3

Time Reversal in Physical Theory

Précis. *The meaning of time reversal on state space is determined by a representation of time translations.*

When the time reversal transformation T appears on the state space of a physical theory, it is usually required to do a number of idiosyncratic things, such as those listed in Figure 3.1. For example, it conjugates wavefunctions in quantum mechanics, and it reverses magnetic fields in electromagnetism. To explain this, its advocates often refer to it as 'motion reversal' and appeal to a wide variety of physical facts about how this affects instantaneous states. This led Albert (2000, p.18) to remark that in the textbooks, "for one physical situation to be the time reverse of another is (not surprisingly!) an obscure and difficult business". Even Earman (2002b) admits that the going gets tough:

> I do not mean to suggest by the above examples that fixing the properties of the time reversal operation is always such an easy or straightforward matter. It is not, and in some instances the quest may not end in any clear answer. (Earman 2002b, p.249)

This chapter aims to make the meaning of time reversal in state space a more straightforward matter. We now have the results of Chapter 2 in hand: on the Representation View (Section 2.3), a dynamical theory is by definition a representation of time translations on state space. We have seen how to construct a time reversal group element $\tau : t \mapsto -t$ that reverses those time translations (Section 2.6); and, when the representation of time translations extends to this larger structure, then we can immediately understand the time reversal transformation T to be the image of τ in state

https://doi.org/10.1017/9781009122139.004 Published online by Cambridge University Press

Reversed		**Preserved**	
Momentum:	$p \mapsto -p$	Position:	$q \mapsto q$
Magnetic Field:	$B \mapsto -B$	Electric Field:	$E \mapsto E$
Spin:	$\sigma \mapsto -\sigma$	Kinetic Energy:	$p^2/2m \mapsto p^2/2m$
Position wavefunction:	$\psi(x) \mapsto \psi(x)^*$	Transition probability:	$\lvert\langle\psi,\phi\rangle\rvert^2 \mapsto \lvert\langle\psi,\phi\rangle\rvert^2$

Figure 3.1 Some effects of the time reversal operator T on state space.

space (Section 2.7). So, the meaning of time reversal is not so haphazard after all: it is determined by a representation of time translations. What remains is to verify that the familiar time reversal operator in textbooks is what we recover by this procedure.

This chapter aims to make good on that, by showing how various expressions of time reversal on a physical state space now follow automatically. I will make this argument case by case, for some common frameworks for dynamical systems in physics. I hope that by illustrating the general technique, it will be clear how to apply it in other physical theories too. In each case, our task will be to determine two things:

1. the *general* character of the time reversal group element in a representation;
2. its *specific* character when it is required to reverse time and 'do nothing else'.

The second task arises because, in general, there are a great number of ways to represent time reversal on state space: we can reverse time and space; we can reverse time and rotate 180 degrees; we can reverse time and exchange matter and antimatter; and so on. Recognising this will be important for our development of the CPT operator in Chapter 8. But, the specific task for us now will be to distinguish 'bare time reversal' from the many other time-reversing transformations.

Our focus in this chapter will be on state space representations. Nearly every theory in modern physics has an expression of this kind; in particular, the framework for analytic mechanics that we discuss in Section 3.3 is so robust that it can be viewed as including both general relativity and quantum theory as special cases.[1] We begin Section 3.1 with some general remarks about state space representations of time translations, which will be applied in all the theories to follow. This includes precise statements of two physical

[1] Wald (1984, Appendix E) gives a classic introduction to general relativity in the Hamiltonian and Lagrangian frameworks. The fact that quantum theory can be viewed as a special case of analytic mechanics is less well-known, but was pointed out independently by Ashtekar and Schilling (1999), Gibbons (1992), Kibble (1979).

postulates about the experience of time's passage in general: that it is a symmetry of local physics, and that the energy is half-bounded. We then turn to deriving the meaning of time reversal in various physical theories: Newtonian mechanics (Section 3.2), analytic mechanics (Section 3.3), and quantum theories (Section 3.4).

3.1 State Space Representations

3.1.1 State Space

State spaces in physics are used to represent possible states of affairs in nature, like the possible locations of the planets in our solar system (Figure 3.2). However, when one spends a little time getting to know state spaces, some patterns among them emerge: many share common structure; and some of them seem perennially useful, while others do not. What makes a structure reasonable, appropriate, or fruitful for use as a state space?

Philosophers of physics have made some comments in this regard, such as the proposal of Albert (2000, p.9) that a state is (at least) a "genuinely instantaneous" and "complete" description of physical facts.[2] Butterfield (2006b,d), inspired by Lewis (1986), characterises this view as one in which states are "temporally intrinsic" facts: when obtaining at a moment, a state does not by itself imply any contingent facts about other times. Thus, physicists sometimes refer to configuration space or phase space in mechanics as an 'instantaneous state space'. However, as Butterfield (2006a,b, 2011) goes on to argue at length, even the state space of classical mechanics cannot possibly be just that; for example, "[m]echanics needs of course to refer to the instantaneous velocity or momentum of a body; and this is temporally extrinsic to the instant in question" (Butterfield 2006a, p.193). I agree. The state spaces to be considered in this chapter are highly structured objects – Hilbert spaces, phase spaces, jet bundles, and the like – and all this structure is needed to understand a symmetry transformation like time reversal.[3]

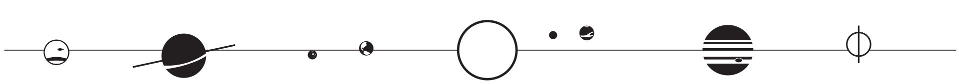

Figure 3.2 A state in the space of the planetary locations.

[2] See Section 2.5.1.
[3] Barrett (2018b, 2019, 2020a, 2020b), Dewar (2022, Part IV), Halvorson (2012, 2013, 2019), and Weatherall (2016) have argued on similar grounds that a physical theory can be viewed as a category. I have reservations about some of the remarks in this literature (cf. Roberts 2020), but agree with the general structuralist spirit.

https://doi.org/10.1017/9781009122139.004 Published online by Cambridge University Press

This might seem to contradict some remarks of Albert (2000) about dynamical theories, but I think that it need not. Albert (2000, pp.9–10) distinguishes between states representing complete descriptions of the world at an instant, and the *dynamical conditions* needed to make use of "the full predictive resources" of the laws of physics. In Newtonian mechanics, he takes the states to be particle positions in $\mathbb{R}^{3n}$, while the dynamical conditions also include velocities. I have no problem with this, but will propose a liberal reading of Albert: nothing about his remarks preclude state space from having a great deal of structure beyond this. It can be a manifold, like a tangent bundle of vectors, or be equipped with a symplectic form or a Euclidean metric, among other things.

In the remainder of this section, I will identify two key structural facts associated with general state spaces: that time translations can be viewed as symmetries in isolated systems, and that energy is bounded from below but not from above. This will set the stage for the analysis of time reversal on particular examples of state spaces in the remainder of the chapter.

3.1.2 Time Translations Are Automorphisms

A representation is a homomorphism from a symmetry structure to the automorphisms of a state space. Each state space structure has its own standard of automorphism or 'structure preserving map': for example, a differentiable manifold has diffeomorphisms, while a vector space has linear maps. So, by choosing a state space structure, and therefore a notion of state space automorphisms, we constrain what sorts of representations are possible.

Representing the group of time translations amongst the transformations of a state space is what justifies referring to the representing state space map as 'time evolution' as opposed to something else; this was the thesis of Section 2.3. And, not just any transformations: time translations should be represented among the state space maps that are *automorphisms*, at least for isolated systems. This is because most physical theories are built to capture the repeatability of local experiments, in the following sense.

We seek theories that can be supported by experimental evidence, but also – where possible – theories for which that experimental evidence can be supplied again at a later time. Local dynamical theories are generally designed in this way: the modelling of time evolution is compatible with setting up an isolated experiment today, collecting the results, and then confirming those same results tomorrow when a structurally equivalent experiment is repeated. In more precise terms, a time translation is an automorphism of a dynamical theory. Of course, whether or not this kind of 'homogeneity in time' accurately describes reality is a matter of experience. But as Jauch points out, it turns out to be central to our experience:

> Whether there are systems which are, in this sense, homogeneous in time is of course a matter of experience, and it is indeed one of the fundamental experiences about the physical world that this is the case. (Jauch 1968, p.152)

Thus, we represent time translations amongst the automorphisms of a state space, as postulated by the Representation View. This by itself does not guarantee that the representation can be extended to include time reversal; as we will see in Chapter 4, a representation of time reversal may still fail to arise, giving rise to the failure of temporal symmetry. Our study here will thus be predicated on the assumption that a representation of time reversal exists.

3.1.3 Half-Bounded-Energy Representations

A second structural fact about typical representations of time translations, in stark contrast to representations of spatial translations, is that they are associated with a conserved quantity called 'energy' that is bounded from below. To get a more precise sense of what this means, recall that a Lie group (Definition 2.1) is a group with a manifold structure. The continuous transformations in a Lie group, which describe familiar symmetries like rotations or time translations, turn out to be 'locally generated' by an object called a Lie algebra:

> **Definition 3.1** The *Lie algebra* $\mathfrak{g}$ of a Lie group G is the algebra of right-invariant vector fields on G, where a 'right-invariant' vector field X is defined by the condition that for all $g, h \in G$, if $\rho_g : G \to G$ is the right-multiplication map on group elements defined by $h \mapsto h \cdot g$, then $d\rho_g(X)\big|_h = X\big|_{\rho_g(h)} = X\big|_{h\cdot g}$.

In more picturesque terms (Figure 3.3), an element of this Lie algebra is any vector field X on the Lie group manifold G with the property that group multiplication 'traces along' the vector field. As this picture suggests, elements of a Lie algebra stand in one-to-one correspondence with the one-parameter subgroups of a Lie group (cf. Olver 1993, Proposition 1.48). In particular, a one-parameter group of time translations $(\mathbb{R}, +)$ is associated with a single Lie algebra element X, called the *generator* of the one-parameter group.

