The Physics and the Philosophy of Time Reversal in Standard Quantum Mechanics

Cristian López
University of Louvain – Institut Supérieur de Philosophie
Louvain-la Neuve, Belgium
cristian.lopez@uclouvain.be

Abstract

A widespread view in physics holds that the implementation of time reversal in standard quantum mechanics must be given by an anti-unitary operator. In foundations and philosophy of physics, however, there has been some discussion about the conceptual grounds of this orthodoxy, largely relying on either its obviousness or its mathematical-physical virtues. My aim in this paper is to substantively change the traditional structure of the debate by highlighting the philosophical commitments underlying the orthodoxy. I argue that the persuasive force of the orthodoxy can benefit from a relational metaphysics of time and a by-stipulation view on symmetries. Within such philosophical background, I submit, the orthodoxy of time reversal in standard quantum mechanics could find a fertile terrain to lay the groundwork for a more thorough conceptual justification.

Keywords: quantum mechanics, time reversal, anti-unitarity, relationalism, symmetry.

1. Introduction

What grounds the claim that a quantum physical system has been *genuinely* time reversed? Answers to this question fall in either of two sides. On the one hand, the overarching attitude (the orthodoxy, henceforth) points out that in order to reverse the dynamical evolution of a quantum system, an *anti-unitary* time-reversal operator must be given. On the other hand, some philosophers have lately argued that such orthodoxy might be challenged (Albert 2000, Callender 2000, Lopez 2019), which has paved the way for non-standard representations of time reversal in quantum mechanics (QM, henceforth), in general, in terms of a *unitary* implementation.

In one way or another, a thorough response to the opening question amounts to spelling the notion of time reversal out. This has led to a reinforcement of the orthodoxy by providing a precise mathematical tailoring of the implementation of time reversal as well as a more attentive philosophical refinement of its foundations (see, for instance, Sachs 1987, Earman 2002, Roberts 2017, 2018). The orthodoxy has generally been defended as the *only* view that is philosophically and physically viable, centering its defense in showing that a non-standard implementation of time reversal (i.e., that provided by a unitary time-reversal operator) fails to deliver a workable mathematical transformation as well as a conceptually defensible notion of time reversal. So, the dispute has been mostly set in terms of a *unitary* versus an *anti-unitary* implementation of time reversal in QM, where the orthodoxy champions the latter, while non-standard views (or 'heretic' views), the former.

Yet, I believe that the dispute between a unitary versus an anti-unitary implementation of time reversal is just the top of the iceberg in a series of philosophical and physical decisions that have to be made in order to conceptualize the idea of time reversal and to formally represent it. The bone of contention is not whether an anti-unitary implementation more genuinely represents time reversal simpliciter, but *which* concept of time reversal it intends to model mathematically, *which* are its philosophical assumptions, and *whether* they are tenable. In this sense, much of the persuasive force of the orthodoxy actually depends upon a philosophical background within which the anti-unitary implementation of time reversal makes sense more naturally. To thoroughly comprehend the nature of time reversal in QM is to bring to light such background.

The aim of this paper is to substantively change the traditional structure of the debate by bringing to the forefront the philosophical background that more straightforwardly can support the orthodoxy. My approach to the debate seeks to identify such background and follow the trail of the series of assumptions that make the orthodoxy a defensible and attractive approach. This would not only strengthen its legitimacy but would also provide a more accurate picture of how complex the notion of time reversal is. Part of this complexity consists in the many unnoticed substantial philosophical assumptions that form the conceptual environment within which it has been developed. To begin, I will distinguish three steps in building up the notion of time reversal, both conceptually and formally.

1. The *mathematical tailoring*, whereby the anti-unitary implementation comes out as the only fair modeling of time reversing a quantum system.
2. The *physical justification*, which defends the anti-unitary implementation by stipulating the physical requirements for time reversing a quantum system. This step sets a physics-based concept of time reversal primarily grounded in the idea of 'backtracking'.
3. The *philosophical background*, which lays the philosophical groundwork for such a physics-based concept of time reversal. This step justifies why the concept of time reversal is to be understood as the orthodoxy says it must be understood.

