3.1 Velocity-momentum proportionality and TRI

Let me summarize a few facts about the Hamiltonian formulation that we have discussed so far. The state of many classical systems can be given by a point $(\mathbf{q}, \mathbf{p}) \in \mathbb{R}^{2n}$. The motion of such systems is typically given by a Hamiltonian function $h(\mathbf{q}, \mathbf{p})$, which satisfies Hamilton's equations,

$$\dot{q}_i = \frac{\partial h(\mathbf{q}, \mathbf{p})}{\partial p_i} \qquad \dot{p}_i = -\frac{\partial h(\mathbf{q}, \mathbf{p})}{\partial q_i} \tag{4.2}$$

for each $i = 1, \ldots, n$. We presume that the Hamiltonian is independent of time. As we saw in Section 4.2.2, this implies that such a system is "conservative" in the sense that $dh/dt = 0$, but not necessarily that it is time reversal invariant. The latter holds only if and only if $h(\mathbf{q}, \mathbf{p}) = h(\mathbf{q}, -\mathbf{p})$.

Is there an interesting kind of classical system for which time reversal invariance is guaranteed? One such class, I claim, is the following.

Claim 3. *If the momentum of a system is proportional to its velocity, $p = m\dot{q}$, then the system is time reversal invariant.*

In Section 4.1.2, I argued that the proportionality $p = m\dot{q}$ is an essential part of what it means to be a "classical bob on a spring." Indeed, it is essential to an extremely broad class of classical systems, including virtually all conservative (meaning $dh/dt = 0$) systems considered before the introduction of electromagnetic vector potentials in the 19th century. It is also straightforward to show that such systems are time reversal invariant.

Proposition 4.1. *Let $(q(t), p(t))$ satisfy Hamilton's equations with the Hamiltonian $h(q,p)$. Suppose that $p = m\dot{q}$ for some constant m. Then time reversal invariance holds, in that $(q(-t), -p(-t))$ also satisfies Hamilton's equations.*

Proof. Taking partial derivatives of $p = m\dot{q}$ with respect to $\dot{q}$ gives $\partial p = m\partial\dot{q}$. We substitute this expression in for the ∂p appearing in hamilton's equations,

$$\dot{q} = \frac{\partial h(q,p)}{\partial p} = \frac{\partial h(q,p)}{m\partial\dot{q}}.$$

Multiplying by $m\partial\dot{q}$ and integrating for h gives

$$h(q,p) = m\int \dot{q}\partial\dot{q} = \frac{m}{2}\dot{q}^2 + v(q) \tag{4.3}$$

for some function $v(q)$ of q alone. Finally, we substitute $p = m\dot{q}$ to get $h(q,p) = \dfrac{1}{2m}p^2 + v(q)$. This Hamiltonian obviously satisfies $h(q,p) = h(q,-p)$, which is a sufficient condition for time reversal invariance. □

This provides a first step toward understanding the extent to which classical mechanics is time reversal invariant. It may be summarized as follows. Classical mechanics does allow a variety of "anomalous" systems that are not time reversal invariant, even among those systems that conserve energy. But, if the momentum of a particle is proportional to its velocity, then none of these anomalous systems are allowed. Time reversal invariance is guaranteed.

4.3.2 Galilei invariance and TRI

There is another, even more general statement about the broad set of classical systems that are time reversal invariant. It is the classical analogue of the theorem discussed in Section 3.5.3 of the previous chapter. That statement is given precisely in Proposition 4.2 below, but may be stated roughly as follows.

Claim 4. *If a classical Hamiltonian system is such that (1) half of the degrees of freedom represent "position" in an appropriate sense, and (2) the motion of the system is covariant under spatial translations and Galilei boosts, then the system is time reversal invariant.*

I will formulate these conditions in the general geometric framework for Hamiltonian mechanics, in which a global Cartesian coordinate system $(\mathbf{q}, \mathbf{p})$ is not presumed. Part (1) is thus required in order to give precise meaning to "spatial translations" and "Galilei boosts"in part (2). By invariance under spatial translations, I aim to capture the assumption that a system does not distinguish a preferred point in space. By invariance under Galilei boosts, I aim to capture the assumption that a system does not distinguish a preferred reference frame.

This account of time reversal invariance, unlike the previous one, requires a certain amount of mathematical machinery in order to formulate precisely. In the remainder of this section, I will set out the required notation and definitions, and then show the sense in which the Claim is true.

4.3.2.1 Notation My notation for Hamiltonian mechanics will now roughly follow that of Geroch (Geroch 1974, §1-2). Let $\mathcal{P}$ (for "$\mathcal{P}$hase space") be a smooth $2n$-dimensional manifold. Each point $x \in \mathcal{P}$ will be interpreted as a "possible state" of a classical system. A function $f : \mathcal{P} \to \mathbb{R}$ will be interpreted as an "observable." Observables assign real values to each possible state of our system, and can represent physical quantities such as the energy or position of that state.

I adopt the "abstract index" notation of Penrose, and accordingly denote a vector v^a with an index upstairs, and a covector w_a with an index downstairs. The operation of contraction (sometimes called "interior multiplication" or "index summation") between tensors will be indicated by a common index in both upper and lower positions, such as $w^a v_a$. The unique exterior derivative on k-forms of a manifold will be denoted d_a.

The central features of Hamiltonian mechanics are captured by a *symplectic form* on $\mathcal{P}$. Mathematically, a symplectic form is a 2-form on $\mathcal{P}$, denoted Ω_{ab}; that is, Ω_{ab} is a skew-symmetric ($\Omega_{ab} = -\Omega_{ba}$), bilinear mapping from pairs of vectors in $T\mathcal{P}$ to the reals, $\Omega_{ab} : v^a w^b \mapsto r \in \mathbb{R}$. It is also closed ($d_a\Omega_{bc} = \mathbf{0}$) and non-degenerate ($\Omega_{ab}v^a = \mathbf{0} \Rightarrow v^a = \mathbf{0}$). This implies that Ω_{ab} is a bijection from vectors to covectors, and thus has an inverse; we denote its inverse by Ω^{ab}.

The interpretive significance of the symplectic form is that it allows us to input an observable h, and output a unique smooth vector field $H^a := \Omega^{ba} d_b h$, such that the value of h is conserved along the trajectories that thread the vector field H^a. This generalizes the traditional role that Hamilton's equations play, in providing a space of deterministic trajectories along which energy is conserved. There do exist classical descriptions that fail to satisfy these conditions, and thus that fail to admit a symplectic form. However, the scope of our discussion will be restricted to the broad class of classical descriptions that do.

Given a manifold and a symplectic form $(\mathcal{P}, \Omega_{ab})$, it will be convenient to define the *Poisson bracket* $\{\cdot\,,\cdot\}$ on smooth functions $f, h : \mathcal{P} \to \mathbb{R}$, given by

$$\{f, h\} := \Omega^{ab}(d_a h)(d_b f).$$

The right hand side is itself a smooth function on $\mathcal{P}$. So, the Poisson bracket takes a pair of scalar fields to a scalar field. From Ω_{ab} and d_a, the Poisson bracket inherits the properties of being antisymmetric, linear in both terms, satisfying the Leibniz rule in both terms, and vanishing for constant functions. If f, h generate vector fields F^a and H^a by the prescription above, let φ^f_α and φ^h_β denote the diffeomorphism flows with tangent fields F^a and H^a, respectively. It will be useful in what follows to observe that, by our definitions,

$$\{f, h\} := \Omega^{ab}(d_a h)(d_b f) = H^b d_b f = \frac{d}{d\beta}\left(f \circ \varphi^h_\beta\right)\Big|_{\beta=0} \tag{4.4}$$

where the last equality is an expression of the chain rule. In other words, the Poisson bracket $\{f, h\}$ is equal to the directional derivative of the scalar field f, in the direction of the vector field H^a determined by h.

We will take a classical system to consist of a $2n$-dimensional symplectic manifold $(\mathcal{P}, \Omega_{ab})$, together with a smooth function $h : \mathcal{P} \to \mathbb{R}$ that we refer to as the "Hamiltonian." The interpretive significance of h will be (1) that we take the quantity it assigns to states in $\mathcal{P}$ to be their *energy*, and (2) that the trajectories h generates (the integral curves that thread H^a) are the possible motions of the classical system in time.

4.3.2.2 Symmetries of position and velocity We will now impose some additional structure on a classical system $(\mathcal{P}, \Omega_{ab}, h)$. Our classical systems will be taken to have have a certain property that can be thought of as "position," and will satisfy certain symmetries with respect to that property.

The "position in space" of a classical system will be defined in terms of what is sometimes called a "maximal orthogonal set" or a "real polarization" on $\mathcal{P}$.

Definition 4.2. A *maximal orthogonal set* for a $2n$-dimensional manifold $\mathcal{P}$ is a set $\{\overset{1}{q}, \overset{2}{q}, \ldots, \overset{n}{q}\}$ of n smooth functions $\overset{i}{q} \colon \mathcal{P} \to \mathbb{R}$ such that (i) $\{\overset{i}{q}, \overset{j}{q}\} = 0$ for each $i, j = 1, \ldots, n$, and (ii) if f is another smooth function satisfying $\{f, \overset{i}{q}\} = 0$ for all i, then $f = f(\overset{1}{q}, \ldots, \overset{n}{q})$ is a function of the $\overset{i}{q}$.

It makes sense to think of position as forming such a set, for example, if we represent possible positions as points in $\mathbb{R}^n$, and represent phase space by the cotangent bundle $\mathcal{P} = T^*\mathbb{R}^n$. Then, for any Cartesian coordinate chart $\{\overset{1}{q}, \overset{2}{q}, \ldots, \overset{n}{q}\}$ on $\mathbb{R}^n$, the set $\{\overset{1}{q} \circ \pi, \overset{2}{q} \circ \pi, \ldots, \overset{n}{q} \circ \pi\}$ is a maximal orthogonal set for $\mathcal{P}$ (where π is the canonical projection, $\pi : (q, p_a) \mapsto q$). This maximal orthogonal set is one typical way of representing position in classical mechanics[5]. However, our more abstract formulation has the advantage of allowing us to speak more generally about the spatial position of a classical system. Indeed, we follow Woodhouse (Woodhouse 1981, p.121) in observing that a maximal orthogonal set is the natural classical analogue of a complete set of a commuting observables in quantum mechanics. In this sense, the assumption that "classical position" is a maximal orthogonal set is analogous to the assumption that "quantum position" is a complete set of commuting observables, and hence, that there are no internal degrees of freedom like spin or charge.

Given a classical system $(\mathcal{P}, \Omega_{ab}, h)$ with a maximal orthogonal set $\{\overset{1}{q}, \overset{2}{q}, \ldots, \overset{n}{q}\}$, we can define the "velocity" or instantaneous change in this set over time. Since change over time is given by the phase flow φ_t^h generated by h, the velocity of a function q is given by

$$\dot{q}(t) := \frac{d}{dt}(q \circ \varphi_t^h),$$

[5] This particular set is sometimes called the *vertical polarization* over $\mathbb{R}^n$. The "polarization" language comes from the fact that a maximal orthogonal set induces a foliation on $\mathcal{P}$, consisting of n-dimensional surfaces on which the values of the functions in $\{\overset{1}{q}, \overset{2}{q}, \ldots, \overset{n}{q}\}$ are constant. In the vertical polarization, each of these surfaces corresponds to the cotangent space at a point in $\mathbb{R}^n$.

In what follows, we will make use in particular of the *initial velocity* $\dot{q}$ of a classical system, defined by

$$\dot{q} := \dot{q}(0) = \frac{d}{dt}(q \circ \varphi_t^h)\bigg|_{t=0} = \{q, h\}, \tag{4.5}$$

where the last equality follows from our observation in Equation (4.4).

In Galilean physics, spatial translations and Galilei boosts are transformations that involve the simple "linear addition" of a vector to the value of position and velocity, respectively.

Definition 4.3 (Translations and Boosts)**.** We take a *translation and boost group* for a classical system $(\mathcal{P}, \Omega_{ab}, h)$ to be a $2n$-parameter family of diffeomorphisms $\Phi(\sigma, \rho) : \mathcal{P} \to \mathcal{P}$, which forms a representation of $\mathbb{R}^{2n}$, and such that

1. $q \circ \Phi(\sigma, \rho) = q + \sigma$
2. $\dot{q} \circ \Phi(\sigma, \rho) = \dot{q} + \rho$

where $q = \{\overset{1}{q}, \ldots, \overset{n}{q}\}$ is a maximal set of orthogonal functions, and $\dot{q}$ is the corresponding initial velocity. We define two associated diffeomorphism groups $\varphi_\sigma^s := \Phi(\sigma, 0)$ and $\varphi_\rho^r := \Phi(0, \rho)$, and refer to them as the *translation group* and the *boost group*, respectively. When these groups have a generator, we denote those generators by $s : \mathcal{P} \to \mathbb{R}$ and $r : \mathcal{P} \to \mathbb{R}$, respectively.

The classical systems of interest to us are "covariant" under translations and boosts, in the following sense. Let $H^a := \Omega^{ba} d_b h$ be the vector field representing a set of dynamical trajectories, corresponding to the Hamiltonian function h. Let $\Phi : \mathcal{P} \to \mathcal{P}$ be a diffeomorphism, and let us use a starred Φ^* to denote its pullback. For a classical system to be *covariant* under Φ means that $\tilde{H}^a := \Phi^* H^a$ is also a set possible dynamical trajectories with respect to some Hamiltonian function $\tilde{h}$; that is, $\tilde{H}^a = \Omega^{ba} d_b \tilde{h}$ for some smooth $\tilde{h} : \mathcal{P} \to \mathbb{R}$. This expression of covariance captures the idea that the "form" of a dynamical equation is preserved.