A fundamental result in the representation theory of Lie groups is that every homomorphism between Lie groups induces a unique homomor-

https://doi.org/10.1017/9781009122139.004 Published online by Cambridge University Press

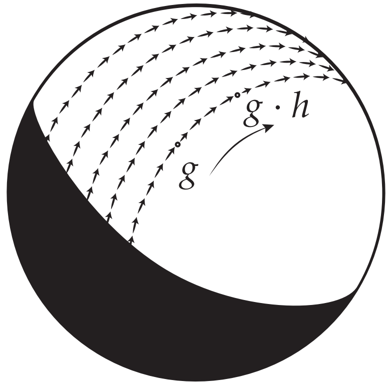

Figure 3.3 A Lie algebra element.

phism between the Lie algebras of those groups.[4] Thus, every state space representation of the Lie group defines a unique representation of its Lie algebra. For example, in a representation of spacetime symmetries, this group will include a generator for each one-parameter group of time translations $t \mapsto \phi_t$. When this generator X can be viewed as the gradient of a smooth function h, then this function h is called the 'Hamiltonian'. Since the Hamiltonian generates time translations, its value is also 'preserved' by time translations, and so in physical terms it can be used to define the conserved quantity known as *energy*. Similarly, each conserved function associated with a one-parameter group of spatial translations is called *three-momentum* or *momentum*, and traditionally denoted p.

Although conserved functions h for time translations and p for spatial translations are formally similar at the level of Lie groups and algebras, experience shows that they display a fundamental difference when interpreted as physical quantities. Namely, the possible values of momentum are unbounded above and below, while the possible values of energy are *half-bounded*: they are bounded from below and unbounded from above.[5] This is one of the truly remarkable differences between space and time, although it has received little attention from philosophers. And, it appears to represent an elementary fact about the local structure of our world in everything from the simple harmonic oscillator to quantum electrodynamics. Energy is typically unbounded from above, owing to the fact that velocity can be boosted arbitrarily close to the speed of light, and its lower bound is usually

4 Cf. Hall (2003, Theorem 2.21) or Landsman (2017, Theorem 5.42).
5 A well-known exception in the history of physics is the unbounded 'negative energy sea' in Dirac's hole theory of electrodynamics (Dirac 1930). This theory encountered irreparable problems, and half-bounded energy was restored in the modern Fock space formulation of quantum electrodynamics (cf. Duncan 2012, §2.1).

interpreted as defining a 'stable ground state'. The common wisdom about why this lower bound exists has been captured by Malament (1996):

> If it failed, the particle could serve as an infinite energy source (the likes of which we just do not seem to find in nature). Think about it this way. We could first tap the particle to run all the lights in Canada for a week. To be sure, in the process of doing so, we would lower its energy state. Then we could run all the lights for a second week, and lower the energy state of the particle still further. And so on. If the particle had no finite ground state, this process could continue forever. There would never come a stage at which we had extracted all available energy. (Malament 1996, p.5)

Whatever the reason for this fact, we will use it as the basis for assuming that in a representation of time translations, the generator h must be half-bounded.

This completes our general discussion of state space and representation theory. In the next sections, we will apply it to a number of different theories. In each case, the procedure will be similar:

1. Represent time translations $(\mathbb{R}, +)$ amongst the automorphisms on a state space.
2. Interpret $\tau : t \mapsto -t$ on time translations as time reversal.
3. Extend the representation to include τ as a group element, interpreted in the representation as an instantaneous time reversal operator T.
4. Where possible, use the half-bounded energy constraint on the Hamiltonian h to determine the general character of T.
5. When further group structure is available, use it to try to determine the unique definition of T.

I hope the result will be a novel, unified approach to the meaning of time reversal: not 'motion reversal', nor a bag of tricks for reversing instants, but as the reversal of time translations in each physical theory where it can be applied.

3.2 Newtonian Time Reversal

3.2.1 Why Time Translations Are Needed

The world according to atomists like Democritus, Bošković, and Boyle consists of a finite number of indistinguishable particles "variously configured and moved", whose only properties are their locations in smooth three-dimensional Euclidean space (cf. Boyle 1772, p.355). When this is the case, the locations of n particles can be represented by the C^∞ (I will usually say 'smooth') real manifold with Euclidean metric $(\mathbb{R}^{3n}, \cdot)$. Each point $x \in \mathbb{R}^{3n}$

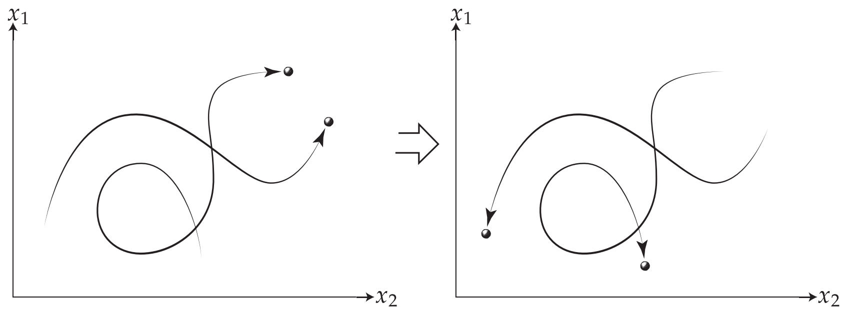

Figure 3.4 The 'folk' view of Newtonian time reversal without time translations.

represents a possible particle configuration at an instant, and the motion of the particles through space is represented by a smooth curve $x : \mathbb{R} \to \mathbb{R}^{3n}$ – in my notation, $x = (x_1, \ldots, x_{3n})$ is a coordinate in a manifold representing total configuration, not a spatial coordinate – with $\mathbb{R}$ interpreted as the time axis. So, we write $x(t)$ to represent the configuration $x \in \mathbb{R}^{3n}$ of the system at given time $t \in \mathbb{R}$.

In this general context, Callender (2000) has set out the common 'folk' view about time reversal:

> Relative to a co-ordinisation of spacetime, the time reversal operator takes the objects in spacetime and moves them so that if their old co-ordinates were t, their new ones are $-t$, assuming the axis of reflection is the co-ordinate origin. (Callender 2000, p.253)

In other words, the time reverse of a smooth curve $x(t)$ is $x(-t)$, as in Figure 3.4.

I agree with this description of Newtonian time reversal, at least when I put on a substantivalist hat.[6] But it should be viewed as incomplete. We often say, 'Let $x(t)$ represent the trajectories of particles over time'. However, as we have seen in Section 2.5.2, a more complete analysis of time requires including its structural properties: not just the time coordinates, but the time *translations*.

An adequate representation of time in Newtonian mechanics is not possible without this extra structure. In particular, $(\mathbb{R}^{3n}, \cdot)$ cannot be a complete Newtonian state space, since its automorphisms include only the (spatial)

[6] Newtonian substantivalism is traditionally justified using examples like Newton's bucket (Newton 1999, Scholium to the Definitions, pp.58–9). Mach's principle is a standard response (Earman 1989, Chapter 4), but this remains an active research area: see Rynasiewicz (1995a,b, 2014), and the novel recent alternative of Gomes and Gryb (2020) using techniques from Kaluza–Klein theory.

https://doi.org/10.1017/9781009122139.004 Published online by Cambridge University Press

transformations of rigid rotation, translation, and reflection. Time translations for this structure cannot even describe an elliptical orbit. So, beware that one's intuitions here can fail: the map $x(t) \mapsto x(-t)$ on smooth curves through a Euclidean manifold is not a complete description of time reversal. It is incomplete because it includes no representation of Newtonian time translations, and so it has no way to say that time translations are reversed. So, let us begin by discussing a more complete state space for Newtonian mechanics.

3.2.2 Newtonian Time Translations

Begin at the level of spacetime. Time translations are usually introduced by extending the Euclidean manifold to include a temporal dimension $\mathbb{R}$, together with some further spacetime structure. Although there are different ways to do this, most[7] admit a group of time translation symmetries isomorphic to $(\mathbb{R}, +)$. So, let me follow my announced practice of postulating the existence of a time translation group, without taking a position on spacetime structure.

To introduce a representation of these time translations, we now need to say what time translation invariance means in the context of Newtonian mechanics. This introduces Newton's second law. However, a rigorous treatment requires some technicalities regarding the symmetries of a differential equation that are not so often discussed. – *Oh dear: Must we complicate a theory as simple as Newtonian mechanics?* – I'm afraid that someone who comes to Newtonian mechanics for its simplicity was misinformed.[8] The situation is as Bill Burke once remarked, and in the spirit of the way Putnam (1962) describes kinetic energy:

> Be careful of the naïve view that a physical law is a mathematical relation between previously defined quantities. The situation is, rather, that a certain mathematical structure represents a given physical structure. Thus Newtonian mechanics does not assert that $F = ma$, with F, m, and a separately defined. Rather, it asserts that the structure of second-order differential equations applies to the motion of masses. (Burke 1985, p.37)

7 This includes Leibnizian, Maxwellian, Galilean, and Newtonian spacetime (Earman 1989, Chapter 2). An exception is Machian spacetime, associated with the work of Barbour (1974); but, insofar as this approach demotes time translations to a 'coordinate relabelling', the question of the meaning of time reversal does not appear to arise.

8 See Butterfield (2006a,b,c, 2007), Wallace (2022), and Wilson (2013) for some subtleties in the philosophical foundations of classical mechanics. The call-to-arms of Butterfield (2006a,b, 2011) is particularly apt for this discussion: he argues that the properties of classical physics cannot be viewed as defined only at points.

The mathematical structure required for Newtonian mechanics is the following. Observe that, given a smooth curve $x(t)$ through $\mathbb{R}^{3n}$, each point on the curve is associated with a set of values $(x, \dot{x}, \ddot{x}, \ldots)$ corresponding to its successively higher derivatives with respect to t. These values are elements of the *jet space* of $\mathbb{R}^{3n}$, associated with the smooth manifold $\mathbb{R}^{3n} \times \mathbb{R}^{3n} \times \cdots$ (see Olver 1993; Saunders 1989). In Newtonian mechanics, attention is restricted to the 'two-jet' associated with $M = \mathbb{R}^{3n} \times \mathbb{R}^{3n} \times \mathbb{R}^{3n}$, whose elements have the form $s = (x, \dot{x}, \ddot{x})$. Given a smooth function $F : \mathbb{R} \times \mathbb{R}^{3n} \times \mathbb{R}^{3n} \to \mathbb{R}^{3n}$ of only $(x, \dot{x})$, called a 'force', we say that a *solution to Newton's equation* is a smooth curve $x : \mathbb{R} \to M$ satisfying Newton's second law,[9]

$$\ddot{x} = F(t, x, \dot{x}), \tag{3.1}$$

at each point along the curve. Where are the masses? To simplify the notation, I interpret them as assigned by the function F. So, for a harmonic oscillator, $F(x) = -(k/m)x$, where k is called the 'spring constant'.

Viewing the solution $x(t)$ as a set of points $(t, s) \in \mathbb{R} \times M$, it is also convenient to define a function $f : \mathbb{R} \times M \to \mathbb{R}^{3n}$ by

$$f(t, x, \dot{x}, \ddot{x}) := \ddot{x} - F(t, x, \dot{x}). \tag{3.2}$$

Then each point $(t, s) = (t, x, \dot{x}, \ddot{x})$ on a curve is in the solution space if and only if $(t, s) \in \ker f$, meaning that $f(t, s) = \mathbf{0}$. The space of all such curves is called the *solution space* $\mathcal{S}_F$.