I will next show that there are two major arguments that physically support the orthodox understanding of time reversal in terms of backtracking, namely, (a) the two-time-evolution argument (or Wigner's general criterion for time reversal), and (b) the Hamiltonian's spectrum argument. The central claim of this article comes in after the physical justification is presented. I will argue that the philosophical background primarily consists in responding the two following questions:

- What do we mean by *time*?
- What *status* do we suppose that symmetries have in physics?

Neither of these questions admits a univocal answer. The first question opens a metaphysical dimension in our understanding of time reversal. At this point, I will argue that temporal relationalism is a friendly metaphysical environment for the orthodoxy, motivating a functional reductionist approach to time reversal, whereby time reversal ought to be functionally reduced to motion reversal. The second question concerns whether time-reversal symmetry is to be conceived either as a by-stipulation or a by-discovery symmetry. I will argue that both the physical justification and the mathematical tailoring of the orthodoxy can conceptually benefit from taking time-reversal invariance as a by-stipulation symmetry.

The upshot of the paper is, hence, that the orthodoxy should not be embraced because it is self-evident or analytically true, but because its persuasive force comes from a philosophical background that articulates extraordinarily well with the physics and the mathematics. This claim might be seen as a double-edged sword, since the philosophical background could now be challenged. Even though it might be seen as a more thorough justification of the orthodoxy, it might be also welcomed among those holding a heretic attitude, causing in the same proportion some discomfort among those defending the orthodoxy. It must be made clear, nonetheless, that it is not my intention to address this controversy here. Although I will occasionally bring the heretic view up as a counterpoint, I will align myself with the orthodoxy without putting it much into question. The philosophical exploration I pursue here homes in on why it is reasonable to side the orthodoxy, instead of showing it wrongheaded.

The structure of the article is as follows. In Section 2, I will begin by briefly introducing the mathematical tailoring of the orthodoxy of time reversal in QM. I will also show here how a non-standard account can come up. In Section 3, I will expose in detail the two major arguments that physically supports the orthodoxy. In Section 4, I will provide the main arguments of the paper by offering the philosophical background the orthodoxy is framed within. It chiefly consists of two pillars –a metaphysical pillar (Section 4.1) and a heuristic-epistemic one (Section 4.2.). Finally, concluding remarks.

2. The Mathematical Tailoring of the Orthodoxy

In the Hamiltonian formulation of classical mechanics, the main features of the time-reversal transformation stem from an analysis of the physics of the simplest cases. So, the starting point is typically a particle moving on a line in a conservative field force. The state of the particle is given by two variables: the generalized coordinates q_i and the conjugate momenta p_i. A trajectory in the phase space will be described through a set of functions $q_i(t), p_i(t)$, which is given by the Hamiltonian:

$$H = \frac{p^2}{2m} + V(x) \tag{1}$$

As $V(x)$ is constant and independent of time, it plays no role, and we can disregard it. In their most general expression, the Hamilton's equations follow from a system's Hamiltonian as function of the q_is and p_is

$$\dot{p}_i = -\frac{\partial H}{\partial q_i} \;\; ; \;\; \dot{q}_i = \frac{\partial H}{\partial p_i} \tag{2}$$

In which way can the time-reversal transformation be implemented? The answer mostly depends on what time reversing a classical system means, conceptually. Even though there would be much to say here, the most common answer, and one that is quite easy to grasp, is that of a film played backward. So, by time reversing a classical system we mean to generate a transformation that *retraces* the trajectory of a system This is the guiding concept that we want to formally implement. Such an implementation is what I will call the 'mathematical tailoring', that is, the process whereby such a concept is *formally* modeled within a theory. To put it more accurately, it is the process whereby a mathematical representation is given the right sort of properties to capture what we conceptually mean by 'time reversing' within a physical theory. Canonically, the mathematical tailoring of time reversal in Hamiltonian classical mechanics involves a transformation T such that reparametrizes the time coordinate, changes the sign of the p_is, and leaves the q_is unchanged.

$$T: t \to -t \; ; \; p_i \to -p_i \; ; \; q_i \to q_i \tag{3}$$

In consequence, T transforms the set of all smooth curves $(q(t), p(t))$ through phase space. This transformation is directly related to a symmetry property of the Hamiltonian

$$H(q_i, p_i) = H(q_i, -p_i) \tag{4}$$

If a system's Hamiltonian satisfies (4), then the equations of motion are invariant under *T*.