A classical system is covariant under a transformation if and only if the transformation is *symplectic*, meaning that it preserves the symplectic form: $\Phi^* \Omega_{ab} = \Omega_{ab}$ (Marsden and Raiu 1999, Proposition 2.6.1). So, requiring classical systems to be covariant under translations and boosts can be expressed by the requirement that translations and boosts be symplectic. This motivates the following.

Definition 4.4 (Translation and Boost Covariance)**.** A classical system $(\mathcal{P}, \Omega_{ab}, h)$ is *covariant under translations and boosts* if there exists a translation and boost group $\Phi(\sigma, \rho)$ on $\mathcal{P}$ such that each element of the group is symplectic, in that $\Phi^*(\sigma, \rho)\Omega_{ab} = \Omega_{ab}$ for all σ, ρ, and such that the group is *complete*, in that the only functions f that commute with both generators r and s are the constant functions.

4.3.2.3 Establishing time reversal invariance

With these definitions in hand, we may now formulate our main result. We take the *time reversal operator* to be a transformation $\tau : \mathcal{P} \to \mathcal{P}$ such that $\tau^* q = q$ and $\tau^* \dot{q} = -\dot{q}$. The time reverse of a classical system $(\mathcal{P}, \Omega_{ab}, h)$ with Hamiltonian vector field H^a is then the transformation that takes each integral curve $c(t)$ of H^a to $\tau \circ c(-t)$.

Proposition 4.2. *If a classical system $(\mathcal{P}, \Omega_{ab}, h)$ is covariant under translations and boosts, then it is time reversal invariant, in that if $c(t)$ is an integral curve of the Hamiltonian vector field generated by h, then so is $\tau \circ c(-t)$.*

The proof of this proposition hinges mainly on a lemma inspired by Josef Jauch's work on quantum mechanics. Jauch (1968) showed that in quantum theory, if a self-adjoint "position" operator Q forms a complete set of commuting observables, the translation and boost covariance implies that the Hamiltonian H "looks typical," in that $H = \frac{\mu}{2}\dot{Q}^2 + v(Q)$ for some function v and some nonzero real μ; that is, the Hamiltonian looks like the sum of a kinetic energy term and a potential term in Q alone. Lévy-Leblond (1970) attempted to prove a similar conclusion from rather different assumptions. However, his work was found to be inconclusive at best (Kraus 1980).

The following lemma is a more direct analogue of Jauch's result in the context of classical mechanics, which is perhaps more conclusive. The proof is given in the appendix.

Lemma 4.3 (Classical Jauch)**.** *If $(\mathcal{P}, \Omega_{ab}, h)$ is translation and Galilei boost covariant with respect to a maximal orthogonal set $\{\overset{1}{q}, \dots, \overset{n}{q}\}$, then $\{q, \mu\dot{q}\} = 1$ for some (non-zero) $\mu \in \mathbb{R}$, and $h = (\mu/2)\dot{q}^2 + v(q)$ for some function v of q alone.*

From this lemma our result follows straightforwardly.

Proof of Proposition 4.2. The classical analogue of Jauch's lemma implies that $\{q, \mu\dot{q}\} = 1$ (Appendix A, Proposition A.1). So, $(q, \mu\dot{q})$ forms a local orthonormal coordinate chart. This implies that the symplectic form Ω_{ab} can be expressed as the product $\Omega_{ab} = (d_a q)(d_b \mu\dot{q})$. Let $\tau : \mathcal{P} \to \mathcal{P}$ be the mapping such that $\tau^*\dot{q} = -\dot{q}$ and $\tau^* q = q$. Then,

$$\tau^*\Omega_{ab} = \tau^*(d_a q)(d_b \mu\dot{q}) = (d_a \tau^* q)(d_b \tau^* \mu\dot{q}) = -(d_a q)(d_b \mu\dot{q}) = -\Omega_{ab}.$$

Moreover, since the Jauch lemma guarantees that $h = (\mu/2)\dot{q}^2 + v(q)$, we have $\tau^* h = (\mu/2)(-\dot{q})^2 + v(q) = h$. But if $\tau^*\Omega_{ab} = -\Omega_{ab}$ and $\tau^* h = h$, then it follows from Proposition 4.3.13 of Abraham and Marsden (1978, p.308) that $(\mathcal{P}, \Omega_{ab}, h)$ is time reversal invariant in the sense above. □

4.4 Conclusion

We began by discussing a sense in which the claim that classical mechanics is time reversal invariant 'full stop' or 'for conservative systems' is insufficient. But these worries can be absolved by restricting the scope of classical mechanics. One might have thought such an approach would become stuck in the muck around the difficult questions of what counts as a 'physically reasonable' Newtonian system. Instead, it appears that time reversal invariance can be established by a condition as plausible as momentum and velocity proportionality, or as plausible as Galilei invariance.

5.0 THREE ROUTES TO T-VIOLATION

5.1 Introduction

At the level of fundamental physics, it is surprisingly difficult to produce systems that are not time reversal invariant. So difficult that, when James Cronin and Val Fitch produced the first evidence that such systems exist in nature, they were soon awarded the Nobel Prize[1]. Over 50 years later, an assortment of curious and complex examples of T-violation have been produced and reviewed[2]. In this chapter, I would like to clarify the nature of these examples, by describing three general ways that a physical system can violate time reversal invariance. I will refer to such systems as *T-violating.*

One of the general paths to T-violation is the one taken by Cronin and Fitch, in studying weakly interacting neutral kaon decay. It rests, I claim, on the following fact.

- *T-Violation by Curie's Principle.* If an initial state evolves unitarily to some final state, and if one of those two states is preserved by a linear transformation while the other is not, then the system is fails to be invariant under that transformation.

A second path to T-violation makes use of a principle, which is similar to Curie's, but "probabilistic" in form.

- *T-Violation by the Reversal Principle.* If an initial and final state are both preserved by the time reversal operator, then the probabilities of transitioning from one to the other after some time t must be the same.

[1]The experiment was published in (Christenson et al. 1964). Cronin and Fitch received the 1980 Nobel Prize in physics for this work.

[2]Book-length reviews can be found in (Sachs 1987), (Khriplovich and Lamoreaux 1997), (Sozzi 2008), (Kleinknecht 2003), and (Bigi and Sanda 2009).

The third path to T-violation that I would like to discuss is the one pursued by those involved in the search for electric dipole moments. I argue that it can be broadly characterized as follows.

- *T-Violation by non-degeneracy.* Under appropriate circumstances, a system described by a Hamiltonian that is non-degenerate[3] will be T-violating.

The plan of the paper is as follows. In Section 5.2, I will discuss the first route to T-violation, and in particular the experiment of Cronin and Fitch, while pointing out its limitations. I then show how these limitations are overcome by a second "probabilistic" route to T-violation. Section 5.3 describes a third route to T-violation, suggested by Wigner's derivation of Kramers degeneracy. In Section 5.4, I show how Wigner's result can be generalized to provide a general template for T-violation, and illustrate how two known examples of T-violation arise as special cases. Section 5.5 further generalizes this approach, by proving a generalization of Wigner's derivation of Kramers degeneracy in the case that the Hamiltonian has a continuous spectrum.

5.2 T-violation and Curie's Principle

A quantum system is time reversal invariant whenever it is the case that, if a trajectory $\psi(t)$ is a solution to the Schrödinger equation with the Hamiltonian H, then so is its time reverse $T\psi(-t)$, where T is the antiunitary time reversal operator. This is equivalent[4] to the statement that $[T, H] = 0$, and it is this statement that we will refer to when using the phrase *time reversal invariance*. If a system is not time reversal invariant, meaning that $[T, H] \neq 0$, then we will say that it is *T-violating*.

There is a folk-proverb in physics sometimes referred to as "Curie's Principle." Although it is often formulated with rueful imprecision, Earman (2004) helpfully stated it in a precise form.

[3] A self-adjoint operator A in finite dimensions is *degenerate* if it has two orthogonal eigenvectors with the same eigenvalue. I will discuss this property in more detail below.

[4] I give the proof of this equivalence in Section 3.5.2 of Chapter 3. For a discussion of these definitions, see Chapter 2.

Fact 1 (Curie's Principle 1). *Let H be a self-adjoint operator on a Hilbert space $\mathcal{H}$, and let $S : \mathcal{H} \to \mathcal{H}$ be a linear bijection. Let $\psi_i \in \mathcal{H}$ (an "initial state") and $\psi_f = e^{-itH}\psi_i$ (a "final state") for some $t \in \mathbb{R}$. If,*

1. *(S-invariance)* $[S, H] = 0$

then,

2. *(equal preservation)* $S\psi_i = \psi_i$ *if and only if* $S\psi_f = \psi_f$.

This formulation of the "principle" is actually a simple mathematical truth[5]. For our purposes, it will be useful to formulate explicitly in the equivalent contrapositive form; it is this that I will refer to as "Curie's Principle" in the remainder of the discussion.

Fact 2 (Curie's Principle 2). *Let H be a self-adjoint operator on a Hilbert space $\mathcal{H}$, and let $S : \mathcal{H} \to \mathcal{H}$ be a linear bijection. Let $\psi_i \in \mathcal{H}$ (an "initial state") and $\psi_f = e^{-itH}\psi_i$ (a "final state") for some $t \in \mathbb{R}$. If either*

1. *(initial but not final state preserved)* $S\psi_i = \psi_i$ *and* $S\psi_f \neq \psi_f$, *or*
2. *(final but not initial state preserved)* $S\psi_f = \psi_f$ *and* $S\psi_i \neq \psi_i$

then,

3. *(S-violation)* $[S, H] \neq 0$.

The advantage of Curie's Principle is that it provides a "phenomenological" way to test for symmetry violation. We need only prepare an initial state that is not preserved by S, and observe that it evolves unitarily to a state that is preserved by S (or vice versa). Curie's Principle then guarantees that the system will be S-violating, even if we are totally ignorant about the explicit form of the Hamiltonian.

The disadvantage of this approach is that Curie's Principle fails to apply directly to time reversal, because T is not a linear operator. For this reason, the Curie's Principle approach normally proceeds by first testing for PC-violation (where PC is a linear operator representing "Parity" and "Charge conjugation," to be discussed more below). One then invokes the powerful CPT theorem to argue that PC-violation implies T-violation.

[5] Its proof is trivial: suppose that $[S, H] = 0$, and hence (since S is linear) that $[S, e^{-itH}] = 0$ (Blank et al. 2008, Corollary 5.9.4). Then $S\psi_i = \psi_i$ if and only if $S\psi_f = S\psi_f$, since $S\psi_i = \psi_i$ if and only if $S\psi_f = Se^{-itH}\psi_i = e^{-itH}S\psi_i = e^{-itH}\psi_i = \psi_f$.

In what follows, I will illustrate both the advantage and disadvantage of this route to T-violation in turn.

5.2.1 The advantage: T-violation and kaon decay

The theory of weak interactions that Cronin and Fitch applied is "phenomenological." They did not know the formal Hamiltonian describing neutral kaon decay. Nevertheless, their observation did provide evidence for T-violation, through a combination of Curie's Principle together with the CPT theorem. Here is how that argument works.

The neutral kaon is a massive particle that has no electromagnetic charge, but which does have an unusual degree of freedom called "strangeness" or "hypercharge." The strangeness observable S for the neutral kaon has two eigenstates, κ and $\bar{\kappa}$, with eigenvalues ± 1:

$$S\kappa = \kappa, \quad S\bar{\kappa} = -\bar{\kappa}.$$

Since κ and $\bar{\kappa}$ have opposite strangeness charges, and no other charges, they are each others' antiparticle.

The phenomenological theory describing the decay of neutral kaons has a complex history[6]. But for our purposes, it is enough to note the few simple facts that give rise to what is called "PC-violation." P denotes the unitary "Parity" transformation, which has the effect of reversing orientation in space. C denotes unitary "Charge Conjugation," which has the effect of mapping a particle to its antiparticle. Since PC is linear, we can use Curie's Principle to determine circumstances under which we have PC-violation.

Here is how. The strangeness eigenstates of the neutral kaon satisfy the following relations with respect to P and C:

$$\begin{aligned} P\kappa &= -\kappa & \qquad C\kappa &= e^{-i\theta}\bar{\kappa} \\ P\bar{\kappa} &= -\bar{\kappa} & \qquad C\bar{\kappa} &= e^{i\theta}\kappa. \end{aligned}$$

[6] For a short overview, see (Bigi and Sanda 2009, §5). The characterization I give here was first introduced by Gell-Mann and Pais (1955).

That is: reversing spatial orientation reverses the sign of both κ and $\bar{\kappa}$, while charge conjugation maps one to the other (up to a phase factor $e^{i\theta}$). But the states that are normally prepared in the study of kaon decay are superpositions of kaon-antikaon pairs, defined by

$$k_{(\pm)} = \frac{1}{\sqrt{2}}\left(e^{i\frac{\theta}{2}}\kappa \mp e^{-i\frac{\theta}{2}}\bar{\kappa}\right).$$

It is easily checked[7] that one of these superpositions is preserved by PC, while the other is reversed,

$$PCk_{(+)} = k_{(+)},$$
$$PCk_{(-)} = -k_{(-)}.$$

On the other hand, it it is known that a state $\pi_+\pi_-$ consisting of a pair of oppositely charged pions is preserved under the same transformation,

$$PC\pi_+\pi_- = \pi_+\pi_-$$

The simple, bizarre fact about neutral kaons verified by Cronin and Fitch is that one can prepare[8] the superposition $k_{(-)}$, and observe its decay into a pair of oppositely charged pions $\pi_+\pi_-$. The former state is preserved by PC, while the latter state is not. So, it follows from Curie's Principle that the interaction describing how $k_{(-)}$ evolves unitarily to $\pi_+\pi_-$ is PC-violating, $[PC, H] \neq 0$.