On reflection, it should be clear that Newton's equation in this form implies almost nothing about the physical world: without specifying the functional form of F, it is just the statement that motion is guided by a second-order differential equation. However, one of the few things it does imply is that the state space of Newtonian mechanics cannot possibly be $\mathbb{R}^3$, nor even $\mathbb{R}^{3n}$. Without an expression of higher derivatives, these spaces do not have enough structure to support Newton's equation. The state space of Newtonian mechanics is rather the state space of a typical differential equation, which has the richer structure of the jet space $\mathbb{R} \times M$. Careful treatments of the philosophy of Newtonian mechanics like that of Wallace (2022, §2) do correctly represent the structure of Newtonian state space, but

[9] The restriction to smooth force functions of at most first derivatives ensures the existence and uniqueness of solutions (see e.g. Arnol'd 1992, §7.2). This rules out pathological examples of local indeterminism like 'Norton's dome', although the latter is (rightly) the subject of much philosophical debate (cf. Earman 1986; Fletcher 2012; Gyenis 2013; Malament 2008; Norton 2008a; Wilson 2009).

this is not often discussed in physics or philosophy,[10] and it is a crucial observation for us here.

We can now develop a more appropriate understanding of 'symmetry' in Newtonian mechanics: we say that an *automorphism* of a Newtonian system with solution space $\mathcal{S}_F$ is a diffeomorphism ϕ on the state space of higher derivatives,

$$\phi : \mathbb{R} \times M \to \mathbb{R} \times M, \tag{3.3}$$

such that ϕ preserves the kernel of f: if $(t, s_t) \in \ker f$, then $\phi(t, s_t) \in \ker f$. This does what one would expect an automorphism of a solution space to do: if $x(t)$ is a solution to Newton's equation, then so is the curve produced by the transformation ϕ. As Belot (2013, §3) has emphasised, a Newtonian symmetry is not just a bijection on curves $x(t)$, but a diffeomorphism on $\mathbb{R} \times M$; indeed, a bijection alone does not necessarily preserve all the local structure of a differential equation, and if one exists that does, then it leads to a pathological perspective on symmetries.

This discussion should make clear that the structure of Newtonian state space is not a trivial matter. However, once it is in place, a *representation of time translations* in Newtonian mechanics can be concisely defined: it is a homomorphism $t \mapsto \phi_t$ from the time translations $(\mathbb{R}, +)$ to the automorphism group of $\mathcal{S}_F$, such that for each point $(t_0, s_0) \in \mathbb{R} \times M$ and for each time translation t,

$$\phi_t : (t_0, s_0) \mapsto (t_0 + t, s_t), \tag{3.4}$$

for some $s_t \in M$. Writing $\phi_t(s_0) = s_t$ for the restriction of this map to M, a representation of $(\mathbb{R}, +)$ thus defines a smooth curve $t \mapsto x(t)$ associated with each initial point (t_0, s_0), which is given by $x(t_0 + t) := \phi_t(s_0)$. Since an automorphism ϕ_t of the solution space $\mathcal{S}_F$ by definition preserves $\ker f$, this implies that $x(t)$ is a solution to Newton's equation. As one would expect, each time translation by a duration t maps a point on a curve $x(t_0)$ to a different point on the same curve, $x(t_0) \mapsto x(t_0 + t)$.

Let me emphasise that time translations are represented here by symmetries, as discussed in Section 3.1.2. If the map defined by Eq. (3.4) were not an automorphism of $\mathcal{S}_F$, then it would not provide a representation. This is standard practice in Newtonian mechanics: in locally isolated systems without any hidden degrees of freedom, forces do not seem to display any time dependence, and so time translations are symmetries of the theory; or,

10 Recognising the fact that state space must have higher derivatives has implications for much of the philosophical literature on the dimensionality of state space following Albert (1996), such as the contributions in Ney and Albert (2013), which I leave as an invitation to the reader.

in the language of this section, time translations are automorphisms of the solution space.[11] We will return to this assumption in Chapter 4, since it is related to the question of time reversal symmetry and symmetry violation.

3.2.3 The Newtonian Time Reversal Operator

Having clarified the structure of Newtonian time translations, we can simply follow the procedure of Chapter 2 to construct the time reversal transformation on Newtonian state space. Here is the fundamental idea, which we will repeat throughout this chapter.

Given a representation $t \mapsto \phi_t$ of the time translation group $(\mathbb{R}, +)$, recall that in Section 2.5.2, we saw that time reversal can be uniquely determined to transform this group as $\tau : t \mapsto -t$. And, in Section 2.6, we showed that it can be 'added in' as a group element to an extension G of those time translations, which satisfies $\tau t \tau^{-1} = -t$. If our representation of $(\mathbb{R}, +)$ can be extended to a representation of G, then there will be a state space transformation $T := \phi_\tau$ corresponding to time reversal that reverses each time translation, $T\phi_t T^{-1} = \phi_{-t}$ for all $t \in \mathbb{R}$. There is nothing mysterious about this transformation: it is just the representative of a temporal reflection on state space.

In the previous section, we saw how time translations give rise to curves in Newtonian mechanics. So, by reversing the time translations, time reversal in Newtonian mechanics gives rise to a transformation on curves, whereby the curve $x(t_0+t) = \phi_t(s_0)$ is replaced with a curve $x^T(t_0+t) = \phi_{-t}(s_0) = x(t_0-t)$. In other words, time reversal transforms curves through Newtonian state space as

$$x(t_0 + t) \mapsto x(t_0 - t). \tag{3.5}$$

This makes clear how the 'folk' account of time reversal discussed by Callender (2000) arises! If we parametrise $x(t)$ so that $t_0 = 0$, then we get exactly the transformation $x(t) \mapsto x(-t)$. But, it arises as a reflection of time translations rather than of time coordinates, and so there is no longer any question about 'which coordinate origin' to reflect about of the kind that concerns North (2008).[12] By reversing the whole structure of time, including time translations, that problem is dissolved.

[11] Of course, damped systems are often treated heuristically as if they do not satisfy this requirement, but this is generally because they are in reality not isolated; see Section 7.2.
[12] See Section 2.5.2.

https://doi.org/10.1017/9781009122139.004 Published online by Cambridge University Press

Now we can also see the origin of the time reversal transformation $T := \phi_\tau$ on state space, debated by the two camps of Section 2.2: it is just the image of the 'true' time reversal transformation $\tau : t \mapsto -t$ on state space. When a representation of it exists, it is guaranteed to reverse time translations when it acts on them by conjugation:

$$T\phi_t T^{-1} = \phi_{-t}. \tag{3.6}$$

For example, suppose we adopt a time-independent force $F = F(x, \dot{x})$, which has the property of being independent of the sign of velocity, $F(x, -\dot{x}) = F(x, \dot{x})$. This might be the force $F = (k/m)x$ for the harmonic oscillator, or $F = kq/x$ for a Coulomb force. Let me drop the time parameter t from state space, since this force does not depend on time. Then a representation of the time reversal operator is defined by transforming each state $s = (x, \dot{x}, \ddot{x}) \in M$ in Newtonian state space by

$$T : (x, \dot{x}, \ddot{x}) = (x, -\dot{x}, \ddot{x}). \tag{3.7}$$

To confirm that this provides a representation of the time reversal group element, we need only check that Eq. (3.6) is satisfied. First, if $(x, \dot{x}, \ddot{x})$ is the initial ($t = 0$) state on a curve $x(t)$, then we have

$$T\phi_t(x, \dot{x}, \ddot{x}) = T\,(x(t), \dot{x}(t), \ddot{x}(t)) = (x(t), -\dot{x}(t), \ddot{x}(t))\,. \tag{3.8}$$

Our assumptions imply that $x(t)$ is a solution only if $x(-t)$ is too,[13] meaning that it evolves from $x(0)$ under the same time translations φ_t to $x(-t)$. This curve has initial state $(x, -\dot{x}, \ddot{x})$, which φ_t transforms as $\varphi_t(x, -\dot{x}, \ddot{x}) = (x(-t), -\dot{x}(-t), \ddot{x}(-t))$. Therefore, applying the reverse time translation φ_{-t} produces $(x(t), -\dot{x}(t), \ddot{x}(t))$, so that we get

$$\phi_{-t}T(x, \dot{x}, \ddot{x}) = \phi_{-t}(x, -\dot{x}, \ddot{x}) = (x(t), -\dot{x}(t), \ddot{x}(t))\,. \tag{3.9}$$

Setting these two equations equal, we thus find that $T\phi_t = \phi_{-t}T$, which means that T is a representation of time reversal, $T\phi_t T^{-1} = \phi_{-t}$.

Thus, a Newtonian time reversal operator really does exist! Folk wisdom suggests that it is the identity or non-existent, and there is a grain of truth in that: the construction of T above is indeed the identity when restricted to the submanifold of particle positions $\mathbb{R}^{3n}$. However, time reversal is *not* the identity on true Newtonian state space, the jet space of higher

[13] For $x(t)$ to be a solution means that $\frac{d^2}{dt^2}x(t) = F\left(x(t), \frac{d}{dt}x(t)\right)$ for all $t \in \mathbb{R}$, and so it is true for each $-t \in \mathbb{R}$. Thus, $\frac{d^2}{dt^2}x(-t) = F\left(x(-t), -\frac{d}{dt}x(t)\right) = F\left(x(-t), \frac{d}{dt}x(t)\right)$, where the last equality applies the fact that F does not depend on the sign of velocity. Therefore $x(-t)$ is a solution as well.

derivatives $(x, \dot{x}, \ddot{x}) \in M$. This follows deductively from the fact that it is a representation of the reversal of time translations. So, when Callender (2000, p.254) proposes that time reversal does nothing but reverse the sign of t and "anything logically supervenient" upon it, we can agree! But, this does not mean that there is no instantaneous time reversal operator, as the Time Reflection Camp seems to suggest: on the contrary, it means that there must be.

How are we to interpret the time reversal transformation T on Newtonian state space? The Instantaneous Camp of Section 2.2 proposed to view it as part of a two-step description of time reversal, in which we first reverse time order and then reverse instants. I would not recommend that it be viewed this way either; at least, not at its foundation. Time reversal is fundamentally the reversal of time translations $\tau t \tau^{-1} = -t$, and the transformation T on state space is just its representative on state space, in that $T\phi_t T^{-1} = \phi_{-t}$. Nor is this somehow really 'motion reversal': since it is an element of a representation, which preserves the essential structure of time, the transformation T has all that is needed to justify viewing it as the reversal of temporal structure.