The shift to QM is not straightforward and requires some further formal work. Even though some features of time reversal will remain (or, better, will be required to remain), others will significantly change. The *mathematical tailoring* of the quantum time-reversal operator, hence, demands a series of assumptions and techniques that need to be detailed carefully. For this purpose, I will follow the traditional approach to the topic[1], though my intention is to be crystal clear about the assumptions and the rationale upon which the mathematical tailoring relies. The justification of such assumptions is partially mathematical, but also physical and, ultimately, philosophical. I will exclusively focus on the mathematical aspects here. The physical and the philosophical aspects will have to wait until the next sections.

To begin, most of the introductions to the mathematical tailoring of time reversal in QM resort on, at least, three interrelated assumptions. Let us use Θ to denote a general, still-unspecified time-reversal transformation.

A1. The Hamiltonian of the system is required to remain invariant under time reversal, $\Theta H \Theta^{-1} = H$ (see Ballentine 1998: 380)

A2. If the time evolution of a quantum state obeys time-reversal symmetry, then it is expected that if the state $|\alpha, t\rangle$ is a solution of the Schrödinger equation, then the time-reversed state $\Theta|\alpha, t\rangle$ will also be a solution (Ballentine 1998: 280, Sakurai 2011: 290). This also means that the Θ-transformation must be such that the time-reversed state belongs to the same unitary function space (Sachs 1987: 36).

A3. The time-reversal transformation is required to generate a reversal of motion (Bigi and Sanda 2016: 27), which imposes that Θ fits with the correspondence principles (Sachs 1987: 34).

Any operator that meets these requirements will be a good candidate for a time-reversal transformation in QM. However, in order to give the *right* form of the transformation, we need to impose some further structure. It is worth stressing that the explicit form of Θ will depend upon the basis of the Hilbert space used to represent the state, so it must be considered separately in each case. To keep things as simple as possible, I will circumscribe myself to the coordinate representation, but I will occasionally introduce more general remarks when needed.

One of the most intuitive features that our time-reversal transformation should possess is that it transforms the time coordinate as $t \rightarrow -t$, which might fairly be seen as representing an intuitive inversion of the direction of time. However, in QM the association between such a transformation and the inversion of the direction of time is not so straightforward. To see why let us stick to this intuition to characterize Θ and see how far it takes us. Suppose a

[1] I will be mainly following Sakurai (2011: 289-293), Ballentine (1998: 380-381), Gasiorowicz (1966: 25-30) Gibson and Pollard (1978: 179-180). I will also introduce some insights from Bigi and Sand (2016: 27-30), Jauch (1959: 88-91), Sachs (1987: 32-36), and Messiah (1966: 664-674)

quantum state in the position basis, $\psi(x,t)$, whose evolution is given by the Schrödinger equation:

$$H\psi(x,t) = i\hbar\frac{\delta}{\delta t}\psi(x,t) \tag{5}$$

Suppose, too, that time-reversal is implemented by an operator such that

$$\begin{gathered}\Theta: t \to -t \\ \Theta: x \to x\end{gathered} \tag{6}$$

If Θ is a well-behaved time-reversal transformation, then it must satisfy the requirements **A1-A3**. To start, it transforms the terms of the equation as follows:

$$\Theta H\psi(x,t) = \Theta i\hbar\frac{\delta}{\delta t}\psi(x,t) \tag{7}$$

Θ leaves the *i* and $\hbar$ unchanged. The operator $\frac{\delta}{\delta t}$ will change sign under Θ, $\Theta\frac{\delta}{\delta t}\ \Theta^{-1} = -\frac{\delta}{\delta t}$, since $\Theta: t \to -t$. The wavefunction $\psi(x,t)$ changes to $\psi'(x,t)\ \Theta$. These transformations yield the following equation:

$$\Theta H\Theta^{-1}\psi'(x,t) = -i\hbar\frac{\delta}{\delta t}\psi'(x,-t) \tag{8}$$