Through a little mathematical juggling, one can now invoke the "CPT theorem" to derive T-violation. Of course, we still do not have an explicit Hamiltonian describing neutral kaon decay at the fundamental level. But suppose that *at least in principle*, the interaction is describable in terms of an appropriate axiomatization of quantum field theory, such as the Haag/Wightman axioms (see Haag 1996, §II.1). They can be summarized roughly:

1. QFT systems a unitary representation of the Poincaré group.
2. QFT systems can be described by smooth local Wightman fields.

[7] $PCk_{(\pm)} = \frac{1}{\sqrt{2}}\left(e^{i\theta/2}PC\kappa \mp e^{-i\theta/2}PC\bar{\kappa}\right) = \frac{1}{\sqrt{2}}\left(-e^{i\theta/2}e^{-i\theta}\bar{\kappa} \pm e^{-i\theta/2}e^{i\theta}\kappa\right) = \frac{1}{\sqrt{2}}\left(\pm e^{i\theta/2}\kappa - e^{-i\theta/2}\bar{\kappa}\right) = \pm k_{(\pm)}$.

[8] In fact, what one prepares is a "kaon mixture," which is a mixture of $k_{(+)}$ and $k_{(-)}$ states. The former turn out to have a much shorter half life than the latter. So, by waiting a sufficiently long time, one is assured that a kaon mixture will consist mostly of $k_{(-)}$ states. For this reason, $k_{(-)}$ is sometimes referred to as the "long-life" kaon state.

When these axioms are made appropriately precise, there is a CPT theorem[9] that any quantum field theory satisfying the Haag/Wightman axioms is CPT invariant, $[CPT, H] = 0$. This means that a PC-violating quantum field theory is a T-violating one. For, if it were the case that $[T, H] = 0$, then it would follow from $[CPT, H] = 0$ that $[CP, H] = 0$, a contradiction. The premises of the CPT theorem thus guarantee that a PC-violating system is also T-violating.

This route to T-violation may thus be summarized as follows. First, use Curie's Principle to determine that the unitary evolution from one state to another is CP-violating. Second, convince oneself that this system can in principle be captured with appropriate axiomatically. It follows from the CPT theorem that the system is T-violating.

5.2.2 The disadvantage: No direct test for T-violation

The route to T-violation described here is an uncomfortable marriage of two conflicting perspectives. One begins by arguing for PC-violation on the basis of a phenomenological theory, avoiding substantial assumptions about the form of the unitary evolution describing the interaction. One then turns to a highly abstract theoretical framework to reformulate this statement, without making any explicit connections to the phenomenological theory, and then argues that T-violation follows as a corollary. The disconnect between the phenomenological weak theory on the one hand, and the circumstances for which the CPT theorem holds on the other, is a serious challenge to the current evidence for T-violation.

To avoid this challenge, one might hope to restrict attention to just one of the two perspectives. For example, one might try to derive T-violation directly using Curie's Principle. Unfortunately, the principle does not apply to the case of time reversal, because T is not linear. In particular, one can check that $[T, H] = 0$ and $T\psi = \psi$ do not in general imply that $T\psi(t) = \psi(t)$, where $\psi(t) = e^{-itH}\psi$. A simple counterexample is the following. Take a particle in a box of width π. Its wavefunction evolves according to $\psi(x, t) = e^{it}\sin(x)$. Since $T = K$ is just conjugation in this context, and since $\psi(x, 0) = \sin(x)$ is real, the initial state is preserved: $T\psi(x, t) = \psi(x, t)$. But at time $t = \pi/2$, we have $\psi(x, t) = i\sin(x)$, so

[9] The theorem in this form was first proved by Jost (1957); Haag (1996, Theorem 5.1.4) also provides an elegant statement and proof.

$T\psi(x,t) = -\psi(x,t)$, and the final state is not preserved. Thus, Curie's Principle fails for time reversal[10].

Alternatively, we could try to improve our confidence that the neutral kaon decay Hamiltonian is CPT invariant. For example, one could simply write down the standard model's Hamiltonian for the weak interaction H_w (see Bigi and Sanda 2009, §5.4), and computing that $[T, H_w] \neq 0$. However, it is precisely this Hamiltonian that tests of T-violation seek to confirm or disconfirm; indeed, there is an extensive literature on "Extended Standard Models" in which the form of H_w changes dramatically, lifting requirements like Poincaré and CPT invariance[11].

To obtain a rigorous guarantee that CPT invariance holds generally, one must turn to a general framework in which we have a CPT theorem[12]. So, another hope for T-violation might be to provide an axiomatically rigorous model of the Hamiltonian appearing in neutral kaon decay, and proceeding as before to compute whether or not it is T-violating. In a certain sense, we know how this would turn out: if the model satisfies the Haag/Wightman axioms, and also captures PC-violation, then the CPT theorem guarantees that it will be T-violating. This is a promising approach, although we are not yet close to producing such a model. There is also a further limitation on this route to T-violation. Namely, it may not apply to theories of quantum fields on generic spacetimes, since the usual proofs of the CPT theorem draw on unique facts about the symmetries of Minkowski spacetime[13]. As a consequence, Curie's Principle and the CPT theorem cannot be used to test for T-violation on scales in which gravitation is relevant. In order to understand time asymmetry on the cosmological, and in other contexts in which the CPT theorem need not hold, a different measure of T-violation appears to be needed. In the next subsection, we will discuss one such measure.

[10]There is much that can be said about this counterexample; for example, Curie's Principle can also be seen to fail for time reversal in classical Hamiltonian mechanics. I reserve further discussion of this topic for a future paper.

[11]For example, see (Colladay and Kostelecký 1997).

[12]A CPT theorem of a different sort was originally proved by Pauli; his proof relied on the conclusion of the spin-statistics theorem. However, a rigorous general proof of spin-statistics itself requires assumptions like those adopted by Jost in the axiomatic framework.

[13]See (Greaves 2010) for a creative elucidation.

5.2.3 Direct T-Violation and a "Reversal Principle"

Recently, Costa and Fogli (2012, p.94) have remarked that, "[a]t present, the only direct detection of a departure from time-reversal invariance comes from the analysis of the $K^0\bar{K}^0$ meson system." They are referring to the groundbreaking experiment of Angelopoulos et al. (1998), performed at the CPLEAR particle detector at CERN. Is this a true direct detection of T-violation? If so, how is this possible in the face of the difficulties discussed above?

The CPLEAR experiment is indeed a direct detection of T-violation. It is made possible, I will argue, by a dramatic modification of Curie's Principle, which is similar to the original only in the initial premise, that time reversal preserves the initial state ($T\psi_i = \psi_i$). The modified principle applies to antilinear transformations like time reversal, but not to linear transformations like PC. It is also essentially probabilistic in form, instead of restricting attention only to unitary evolution. To begin, let me introduce the notation,

$$\Pr(\psi_i \xrightarrow{t} \psi_f) := \left|\langle \psi_f, e^{-itH}\psi_i\rangle\right|^2$$

to represent the probability of "transitioning" from an initial state ψ_i to a final state ψ_f after a duration of time t. In this notation, here is the revised principle available for antiunitary operators.

Fact 3 (Reversal Principle)**.** *Let H be a self-adjoint operator on a Hilbert space $\mathcal{H}$, and let $T : \mathcal{H} \to \mathcal{H}$ be a unitary or antiunitary bijection. Let $\psi_i \in \mathcal{H}$ (an "initial state") and $\psi_f = e^{-itH}\psi_i$ (a "final state") for some $t \in \mathbb{R}$. If*

1. *$T\psi_i = e^{i\theta}\psi_i$ and $T\psi_f = e^{i\theta'}\psi_f$ (for some phase $e^{i\theta}$, $e^{i\theta'}$), but*
2. *$\Pr(\psi_i \xrightarrow{t} \psi_f) \neq \Pr(\psi_f \xrightarrow{t} \psi_i)$,*

then,

3. *(T-violation) $[T, H] \neq 0$.*

To summarize this principle: if two states of a system are preserved by time reversal (up to a phase), but the probabilities of transitioning from one to the other after a time t differ, then the system is T-violating. This fact is checked as easy to check as Curie's Principle[14], and provides a direct test for T-violation.

This technique for verifying T-violation applies neatly to above description[15] of the neutral kaon state κ and its aniparticle state $\bar{\kappa}$. Both of these states are preserved by the time reversal operation,

$$T\kappa = \kappa, \quad T\bar{\kappa} = \bar{\kappa}.$$

Moreover, if we prepare κ as our initial state, then there is some probability that it will transition to its antiparticle $\bar{\kappa}$ after a length of time t, and vice versa. This behavior is called "kaon oscillation," and it provides the following simple test for T-violation. By the Reversal Principle, we need only show that for some length of time t,

$$\Pr(\kappa \xrightarrow{t} \bar{\kappa}) \neq \Pr(\bar{\kappa} \xrightarrow{t} \kappa).$$

Since T preserves the initial and final states in these transitions, it follows that such a detection would confirm directly that neutral kaon oscillations are T-violating. This is precisely what was shown by the CPLEAR experiment in 1998.

Thus, like Cronin and Fitch's confirmation of PC-violation, the CPLEAR experiment provides a confirmation of T-violation without any assumption about the particular form of the Hamiltonian H. But whereas the former provided only indirect evidence for T-violation by way of the CPT theorem, the latter provides a direct confirmation. The key to this result was the abandonment of Curie's Principle, in favor of a principle that is similar in form, but more relevant to the study of time reversal.

5.3 T-violation and Kramers degeneracy

There is another route to T-violation, which is much more direct, but perhaps less well-known. This route involves the search for exotic new kinds of matter. Here is a rough

[14] We argue the contrapositive. Suppose $[T, H] = 0$. Then $\left|\langle \psi_2, e^{-itH}\psi_1\rangle\right|^2 = \left|\langle T\psi_2, Te^{-itH}\psi_1\rangle\right|^2 = \left|\langle T\psi_2, e^{itH}T\psi_1\rangle\right|^2 = \left|\langle T\psi_1, e^{-itH}T\psi_2\rangle\right|^2$, where the first equality follows from the unitarity/antiunitarity of T, the second from the fact that $[T, H] = 0$ and T is antiunitary, and the third from the properties of the norm. So, $T\psi_1 = e^{i\theta}\psi_1$ and $T\psi_2 = e^{i\theta'}\psi_2$ implies that $\left|\langle \psi_2, e^{-itH}\psi_1\rangle\right|^2 = \left|\langle \psi_1, e^{-itH}\psi_2\rangle\right|^2$.

[15] This was apparently first recognized by Kabir (1970).

sketch of how this can work (with more details to follow), using the example of an electric dipole moment. An electric dipole moment typically describes the displacement between two opposite charges, or within a distribution of charges. But suppose that, instead of describing a distribution of charges, we use an electric dipole moment to characterize a property of just one elementary particle. This particle might be referred to as an "elementary" electric dipole moment.

Such particles have indeed been described, though not yet detected. Let H_0 be the Hamiltonian describing the particle in the absence of interactions; let S represent its angular momentum; and let E represent an electromagnetic field. Then these "elementary" electric dipoles have been[16] characterized by the Hamiltonian,

$$H = H_0 + S \cdot E.$$

Since time reversal preserves the free Hamiltonian H_0 and the electric field E, but reverses angular momentum S, this Hamiltonian is manifestly T-violating: $[T, H] \neq 0$. Therefore, an elementary electric dipole of this kind would constitute a direct detection of T-violation. No need for Curie's Principle. No need for the CPT theorem.

There are general principles underpinning this example of T-violation, too. In this and the next two sections, I will characterize those general principles. In particular, I will show how the electric diopole moment (and other kinds of exotic matter) are examples of a general template for T-violation, related to the non-degeneracy of the Hamiltonian, and to a condition I will call the "distinct ray condition."

5.3.1 Kramers degeneracy and the electron

An early success of modern quantum mechanics was the explanation of why the hydrogen atom has two distinct energy states corresponding to the same energy value. Kramers (1930) showed that this phenomenon was a direct consequence of properties of the electron's spin, and his name remains attached to the effect: "Kramers degeneracy." But for our purposes, the more significant derivation was given two years later by Wigner (1932), who showed that there is a deep connection between this kind of degeneracy and time reversal invariance.

[16] (See Khriplovich and Lamoreaux 1997)

Wigner's derivation made use of an important background assumption about the (spin-$\frac{1}{2}$) electron, that if T is the time reversal operator acting on that system, then $T^2 = -I$. That is: applying time reversal twice does not exactly bring an electron back to where we started, but adds a phase factor of -1. Only by applying time reversal twice more can we return an electron to its original vector state.