That said, we can still make sense of the Instantaneous Camp's proposal, as a shorthand way to answer a particular question:

Given that an initial state $s = (x, \dot{x}, \ddot{x})$ evolves along $x(t)$ according to Newton's equation, what is the solution associated with the 'reversed' initial state Ts?

The answer is: $x(-t)$, which is really $(Tx)(-t)$ together with the fact that $T(x) = x$. This follows immediately from our basic construction: a representation of time reversal must satisfy $T\phi_t T^{-1} = \phi_{-t}$, which is equivalent to $\phi_t T = T\phi_{-t}$; so the evolution $\phi_t(Ts)$ of the initial state Ts by ϕ_t can equivalently be written as the evolution $T\phi_{-t}(s)$ of the initial state s. Thus, Ts is associated with the solution, $(Tx)(-t) = x(-t)$. The controversy over whether Ts is really a 'time reversed' instantaneous state is, as far as I can tell, only an issue of terminology. One can call it 'motion reversal' or whatever one wishes. But, the basic interpretation of time reversal is still just as the reversal of time translations.

3.2.4 *Uniqueness of the Newtonian Time Reversal Operator*

When a representation of the time reversal operator exists, it is usually not unique. For example, in the case above where force does not depend on the sign of velocity, the transformation $\tilde{T}(x, \dot{x}, \ddot{x}) = (-x, \dot{x}, \ddot{x})$ provides another representation of time reversal, in that it implies that $\tilde{T}\phi_t\tilde{T}^{-1} = \phi_{-t}$.

It is instructive to see this in an example: consider the 'free' particle system for which $F = \mathbf{0}$, and where solutions to Newton's equation are curves of the form $x(t) = \dot{x}t + x$ and time translations can be written as $\varphi_t(x, \dot{x}, \ddot{x}) = (\dot{x}t + x, \dot{x}, \ddot{x})$ for all $t \in \mathbb{R}$. Then,

$$\begin{aligned}\tilde{T}\phi_t\tilde{T}^{-1}(x, \dot{x}, \ddot{x}) &= \tilde{T}\phi_t(-x, \dot{x}, \ddot{x}) = \tilde{T}(\dot{x}t - x, \dot{x}, \ddot{x}) = (-\dot{x}t + x, \dot{x}, \ddot{x}) \\ &= \phi_{-t}(x, \dot{x}, \ddot{x}),\end{aligned} \tag{3.10}$$

satisfying our requirement that $\tilde{T}$ is a representation of the temporal reflection $\tau t \tau^{-1} = -t$. So, there are multiple representatives of time reversal in this sense, which have the same effect of transforming curves, as $x(t) \mapsto x(-t)$.

However, the transformation $\tilde{T}$ also reverses 'spatial translations', in that if we define a translation in space L_a by the statement $L_a(x, \dot{x}, \ddot{x}) := (x + a, \dot{x}, \ddot{x})$ for each $a \in \mathbb{R}^{3n}$, then it follows[14] that $\tilde{T}L_a\tilde{T}^{-1} = L_{-a}$. That is what most would expect of 'space and time reversal', but not time reversal alone. So, there is reason to think that not every reversal of a curve $x(t) \mapsto x(-t)$ is 'really' time reversal. But, to distinguish time reversal from space-time reversal, it turns out that we must look beyond the folk wisdom that time reversal just 'reverses little-t', and make essential use of the time reversal transformation T on Newtonian state space.

In particular, the Galilei group is a common choice for the symmetries of pre-relativistic spacetime, and so we often suppose that we have a representation of it on Newtonian state space.[15] In this context, the transformation $\tilde{T}$ would normally be defined as a representation of the 'parity and time' transformation $p\tau$. Just as time reversal can be defined as the reversal of time translations, so p can be defined as the reversal of spatial translations, in the sense that $pap^{-1} = -a$ for each spatial translation s. In contrast, the time reversal group element τ commutes with each spatial translation, $\tau a \tau^{-1} = a$. This expresses the 'homogeneity' of time's direction across space: applying time reversal in different spatial locations produces exactly the same result.

Thus, in a representation of the Galilei group with $T = \phi_\tau$ representing time reversal and $L_s = \phi_s$ representing spatial translations, the homomorphism property implies that

$$TL_aT^{-1} = L_a. \tag{3.11}$$

The requirement of homogeneity expressed by Eq. (3.11) is satisfied by the ordinary time reversal transformation $T(x, \dot{x}, \ddot{x}) = (x, -\dot{x}, \ddot{x})$. But, it is not

14 Namely, $\tilde{T}L_a\tilde{T}^{-1}(x, \dot{x}, \ddot{x}) = \tilde{T}L_a(-x, \dot{x}, \ddot{x}) = \tilde{T}(-x + a, \dot{x}, \ddot{x}) = (x - a, \dot{x}, \ddot{x}) = L_{-a}(x, \dot{x}, \ddot{x})$.

15 For a discussion of the geometry underlying this, see Abraham and Marsden (1978, Theorem 5.4.21 and commentary thereafter) or Marle (1976); a philosophical discussion can be found in Belot (2000) and Earman (1989), or more recently Dewar (2022, Chapter 6).

https://doi.org/10.1017/9781009122139.004 Published online by Cambridge University Press

satisfied when T is replaced with the space-time reversal $\tilde{T}$. In this sense, T is the preferred time reversal transformation in Newtonian mechanics.[16] This technique for identifying the time reversal operator turns out to be remarkably general, and we will make use of it in other contexts below.

Here is a general lesson from our discussion: since a state space is a highly structured object, one should expect a representation of time translations to contain significantly more structure than the group of time translations itself. This, in the end, is where the instantaneous time reversal operator gets its 'bells and whistles': a representation $T = \phi_\tau$ of the time reversal group element τ is a highly structured thing. The same moral applies in nearly every approach to dynamical systems, and so I will appeal to similar arguments in the remainder of this chapter.

What the reader should *not* conclude from this is that Newtonian mechanics is necessarily time reversal invariant. Although we can always extend the time translation group $(\mathbb{R}, +)$ to one that includes time reversal, it is a significant assumption to say that a representation of time translations can be appropriately extended to this larger group too. As we will see in Chapter 4, the latter assumption can fail, even in the context of Newtonian mechanics.[17] Indeed, even more interesting things happen outside the austere world of Democritus, Bošković, and Boyle: if we allow matter to have intrinsic properties besides position (as even Aristotle did[18]), such as a fluid's instantaneous viscosity field in Navier–Stokes theory, then the state space becomes more interesting too. This allows for all sorts of possibilities regarding the failure of time reversal symmetry, even in Newtonian physics.

3.3 Analytic Mechanics

There are three great approaches to analytic mechanics: Hamiltonian, Lagrangian, and Hamilton–Jacobi. Folklore has it that these frameworks are all equivalent to Newtonian mechanics, and there are senses in which this is true.[19] But, each adopts a slightly different structure for state space. And, these frameworks are so general as to provide a framework for virtually every area of modern physics. So, to understand the meaning of time

[16] This suggests that a uniqueness theorem is possible in this context; however, stating this appears to be much simpler in the context of analytic mechanics, and so I reserve this for Section 3.3.
[17] Cf. Roberts (2013b).
[18] In *On Generation and Corruption* Book I, Part 8, 326a.
[19] Philosophical challenges to the folklore are given by Butterfield (2004), Curiel (2013), and North (2009). It has been defended by Barrett (2015, 2019).

reversal in these theories, they should really each be treated independently. As Butterfield observed,[20] the analytic framework,

> helps rebut the false idea that classical mechanics gives us a single matter-in-motion picture. . . . [T]hese equivalences are subtler than is suggested by textbook impressions, and folklore slogans like 'Lagrangian and Newtonian mechanics are equivalent'. (Butterfield 2004, p.29)

Nevertheless, to avoid dragging the reader through the derivation of time reversal in each of these sophisticated frameworks, I hope to be forgiven for analysing just one of them in detail, the symplectic formulation of Hamiltonian mechanics. I will then just give a few brief comments on how the analysis applies to Lagrangian mechanics in Section 3.3.4.[21]

3.3.1 State Space for Symplectic Mechanics

With suitable definitions, if $F = -\nabla U$ for some smooth function U of position alone, then with an appropriate definition of a smooth 'Hamiltonian' function h, Newton's equation is equivalent to *Hamilton's equations*,[22]

$$\frac{d}{dt}q_i(t) = \frac{\partial h}{\partial p_i} \qquad\qquad \frac{d}{dt}p_i(t) = -\frac{\partial h}{\partial q_i} \tag{3.12}$$

for each $i = 1, \ldots, n$ and for all $t \in \mathbb{R}$. Hamilton's equations are invariant under time translations, and so they provide a natural context for a representation of time translations. This can be stated most clearly in the language of symplectic mechanics.

One can view symplectic mechanics as a framework for producing Hamilton's equations in 'local' form by elevating time translation invariance to the status of an axiom. The state space is a pair (M, ω) called a *symplectic manifold*, where M is a smooth $2n$-dimensional real manifold, and ω is a closed, non-degenerate two-form on M called a *symplectic form*. The central axiom of the theory can be stated as follows:

Time evolution is along a vector field X that preserves the structure of state space.

[20] Butterfield, J. (2004). "Between laws and models: Some philosophical morals of Lagrangian mechanics". In: Unpublished manuscript, http://philsci-archive.pitt.edu/1937

[21] An even more general framework for Hamiltonian mechanics is Poisson mechanics, although I will not discuss this here; see Landsman (1998) for a treatment of mechanics in this framework.

[22] Namely, given a basis $(x_i, \dot{x}_i) \in \mathbb{R}^n \times \mathbb{R}^n$, define $q_i = x_i$ and $p_i = m_i \dot{x}_i$, and let $h : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a function is a sum of all the energy, both 'kinetic' and 'potential', $h(q_1, \ldots, q_n, p_1, \ldots, p_n) := \sum_{i=1}^n \left(\frac{1}{2m_i} p_i \cdot p_i\right) + U(q_1, \ldots, q_n)$. Now check that $-\nabla U = m_i \ddot{x}_i$ is equivalent to Hamilton's equations.