Now, we have to figure out how the Hamiltonian transforms under Θ**.** Given **A2**, eq. 8 should be symmetric, which formally amounts to rendering both sides of the equation equal. Given **A1**, ψ' must also be a solution of the Schrödinger equation. Both requirements, then, need that the Hamiltonian transforms as follows:

$$\Theta H\Theta^{-1} = -H \tag{9}$$

This suggests that the Hamiltonian should transform its sign under Θ. In the literature, there are at least two ways to motivate the transformation (eq. 9). One of them has been given by Craig Callender (2000). According to him, the Hamiltonian is a first-time derivative in the Schrödinger equation (in its simplest form), so it is natural (or logical) that transforms its sign under time reversal[2]. The other answer is the one given in the previous paragraph: in order to keep the equation invariant (i.e., requirement **A2**) the Hamiltonian should change sign under time reversal (see Gasiorowicz 1966: 27), so that both sides on the eq. 8 have negative signs.

[2] This argument is troublesome because it somehow assumes that the Schrödinger equation *defines* the Hamiltonian, when in general it "is defined independently as an operator that acts on the *x* dependence of a state function" (see Laue 1996 for discussion).

So far, this is formally correct (for a more thorough proof see Sakurai 2011: 291; also, Bigi and Sanda 2016: 27). But *if* time-reversal symmetry is to make *physical* sense while accomplishing requirements **A1-A3**, eqs. 8 and 9 are unacceptable. The crucial problem here is that they imply that Θ changes the sign of the Hamiltonian. And these minus signs on both sides of eq. 8 necessarily appeared in there *because* we started off by assuming that time reversal was represented by Θ, which, when looked closely, is *unitary* and *linear*. So, the conclusion we have reached can be put in the following conditional form: *if* time-reversal symmetry is to make *physical* sense while accomplishing requirements **A1**-**A3**, then Θ cannot be unitary (Sakurai 2011: 291). Let us analyze the argument a bit more carefully.

Basically, the argument is a reductio ad absurdum. It begins by assuming that Θ exists and satisfies **A1**-**A3**. If Θ is only given by eq. 6, then it is unitary and linear. If Θ is unitary and linear, then we are allowed to eliminate any *c*-number:

$$-iH\Theta|\ \rangle = \Theta iH|\ \rangle \rightarrow -H\Theta|\ \rangle = \Theta H|\ \rangle \tag{10}$$

Consider now an eigenstate $|\alpha\rangle$ of the Hamiltonian with eigenvalue E_{α}. The time-reversed state of $|\alpha\rangle$ would be $\Theta|\alpha\rangle$. Plugging this into eq. 10, we obtain

$$H\Theta|\alpha\rangle = -\Theta H|\alpha\rangle = (-\mathrm{E}_{\alpha})\Theta|\alpha\rangle \tag{11}$$

This means that $\Theta|\alpha\rangle$ is an eigenstate of the Hamiltonian with eigenvalue $-\mathrm{E}_{\alpha}$, that is, with negative energies. This is problematic for many reasons, but the most important one (which I will develop a bit further in the next section) is that it explicitly violates **A1**, conforming to which the Hamiltonian is required to remain bounded-from-below but unbounded-from-above after the transformation. So, no time-reversal transformation should change the Hamiltonian's spectrum from positive to negative. So, the result that has been reached by assuming that Θ exists is unacceptable (or "nonsensical", to borrow Sakurai's wording). To put it differently, what this proof shows is that there exists no unitary time-reversal transformation that satisfies the requirements I numbered previously (see Jauchs 1959: 88), which strongly suggests that time reversal must be implemented by an operator that does not generate eq. 8 and 9.