This is a curious property indeed. But it follows straightforwardly from the standard definition of the time reversal operator. Start by describing the electron using the 2-dimensional Pauli operators, $\sigma_1 = \left(\begin{smallmatrix} & 1 \\ 1 & \end{smallmatrix}\right)$, $\sigma_2 = \left(\begin{smallmatrix} & -i \\ i & \end{smallmatrix}\right)$, $\sigma_3 = \left(\begin{smallmatrix} 1 & \\ & -1 \end{smallmatrix}\right)$. In this context, the standard time reversal operator is $T = \sigma_2 K$, where K is the antilinear conjugation operator in the $\{\binom{1}{0}, \binom{0}{1}\}$ basis[17]. It follows by direct calculation that,

$$T^2 = \left(\begin{smallmatrix} & -i \\ i & \end{smallmatrix}\right) K \left(\begin{smallmatrix} & -i \\ i & \end{smallmatrix}\right) K = \left(\begin{smallmatrix} & -i \\ i & \end{smallmatrix}\right) \left(\begin{smallmatrix} & i \\ -i & \end{smallmatrix}\right) K^2 = -\left(\begin{smallmatrix} & -i \\ i & \end{smallmatrix}\right)^2 = -I.$$

One might respond by challenging the definition of the standard time reversal operator. This does not help. There is a uniqueness result in this context, which guarantees that any reasonable time reversal operator T will be proportional to $\sigma_2 K$ up to an arbitrary phase, and will thus satisfy $T^2 = -I$ regardless. I provide this result in the appendix. So, in the description of the electron, it is safe to assume that $T^2 = -I$.

5.3.2 Wigner's derivation of Kramers degeneracy

Whereas Kramers took several pages to derive the degeneracy of the hydrogen atom's energy spectrum, Wigner's remarkable derivation took only a few lines. A version of Wigner's argument can be stated as follows. (The proof is well-known, but I provide it here for convenience.)

Proposition 5.1 (Wigner)**.** *Let H be a self-adjoint operator on a finite-dimensional Hilbert space, which is not the zero operator. Let $T : \mathcal{H} \to \mathcal{H}$ be an antiunitary bijection. If*

1. *(electron condition) $T^2 = -I$, and*
2. *(T-invariance) $[T, H] = 0$*

[17] By definition, K maps each vector $\psi = \alpha\binom{0}{1} + \beta\binom{1}{0}$ to its complex conjugate, $K\psi = \psi^* = \alpha^*\binom{0}{1} + \beta^*\binom{1}{0}$. Notably, K satisfies $K^2 = I$. See (Jauch 1968, §14-5) for a statement of the standard T.

then,

3. *(degeneracy) H has two orthogonal eigenvectors with the same eigenvalue.*

Proof. Since H is non-zero, there is some vector ψ such that $H\psi = \lambda\psi$. Let $\psi' = T\psi$. By (2), $H\psi' = HT\psi = TH\psi = \lambda T\psi = \lambda\psi'$. The eigenvectors ψ and ψ' thus have the same eigenvalue, λ. They are moreover orthogonal, since

$$\langle\psi, T\psi\rangle = \langle T\psi, T^2\psi\rangle^* = -\langle T\psi, \psi\rangle^* = -\langle\psi, T\psi\rangle,$$

where the first equality follows from the fact that T is antiunitary, and the second from (1). Therefore, $\langle\psi, T\psi\rangle = 0$. □

Of course, we have seen that there are interactions in which time reversal invariance may fail. This suggests that Wigner's result might be more helpfully formulated in the equivalent, contrapositive form. That is the following.

Corollary 1 (Reverse Wigner). *Let H be a self-adjoint operator on a finite-dimensional Hilbert space, which is not the zero operator. Let $T : \mathcal{H} \to \mathcal{H}$ be an antiunitary bijection. If*

1. *(electron condition) $T^2 = -I$, and*
2. *(non-degeneracy) H has no two orthogonal eigenvectors with the same eigenvalue*

then,

3. *(T-violation) $[T, H] \neq 0$.*

Although this corollary is mathematically equivalent to Wigner's statement, formulating it explicitly in this way suggests a template for T-violation. Suppose we discover that some interaction of the electron requires a non-degenerate Hamiltonian. Then this corollary assures us that we have achieved a direct detection of T-violation.

There are more general circumstances in which non-degeneracy implies T-violation, which have further empirical significance. In the next section, I would like to discuss those more general circumstances. I will then discuss two experimental research programs, and show that they are both examples of this route to T-violation.

5.4 A template for T-violation

In this section, I will point out an easy way to generalize Corollary 1, and then go over some physically relevant examples that instantiate it.

5.4.1 Generalization of Wigner's argument

The condition that $T^2 = -I$ is rather specific to the electron. But its role in Wigner's proof is just to show that there is some eigenstate ψ of the Hamiltonian that is orthogonal to its time reverse, in that $\langle\psi, T\psi\rangle = 0$. In fact, a similar result can be formulated with only the weaker statement that $\langle\psi, T\psi\rangle \neq e^{i\theta}$ for all real θ. This provides a small generalization of Wigner's theorem, which may be stated as follows.

Proposition 5.2. *Let H be a self-adjoint operator on a finite-dimensional Hilbert space, which is not the zero operator. Let T be an antiunitary bijection. If*

1. *(distinct ray condition) $T\psi \neq e^{i\theta}\psi$ for some eigenvector ψ of H, and*
2. *(non-degeneracy) H has no two orthogonal eigenvectors with the same eigenvalue*

then,

3. *(T-violation) $[T, H] \neq 0$*

Proof. We prove the contrapositive, by assuming (3) fails, and proving that either (1) or (2) fails as well. Let $H\psi = h\psi$ for some $h \neq 0$ and some eigenvector ψ of unit norm. Since T is antiunitary, $T\psi$ will also have unit norm.

Suppose (3) fails, and hence that $[T, H] = 0$. As we saw in the proof of Proposition 5.1, this implies that if ψ is an eigenvector of H with eigenvalue h, then so is $T\psi$. By the spectral theorem, the eigenvectors of H form an orthonormal basis set. So, since ψ and $T\psi$ are both unit eigenvectors, either $T\psi = e^{i\theta}\psi$ or $\langle T\psi, \psi\rangle = 0$. The latter violates non-degeneracy (2). And, since ψ was arbitrary, the former violates the distinct ray condition (1). Therefore, either (1) or (2) must fail. □

Mathematically, this is an easy generalization of Wigner's argument. However, there is a significant interpretive payoff to formulating Proposition 5.2 in this way. If a system (like a single electron) happens to satisfy the distinct-ray condition, then a non-degenerate Hamiltonian is a T-violating one. This provides us with the following general template for T-violation.

1. Verify that some system would satisfy the distinct ray condition, that $T\psi \neq e^{i\theta\psi}$ for some eigenvector ψ of H.
2. Determine that the Hamiltonian H is non-degenerate.

The result is our second route to T-violation.

5.4.2 Two Empirical Examples

We observed in Section 5.3 that certain exotic particles, such as an elementary electric dipole moment, can provide examples of T-violation. What I would now like to point out is that such examples are special cases of the route to T-violation just described. The first example that I will discuss is a "non-zero current density." I will then return to example of the elementary electric dipole moment. Both are often cited as examples of T-violating systems. In this section, I will verify that both must satisfy the distinct ray condition. Determining that such matter admits a non-degenerate Hamiltonian thus provides a direct route to T-violation.

5.4.2.1 Example: Permanent Current Density Intuitively, a system has a current density if some density (of matter, say) undergoes a non-zero rate of change. That density is "permanent" if its expectation value does not change over time.

To make this mathematically precise, we say that a self-adjoint operator P is *permanent* if it has a non-zero expectation value for some eigenstate ψ of the Hamiltonian:

$$\langle\psi, P\psi\rangle = a \neq 0.$$

To verify that this really does mean a does not change over time, suppose ψ is an eigenvector of the Hamiltonian H with eigenvalue h. Its time evolution is given by $\psi(t) = e^{-ith}\psi$. So, for all times $t \in \mathbb{R}$, we find the same expectation value,

$$\langle\psi(t), P\psi(t)\rangle = \langle(e^{-ith})\psi, P(e^{-itH})\psi\rangle = e^{ith}e^{-ith}\langle\psi, P\psi\rangle = a.$$

A current density has a somewhat more involved definition; for example, see (Jauch 1968, §13-6). However, for us, it is sufficient characterize it as a self-adjoint operator J that reverses sign under time reversal, $TJT^{-1} = -J$. That is, we will treat current just as one would treat velocity: when we film a motion involving current, and then reverse the film, the current (and velocity) vectors point in the reverse direction.

One can now verify that a permanent current density satisfies the distinct ray condition, $T\psi \neq e^{i\theta}\psi$ for some energy eigenvector ψ. For suppose it were not satisfied, and $T\psi = e^{i\theta}\psi$ for all energy eigenvectors. The assumption of a permanent current density says that $\langle\psi, J\psi\rangle = a > 0$, for some energy eigenvector ψ. So,

$$\langle\psi, J\psi\rangle = \langle T\psi, TJ\psi\rangle = -\langle T\psi, JT\psi\rangle = -\langle e^{i\theta}\psi, e^{i\theta}J\psi\rangle = -\langle\psi, J\psi\rangle,$$

where the first equality follows from the antiunitarity of T and the self-adjointness of J, the second equality from the fact that $TJT^{-1} = -J$, and the third equality by our rejection of the distinct ray condition. Therefore, $\langle\psi, J\psi\rangle = 0$. The failure of the distinct ray condition thus implies the failure of $\langle\psi, J\psi\rangle = a > 0$, contradicting the assumption that J is permanent. So, the distinct ray condition must be satisfied here.

If we have a permanent current density, then determining its Hamiltonian to be non-degenerate would provide a determination of T-violation. Thus we have another example of the second route to T-violation.

5.4.2.2 Example: Elementary Electric Dipole Moment In Section 5.3, we saw that there are T-violating Hamiltonians that describe elementary electric dipole moments. In this section, I will point out a sense in which *all* reasonable Hamiltonians describing such a particle will satisfy the distinct spin condition. It is in this sense that, if an elementary electric dipole were discovered, then its Hamiltonian would be non-degenerate only if it were T-violating.

In the textbooks, an elementary electric dipole moment is typically defined[18] to have the following three properties.

[18] For example, see the definitions of (Ballentine 1998, §13.3), (Messiah 1999, §XXI.31), or (Sachs 1987, §4.2).

- *(permanent)* The observable D representing the dipole moment is "permanent" in the sense discussed above: $\langle\psi, D\psi\rangle = a > 0$ for some eigenvector ψ of the Hamiltonian H. That is, the dipole is a permanent feature of the particle, like its charge or spin-type.
- *(isotropic dynamics)* Since it is an elementary particle, its simplest interactions are assumed to be isotropic, in that time evolution commutes with all rotations, $[e^{-itH}, R_\theta] = 0$. Note that if J is the "angular momentum" observable that generates the rotation $R_\theta = e^{i\theta J}$, then this is equivalent to $[H, J] = 0$ (Blank et al. 2008, Corollary 5.9.4).
- *(time reversal properties)* Time reversal is an antiunitary operator that has no effect on the electric dipole observable ($TDT^{-1} = D$), but reverses angular momentum ($TJT^{-1} = -J$).

A particle with these three properties turns out to satisfies the distinct ray condition. To see why, assume (for reductio) that the distinct ray condition fails, and thus that for each eigenvector ψ of the Hamiltonian, there is a unit $e^{i\theta}$ such that $T\psi = e^{i\theta}\psi$. We will show that the assumption that the dipole moment is "permanent" also fails, contradicting our hypothesis.

Since $[H, J] = 0$, there is a common eigenvector for H and J, which we will denote ψ. By the Wigner-Eckart Theorem[19], each eigenvector of and J will satisfy,

$$\langle\psi, D\psi\rangle = c_\psi\langle\psi, J\psi\rangle \tag{5.1}$$

for some $c_\psi \in \mathbb{R}$ (called the Clebsh-Gordon coefficient). Now, an antiunitary operator T satisfies $\langle T\psi, T\phi\rangle = \langle\psi, \phi\rangle^*$ for any ψ, ϕ. And a self-adjoint operator satisfies $\langle\psi, A\psi\rangle^* = \langle\psi, A\psi\rangle$ for any ψ. Applying these two facts to Equation (5.1), we get that $\langle T\psi, TD\psi\rangle = c\langle T\psi, TJ\psi\rangle$. But T commutes with D and anticommutes with J, so this equation may be written,

$$\langle T\psi, D(T\psi)\rangle = -c\langle T\psi, J(T\psi)\rangle \tag{5.2}$$

[19] Let $\mathbf{A} = (A_1, A_2, A_3)$ be any vector observable, and let $\mathbf{J}$ be the angular momentum observable. A straightforward consequence of the Wigner-Eckart theorem is that the matrix elements of $\mathbf{A}$ are proportional to those of $\mathbf{J}$, as in Equation 5.1, where the c_ψ is the Clebsch-Gordon coefficient for ψ (see Ballentine 1998, §7, Equation (7.125)).

Finally, we assume the distinct ray condition fails, so $T\psi = e^{i\theta}\psi$ for some $e^{i\theta}$. Applying this to Equation (5.2), we get

$$\begin{aligned} & (e^{-i\theta}e^{i\theta})\langle\psi, D\psi\rangle = -(e^{-i\theta}e^{i\theta})c\langle\psi, J\psi\rangle \\ \Rightarrow\quad & \langle\psi, D\psi\rangle = -c\langle\psi, J\psi\rangle. \end{aligned}$$

Combined with Equation (5.1), this implies that $\langle\psi, D\psi\rangle = 0$, contradicting our hypothesis that D is permanent. So, the distinct ray condition cannot fail in this system. The elementary electric dipole is therefore an example of our template for T-violation.

There is in fact an active search for T-violation in elementary electric dipole moments. A recent book-length overview has been given by Khriplovich and Lamoreaux (1997). However, the close relationship between this empirical search and Wigner's derivation of Kramers degeneracy does not seem to have been recognized. Namely, experimental search proceeds by trying to show that there is an elementary electric dipole moment with a non-degenerate energy spectrum. But this system satisfies the distinct spin requirement. So, the elementary electric dipole is an instantiation of the general template for T-violation provided by Proposition 5.2.