This axiom can be interpreted to mean that X is a smooth vector field with Lie derivative satisfying $\mathcal{L}_X\omega = 0$, called a *symplectic vector field*. That is equivalent[23] to the statement that $\iota_X\omega$ is a closed one-form, which implies that it is 'locally' exact by the Poincaré lemma: in a neighbourhood of every point,

$$\iota_X\omega = dh \tag{3.13}$$

for some smooth function $h : M \to \mathbb{R}$. We can now confirm that Eq. (3.13) is a local geometric expression of Hamilton's equations by an application of Darboux's theorem, which ensures that there is a local coordinate system $(q_1, \ldots, q_n, p_1, \ldots, p_n)$ in which $\omega = \begin{pmatrix} & \mathbb{1} \\ -\mathbb{1} & \end{pmatrix}$, which reproduces[24] the Eqs. (3.12). However, the framework of symplectic mechanics is much more general: (M, ω) can be any symplectic manifold, and a state $s \in M$ can be used to represent a much more general state of affairs than Boyle or Bošković intended for Newtonian mechanics.

We begin by reviewing the nature of time translations in this framework. A symplectic vector field X is threaded by a one-parameter set of diffeomorphisms called the *symplectic flow* $t \mapsto \phi_t$ along X. In general, an automorphism of (M, ω) is a diffeomorphism that either preserves ω or reverses its sign; the latter possibility arises because the symplectic form introduces an orientation on M, which is arbitrary from a physical perspective. The former is called a *symplectomorphism* and the latter an *anti-symplectomorphism*.

By construction, each ϕ_t in a symplectic flow is an symplectomorphism such that $\phi_{t_1}\phi_{t_2} = \phi_{t_1+t_2}$. Thus, $t \mapsto \phi_t$ forms a representation of the Lie group of time translations. Its local Lie algebra generator in the representation is the smooth function h, called the 'Hamiltonian' or 'energy' of the representation. Following the discussion of Section 3.1.3, we will require that the Hamiltonian h be half-bounded, meaning that there exists some fixed lower bound $b \in \mathbb{R}$ such that $h(s) \geq b$ for all $s \in M$, but no such upper bound.

3.3.2 Time Reversal in Symplectic Mechanics

Let $t \mapsto \phi_t$ be a representation of time translations amongst the symplectomorphisms of (M, ω). In other words, ϕ_t is a symplectic flow that threads

23 Cartan's 'magic formula' for a two-form ω and a vector field X states, $\mathcal{L}_X\omega = d\iota_X\omega + \iota_X d\omega$. But $d\omega = 0$, so $\mathcal{L}_X\omega = d\iota_X\omega$. Hence, $\mathcal{L}_X\omega = 0$ if and only if $d\iota_X\omega = 0$, where the latter says $\iota_X\omega$ is closed.

24 That is, writing an integral curve of X as $q_i(t), p_i(t)$ so that $X = \left(\frac{d}{dt}q_1(t), \ldots, \frac{d}{dt}p_n(t)\right)$. Thus, since for smooth functions $dh = \nabla h = \left(\frac{\partial h}{\partial q_1}, \ldots, \frac{\partial h}{\partial p_n}\right)$, Eq. (3.13) produces Hamilton's equations.

a symplectic vector field X. Let h denote a half-bounded local generator of ϕ_t in some neighbourhood. Extending the time translations to a group that includes a time reversal element τ, suppose that an extension of the representation ϕ exists too. Then, the *time reversal transformation in symplectic mechanics* is just the representative $T := \phi_\tau$ of time reversal on state space. As before, $\tau t \tau^{-1} = -t$ implies that $T\phi_t T^{-1} = \phi_{-t}$, since a representation is a homomorphism.

To verify that this notion of time reversal transforms dynamical trajectories in an appropriate way, let $s \in M$ be an initial state, and let us write $t \mapsto s(t) = \phi_t s$ to denote a curve representing the time evolution that begins at s. The application of time reversal to each time translation $T\phi_t T^{-1} = \phi_{-t}$ implies that the curve $s(t) = \phi_s(t)$ is transformed to $s(-t) = \phi_{-t}s$, and so time reversal induces a transformation on curves of the form

$$s(t) \mapsto s(-t). \tag{3.14}$$

However, as before, more structure is needed in order to determine what this transformation T is like.

By adopting the (Darboux) coordinates of Hamilton's equations (3.12), one might guess that T should be defined by $T : (q, p) \mapsto (q, -p)$, so as to 'reverse instantaneous momentum'. Writing the symplectic form in terms of the wedge product as $\omega = dp \wedge dq$, this would imply that time reversal also reverses the symplectic form $\omega \mapsto -\omega$, and so is an anti-symplectomorphism.[25] This is indeed the right result. Unfortunately, in symplectic mechanics we do not always have assurance that the coordinates (q, p) represent physical 'position' and 'momentum', respectively. Indeed, if one coordinate system (q, p) happens to represent position and momentum, then there typically remain many coordinate systems (Q, P) related to it by a symplectomorphism, which thus satisfy Hamilton's equations, but which do not represent position and momentum. So, in symplectic mechanics, we will need a more general perspective on the meaning of the time reversal operator.

Happily, we have already done the work of building that general perspective and can prove the general result that time reversal is antisymplectic without this appeal to coordinate systems. Our only substantial assumption will be that the energy associated with time translations is half-bounded. And, to preserve the local structure of symplectic mechanics, we will take our time translations to be associated with a *local* Lie group, defined by the $(\mathbb{R}, +)$ in some neighbourhood of the identity. Since an automorphism of

[25] This is analogous to an antiunitary time reversal operator in quantum theory; see Section 3.4.

a symplectic manifold (M, ω) is by definition either symplectic or antisymplectic (as discussed in the end of Section 3.3.1), it will be enough to show that time reversal cannot be symplectic. Here is the result.

> **Proposition 3.1** *Let $t \mapsto \phi_t$ be a representation of a local Lie group of $(\mathbb{R}, +)$ in some neighbourhood of the identity, amongst the symplectic and antisymplectic transformations of (M, ω), with h a half-bounded generator. If the representation extends to one that includes a time reversal operator $T := \phi_\tau$ such that $T\phi_t T^{-1} = \phi_{-t}$, then T is not symplectic, and so it must be antisymplectic.*

Proof Assume for reductio that T is symplectic, and let $s(t) := \phi_t s$ for some $s \in M$. Then $s(t)$ is an integral curve with tangent vector field X_h, where h is the half-bounded local generator of time translations. We assumed $T \circ \phi_t \circ T^{-1} = \phi_{-t}$, which is equivalent to $\phi_t \circ T = T \circ \phi_{-t}$ using the fact that $T = T^{-1}$. This implies that $T \circ s(-t) = T \circ \phi_{-t}s = \varphi_t \circ Ts$. So, $(Ts)(-t) = Ts(-t)$ is an integral curve of X_h. Moreover, by Hamilton's equations, $s(-t)$ has a Hamiltonian vector field given by $-X_h = X_{-h}$. Combining these two facts implies that

$$X_h = T_* X_{-h}, \tag{3.15}$$

where T_* is the push-forward of T on vector fields. But we have assumed T is symplectic, so Jacobi's theorem can be applied (Abraham and Marsden 1978, Theorem 3.3.19), which says that for symplectic maps, Eq. (3.15) is true if and only if $X_h = X_{-h \circ T}$. Therefore: $h(x) = -h \circ T(x) + c$ for some $c \in \mathbb{R}$ and for all $x \in M$. But h is half-bounded, so we can write $m \leq h \circ T(x)$ for all $x \in M$, which is equivalent to: $-h \circ T(x) + c \leq m + c$. Combining these two thus entails

$$h(x) = -h \circ T(x) + c \leq m + c$$

for all $x \in M$, contradicting the assumption that h is unbounded from above. ∎

One immediate corollary is that, like in Newtonian mechanics, time reversal cannot be the identity transformation: if a representation of time reversal exists, then it must be antisymplectic, whereas the identity is symplectic.

Another interesting application is in electromagnetism. Let $(\mathbb{R}^4, g_{ab})$ be a relativistic spacetime, where g_{ab} is a Lorentzian metric. A two-form F on $\mathbb{R}^4$ that is closed, $dF = 0$, is called a *Maxwell–Faraday field* and is a common way to represent the electromagnetic field. To analyse this system in symplectic mechanics, let $M = T^*\mathbb{R}^4$ be the cotangent bundle over $\mathbb{R}^4$, whose canonical coordinates can be written (q, p) with $q \in \mathbb{R}^4$ representing position in spacetime and p a one-form on $\mathbb{R}^4$ at the point q. Let θ be the canonical one-form on T^*M, and let $\omega_0 = d\theta$ be its canonical symplectic form.[26] Using the projection $\pi : T^*M \to M$, we can pull F back to a two-form on T^*M, which we will again denote by F. Now, the two-form defined by

$$\omega_F := \omega_0 + eF, \tag{3.16}$$

where $e \in \mathbb{R}$ represents 'electric charge', is again a symplectic form; and, the symplectic flow on (M, ω_F) generated by the (half-bounded) free Hamiltonian $h(q, p) := \frac{1}{2}g^{ab}p_a p_b$ gives rise to the ordinary equations of motion for a charged particle in an electromagnetic field.[27]

What is the effect of time reversal on this system? From Proposition 3.1, we know that its representative T on the symplectic manifold (M, ω_F) is antisymplectic, and so reverses the sign of ω_F. If we suppose also that $T : (q, p) \mapsto (q, -p)$, given that we know the coordinates we have adopted represent a particle's position and momentum, then the one-form θ reverses sign as well, and so $\omega_0 = d\theta$ does too. This implies that T reverses the electromagnetic field, $F \mapsto -F$, since $F = (\omega_F - \omega_0)/e$. Similarly, given an electromagnetic 'four-potential', which is a one-form A satisfying $F = dA$, it follows from this that $A \mapsto -A$. This is a different way of getting to the account of Malament (2004), who instead uses the reversal of a temporal orientation to define time reversal (see Section 2.5.3) and then observes that this induces the transformation $F \mapsto -F$.

3.3.3 Uniqueness of the Hamiltonian Time Reversal Operator

Like in Newtonian mechanics, there are generally many representations of time reversal on a symplectic manifold: if T satisfies $T\phi_t T^{-1} = \phi_{-t}$, and if α is any symplectomorphism such that $\alpha\varphi_t = \varphi_t\alpha$ for all t (called an 'integral' of time translations), then one can check[28] that this is also satisfied by $\tilde{T} := \alpha \circ T$. But again, the broader context of a symmetry group will often help to determine the meaning of T. Let me indicate one scenario in which a uniqueness result can be obtained.

As discussed in Section 3.2.4, the structure of the Galilei group (and the Lorentz group for that matter) guarantees that time reversal, in addition

[26] The *canonical one-form* θ on a cotangent bundle T^*M is a standard construction given in coordinates (q, p) by $\sum_i p_i dq^i$. It has the property that $\omega = -d\theta$ is a symplectic form, called the *canonical symplectic form* (cf. Arnol'd 1989, §37).