What to do then? We need to formally redefine the transformation that is to implement time reversal. We can do it by defining an *anti-unitary operator*, $\mathcal{T}$, which can be easily done by involving complex conjugation $KzK = z^{*}$, where z is a complex number and z^* its complex conjugate[3]. If we come back to eq. 7, but we apply $\mathcal{T}$, we see that the right-hand side of the equation becomes

[3] I am borrowing the notation '*' to refer to complex conjugation from Ballentine 1998, Sachs 1987 and Gibson and Pollard 1978.

$$\mathcal{T}H\psi(x,t) = \mathcal{T}ih\mathcal{T}\frac{\delta}{\delta t}\mathcal{T}\psi(x,t) = $$
$$\mathcal{T}H\psi(x,t) = ih\frac{\delta}{\delta t}\psi^*(x,-t) \tag{11}$$

$\mathcal{T}$ changes the sign of the operator $\frac{\delta}{\delta t}$ as Θ did, $\mathcal{T}\frac{\delta}{\delta t}\mathcal{T}^{-1} = -\frac{\delta}{\delta t}$. But now $\mathcal{T}$ does transform the sign of *i*, since it is anti-unitary, $\mathcal{T}i\mathcal{T}^{-1} = -i$. This eliminates the minus sign on the right side of the equation. Also, $\mathcal{T}$ takes the complex conjugate over the wavefunction, $\mathcal{T}$: $\psi(x,t) \rightarrow \psi^*(x,-t)$. We may notice now that, by simplifying and cancelling the *i*s and the kets, the eq. 11 suggests that the Hamiltonian should transform as

$$\mathcal{T}H\mathcal{T}^{-1} = H \tag{12}$$

To satisfy **A1** and **A2**. Now we are getting somewhere, since eq. 12 leaves the time-reversed Hamiltonian bounded-from-below (all its possible eigenvalues will be $E_i > 0$). And this makes perfectly physical sense.

This is, basically, the core of the mathematical tailoring of time reversal in QM, which comes down to the fact that it must be implemented by an antiunitary operator. This might be bit a surprise since we are accustomed to associating *unitary* transformations with physically interesting symmetry transformations. This is clearly not the case here where the rationale has led us to an *antiunitary* time-reversal transformation. But the reasons are quite strong. On the one hand, the preservation of transition probabilities allows an antiunitary operator to implement a symmetry transformation. On the other, if we want time reversal to be consistent with the kinematics and the dynamics of non-relativistic quantum mechanics, it seems we are forced to opt for an antiunitary implementation (Bigi and Sanda 2016: 27).

To complete the section, let me review the requirements **A1-A3** for $\mathcal{T}$. As I mentioned previously, the explicit form will depend ultimately on the basis of the Hilbert space, but if we express the Schrödinger equation in the coordinate basis as in eq. 5, we deduce from eqs. 11 and 12 that $\mathcal{T}$ delivers the following (time-reversed) Schrödinger equation

$$H\psi^*(x,-t) = ih\frac{\delta}{\delta t}\psi^*(x,-t) \tag{13}$$

Which not only satisfies **A1** (as previously shown), but also **A2**. If we look closely at $\mathcal{T}$ in the coordinate representation, we will find that $\mathcal{T}$ is just the complex conjugation for the general equation of a one-particle structureless system (see Ballentine 1998: 381; for a proof, see Sachs 1987: 39 and Bigi and Sanda 2016: 28) plus a re-parametrization of the *t* coordinate.

$$\mathcal{T} = K_0 \tag{14}$$

Finally, the antiunitary operator $\mathcal{T}$ also satisfies **A3** since it leads to inverting the sign of the momentum operator and to leaving the position operator unchanged

$$\begin{gathered} \mathcal{T}P\mathcal{T}^{-1} = -P \\ \mathcal{T}X\mathcal{T}^{-1} = X \end{gathered} \tag{15}$$

To sum up, the mathematical tailoring of time reversal in QM is guided by the preservation of some classical features in the transformation, but also by satisfying some formal constraints (**A1**-**A3**) that make time reversal formally well-behaved in QM. Despite the soundness of the argumentation for an anti-unitary implementation of time reversal, some philosophers have casted some doubts on it. This "heretic" attitude basically comes down to a positive defense of a "more natural" way to formally represent time reversal in physics (see Albert 2000, Callender 2000. Costa de Beauregard 1980[4] also defends such a view in quantum field theory), which would consist in giving a unitary implementation, like Θ in eqs. 6, 7 and 8.