5.5 Generalization to continuous observables

In the previous discussion, we assumed that the Hamiltonian H has eigenvectors. This is not generally the case when H has a continuous spectrum. But H *does* have a continuous spectrum in many interesting physical descriptions, such as when an energy-momentum observable is in play. In this section, I will give a generalization of the template for T-violation to the continuous spectrum case. In particular, I will show that Wigner's derivation of Kramers degeneracy can be generalized to continuous observables.

Let us return to the premises of Wigner's result (Proposition 5.1), which has as its conclusion:

(degeneracy) H has no two orthogonal eigenvectors with the same eigenvalue.

There is a way to generalize this statement, which is equivalent for finite dimensions, but which also makes sense when H has a continuous spectrum. I will begin by describing that generalization. I will then derive the Kramers degeneracy result in this more general context, and finally discuss the consequences for T-violation.

5.5.1 Degeneracy and simple spectra

The natural analogue of non-degeneracy for observables with a continuous spectrum is called the *simple spectrum property*. Some readers may find this generalization unfamiliar, and so I will recapitulate it here. Those very familiar with non-degeneracy and simple spectra may safely skip to the next subsection.

To begin, here is another way to think about the "eigenvector definition" of non-degeneracy. In finite dimensions, an operator A has a degenerate spectrum if two orthogonal eigenvectors ψ and ψ' of A have the same eigenvalue λ. In that case, their eigenvalue equations allow us to write $(A - \lambda I)\psi = (A - \lambda I)\psi' = 0$. Then the equation,

$$(A - \lambda I)\phi = 0 \tag{5.3}$$

will be satisfied by any vector ϕ that is a linear combination of ψ and ψ'. Since ψ and ψ' are orthogonal, another way of saying this is: the set of vectors that are mapped to 0 by the operator $(A - \lambda I)$ forms a subspace of dimension at least 2. More generally, the dimension of this subspace will be equal to the number of orthogonal eigenvectors satisfying Equation (5.3). This number is referred to as the *multiplicity* of λ. So, another way of saying that A is *non*-degenerate is just to say that all its eigenvalues have multiplicity 1.

There is a theorem that provides us with yet another way to express non-degeneracy. Let $\{H\}'$ be the *commutant* of H, meaning the set of bounded Hilbert space operators that commute with H. Let $\{H\}''$ be the *double-commutant* of H, meaning the set of closed[20] operators that commute with all the elements of $\{H\}'$. Then the theorem may be stated as follows (for a proof, see Blank et al. 2008, Theorem 5.8.6).

[20] An operator A is *closed* if, whenever a sequence ψ_i in the domain of A is such that $\psi_i \to \psi$ and $A\psi_i \to \phi$, then it follows that ψ is in the domain of A, and $A\psi = \phi$. Closed operators are continuous on their domain, but need not be bounded.

Theorem 5.1 (Blank, Exner, Havlíček). *Let A be a self-adjoint linear operator on a finite-dimensional Hilbert space. Then all the eigenvectors of A have multiplicity 1 if and only if $\{A\}' = \{A\}''$.*

So, in finite dimensions, non-degeneracy is equivalent to the condition that $\{A\}' = \{A\}''$. Although the former statement is perhaps more familiar and intuitive, the latter is more general. In particular, we have noted that in Hilbert spaces of infinite dimensions, non-degeneracy does not in general make sense, because an self-adjoint operator A does not in general eigenvectors. However, the statement that $\{A\}' = \{A\}''$ does make sense. It can be satisfied even if A has a continuous spectrum or is unbounded[21]. This is generally referred to as the "simple spectrum" property, and it provides a generalization of the eigenvector definition of non-degeneracy.

Definition 5.1 (Simple Spectrum). A self-adjoint operator A has a *simple spectrum* if $\{A\}' = \{A\}''$.

This definition will form the basis for the generalization that follows.

5.5.2 Continuous spectrum analogue of Kramers degeneracy

The only statement appearing in Wigner's derivation of Kramers degeneracy that precludes continuous-spectrum operators is the statement of degeneracy itself. So, let us adopt the simple spectrum condition instead, since this is the natural generalization of that statement. Then one can prove a generalized derivation of Kramers degeneracy, where "degeneracy" is now taken to mean that the Hamiltonian fails to have a simple spectrum. This derivation is of some independent interest, as it mean we can expect the Kramers degeneracy of the hydrogen atom to appear even when the interactions depend on continuous-spectrum observables, such as energy-momentum. It also provides another way to characterize of our template for T-violation.

Proposition 5.3 (Continuous Kramers). *Let H be a (possibly continuous-spectrum) self-adjoint operator on a separable Hilbert space, which is not the zero operator. Let $T : \mathcal{H} \to \mathcal{H}$*

[21] For example, the statement that $\{Q\}' = \{Q\}''$ is one of the defining properties of the "position operator" Q appearing in the Schrödinger representation (see Blank et al. 2008, Example 5.8.2). This operator is both unbounded and has a continuous spectrum.

be an antiunitary bijection. If

1. *(electron condition)* $T^2 = -I$
2. *(time reversal invariance)* $[T, H] = 0$

then,

3. *(no simple spectrum)* $\{H\}' \neq \{H\}''$

Proof. We show that the supposition $\{H\}' = \{H\}''$ leads to a contradiction. Let K be the conjugation operator in the H basis, so that $[H, K] = 0$. Since T is antiunitary, there exists a unitary operator U such that $T = UK$. By time reversal invariance,

$$0 = [T, H] = (UK)H - H(UK) = (UH)K - (HU)K,$$

which implies that $UH - HU = 0$. So, U is a bounded linear operator that commutes with H, and hence $U \in \{H\}'$. Our hypothesis $\{H\}' = \{H\}''$ now implies that $U \in \{H\}''$. By von Neumann's double-commutant theorem, it follows that we can write $U = f(H)$ for some function in the weak closure of H. But U is unitary, and so can be expressed as $U = e^{iS}$ for some self-adjoint operator S (Blank et al. 2008, Proposition 5.3.8). Combining these two facts, we have that $U = e^{ig(H)}$, where $g(H) = S$ is self-adjoint, and is thus a real-valued function. But for real-valued g, $[g(H), K] = 0$, and so $KUK^{-1} = e^{Kig(H)K} = e^{-ig(H)K^2} = U^*$ (where we have applied the fact that $K^2 = I$ in the last equality). Therefore,

$$T^2 = UKUK = UU^* = I.$$

This contradicts the electron condition, so we are done. □

Proposition 5.3 entails Wigner's derivation of Kramers degeneracy (Proposition 5.1) as a special case, since in finite dimensions, $\{H\}' \neq \{H\}''$ is equivalent to degeneracy in the usual sense. It can thus be similarly reformulated to provide a template for T-violation. That is, Proposition 5.3 is equivalent to the following.

Corollary 2 (Reverse Continuous Kramers)**.** *Let H be a (possibly continuous-spectrum) self-adjoint operator on a separable Hilbert space, which is not the zero operator. Let $T : \mathcal{H} \to \mathcal{H}$ be an antiunitary bijection. If*

1. *(electron condition)* $T^2 = -I$
2. *(simple spectrum)* $\{H\}' = \{H\}''$

then,

3. *(T-violation)* $[T, H] \neq 0$.

We thus have an example of T-violation that applies to continuous spectrum observables, which is a natural generalization of the example of T-violation introduced in Section 5.3.

Thus, both Kramers degeneracy and our discussion of T-violation have analogues when H has a continuous spectrum. Of course, it is true that it is a very unusual Hamiltonian that satisfies $\{H\}' = \{H\}''$. Indeed, the "usual" Hamiltonians are time reversal invariant (see Chapter 3), so Proposition 5.3 *guarantees* that they will not satisfy $\{H\}' = \{H\}''$. However, this property is satisfied[22] by Hamiltonians such as the electric dipole moment discussed at the beginning of Section 5.3, and thus remains a topic of open investigation.

5.6 Conclusion

We have seen three routes to T-violation, of distinctly different forms. The first route, which employs Curie's Principle and the CPT theorem, is by necessity indirect. The reason is the curious result that Curie's Principle fails for time reversal in quantum mechanics. As a consequence, one can only use this principle to test for linear symmetries like PC-violation. Insofar as the premises of the CPT theorem are correct, T-violation can then be derived as a consequence of PC-violation. The second route replaces Curie's Principle with a probabilistic "Reversal Principle," which restores the possibility of a direct detection of T-violation. The third route makes use of a non-degenerate Hamiltonian, together with the distinct ray condition. This allows for a direct test of T-violation, which is not contingent

[22] To see why, first notice that if we have a representation of the Canonical Commutation Relations with $H = P^2$, then H is "degenerate" in that $\{H\}' \neq \{H\}''$. We verify using the parity operator Π, given by $\Pi = e^{(\pi/2)(P^2+Q^2)}$. It satisfies $\Pi P \Pi^* = -P$, so $\Pi(P^2)\Pi^* = (P^2)$, and we have that Π is in $\{H\}'$. But Π is not a function of P because it has Q terms in it, so Π is not in $\{H\}''$. We can break the degeneracy by adding in a term that is linear in P, such as angular momentum. That's precisely what happens with the electric dipole. If the Hamiltonian is $H = P_1^2 + L$, where $L = Q_1P_2 - Q_2P_1$, then reversing momentum also reverses L. So, the parity operator is no longer in the commutant $\{H\}'$, and the degeneracy is broken.

on the premises of the CPT theorem, although it requires knowing more about the form of the Hamiltonian.

It is worth observing, as a final note, that these three routes to T-violation are perfectly compatible with our previous results about time reversal invariance. Time reversal invariance is guaranteed in the presence Galilei invariance and the absence of internal degrees of freedom. In the first route to T-violation, both of these assumptions are violated. The example hinges on the "strangeness" internal degree in the description of CP-violation, and Poincaré invariance in applying the CPT theorem to get to T-violation. The second route draws on systems that, if they do not admit internal degrees of freedom, turn out to violate Galilei invariance by force of the T-theorem. An next step of interest would be to attempt to close the gap between these results about T-invariance and T-violation, by producing interesting necessary and sufficient conditions for T-invariance and T-violation. For now, this will remain a project for future work.

6.0 CONCLUDING REMARKS

6.1 Summary

This work has grounded the meaning of time reversal invariance in quantum theory, established generic circumstances under which it holds, pointed out an analogous discussion in classical physics, and presented three general ways that time reversal invariance can fail. The reader who has made it with me this far has my thanks.

We have seen that the meaning of time reversal in quantum mechanics is not arbitrary. In Chapter 2, we built it up in three stages. In the first stage, we assumed that whether or not two states are orthogonal does not depend on the direction of time. This was enough to imply that time reversal is either unitary or antiunitary. In the second stage, we assumed that one can always describe at least one interesting quantum system that is time reversal invariant, and proved that time reversal must be antiunitary as a consequence. In the third stage, we saw two distinct ways in which the time reverse of position and momentum can be grounded by assumptions about the way time reversal interacts with space. All this was done with no assumption whatsoever about the symmetries of the background spacetime; it is completely general in this sense.

With the meaning of time reversal in hand, we found in Chapter 3 that there is a generic class of quantum systems that are time reversal invariant. Namely, the quantum systems that are both bereft of any internal degrees of freedom, as well as covariant under spatial translations and Galilei boosts, are time reversal invariant. These assumptions induce a tight constraint on the form of the Hamiltonian. Namely, by Jauch's lemma, the Hamiltonian is constrained to have the form of kinetic energy, plus a potential term in position alone. The consequences for time reversal have as much significance the other way around: if a Galilei covariant quantum system is found to violate time reversal invariance, then it *must* admit internal degrees of freedom. This represents a somewhat unexpected connection between asymmetry in time, and a certain richness in the ontology of Galilei covariant quantum theory. One cannot have the temporal asymmetry without a sufficiently rich ontology.

Chapter 4 showed that a similar discussion can be carried out in the context of classical mechanics. After settling what time reversal means, we saw that many of the common dogmas about classical time reversal invariance are incorrect. There turn out to be many classical systems, even systems that conserve energy in some sense, that nevertheless fail to be time reversal invariant. However, we then showed two ways to describe a generic set of classical systems that *are* time reversal invariant. The simplest is the class of systems for which momentum is proportional to velocity. The more interesting is the class of systems that are Galilei covariant. This result is made available by a new classical analogue of Jauch's lemma.

We finally returned to the violation of time reversal invariance in quantum theory. Chapter 5 showed that there are at least three general ways that time reversal can be violated. The first requires an application of Curie's Principle, together with the CPT theorem. This route is necessarily indirect. The second makes use of a modified Curie-type principle, but which is probabilistic. This is what allowed the only existing direct detection of T-violation so far. The third makes use of a different property altogether, which is the non-degeneracy of the Hamiltonian. This last route finds its origin in Wigner's work on Kramers degeneracy, although we showed that it can be made considerably more general.

6.2 Open Questions

There remain a number of open questions about time reversal in quantum theory, some of which I have mentioned above. I will summarize a few of the main open problems here.

6.2.1 How to time reverse a non-classical observable

Chapter 2 produced an account of the meaning of time reversal, including a derivation of the transformation rules for the position and momentum observables. However, this account stopped short of describing how time reversal transforms *arbitrary* observables, and a description of this more general sort remains an open project.

A first step forward is given in Appendix B, where I show that the meaning of time reversal in a spin-1/2 system can be derived from the assumption that time reversal commutes with rotations. It seems plausible that a similar argument might be made for any observable that generates a global symmetry group. However, this is work that remains to be done.