[27] Cf. Guillemin and Sternberg (1984, p.140).

[28] Namely, $\tilde{T}\phi_t\tilde{T}^{-1} = \alpha(T\phi_t T^{-1})\alpha^{-1} = \alpha(\phi_{-t})\alpha^{-1} = \phi_{-t}\alpha\alpha^{-1} = \phi_{-t}$.

https://doi.org/10.1017/9781009122139.004 Published online by Cambridge University Press

to reversing time translations $T\varphi_t T^{-1} = \varphi_t$, also commutes with spatial translations, $TL_sT^{-1} = L_s$. Moreover, time reversal transforms a velocity boost to the reverse boost, in that $TB_vT^{-1} = B_{-v}$, capturing an intuitive sense in which it reverses velocities. Time reversal similarly reverses rotations, $TR_\theta T^{-1} = R_{-\theta}$. These transformations are the 'generating elements' of the Galilei group, in that all its elements are compositions of them. So, in summary, we have that for each generating element F of the Galilei group, time reversal either commutes with the element, $TFT^{-1} = F$, or maps it to its inverse, $TFT^{-1} = F^{-1}$.

Suppose now that we are concerned with a linear symplectic manifold (M, ω), meaning that the manifold M is also a vector space, for example in the context of a classical field theory (cf. Section 8.2.2). Suppose further that we have a representation of the Galilei group that is *irreducible*, in the sense that it is a non-trivial representation with no non-trivial subspaces that are also representations. One often uses this condition to characterise 'elementary' systems that cannot be decomposed into further component parts. And, when this is the case, the properties above guarantee that T is unique, up to a multiplicative constant.

This follows immediately from one of the cornerstones of linear representation theory known as *Schur's lemma*, which says that if a transformation K commutes with every element of a linear representation, then it must be a constant multiple of the identity (cf. Blank, Exner, and Havlíček 2008, Theorem 6.7.1). In particular, suppose that given an irreducible representation, there are two transformations T and $\tilde{T}$ that both either commute with each generating element F, or map it to its inverse. Then $T\tilde{T}$ commutes with each generating element in the representation. Thus, $T\tilde{T} = c$ for some constant by Schur's lemma, and so applying the fact that $T = T^{-1}$, we get that $\tilde{T} = cT$. We can summarise this general strategy for determining uniqueness in the following.

Proposition 3.2 *Let $T = T^{-1}$ be a linear bijection such that, for all the generating elements F of an irreducible linear representation, either $TFT^{-1} = F$ or $TFT^{-1} = F^{-1}$. Then T is the unique transformation with these properties up to a multiplicative constant, in that any other $\tilde{T}$ with these properties satisfies $\tilde{T} = cT$ for some constant c.*

As one might expect, this technique finds its most natural home in theories with a built-in linear structure, such as quantum theory. Time reversal in that context is the subject of Section 3.4. But before we turn to the *verdammten Quantenspringerei*, let me offer a few brief remarks on time reversal in Lagrangian mechanics.

3.3.4 Lagrangian Mechanics

Lagrangian mechanics has an elegant geometric formulation due to Klein (1962), which makes its state space structure particularly clear.[29] This formulation reveals that much of Lagrangian mechanics is in fact just a special case of symplectic mechanics. So, let me conclude our discussion of analytic mechanics with a sketch of this fact, which allows one to analyse time reversal using the results above.

Lagrangian mechanics is formulated on the tangent bundle TM of a smooth real manifold M. A point in TM is often written $s = (x, \dot{x})$, with $x \in M$ and $\dot{x}$ a vector at x, representing the 'configuration' and 'velocity' of a physical system at an instant. Motion in this framework is given by a solution to the *Euler–Lagrange equations*,

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}_i} - \frac{\partial L}{\partial x_i} = 0 \tag{3.17}$$

for each $i = 1, 2, \ldots, n$ and for all $t \in \mathbb{R}$.

The state space structure underpinning this framework can often be formulated as a symplectic manifold, and in this sense it is a special case of symplectic mechanics. This applies specifically to systems with a 'regular' Lagrangian, or one for which the Hessian is invertible.[30] For each smooth Lagrangian $L : TM \to \mathbb{R}$, there exists a *canonical closed two-form* ω_L on TM, which can be 'pulled back' from the canonical symplectic form ω_0 on T^*M, its 'partner' cotangent bundle; this two-form turns out to be a symplectic form if and only if the Lagrangian is regular.[31] There is also a *canonical energy function* $h_L : TM \to \mathbb{R}$ defined[32] by L. For regular Lagrangians, the Euler–Lagrange equations can then be expressed in the same form as Hamilton's equations, but on the manifold TM:

$$\iota_X \omega_L = dh_L. \tag{3.18}$$

In coordinates $(x, \dot{x})$, this reduces to the familiar expression of the Euler–Lagrange equations above (De León and Rodrigues 1989, p.304). And, just

[29] For details on Klein's approach, see Nester (1988) and Curiel (2013), De León and Rodrigues (1989, §7), or Woodhouse (1991, §2).
[30] This condition is required for the existence of the Legendre transformation; for details, see Abraham and Marsden (1978, §3.5). An excellent analysis of the relationship between Hamiltonian and Lagrangian structures can be found in Barrett (2015, 2019).
[31] See Abraham and Marsden (1978, Proposition 3.5.9) or De León and Rodrigues (1989, §7.1).
[32] This makes use of canonical vertical (or 'Liouville') vector field V on TM; it is defined in the standard way by $h_L := V(dL) - L$ (cf. Nester 1988).

as in Hamiltonian mechanics, the time translations here are given by the symplectic flow $t \mapsto \phi_t$ on TM that 'threads' the symplectic vector field X.

The state space for 'regular' Lagrangian mechanics can thus be viewed as a symplectic manifold (TM, ω_L). This means that these models of Lagrangian mechanics are not just equivalent to models of Hamiltonian mechanics: the former are a subset of the latter![33] As a result, the conclusion of Proposition 3.1 applies to this sector of Lagrangian mechanics as well, that time reversal is represented by an antisymplectic transformation T. Of course, this conclusion does not apply to the interesting case of non-regular or *singular* Lagrangians, which are ubiquitous in gauge physics and for which the two-form ω_L in Eq. (3.18) is not symplectic. However, nothing prevents the analysis of time reversal on the Representation View from being carried out in this more general context, as a transformation that reverses time translations associated with singular Lagrangians. I leave this analysis as an invitation to the reader.

3.4 Quantum Theories

3.4.1 Two Camps on Hilbert Space Quantum Theory

Recall that in Chapter 2, we discussed two camps regarding the meaning of time reversal. The Time Reflection Camp argued that time reversal has the form $t \mapsto -t$, while the Instantaneous Camp argued that time reversal involves this reflection plus a richly structured transformation T on state space (Section 2.2). My thesis was that, once we adopt the Representation View, we find a sense in which both these camps are correct: time reversal does transform time translations as $\tau : t \mapsto -t$; and, its representative on state space is a richly structured operator T, of the kind we have now seen in the previous sections. Since a significant portion of this debate has taken place in the context of Hilbert space quantum theory, it is worth briefly reviewing what they say in this context, before we turn to the new derivation of the time reversal operator that the Representation View affords.

Modern quantum theory on Hilbert space was set out by Von Neumann (1932), who took experimental outcomes to be represented by elements of a lattice of closed subspaces of a separable Hilbert space $\mathcal{H}$, or equivalently, by the lattice $L(\mathcal{H})$ of projections onto those subspaces. Given a suitably

33 Compare this to Barrett (2018a).

https://doi.org/10.1017/9781009122139.004 Published online by Cambridge University Press

Unitary Operator U	Antiunitary Operator A
(U1) $UU^* = U^*U = I$ (U2) $U(a\psi + b\phi) = aU\psi + bU\phi$	(A1) $AA^* = A^*A = I$ (A2) $A(a\psi + b\phi) = a^*A\psi + b^*A\phi$
(U3) $\langle U\psi, U\phi\rangle = \langle\psi,\phi\rangle$	(A3) $\langle A\psi, A\phi\rangle = \langle\psi,\phi\rangle^*$

Figure 3.5 Defining properties of unitary and antiunitary operators.

generalised notion of a 'probability measure' p on a suitable lattice, one can always[34] find a density operator ρ that allows p to be expressed in canonical form (the 'Born rule'):

$$p(F) = \mathrm{Tr}(\rho F) \tag{3.19}$$

for each projection $F \in L(\mathcal{H})$.

Symmetry transformations in quantum theory are represented by unitary or antiunitary operators (defined by the properties in Figure 3.5), for reasons that I will explain shortly. Members of the Instantaneous Camp, which has included Earman (2002b) and myself (Roberts 2017), defend the standard practice of including an antiunitary time reversal operator, following the definition set out by Wigner (1931).

The statement that time reversal is antiunitary gives rise to the 'complex conjugation' aspect of time reversal: for example, in the Schrödinger representation on $\mathcal{H} = L^2(\mathbb{R})$ with Q defined by $Q\psi := x\psi$ (for all vectors ψ in its domain D_Q), with Q interpreted as a 'position observable', the instantaneous time reversal operator $T = K$ is the 'conjugation operator' on wavefunctions, defined by $K\psi := \psi^*$ for all $\psi \in L^2(\mathbb{R})$. One can check that this T is antiunitary.[35] We thus take time reversal to transform a trajectory $\psi(t)$ to $T\psi(-t)$. In contrast, as a member of the Time Reflection Camp, Callender (2000) refers to this T as "Wigner reversal" and argues that 'true' time reversal in quantum theory is just the transformation $\psi(t) \mapsto \psi(-t)$, suggesting that T is actually identity transformation, which is unitary.

[34] This is the content of Gleason's theorem. Birkhoff and von Neumann (1936) pointed out that 'probability' in quantum theory cannot be a probability in the ordinary mathematical sense of a bounded measure on a σ-algebra because the distributive axiom is violated in the lattice $L(\mathcal{H})$. A common response is to propose a more general logic, such as the lattice of Hilbert space projections: see Jauch (cf. 1968) for a classic treatment and Rédei (1996, 1998) for a philosophical perspective; a more general operational perspective can be found in Busch, Grabowski, and Lahti (1995) and in Landsman (2017).

[35] The inner product in the Schrödinger representation is $\langle\psi,\phi\rangle := \int_{\mathbb{R}^n} \psi(x)^*\phi(x)dx$, which implies by the linearity of complex conjugation that $\langle\psi^*,\phi^*\rangle = \int_{\mathbb{R}^n} \psi(x)^{**}\phi(x)^*dx = \int_{\mathbb{R}^n}(\psi(x)^*\phi(x))^*dx = \langle\psi,\phi\rangle^*$.