Even though this attitude might seem outrageously absurd from the formal point of view, I think it should be rightly framed: I do not think that the heretic attitude holds that there is some formal argument to back its thesis, or that there is some formal flaw in the orthodox approach. What I believe, and it is a fair point that ought to be seriously considered, is that its defenders rather want to move the discussion to a more conceptual terrain by claiming that a non-standard, heretic account would better capture the idea of time reversal. What they put into question, in brief, is the binding between the *concept* of time reversal and its orthodox *implementation* –even though $\mathcal{T}$ is a well-behaved transformation, it does not implement time reversal, but something different. In fact, to a great extent, the persuasive force of the orthodoxy is grounded in accepting what it means by 'time reversal'. What I submit is that the assumptions that the orthodoxy imposes on an implementation of time reversal (like **A1-A3**) seeks to define time reversal physically and conceptually in terms of motion reversal, that is, in terms of retracing a system's state to the original state. If we now accept this core idea, then the orthodoxy succeeds not simply because it provides the right sort of mathematical tailoring, but because such a core idea (once accepted) can successfully justify the mathematical tailoring on more solid grounds.

My proposal for the rest is to view the orthodoxy as a chain of formal, physical and philosophical assumptions that articulate very well to yield a coherent view of time reversal. The mathematical tailoring is just the last link in the chain, which formally adapts, shapes, and implements a particular conceptualization of time reversal already at work. This

[4] Beauregard refers to "Racah's operator" as opposed to Wigner's (Costa de Beuregard 1980: 524 and further references therein).

conceptualization consists in construing time reversal in terms of *backtracking* (or motion reversal). The mathematical tailoring is sparklingly clean, providing good and sound reasons for time reversal to be anti-unitary. But, if we do not want to take this as self-evident or analytically true, we ought to provide a more careful justification of its assumptions along with the justification of why time reversal should be construed as motion reversal.

3. The Physical Justification of the Orthodoxy

We know that the orthodox understanding of time reversal seems somehow to be guided by the idea of backtracking. Yet, it is still unclear what this means exactly in physical terms within QM, and why the concept of time reversal must be conceived as the orthodoxy claims it must be. Though standard textbooks remain largely silent about these questions, literature on foundations and philosophy of physics has addressed them in some detail. Answers to these questions amounts to *justifying* the orthodox mathematical tailoring as well as its binding to the idea of backtracking. This justificatory task has probably its origins in the work of Eugene Wigner (1932), and it has been re-elaborated in the last decades (see, for instance, Sachs 1987, Earman 2002 and Roberts 2017).

The justificatory task is not simple, though: in general, the justification proceeds in two steps. The first consists of at least two arguments aiming to show not only how the idea of time reversal as backtracking should be physically understood, but also that the mathematical tailoring as presented in Section 2 is the right, and the only possible, implementation. This is the *physical justification* and I will develop it along this section. The second step, which, to the best of my knowledge, has not been sufficiently recognized, consists in the *philosophical* reasons we might have to conceive time reversal as backtracking. This is what I call the *philosophical background* and I will address it in Section 4.

Setting aside any classically rooted intuition on time reversal as backtracking, the details of what we physically mean by time reversal must be given within a theoretical framework. There are at least two key arguments[5] upholding the orthodoxy, namely:

- The 'two time-evolution' argument (or Wigner's general criterion for time reversal)

[5] It can be pointed out that there is a third key argument, to wit, that momentum changes its sign under time reversal. Certainly, this is one of the most salient features of the time reversal implementation in classical mechanics. For the most part though, the reasons *why* the sign of momentum should change under time reversal in QM follow the lines of the other two arguments. In the particular case of momentum, reasons swing back and forth from preserving certain smooth continuity between the classical mechanics and QM to achieving the representation of motion and reversal and appealing to its obviousness. Some authors just claim that the transformation follows by definition (Messiah 1966: 667, Sachs 1987, Ballentine 1998: 377-378). A more philosophically refined discussion can be found in Callender (2000) and Roberts (2018). In addition, it can be argued that the transformation of momentum plays a paramount role in the semi-classical limit, mainly in relation to Ehrenfest's theorem. However, this argument, and various versions thereof, does not add anything substantive to the point I want to make in this paper, so I will set it aside.