6.2.2 Connection to Malament's approach

Malament (2004) developed a different way to understand time reversal from the one advocated in Chapter 2. He began by positing an orientable relativistic spacetime with an orientation τ^a. He then argued that the basic effect of time reversal is to reverse the sign of this orientation, $\tau^a \mapsto -\tau^a$. Understanding the relationship between Malament's approach and the one advocated in this dissertation remains an open project for future research.

One strategy in implementing Malament's approach in quantum theory would be to construe quantum theory as a theory of local nets of algebras on Minkowski spacetime (see Haag 1996). Suppose each algebra admits a representation of the complete Poincaré group (including time reversal) among the operators that are either unitary or antiunitary. Then one can ask: what is the nature of the operators U_τ characterizing Malament's time reversal transformation $\tau^a \mapsto -\tau^a$? One would hope that they can be shown to have the usual properties of antiunitarity; this is a topic that remains to be pursued.

6.2.3 Internal degrees of freedom and the Lorentz group

The T-Theorem formulated in Chapter 3 (Proposition 3.1) suggests that internal degrees of freedom are required for the phenomenon of T-violation, in theories that are Galilei covariant. Is this requirement general enough to hold of theories that are Lorentz covariant as well?

Proposition 3.1 can be formulated straightforwardly with the Lorentz group replacing the Galilei group. The status of that reformulated proposition is an open question.

One way to approach the problem would be to try to develop a version of Jauch's lemma (Appendix A), using the Lorentz group in place of the Galilei group. This problem is of interest independently of the question of time reversal invariance.

Another approach would be to attack the problem of time reversal invariance more directly. The role of Jauch's lemma in the T-Theorem is just to provide a Hamiltonian that is an even function of the velocity operator $\dot{Q}$. Any antiunitary operator T that reverses $\dot{Q}$ while preserving Q will commute with any such Hamiltonian, thus guaranteeing T-invariance. The question is then whether there are any interesting physical conditions that give rise to this property. The hope with such a result would be to ultimately bridge the gap between the conditions of Chapter 3 that guarantee time reversal invariance, and the conditions of Chapter 5 that guarantee its violation. For now, this remains an open problem.

6.2.4 Curie's Principle and time reversal

Curie's Principle, as it was formulated in Chapter 5, was found not to apply to time reversal. This was a passing remark in our argument; however, an open project remains to determine just what went wrong. Is our formulation of Curie's Principle too restrictive? Is there a plausible reformulation that applies to antilinear operators, or to time reversal invariance? Is the failure of Curie's Principle something special to quantum theory, or does it fail in classical mechanics as well?

Curie's Principle does appear to be inapplicable to time reversal in classical Hamiltonian mechanics, but there do seem to be alternative ways to formulate Curie-like principles for which it does apply. Still, the troubled relationship between time reversal and Curie's Principle has not yet been well articulated, and so a discussion of this relationship remains an open topic of research.

APPENDIX A

JAUCH'S LEMMA AND A CLASSICAL ANALOGUE

Jauch (1964, 1968) argued that Galilei invariance tightly constrains the form of the Hamiltonian in quantum theory. However, Jauch appeared to think that this kind of result could not be extended to classical mechanics, writing that the central distinctions of his argument are "only significant for quantum mechanical systems" (Jauch 1964, p.285).

In the first section of this Appendix, I will provide a reconstruction of Jauch's result, which I find to be mathematically clearer than Jauch's original argument. In the second section I will show that, in spite of Jauch's own suggestion, there is indeed a straightforward analogue of the result in classical Hamiltonian mechanics.

A.1 Jauch's lemma

A.1.1 Background

In this section I will adopt the notation of Chapter 3, in which the basic description of a quantum system consists of a triple $(\mathcal{H}, \Delta \mapsto E_\Delta, t \mapsto U_t)$. The position operator Q is the self-adjoint operator corresponding to the measure $\Delta \mapsto E_\Delta$, and the velocity operator is its rate of change, $\dot{Q} = i[H, Q]$. We say that the structure $(\mathcal{H}, \Delta \mapsto E_\Delta, t \mapsto U_t)$ is *Galilei invariant* only if there exist two strongly continuous one-parameter unitary representations

S_a (translations) and R_b (boosts) of the additive group of real numbers, such that

$$S_a Q S_a^{-1} = Q + aI \qquad\qquad S_a \dot{Q} S_a^{-1} = \dot{Q}$$

$$R_b Q R_b^{-1} = Q \qquad\qquad R_b \dot{Q} R_b^{-1} = \dot{Q} + bI$$

for all $a, b \in \mathbb{R}$. Finally, we assume that there are no internal degrees of freedom, by supposing that the position operator Q forms a complete set of commuting observables.

A.1.2 Jauch's Lemma

Jauch's lemma shows that when a system is Galilei invariant and lacks internal degrees of freedom, the Hamiltonian H can only take its "standard" form, $H = \frac{\mu}{2}\dot{Q}^2 + v(Q)$. Here I provide my preferred statement and proof of this result.

Theorem A.1 (Jauch)**.** *Suppose $(\mathcal{H}, \Delta \mapsto E_\Delta, t \mapsto \mathcal{U}_t)$ is Galilei invariant, and that the self-adjoint operator Q associated with $\Delta \mapsto E_\Delta$ forms a complete set of commuting observables. Then $(\mathcal{H}, e^{ibQ}, e^{ia\mu\dot{Q}})$ is an irreducible unitary representation of the canonical commutation relations in Weyl form, and $\mathcal{U}_t = e^{itH}$, where*

$$H\psi = \frac{\mu}{2}\dot{Q}^2\psi + v(Q)\psi$$

for some non-zero real number μ, some Borel function v, and for all ψ in the domain of H.

Proof. Define Q and $\dot{Q}$ as above, and recall our definition of Galilei invariance:

$$S_a Q S_a^{-1} = Q + aI \qquad\qquad S_a \dot{Q} S_a^{-1} = \dot{Q}$$

$$R_b Q R_b^{-1} = Q \qquad\qquad R_b \dot{Q} R_b^{-1} = \dot{Q} + bI$$

where we note that the unbounded operators Q and $\dot{Q}$ act only on their respective domains of definition in $\mathcal{H}$. The essential observation of this proof is that Galilei invariance implies the existence of three representations of the Weyl form of the canonical commutation relations:

(i) $S_a e^{ibQ} S_a^* = e^{iab} e^{ibQ}$

(ii) $R_b e^{ia\dot{Q}} R_b^* = e^{iab} e^{ia\dot{Q}}$

(iii) $S_a R_b^* S_a^* = e^{i\mu ab} R_b^*$

where μ is a fixed non-zero real number. Representations (i) and (ii) follow immediately from our expression of Galilei invariance; we simply exponentiate Q and $\dot{Q}$, and recognize that the unitary operators acting on them may be pulled up into the exponential.

Representation (iii) is somewhat more subtle, being constructed from the fact that the unitary representations S_a and R_b are each defined to be a representation of $\mathbb{R}$, and thus together provide a projective representation of the real plane $\mathbb{R} \times \mathbb{R}$. In particular, it is part of the definition and S_a and R_b that

$$S_a S_b = S_{a+b}, \quad R_a^* R_b^* = R_{a+b}^*. \tag{A.1}$$

Hence, the mapping $(a, b) \mapsto (S_a, R_b^*)$ is a homomorphism from the vectors $v = (a, b)$ of $\mathbb{R} \times \mathbb{R}$ into the pairs of unitary operators. This means that the group (S_a, R_b^*) is a projective representation of the plane $(\mathbb{R} \times \mathbb{R}, +)$ under vector addition. Since the latter group is abelian and hence satisfies $(a, 0) + (0, b) - (a, 0) = (0, b)$, it follows that the same group properties will hold in the projective representation up to a phase factor $e^{if(a,b)}$:

$$S_a R_b^* S_a^* = e^{if(a,b)} R_b^*$$

for some $f(a, b) \in \mathbb{R}$. Moreover, it is known that the representation may always be chosen such that $f(a, b) = \mu ab$ (Blank et al. 2008, §10 Problem 15). Thus we have the representation expressed in (iii): $S_a R_b^* S_a^* = e^{i\mu ab} R_b^*$.

With a bit of work, (i), (ii) and (iii) can now be shown to imply that Q and $\mu\dot{Q}$ satisfy the canonical commutation relation. Showing this involves three steps.

First, we note that the representation expressed in (ii) implies that $e^{i\mu ab} R_b^* = e^{ia\mu\dot{Q}} R_b^* e^{-ia\mu\dot{Q}}$. Plugging this result into (iii), we find that $e^{ia\mu\dot{Q}} R_b^* e^{-ia\mu\dot{Q}} = S_a R_b^* S_a^*$, and hence that

$$R_{b/\mu}^* (S_a^* e^{ia\mu\dot{Q}}) = (S_a^* e^{ia\mu\dot{Q}}) R_{b/\mu}^*. \tag{A.2}$$

where we have substituted b/μ in for b, recognizing that the equation holds for any real value of b.

Keeping Equation (A.2) in mind, we now proceed to the second step, which is to show that $R_{b/\mu}$ is a constant multiple of e^{iaQ}. This draws on the fact that representation (i) contains a term e^{iaQ} for which Q forms a complete set of commuting observables, which implies that the representation is irreducible (Blank et al. 2008, Ex. 6.7.2e). We apply the irreducibility of representation (i) as follows. Rewrite (iii) as

$$S_a R_{b/\mu} S_a^* = e^{-iab} R_{b/\mu}.$$

We know from (i) that $S_a e^{ibQ} S_a^* = e^{iab} e^{ibQ}$. So, multiplying the left sides of these two equations as well as the right sides, we see that

$$(S_a R_{b/\mu} S_a^*)(S_a e^{ibQ} S_a^*) = e^{-iab} e^{iab} (R_{b/\mu} e^{ibQ}),$$

and so since $S_a^* S_a = e^{-iab} e^{iab} = I$, we have that

$$S_a (R_{b/\mu} e^{ibQ}) S_a^* = (R_{b/\mu} e^{ibQ})$$

This says that the operator $R_{b/\mu} e^{ibQ}$ commutes with S_a. But the same operator also commutes with e^{ibQ}, since

$$R_{b/\mu} e^{ibQ} R_{b/\mu}^* = e^{ib(R_{b/\mu} Q R_{b/\mu}^*)} = e^{ibQ}.$$

Schur's lemma (Blank et al. 2008, Thm. 6.7.1) establishes that multiples of the identity are the only operators that commute with both terms in an irreducible representation. So, since $R_{b/\mu} e^{ibQ}$ commutes with both S_a and e^{iaQ}, we may write $R_{b/\mu} e^{ibQ} = cI$, which implies that

$$R_{b/\mu}^* = \frac{1}{c} e^{ibQ}$$

as claimed.

The third step now substitutes this $R_{b/\mu}^*$ into Equation (A.2), to get that

$$\frac{1}{c} e^{ibQ} (S_a^* e^{ia\mu\dot{Q}}) = (S_a^* e^{ia\mu\dot{Q}}) \frac{1}{c} e^{ibQ},$$

or equivalently,

$$e^{ia\mu\dot{Q}} e^{iaQ} = (S_a e^{iaQ} S_a^*) e^{ia\mu\dot{Q}}.$$

Applying (i) to the right-hand side of this equation, we finally see that $e^{ia\mu\dot{Q}} e^{iaQ} = e^{iab} e^{ibQ} e^{ia\mu\dot{Q}}$, which is the desired representation of the commutation relations in Weyl form. It is irreducible for the same reason that representation (i) is, namely because Q forms a complete set of commuting observables (Blank et al. 2008, Ex. 6.7.2e).

It is now straightforward to determine the form of $\mathcal{U}_t$ in this representation. First, we note that $\mathcal{U}_t = e^{itH}$ for a unique self-adjoint H by Stone's theorem. Moreover, the canonical commutation relation in Weyl form implies the "standard" commutation relation $[Q, \mu\dot{Q}]\psi = i\psi$, for all ψ in the common dense domain of Q and $\dot{Q}$. This in turn implies that $[Q, (1/2)(\mu\dot{Q})^2] = i(\mu\dot{Q})$, which we multiply through by $-i/\mu$ to find that

$$i[(\mu/2)\dot{Q}^2, Q] = \dot{Q}.$$

But by definition, $\dot{Q} = i[H, Q]$, where H is the self-adjoint generator of $\mathcal{U}_t$. We may thus equate $i[(\mu/2)\dot{Q}^2, Q]$ and $i[H, Q]$, which implies that

$$[(H - \frac{\mu}{2}\dot{Q}^2), Q] = 0.$$

Since Q forms a complete set of commuting observables, all operators that commute with Q are Borel functions of it. So the fact that $H - \frac{1}{2}\mu\dot{Q}^2$ commutes with Q implies that it is in fact a function of Q. Call that function $v(Q)$. Then we have the desired result,

$$H = \frac{\mu}{2}\dot{Q}^2 + v(Q),$$

which proves the theorem. □

A.2 A classical analogue of Jauch's lemma

A.2.1 Background

In this section I will adopt the notation of Chapter 4, Section 4.3.2. We take our basic object of study to be a symplectic manifold $(\mathcal{P}, \Omega_{ab}, h)$, where $\mathcal{P}$ is a smooth manifold, Ω_{ab} is a symplectic form (with inverse Ω^{ab}), and h is a smooth function (the "Hamiltonian") representing the energy of the system. We define the Poisson bracket of this system by,

$$\{f, h\} := \Omega^{ab}(d_a h)(d_b f),$$

for any smooth functions f and h. From Ω_{ab} and d_a, the Poisson bracket inherits the properties of being antisymmetric, linear in both terms, satisfying the Leibniz rule in both terms, and vanishing for constant functions. If f, h generate vector fields F^a and H^a by the prescription above, let φ^f_α and φ^h_β denote the diffeomorphism flows with tangent fields F^a and H^a, respectively. It will be useful in what follows to observe that, by our definitions

$$\{f, h\} := \Omega^{ab}(d_a h)(d_b f) = H^b d_b f = \frac{d}{d\beta}\left(f \circ \varphi^h_\beta\right)\bigg|_{\beta=0} \tag{A.3}$$

where the last equality is an expression of the chain rule. In other words, the Poisson bracket $\{f, h\}$ at a point p is equal to the directional derivative of the scalar field f at p, in the direction of the vector field H^a determined by h.