I have shown elsewhere on the basis of some well-loved adequacy conditions that the time reversal operator must be antiunitary (Roberts 2017, Proposition 1). That result is the following.

Proposition 3.3 *Let T be a unitary or antiunitary bijection on a separable Hilbert space $\mathcal{H}$. Suppose there is at least one densely-defined self-adjoint operator H on $\mathcal{H}$ that satisfies the following conditions.*

i) (positive energy) $0 \leq \langle\psi, H\psi\rangle$ for all ψ in the domain of H.
ii) (nontrivial) H is not the zero operator.
iii) (invariance) $Te^{itH}\psi = e^{-itH}T\psi$ for all ψ.

Then T is antiunitary.

However, this result did not convince the Time Reflection Camp. Callender (Forthcoming) has responded by rejecting these adequacy conditions. He writes:

Quantum textbooks sometimes address this point and claim that in quantum mechanics time reversal invariance is to be given by *two* operations, a temporal reflection and the operation of complex conjugation $K : \psi \to \psi^*$. This idea can be traced back to Wigner . . . . Roberts (2017) shows that if one assumes that there exists at least one non-trivial time reversal invariant quantum physical system then Wigner's operation follows. But in this context this assumption is a large one for it's up in the air whether quantum mechanics is time reversal invariant. . . . Call it what you like, Wigner's reversal is different from a temporal reflection. Taking the complex conjugation of a state *doesn't follow by logic or definition alone* from a temporal reflection. (Callender Forthcoming, pp.9–10)

Callender has given a reasonable reply. So, let me try to strengthen mine. After a brief primer on symmetries in quantum theory (Section 3.4.2), I will argue that Callender's conclusion that "Wigner's reversal is different from temporal reflection" is too quick: Wigner's antiunitary time reversal transformation is nothing more than the representation of temporal reflection $\tau : t \mapsto -t$, where each t is a time translation. As a result, effects like the conjugation of wavefunctions really do follow from a temporal reflection – indeed, by logic and definition alone.

3.4.2 Symmetries of Quantum Theory

As in earlier sections, we begin with the automorphisms of our state space. Experimental outcomes that cannot both occur at once, such as 'z-spin up' and 'z-spin down' in a Stern–Gerlach apparatus, are represented in quantum theory by projections $E, E' \in L(\mathcal{H})$ that are *orthogonal*, $E \perp E' = 0$,

https://doi.org/10.1017/9781009122139.004 Published online by Cambridge University Press

meaning that they project onto orthogonal subspaces. An *automorphism* of a Hilbert space lattice $L(\mathcal{H})$ is a bijection $\mathbf{U}$ on one-dimensional projections that preserves orthogonality:

$$\mathbf{U}(E)\perp\mathbf{U}(E') \text{ if and only if } E\perp E'. \tag{3.20}$$

The group of automorphisms Aut $L(\mathcal{H})$ provides a sensible definition of the symmetries in quantum theory: they preserve the facts about which outcomes can and cannot occur together. They are also strikingly classified by *Uhlhorn's theorem*, which assures us that every automorphism $\mathbf{U} \in \text{Aut}\, L(\mathcal{H})$ implementable by a unique Hilbert space operator U (in that $U\psi \in \mathbf{U}(E)$ if and only if $\psi \in E$) is either unitary or antiunitary (Uhlhorn 1963). Uhlhorn's theorem is a more precise and powerful expression of what is commonly called *Wigner's theorem* in quantum theory (cf. Bargmann 1964).

A representation of time translations in quantum theory is thus a map from the group of time translations $(\mathbb{R}, +)$ to the group of unitary and antiunitary operators on a Hilbert space. In quantum theory, it is entirely standard practice to view time evolution in this way: we begin with a strongly continuous representation of $(\mathbb{R}, +)$ amongst the unitary or antiunitary operators, and from this derive the Schrödinger equation. To show this, we first note that the representation must in fact be entirely unitary, since for each $t \in \mathbb{R}$ we have that $U_t = U_{t/2}U_{t/2}$, which produces a unitary operator regardless of whether $U_{t/2}$ is unitary or antiunitary. We can then apply Stone's theorem (Blank, Exner, and Havlíček 2008, Theorem 5.9.2), which says that for such a representation there exists a unique densely-defined self-adjoint operator H such that $U_t = e^{-itH}$ for all $t \in \mathbb{R}$. This leads immediately to the Schrödinger equation, by defining $\psi(t) := e^{-itH}\psi$ for some $\psi \in \mathcal{H}$ and taking derivatives of both sides.[36] So, on this standard reading of time in quantum theory, the 'little t' parameter in the Schrödinger equation does not represent a time coordinate, but a time translation, just as in my general proposal of Chapter 2.

3.4.3 Time Reversal in Quantum Theory

We now have the tools to see where Wigner's antiunitary time reversal operator comes from. Let $(\mathbb{R}, +)$ be the group of time translations; as I have shown in Section 2.6, it can always be extended using a semidirect product to a group that includes a time reversal element τ satisfying $\tau t \tau^{-1} = -t$,

[36] Namely, $\frac{d}{dt}\psi(t) = -iHe^{-itH}\psi = -iH\psi(t)$. A more detailed treatment of the rationale for the Schrödinger equation is Jauch (1968, §10-1 and 10-2) and Landsman (2017, §5.12).

for each time translation $t \in \mathbb{R}$. In any representation ϕ of this larger group amongst the unitary and antiunitary operators on a Hilbert space, there is no mystery about where Wigner's time reversal operator comes from: we define it to be the representative of the temporal reflection τ, in that,

$$T := \varphi_\tau. \tag{3.21}$$

In other words, Wigner's time reversal operator is no different than a temporal reflection: these two have the very same effect on time, guaranteed by the fact that the representation ϕ is a ('structure-preserving') homomorphism.

No further assumptions about the nature of time reversal are needed. We will only restrict our attention to a certain very large class of quantum theories, in which energy is bounded from below but not from above. But, this is an assumption about time *translations*, and not about time reversal. On this basis alone, it is possible to prove that temporal reflection, together with logic and definition alone, give rise to Wigner's antiunitary time reversal operator. Let me state the formal proposition first, before turning to its interpretation.

Proposition 3.4 *Let $t \mapsto U_t$ be a strongly continuous unitary representation from the time translation group $(\mathbb{R}, +)$ to the automorphisms of a separable Hilbert space $\mathcal{H}$, with a half-bounded generator H. Let G be the extension of this group to include time reversal τ satisfying $\tau t \tau^{-1} = -t$. Then:*

1. *the representation of $(\mathbb{R}, +)$ extends to a representation of G, with the representative $\tau \mapsto T$ of time reversal satisfying $TU_tT^{-1} = U_{-t}$; and*
2. *in every such representation, T must be antiunitary.*

Proof To prove that such an extension exists, let H be the self-adjoint generator satisfying $U_t = e^{-itH}$, and let $\mathrm{sp}(H) = \Lambda \subseteq \mathbb{R}$ be its spectrum. Let H_s be its spectral representation on $L^2(\Lambda)$, meaning that $H_s\psi(x) = x$ for all ψ in its domain, and $VHV^{-1} = H_s$ for some unitary $V : \mathcal{H} \to L^2(\Lambda)$ (cf. Blank, Exner, and Havlíček 2008, §5.8). If K is the conjugation operator on $L^2(\Lambda)$, meaning $K\psi = \psi^*$ for all $\psi \in L^2(\Lambda)$, then $[K, H_s] = 0$, since for all ψ in the domain of H_s we have $KH_sK^{-1}\psi(x) = x\psi(x) = H_s\psi(x)$. Thus, $T := V^{-1}KV$ is the desired antiunitary operator, since our definitions imply that $[T, H] = 0$, and hence $T\mathcal{U}_tT^{-1} = e^{T(-itH)T^{-1}} = e^{itTHT^{-1}} = e^{itH} = U_{-t}$.

To prove that every automorphism T (a unitary or antiunitary by Uhlhorn's theorem) in such a representation must be antiunitarity, suppose for reductio that T is unitary. Since $TU_tT^{-1} = U_{-t}$, we have that,

$$e^{itH} = Te^{-itH}T^{-1} = e^{T(-itH)T^{-1}} = e^{-itTHT^{-1}}, \tag{3.22}$$

https://doi.org/10.1017/9781009122139.004 Published online by Cambridge University Press

where the last equality applies unitarity. By Stone's theorem, the generator of the unitary group is unique, so $-H = THT^{-1}$, and hence H and $-H$ have the same spectrum. But since H is bounded from below, $m \leq \text{sp}(H) = \text{sp}(-H) = -\text{sp}(H) \leq -m$, contradicting the assumption that sp(H) is unbounded from above. ∎

I hope this dissolves the remaining mystery surrounding Wigner's antiunitary time reversal operator. Quantum theory is a dynamical theory, and so like any dynamical theory, it admits a representation of time translations with half-bounded energy. Whenever this is the case, Proposition 3.4 shows that time translations extend to include time reversal $\tau : t \mapsto -t$, and, Wigner's time reversal operator is nothing more than its representative on state space, which is always antiunitary. Although a 'unitary time reversal' operator is sometimes associated with the work of Racah (1937), this proposition implies that no such unitary operator reverses time translations in a theory with half-bounded energy.[37]

The Temporal Reflection Camp can also rest reassured that this 'instantaneous' time reversal operator need not be interpreted as 'reversing instants'. Fundamentally, Wigner's time reversal operator is just the representative of a temporal reflection, which has the property that $\tau t \tau^{-1} = -t$, and therefore,

$$TU_tT^{-1} = U_{-t}. \tag{3.23}$$

This T is no more 'instantaneous' than time translations are, in that both are defined as operators on 'instantaneous' state space.

What then of the apparent textbook application of *two* operations, T and $t \mapsto -t$, instead of just $t \mapsto -t$? This is nothing more than a shorthand way to answer the following question:

Given a solution to the Schrödinger equation $\psi(t) := U_t\psi$ with initial state ψ, what is the solution associated with the 'time-reversed' initial state $T\psi$?

The answer is: $T\psi(-t)$. This follows because Eq. (3.23) implies that $U_tT\psi = TU_{-t}\psi = T\psi(-t)$; so, the unitary dynamics U_t starting with $T\psi$ is given by $T\psi(-t)$. In a misleading sense, this gives the appearance of two operations, $\tau : t \mapsto -t$ and $T : \psi \mapsto T\psi$. But, this is not the origin of the time reversal operator. Time reversal is simply a temporal reflection,

[37] This was confirmed explicitly for quantum electrodynamics by Jauch and Rohrlich (1976, pp.88–9). Costa de Beauregard (1980) proposes to restore a unitary reversal of time translations, but at the cost of introducing unbounded negative energy into the theory.

which reverses the group of time translations, $U_t \mapsto U_{-t}$. The antiunitary instantaneous time reversal operator T is the operator that implements time reflection, as in Eq. (3.23).