- The Hamilton's spectrum argument

3.1 The 'two-time-evolution' argument (or Wigner's general criterion for time-reversal)

One of the distinctive properties of the time-reversal operator in general is that it is an *involution*. Mathematically, this means that when time reversal is applied twice, it is equal to the identity. Naturally, this is met by *any* operator that satisfies $X^2 = I$, but this is not enough to get to the idea of backtracking a system's state to its initial state. The *locus classicus* of this requirement is the work of Eugene Wigner. In his 1932 book, *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra*, Wigner imposes a general criterion for time reversal stating that it is a transformation such that, when the following operations are sequentially performed, we obtain the identity. Informally,

time displacement by *t* × time reversal × time displacement by *t* × time reversal = ***I***

And more formally,

$$\mathcal{T}\left[U_{\Delta t_2}\mathcal{T}\left(U_{\Delta t_1}s_0\right)\right] = s_0 \tag{16}$$

Where s_0 is the initial state, and $\Delta t_1 = t_1 - t_2 = t_2 - t_1 = \Delta t_2$

Wigner's general criterion for a time reversal implementation evidently supposes further structure than a simple involution –the time reversal operator is expected to obtain the original state we started with *after* producing a time evolution with *t* increasing and by producing a (formally identical) second time evolution with t decreasing. In other words, the time-reversal operator is not only required to give us the same initial state when applied twice, but also to give us the same initial state after temporally evolving the system twice. This is a stronger requirement since the time-reversal transformation is expected to carry out the right sort of transformations to, at least, generate a time evolution with *t* decreasing. This twofold time evolution is not trivial and defines what a time-reversal transformation is. To put it into a slogan –to be a time-reversal operator *is* to be an operator that yields the identity after two-time evolutions.

Wigner additionally establishes that any candidate for a time-reversal transformation has to preserve transition probabilities.

$$|\langle\psi|\varphi\rangle| = |\langle\mathcal{T}\psi|\mathcal{T}\varphi\rangle| \tag{17}$$

This requirement intuitively makes senses since, if it were not the case, the second time evolution would no longer be possible. More specifically, Wigner postulates that the transition probabilities between two states have an *invariant* physical sense, so any symmetry should preserve them. Therefore, if time reversal is a symmetry, it must preserve transition probabilities. Yet, this justification only follows from the *invariance* under time reversal, and not from the time-reversal *transformation* itself. That is, *if* a symmetry holds, then transition probabilities must be preserved by the symmetry transformation. Clearly, this does not tell us whether a generic time reversal transformation should always preserve transitions probabilities[6].

How does all this relate to an anti-unitary representation of time reversal in QM? The famous Wigner's theorem states that a symmetry transformation is represented either by a unitary or an anti-unitary operator. As the anti-unitary operator is the only one that satisfactorily meets the general criterion (eq. 16) and preserves transition probabilities, the unitary operator is discarded. However, it has been pointed out that Wigner's proof of his theorem was "incomplete" (see Chevalier 2007: 429) and that a correct proof has been given by U. Uhlhorn in 1962, who also generalizes the condition of preserving the probabilities. In a nutshell, Uhlhorn's proof replaces the preservation of the transition probabilities by the preservation of orthogonality: any pair of orthogonal states $\langle\psi,\varphi\rangle = 0$ remains orthogonal under a symmetry transformation S, $\langle S\psi, S\varphi\rangle = 0$. It follows from this that $\langle\psi,\varphi\rangle = \langle S\psi, S\varphi\rangle$ (see Chevalier 2007, Section 5, for a proof of Uhlhorn's theorem). For Chevalier, Uhlhorn generalizes Wigner's proof as he shows that a symmetry transformation preserves the logical structure of a quantum theory, as Uhlhorn himself states in the Introduction of his book.

Following the same logic as before, *if* time-reversal invariance holds, then any orthogonal pair of states remains orthogonal under time reversal, that is, if $\langle\psi,\varphi\rangle = 0$, then $\langle T\psi, T\varphi\rangle = 0$. But, once again, the justification hinges upon what we should expect from time-reversal *invariance*. Bryan Roberts (2017) notes that Uhlhorn's theorem provides a general answer to why transition probabilities must be preserved under time reversal and advances a more convincing answer for why the time-reversal *transformation* ought to preserve transition probabilities. His argument is quite simple: orthogonality has nothing to do with time reversal since it simply relates to "what is possible in an experimental outcome, independently of their time development" (Roberts 2017: 321). So, why should we expect that something completely unrelated to time (as two states being mutually exclusive) be modified by time reversal?