We take "position in space" of a $2n$-dimensional symplectic manifold $(\mathcal{P}, \Omega_a b)$ to be defined by a maximal orthogonal set, meaning a set $\{\overset{1}{q}, \overset{2}{q}, \ldots, \overset{n}{q}\}$ of n smooth functions $\overset{i}{q}$: $\mathcal{P} \to \mathbb{R}$ such that (i) $\{\overset{i}{q}, \overset{j}{q}\} = 0$ for each $i, j = 1, \ldots, n$, and (ii) if f is another smooth function satisfying $\{f, \overset{i}{q}\} = 0$ for all i, then $f = f(\overset{1}{q}, \ldots, \overset{n}{q})$ is a function of the $\overset{i}{q}$. We then define the "initial velocity" to be the rate of change,

$$\dot{q} := \dot{q}(0) = \frac{d}{dt}(q \circ \varphi^h_t)\bigg|_{t=0} = \{q, h\}, \tag{A.4}$$

where the last equality follows from our observation in Equation (A.3).

We take a *translation and boost group* for a classical system $(\mathcal{P}, \Omega_{ab}, h)$ to be a $2n$-parameter family of diffeomorphisms $\Phi(\sigma, \rho) : \mathcal{P} \to \mathcal{P}$, which forms a representation of $\mathbb{R}^{2n}$, and such that

1. $q \circ \Phi(\sigma, \rho) = q + \sigma$
2. $\dot{q} \circ \Phi(\sigma, \rho) = \dot{q} + \rho$

where $q = \{\overset{1}{q}, \ldots, \overset{n}{q}\}$ is a maximal set of orthogonal functions, and $\dot{q}$ is the corresponding initial velocity. We define two associated diffeomorphism groups $\varphi^s_\sigma := \Phi(\sigma, 0)$ and $\varphi^r_\rho := \Phi(0, \rho)$, and refer to them as the *translation group* and the *boost group*, respectively. When these groups have a generator, we denote those generators by $s : \mathcal{P} \to \mathbb{R}$ and $r : \mathcal{P} \to \mathbb{R}$, respectively.

A classical system $(\mathcal{P}, \Omega_{ab}, h)$ is *covariant under translations and boosts* if there exists a translation and boost group $\Phi(\sigma, \rho)$ on $\mathcal{P}$ such that each element of the group is symplectic, in that $\Phi^*(\sigma, \rho)\Omega_{ab} = \Omega_{ab}$ for all σ, ρ, and such that the group is *complete*, in that the only functions f that commute with both generators r and s are the constant functions.

A.2.2 A classical analogue of Jauch's lemma

Jauch argued that the if Galilei covariant quantum theory has Q as a complete set of commuting observables, then the Hamiltonian must take its standard form. We have now said what it means for classical to mechanics to be similarly Galilei covariant. And, we have seen that the natural classical analogue of a complete set of commuting observables is a "maximal orthogonal set" (or a "real polarization"), written $\{\overset{1}{q}, \ldots, \overset{n}{q}\}$. For such systems, one can use these analogues to formulate a direct classical analogue of Jauch's lemma

Theorem A.2 (Classical Jauch)**.** *If $(\mathcal{P}, \Omega_{ab}, h)$ is translation and Galilei boost covariant with respect to a maximal orthogonal set $\{\overset{1}{q}, \ldots, \overset{n}{q}\}$, then $\{q, \mu\dot{q}\} = 1$ for some (non-zero) $\mu \in \mathbb{R}$, and $h = (\mu/2)\dot{q}^2 + v(q)$ for some function v of q alone.*

It is convenient to build the proof of the theorem using two lemmas.

Lemma A.1. *If $(\mathcal{P}, \Omega_{ab}, h)$ is translation and Galilei boost covariant, with s generating the translation group and r generating the boost group, then $\{s, r\} = \mu$, where μ is a constant function $\mu(p) = \mu \in \mathbb{R}$ for all $p \in \mathcal{P}$.*

Proof. The translation and boost group $\Phi(\sigma, \rho)$ is defined to be a representation of the additive group of real vectors. Since the latter is abelian, $\Phi(\sigma, \rho) = \Phi(\sigma, 0)\Phi(0, \rho) = \Phi(0, \rho)\Phi(\sigma, 0)$. Thus, the translation group $\varphi^s_\sigma := \Phi(\sigma, 0)$ and the boost group $\varphi^r_\rho := \Phi(0, \rho)$ are commuting diffeomorphism flows, in that $\varphi^s_\sigma\varphi^r_\rho = \varphi^r_\rho\varphi^s_\sigma$. Moreover, the covariance assumption entails that these flows are symplectic. But the symplectic flows generated by s and r commute if and only if $\{s, r\}$ is a constant function (Arnold 1989, p.218 Cor. 9). Therefore, $\{s, r\} = \mu \in \mathbb{R}$. □

Lemma A.2. *If $(\mathcal{P}, \Omega_{ab}, h)$ is translation and Galilei boost covariant, with s generating the translation group and r generating the boost group, then $\{q, s\} = \{\dot{q}, r\} = 1$ and $\{q, r\} = \{\dot{q}, s\} = 0$.*

Proof. The position and initial velocity functions q and $\dot{q}$ transform under translations as $q \circ \varphi^s_\sigma = q + \sigma$ and $\dot{q} \circ \varphi^s_\sigma = \dot{q}$. They transform under Galilei boosts as $\dot{q} \circ \varphi^r_\rho = \dot{q} + \rho$ and

$q \circ \varphi^r_\rho = q$. Thus, applying the definitions summarized in Equation (A.3), we have,

$$\{q, s\} = \frac{d}{d\sigma}\left(q \circ \varphi^s_\sigma\right) = \frac{d}{d\sigma}(q + \sigma) = 1, \qquad \{q, r\} = \frac{d}{d\rho}\left(q \circ \varphi^r_\rho\right) = \frac{d}{d\rho}(q) = 0,$$
$$\{\dot{q}, r\} = \frac{d}{d\rho}\left(\dot{q} \circ \varphi^r_\rho\right) = \frac{d}{d\rho}(\dot{q} + \rho) = 1, \qquad \{\dot{q}, s\} = \frac{d}{d\sigma}\left(\dot{q} \circ \varphi^s_\sigma\right) = \frac{d}{d\sigma}(\dot{q}) = 0.$$

□

The theorem is now established by the following two propositions.

Proposition A.1. *If $(\mathcal{P}, \Omega_{ab}, h)$ is translation and Galilei boost covariant with respect to a maximal orthogonal set $\{\overset{1}{q}, \dots, \overset{n}{q}\}$, then $\{q, \mu\dot{q}\} = 1$ for some (non-zero) $\mu \in \mathbb{R}$.*

Proof. By our covariance assumption, the translation and Galilei boost groups are symplectic. This is a necessary and sufficient condition for each to have a generator (Marsden and Raiu 1999, Proposition 2.6.1), which we denote by s and r, respectively. From Lemmas A.1 and A.2 we obtain the relations,

(i) $\{s, r\} = \mu \in \mathbb{R}$,

(ii) $\{q, s\} = \{\dot{q}, r\} = 1$,

(iii) $\{q, r\} = \{\dot{q}, s\} = 0$.

Since the Poisson bracket is skew-symmetric, relation (i) may be written $\{r, s\} = -\mu$. Also, multiplying both sides of $\{q, s\} = 1$ by μ, we have $\{\mu q, s\} = \mu$. Therefore, using the linearity of the Poisson bracket,

$$\{r + \mu q, s\} = \{r, s\} + \{\mu q, s\} = -\mu + \mu = 0.$$

But the function $(r + \mu q)$ also satisfies $\{r + \mu q, r\} = 0$, since $\{r, r\} = \{r, q\} = 0$ by relation (iii). But by assumption the only functions that commute with both r and s are constant functions. So, $r + \mu q = k$ for some constant k, or equivalently, $r = -\mu q + k$. Substituting this into $\{\dot{q}, r\} = 1$ of relation (ii), we have,

$$1 = \{\dot{q}, r\} = \{\dot{q}, (-\mu q + k)\} = \{\dot{q}, (-\mu q)\} = \{q, \mu\dot{q}\},$$

where the penultimate equality follows from the fact that the Poisson bracket is linear and vanishes for constants in either term, and the last equality is an application of skew symmetry.

□

Proposition A.2. *If* $(\mathcal{P}, \Omega_{ab}, h)$ *is a classical system with* $\{q, \mu\dot{q}\} = 1$ *for some* $\mu \in \mathbb{R}$*, then* $h = \frac{\mu}{2}\dot{q}^2 + v(q)$ *for some function* v *of* q *alone.*

Proof. From the fact that the Poisson bracket satisfies the Leibniz rule,

$$\left\{q, \frac{\mu}{2}\dot{q}^2\right\} = \frac{1}{2}\dot{q}\{q, \mu\dot{q}\} + \frac{1}{2}\{q, \mu\dot{q}\}\dot{q} = \frac{1}{2}\dot{q} + \frac{1}{2}\dot{q} = \dot{q},$$

where the penultimate equality follows from our hypothesis that $\{q, \mu\dot{q}\} = 1$. But by definition, $\dot{q} = \{q, h\}$ (see Equation (A.4)). Subtracting the expression for $\dot{q}$ just calculated from this definition, we see that $\{q, h - \frac{\mu}{2}\dot{q}^2\} = 0$. But q is a maximal orthogonal set (Definition 4.2), and so by definition, $h - \frac{\mu}{2}\dot{q}^2 = v(q)$ for some function v of q alone.

□

APPENDIX B

UNIQUENESS OF THE TIME REVERSAL OPERATOR FOR AN ELEMENTARY PARTICLE WITH SPIN-1/2

In Section 5.3.1 of Chapter 5, I suggested that there is no point in trying to invent a time reversal operator for a spin-1/2 particle that differs from the standard one. One can prove a uniqueness result, which guarantees time reversal can really only mean one thing in this context.

Let $\sigma_1 = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}$, $\sigma_2 = \begin{pmatrix} & -i \\ i & \end{pmatrix}$, $\sigma_3 = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}$ be the 2-dimensional (irreducible) Pauli representation of an elementary particle with spin. The "standard" time reversal operator in this context is $T = \sigma_2 K$, where K is the antilinear conjugation operator in the $\{\binom{1}{0}, \binom{0}{1}\}$ basis[1]. By simple matrix multiplication, one easily verifies that this T has two important properties. First, it has the effect of reversing spin: $T\sigma_i T^{-1} = -\sigma_i$, for each $i = 1, 2, 3$. This is expected, although more on that below. Second, applying time reversal twice is not the same as the identity: $T^2 = -I$. This is unexpected. Roughly speaking: if we could film an electron, reverse the film, and then reverse it again, we would add a phase factor of -1. Only by reversing twice more do we get back to the original film.

But why assume $T = \sigma_2 K$? Would any other option be satisfactory? I argue that the answer is no. This is the only reasonable time reversal operator for a spin-1/2 particle, up to a complex unit. Moreover, even if we play around with the complex unit, the time reversal operator will still have have the unusual property that $T^2 = -I$.

[1]By definition, K maps each vector $\psi = \alpha\binom{0}{1} + \beta\binom{1}{0}$ to its complex conjugate, $K\psi = \psi^* = \alpha^*\binom{0}{1} + \beta^*\binom{1}{0}$. Notably, K satisfies $K^2 = I$. See (Jauch 1968, §14-5) for a statement of the standard T.

My argument assumes only that the following are minimal features of any satisfactory operator T representing "time reversal" in the Pauli representation.

- *T is antiunitary.* This is a feature of time reversal in any context; for an explanation of why, see Chapter 2, Section 2.2.
- *T is an involution, $T^2 = e^{i\theta} I$.* This captures what it means for T to be a "reversal": apply it twice, and you get back to where you started, up to an arbitrary phase factor $e^{i\theta}$.
- *T reverses angular momentum.* Since each σ_i represents a kind of angular momentum, they must each reverse sign under time reversal. A rough way to think about this is to imagine a film of a spinning ball: reversing the film results in a ball spinning in the opposite direction. There are more precise ways to think about this as well[2].

These bare assertions about the nature of time reversal give us the following.