What about Callender's charge that we have assumed time reversal invariance? Here there is a subtlety: Eq. (3.22) does indeed give a standard expression of what it means for a quantum system to be time reversal invariant, which we discuss more in Chapter 4. However, this does not mean that we have assumed time reversal invariance. Although we can always extend time translations to include a time reversal group element, as described in Section 2.6, this does not necessarily mean that we can extend the *representation* of time translations to an appropriate representation of time reversal. If such an extension does exist, then the second part of Proposition 3.4 shows that it must be antiunitary. But, an appropriate representation might fail to exist, in which case we say that a quantum system violates time reversal invariance. I will return to the discussion of time reversal invariance and time reversal violation in Chapter 4.

Note that I am making careful use of the word 'appropriate' when I say that an appropriate representation of time reversal might not exist. Proposition 3.4 is perhaps surprising because the first part shows that, in quantum theory, a representation of time reversal – and hence, of time reversal symmetry – always does exist! This expresses a certain sense in which quantum theory is always symmetric in time, which I will discuss in more detail in Chapter 8. However, it is not always appropriate to call the resulting representative T the 'time reversal operator': there are in general many representatives of time reversal, such as parity–time reversal, just as in the case of classical mechanics (Sections 3.2.4 and 3.3.2).

Fortunately, as in our discussions of classical mechanics, considerations of the more general spacetime symmetry group help to uniquely determine which antiunitary time reversal operator is the 'appropriate' one for time reversal, through the application of Proposition 3.2.

For example, we might adopt the Galilei group, or the Lorentz group, as a more complete group G containing time translations and consider its representation among the automorphisms of a Hilbert space. This representation will be irreducible whenever it describes an 'elementary' system, interpreted as one that cannot be decomposed into component parts. But, time reversal transforms each generating elements of this group to itself or to its inverse, as discussed in Section 3.3.2, and so by Schur's lemma, it follows that any two representatives of time reversal T and $\tilde{T}$ must be related by a multiplicative constant. It is for this reason that in many concrete applications, the choice is uniquely determined as to which antiunitary time reversal operator is appropriate.[38]

3.4.4 The Time Reversal of Spin

A first topic in many quantum mechanics textbooks is the analysis of *spin*, a degree of freedom associated with a system's total angular momentum.[39] Spin-1/2 systems are often studied on a two-dimensional Hilbert space, with an algebra of observables generated by the Pauli 'spin observables',

$$\sigma_x = \begin{pmatrix} & 1 \\ 1 & \end{pmatrix}, \qquad \sigma_y = \begin{pmatrix} & -i \\ i & \end{pmatrix}, \qquad \sigma_z = \begin{pmatrix} 1 & \\ & -1 \end{pmatrix}.$$

Their eigenvectors are usually interpreted as 'spin-up' and 'spin-down' states of the system with respect to three spatial axes x, y, z. The algebraic relations that these observables satisfy are called the 'Pauli relations'.

Time reversal is usually assumed to reverse the sign of the Pauli spin observables, $T\sigma_j T^{-1} = -\sigma_j$ for each $j = x, y, z$. A common explanation of this is that spin is a 'kind' of angular momentum: it is intrinsic angular momentum that does not correspond to rotation in space, but is angular momentum nevertheless. So, the argument goes, since angular momentum is reversed by time reversal, spin must be reversed as well.[40] Of course, one might still wonder *why* a kind of angular momentum that does not actually 'rotate' anything in space must change sign under time reversal.

The Representation View has something to say about this too. To see it, we will first need to get a better grip on how the Pauli spin observables are related to spacetime. Define the 'spin rotations' through $\theta \in (0, 2\pi]$ about each axis $j = x, y, z$ by

$$R_j(\theta) := e^{-(i/2)\theta\sigma_j}. \tag{3.24}$$

Each rotation can be written in explicit form as $R_j(\theta) = \cos(\theta/2)I + \sigma_j \sin(\theta/2)$, where I is the identity operator. The group given by the closure of all these rotations under multiplication is isomorphic to a Lie group called $SU(2)$, and which is *not* isomorphic to the spatial rotation group $SO(3)$. Famously, $SU(2)$ has the unusual property that $R_j(2\pi) = -I$, whereas for

[38] This observation was made explicit in Roberts (2017, Propositions 2 and 3).
[39] See Jauch (1968, Chapter 14) for an elegant introduction.
[40] Cf. Sachs (1987, p.34).

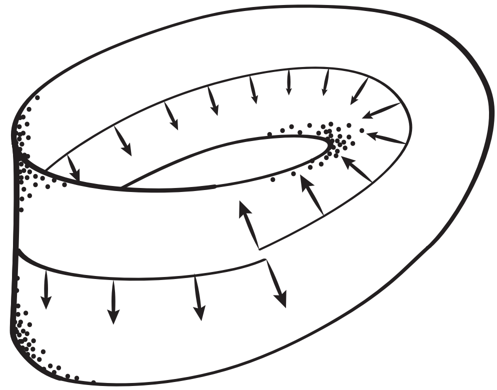

Figure 3.6 On a Möbius strip, a rotation through 2π reverses the arrow, which is returned to its initial state by another rotation through 2π.

$SO(3)$, a rotation through 2π is the identity. So, viewing 'ordinary' angular momentum as defined by the generators of $SO(3)$, it follows that the spin observables are not ordinary angular momentum.

However, we can still think of $SU(2)$ as consisting of spatial rotations in a certain 'degenerate' sense. The group $SU(2)$ is an example of a doubly-degenerate 'covering group' for $SO(3)$. The precise meaning of this is defined in Section 8.3; but, for now, we can visualise a representation of the spin rotation $R_j(\theta)$ using a Möbius strip, as in Figure 3.6. It has the property that, when an arrow is transported through 2π around the loop of a Möbius strip, it is not returned to its original state, but rather reverses: a second 'copy' of rotations through 2π is needed to restore it to its original orientation. There are thus two elements of $SU(2)$ corresponding to each 'ordinary' rotation through θ, given by $R_i(\theta)$ and $-R_i(\theta)$. This is what it means to say that the covering group is 'doubly-degenerate'. So, although $SU(2)$ is not isomorphic to $SO(3)$, we can still unambiguously associate each of its elements with a spatial rotation.

Now, recall that as a group element, time reversal generally leaves each spatial rotation intact, $\tau r \tau^{-1} = r$. This is true by definition in the Galilei and Lorentz groups; and, conceptually, it makes sense of the statement that the meaning of time reversal is 'independent of spatial orientation'. These facts can now be carried over to state space using the Representation View: given a group containing both $SU(2)$ and a time reversal group element τ, and which satisfies $\tau r \tau^{-1} = r$, any representation will satisfy

$$T R_j(\theta) T^{-1} = R_j(\theta), \tag{3.25}$$

https://doi.org/10.1017/9781009122139.004 Published online by Cambridge University Press

for each $j = x, y, z$. Using the fact that T is antiunitary, this in turn implies[41] that $Te^{-i\sigma_j}T^{-1} = e^{iT\sigma_j T^{-1}} = e^{-i\sigma_j}$, and hence that $T\sigma_j T^{-1} = -\sigma_j$ for each $j = x, y, z$. In this way, the fact that time reversal does not change spatial rotations can be used to explain why it reverses the Pauli spin observables.

Another interesting fact about the time reversal in a Pauli spin system is that, like a rotation through 2π, applying time reversal twice does not produce the identity: $T^2 = -I$. Only by two more applications of time reversal do we recover: $T^4 = I$. Our discussion above explains this curious fact as well and indeed determines the unique form of the time reversal operator in a Pauli spin system. The proof is an application of the general technique using Schur's lemma, introduced in the discussion of Proposition 3.2.

Proposition 3.5 *Let $\sigma_x, \sigma_y, \sigma_z$ be an irreducible unitary representation of the Pauli relations, and let K be the conjugation operator in the σ_z basis. If T is any antiunitary operator satisfying $TR_j(\theta)T^{-1} = R_j(\theta)$ for $j = x, y, z$, then $T = c\sigma_y K$ for some complex unit c, and $T^2 = -I$.*

Proof Our assumptions imply that the antiunitary T reverses each of the spin observables. One can straightforwardly check that $\sigma_y K$ does as well. Since both are antiunitary, the composition $-T\sigma_y K$ is unitary and commutes with σ_x, σ_y, and σ_z. These are the generators of an irreducible representation, and so by Schur's lemma, $-T\sigma_y K = cI$ for some $c \in \mathbb{C}$. This c is a complex unit, $c^*c = 1$, because $-T\sigma_y K$ is unitary. So, multiplying on the right by $\sigma_y K$ and recalling that $(\sigma_y K)^2 = -I$, we find that $T = c\sigma_y K$. Moreover, $T^2 = (c\sigma_2 K)^2 = c^*c(\sigma_2 K)^2 = -I$. ■

Thus, given a reflection of time translations $\tau : t \mapsto -t$ on spacetime, any representation of it on Hilbert space must be antiunitary by Proposition 3.4; and, the form of this antiunitary in a spin-1/2 system is uniquely determined (up to a constant) by Proposition 3.5. Although the concept of 'reversing time translations' at the level of spacetime is relatively simple, its representation in a spin system carries some of the interesting structure of that system.

3.5 Summary

This section has been a tour of several different state spaces for modern physics: Newtonian mechanics, analytic mechanics, and quantum theory. All of them include a rich structure for the description of what it means to be

[41] We use the fact that if T is any antiunitary, then $T(\alpha A) = \alpha^* TA$ for each complex constant α. In particular, $Ti\sigma_j = -iT\sigma_j$.

https://doi.org/10.1017/9781009122139.004 Published online by Cambridge University Press

a 'solution' to their equation of motion. Although that structure is not always visible in elementary presentations, we have made it visible here through a representation of time translations. This makes it a more straightforward matter to understand what time reversal means: it is nothing more and nothing less than a temporal reflection, represented on a highly structured state space.

This means that, when it comes to writing down the representative of time reversal on state space, we may find some 'bells and whistles', like the fact that it conjugates wavefunctions. But, the philosopher of time should not be alarmed by this. Time reversal is still ultimately just the automorphism $\tau : t \mapsto -t$ that reverses time translations. But, when it is represented using a highly structured state space, the operator T representing τ will inevitably pick up some of that novel structure.