Proposition B.1. *Let σ_1, σ_2, σ_3 be the spin operators in the Pauli representation, and let K be the conjugation operator in the $\{\binom{1}{0}, \binom{0}{1}\}$ basis. If T is any antilinear involution that satisfies $T\sigma_i T^{-1} = -\sigma_i$ for each $i = 1, 2, 3$, then $T = k\sigma_2 K$ for some complex unit k, and $T^2 = -I$.*

Proof. Define $\tilde{T} := \sigma_2 K$, and let T be any involution that reverses the sign of σ_i for each $i = 1, 2, 3$. Then,

$$(\tilde{T}T)\sigma_i(\tilde{T}T)^{-1} = \tilde{T}(T\sigma_i T^{-1})\tilde{T}^{-1} = \tilde{T}(-\sigma_i)\tilde{T}^{-1} = \sigma_i.$$

$\tilde{T}T$ commutes with all the generators of the representation, and so it commutes with everything. But the Pauli representation is irreducible, so $\tilde{T}T = -kI$ for some $k \in \mathbb{C}$ by Schur's lemma (Blank et al. 2008, Theorem 6.7.1). Multiplying on the left by $-\tilde{T}$ and recalling that $\tilde{T}^2 = -I$,

$$T = k\tilde{T} = k\sigma_2 K.$$

[2]For example: the σ generate spatial rotations, $R_\theta = e^{i\theta\sigma}$. But it is reasonable to assume that the reversal of time does not pick out a preferred direction in space (it is "isotropic"). This is made precise by assuming that the time reversal operator T commutes with R_θ. But T is antiunitary, as discussed in Chapter 2. And if an antiunitary operator T commutes with $R_\theta = e^{i\theta\sigma}$, then it must *anti*commute with the generator σ. (Why? Let $R_\theta = TR_\theta T^{-1}$. Then $e^{i\theta\sigma} = Te^{i\theta\sigma}T^{-1} = e^{Ti\theta\sigma T}$. By Stone's theorem, R_θ has a unique self-adjoint generator, so $i\theta\sigma = Ti\theta\sigma T$. But T is antiunitary, so $i\theta\sigma = Ti\theta\sigma T = -i\theta T\sigma T^{-1}$. Hence, $T\sigma T^{-1} = -\sigma$.)

This T is an involution, so there is a $c \in \mathbb{C}_{unit}$ such that $cI = T^2 = (k\tilde{T})(k\tilde{T}) = kk^*\tilde{T}^2 = -kk^*$. But $-kk^*$ is real and negative, and the intersection $\mathbb{R}^- \cap \mathbb{C}_{unit} = \{-1\}$. Thus $c = -1$, and so $kk^* = 1$. This says that k is a complex unit and $T^2 = -I$. $\square$

In this sense, the mathematical structure of an electron truly forces one to admit that $T^2 = -I$, however uncomfortable that assertion may be.

A final comment on this property: little is lost in choosing the 2-dimensional Pauli representation in characterizing an electron. All finite-dimensional representations of the Pauli spin relations are unitarily equivalent[3]. So, from the fact that $T^2 = -I$ in one representation, it follows that the corresponding time reversal operator in every other representation (given by ${}^{U}T = UTU^{-1}$ for some unitary operator U) has this property as well.

[3]This is the sometimes called the *Jordan-Wigner theorem*; see (Jauch 1968, §14.3) for a discussion of the proof, and (Ruetsche 2011, §3.3.1) for implications and limitations.

APPENDIX C

BIBLIOGRAPHY

Abraham, R. and Marsden, J. E. (1978). *Foundations of Mechanics*, 2nd edn, Addison-Wesley Publishing Company, Inc. (Referred to on page 67)

AIP (1963). Interview of Werner Heisenberg by T. S. Kuhn on February 15, 1963, Session V, Niels Bohr Library & Archives, American Institute of Physics, College Park, MD USA. **URL:** *http://www.aip.org/history/ohilist/4661_5.html* (Referred to on page 30)

Albert, D. Z. (2000). *Time and chance*, Harvard University Press. (Referred to on page 12, 40)

Angelopoulos, A. et al. (1998). First direct observation of time-reversal non-invariance in the neutral-kaon system, *Physics Letters B* **444**(1–2): 43 – 51. (Referred to on page 75)

Arnold, V. I. (1989). *Mathematical methods of classical mechanics*, Springer. (Referred to on page 59, 100)

Arntzenius, F. and Greaves, H. (2009). Time reversal in classical electromagnetism, *The British Journal for the Philosophy of Science* **60**(3): 557–584. (Referred to on page 3)

Ballentine, L. E. (1998). *Quantum Mechanics: A Modern Development*, World Scientific Publishing Company. (Referred to on page 11, 82, 83)

Bigi, I. I. and Sanda, A. I. (2009). *CP Violation*, Cambridge: Cambridge University Press. (Referred to on page 6, 68, 71, 74)

Blank, J., Exner, P. and Havlíček, M. (2008). *Hilbert Space Operators in Quantum Physics*, second edn, Springer Science and Business Media B.V. (Referred to on page 7, 8, 15, 24, 35, 38, 41, 43, 70, 83, 85, 86, 87, 96, 97, 104)

Boyle, R. (1772). *The Works of the Honourable Robert Boyle*, Vol. 1, London: J. and F. Rivington. (Referred to on page 30)

Callender, C. (1995). The metaphysics of time reversal: Hutchison on classical mechanics, *The British Journal for the Philosophy of Science* **46**(3): 331–340. (Referred to on page 55, 56)

Callender, C. (2000). Is Time 'Handed' in a Quantum World?, *Proceedings of the Aristotelian Society*, Aristotelian Society, pp. 247–269. (Referred to on page 12, 40)

Christenson, J. H., Cronin, J. W., Fitch, V. L. and Turlay, R. (1964). Evidence for the 2π decay of the k_2^0 meson, *Phys. Rev. Lett.* **13**(4): 138–140. (Referred to on page 6, 47, 68)

Colladay, D. and Kostelecký, V. A. (1997). CPT violation and the standard model, *Physics Review D* **55**: 6760–6774. (Referred to on page 74)

Costa, G. and Fogli, G. (2012). *Symmetries and Group Theory in Particle Physics: An Introduction to Space-time and Internal Symmetries*, Vol. 823 of *Lecture Notes in Physics*, Springer-Verlag Berlin Heidelberg. (Referred to on page 75)

Dirac, P. A. M. (1949). Forms of relativistic dynamics, *Reviews of Modern Physics* **21**(3): 392–399. (Referred to on page 6)

Dirac, P. A. M. (2001). *Lectures on quantum mechanics*, New York: Dover. (Referred to on page 13)

Earman, J. (2002). What time reversal is and why it matters, *International Studies in the Philosophy of Science* **16**(3): 245–264. (Referred to on page 41)

Earman, J. (2004). Curie's Principle and spontaneous symmetry breaking, *International Studies in the Philosophy of Science* **18**(2-3): 173–198. (Referred to on page 69)

Earman, J. (2008). Superselection Rules for Philosophers, *Erkenntnis* **69**(3): 377–414. (Referred to on page 38)

Earman, J. and Norton, J. D. (1998). Exorcist XIV: The Wrath of Maxwell's Demon. Part I. From Maxwell to Szilard, *Studies In History and Philosophy of Science Part B: Studies In History and Philosophy of Modern Physics* **29**(4): 435 – 471. (Referred to on page 2)

Frigg, R. (2008). A Field Guide to Recent Work on the Foundations of Statistical Mechanics, *in* D. Rickles (ed.), *The Ashgate Companion to Contemporary Philosophy of Physics*, Aldershot, England: Ashgate Publishing Limited, chapter 3, pp. 99–196. (Referred to on page 55, 58)

Galilei, G. (1967). *Dialogue Concerning the Two Chief World Systems*, second revised edn, Berkeley and Los Angeles: University of California Press. (Referred to on page 4)

Gell-Mann, M. and Pais, A. (1955). Behavior of Neutral Particles under Charge Conjugation, *Physical Review* **97**(5): 1387–1389. (Referred to on page 71)

Geroch, R. (1974). Geometric quantum mechanics. Transcribed from the original lecture notes, version of 1/31/2006. http://www.phy.syr.edu/~salgado/geroch.notes/geroch-gqm.pdf. (Referred to on page 62)

Goldstein, H., Poole, C. and Safko, J. (2002). *Classical Mechanics*, 3rd edn, Addison Wesley. (Referred to on page 57)

Greaves, H. (2010). Towards a Geometrical Understanding of the CPT Theorem, *The British Journal for the Philosophy of Science* **61**(1): 27–50. (Referred to on page 74)

Haag, R. (1996). *Local Quantum Physics: Fields, Particles, Algebras*, 2nd ed edn, Springer-Verlag Berlin Heidelberg. (Referred to on page 72, 73, 92)

Halvorson, H. and Clifton, R. (2002). No Place for Particles in Relativistic Quantum Theories?, *Philosophy of Science* **69**(1): 1–28. (Referred to on page 27)

Hutchison, K. (1993). Is classical mechanics really time-reversible and deterministic?, *The British journal for the philosophy of science* **44**(2): 307–323. (Referred to on page 55)

Jauch, J. M. (1964). Gauge invariance as a consequence of Galilei-invariance for elementary particles, *Helvetica Physica Acta* **37**: 284–292. (Referred to on page 7, 42, 94)

Jauch, J. M. (1968). *Foundations of Quantum Mechanics*, Addison-Wesley Series in Advanced Physics, Reading, MA, Menlo Park, CA, London, Don Mills, ON: Addison-Wesley Publishing Company, Inc. (Referred to on page 13, 15, 25, 26, 34, 42, 66, 78, 82, 94, 103, 105)

Jost, R. (1957). Eine Bemerkung zum CTP Theorem, *Helvetica Physica Acta* **30**: 409–416. (Referred to on page 73)

Kabir, P. K. (1970). What Is Not Invariant under Time Reversal?, *Physical Review D* **2**(3): 540–542. (Referred to on page 76)

Khriplovich, I. B. (2005). *General Relativity*, New York: Springer-Verlag+Business Media Inc. (Referred to on page 2)

Khriplovich, I. B. and Lamoreaux, S. K. (1997). *CP Violation Without Strangeness: Electric Dipole Moments of Particles, Atoms, and Molecules*, Springer-Verlag Berlin Heidelberg. (Referred to on page 68, 77, 84)

Kleinknecht, K. (2003). *Uncovering CP Violation: Experimental Clarification in the Neutral K Meson and B Meson Systems*, Springer-Verlag Berlin Heidelberg. (Referred to on page 68)

Kramers, H. A. (1930). Théorie générale de la rotation paramagnétique dans les cristaux, *Proceedings Koninklijke Akademie van Wetenschappen* **33**: 959–972. (Referred to on page 77)

Kraus, K. (1980). Galilei Covariance Does Not Imply Minimal Electromagnetic Coupling, *Annalen der Physik* **492**(2): 81–101. (Referred to on page 66)

Leeds, S. (2006). Discussion: Malament on time reversal, *Philosophy of Science* **73**(4): 448–458. (Referred to on page 3)

Lévy-Leblond, J.-M. (1970). Minimal electromagnetic coupling as a consequence of Lorentz invariance, *Annals of Physics* **57**(2): 481 – 495. (Referred to on page 66)

Mackey, G. W. (1963). *Mathematical foundations of quantum mechanics*, New York: W. A. Benjamin, Inc. (Referred to on page 7, 32)

Malament, D. B. (1996). In Defense Of Dogma: Why There Cannot Be A Relativistic Quantum Mechanics Of (Localizable) Particles, *in* R. Clifton (ed.), *Perspectives on Quantum Reality: Non-Relativistic, Relativistic, and Field-Theoretic*, Kluwer Academic Publishers, pp. 1–10. (Referred to on page 27)

Malament, D. B. (2004). On the time reversal invariance of classical electromagnetic theory, *Studies in History and Philosophy of Modern Physics* **35**: 295–315. (Referred to on page 3, 92)

Marsden, J. and Raiu, T. (1999). *Introduction to mechanics and symmetry: a basic exposition of classical mechanical systems*, Vol. 17, Springer-Verlag. (Referred to on page 65, 101)

Maudlin, T. (2007). *The Metaphysics Within Physics*, New York: Oxford University Press. (Referred to on page 12, 40)

Messiah, A. (1999). *Quantum Mechanics, Two Volumes Bound as One*, New York: Dover. (Referred to on page 22, 42, 82)

North, J. (2008). Two views on time reversal, *Philosophy of Science* **75**(2): 201–223. (Referred to on page 3)

Penrose, R. (2004). *The Road to Reality: A Complete Guide to the Laws of the Universe*, London: Jonathan Cape. (Referred to on page 9)

Rédei, M. (1998). *Quantum Logic in Algebraic Approach*, Dordrecht: Kluwer Academic Publishers. (Referred to on page 8)

Ruetsche, L. (2011). *Interpreting Quantum Theories*, New York: Oxford University Press. (Referred to on page 105)

Sachs, R. G. (1987). *The Physics of Time Reversal*, Chicago: University of Chicago Press. (Referred to on page 12, 68, 82)

Sozzi, M. S. (2008). *Discrete Symmetries and CP Violation*, Oxford: Oxford University Press. (Referred to on page 68)

Streater, R. F. (2007). *Lost Causes in and beyond Physics*, Springer-Verlag Berlin Heidelberg. (Referred to on page 13)

Uffink, J. (2007). Compendium of the Foundations of Classical Statistical Physics, *in* J. Butterfield and J. Earman (eds), *Philosophy of Physics Part B*, Elsevier, pp. 923–1074. (Referred to on page 1)

Uhlhorn, U. (1963). Representation of symmetry transformations in quantum mechanics, *Arkiv för Fysik* **23**: 307–340. (Referred to on page 12, 17, 34)

Varadarajan, V. S. (2007). *Geometry of Quantum Theory*, New York: Springer Science and Business Media, LLC. (Referred to on page 17, 34)

Wigner, E. P. (1931). *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra*, New York: Academic Press (1959). (Referred to on page 3, 4, 19)

Wigner, E. P. (1932). Über die Operation der Zeitumkehr in der Quantenmechanik, *Nachrichten der Gesellschaft der Wissenschaften zu Göttingen Mathematisch-Physikalische Klasse* pp. 546–559. (Referred to on page 4, 77)

Woodhouse, N. (1981). Geometric quantization and the Bogoliubov transformation, *Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences* **378**(1772): 119–139. (Referred to on page 64)

Woodhouse, N. M. J. (1991). *Geometric Quantization*, New York: Oxford University Press.
(Referred to on page 13)