The argument is interesting because it concerns what we should expect from a time-reversal transformation *independently* of whether it yields an invariance or not. It is worth bearing in mind that there is yet an assumption here: such implementation of time reversal is supposed to transform quantum-mechanical states into quantum mechanical states. Even though it is true that orthogonality has nothing to do with time reversal, it does have to do

[6] One could argue that a general time-reversal *transformation* will never change the sign of the position operator, because it is not the right sort of transformation that time reversal is expected to carry out. I do not find any equally stronger argument for transition probabilities, though Uhlhorn's theorem could, after some assumptions, do the work (see below).

with the notion of state. The time-reversal transformation is then required to *preserve* the notion of quantum-mechanical state[7].

To sum up the "two-evolution-based argument". The formal implementation of time reversal might take two forms: either unitary or anti-unitary. Naturally, we have a wide panoply of transformations that fills the bill. So, we need to narrow the possibilities down. Wigner's general criterion is a first step toward such a direction, since it states that whatever the time reversal comes to be, it is a transformation such that it delivers the state we started with after a twofold application and two-time evolutions. As a subsidiary requirement, it is demanded to preserve orthogonality, and thereby, transition probabilities.

3.2. The Hamiltonian's spectrum argument

In Section 2, I mentioned that one of the main virtues of the anti-unitary representation of time reversal is that it leaves the Hamiltonian invariant, $\mathcal{T}\mathbf{H}\mathcal{T}^{-1} = \mathbf{H}$. I will now expand on this requirement. As pointed out above, this requirement is essential for upholding the orthodoxy (in addition to references in Section 2, see also Sachs 1987: 36). As the Hamiltonian represents the energy of the system, its spectrum is supposed to be always positive. Yet, it was shown that a unitary and linear implementation of time reversal (Θ in Section 2) should bring about a minus sign on the right side of eq. 8, in order to accomplish **A1**-**A3**. This led us to the following: if $|\alpha\rangle$ is an eigenstate of the Hamiltonian with energy E, then the temporally reversed eigenstate $\Theta|\alpha\rangle$ should involve negative energies, $-E_{\alpha}$, that is, the quantum state would evolve backwards displaying *negative* energies[8]. The upshot was that if this is so, the time-reversal transformation does not make sense in QM. But we believe that the time-reversal transformation makes sense in QM. Therefore, Θ cannot exist. It is

[7] From a philosophical viewpoint, I think it is not trivial. Even though most symmetries physicists are interested in are required at minimum to transform states into states, we could want to leave some room for metaphysically possible scenarios in which some transformation fails to transform a state into a physical state. Primitivists with respect to time could argue that time is fundamental and defines not only the dynamics of a physical theory, but also is constitutive of its kinematics. Consider, for instance, Maudlin's argument about doppelgängers and mental states (Maudlin 2002: 271): if we suppose that time reversal acts in such a way that the time-reversed mental states are still mental states, we are unjustifiably assuming that the direction of time does not play any role in making a mental state what it is. How do we know that when reversing time, we will still end up with something like mental states, and not something completely different? An analogous argument, I think, can be run here: Why should the substantivalist assume that time does not play a role in defining what is a physical state? Does the requirement of preserving the notion of state when time is reversed not discard, from the outset, the notion of time as fundamental? I am not defending this viewpoint here, but I just want to draw the attention towards the non-obviousness of the assumption from a metaphysical viewpoint.

[8] It is worth clarifying that the predicates "positive" or "negative" for the energy spectrum, or "unbounded from below/from above" for Hamiltonians are conventional. So, the argument could not hinge upon which predicate we adopt to describe the system properly. The problem is not exactly whether the Hamiltonian is unbounded from below. The problem is that if we start with a Hamiltonian unbounded from above (but bounded from below) *and* end up with a Hamiltonian unbounded from below (but bounded from above) after a transformation. A *specific* Hamiltonian must be bounded (either from above or from below), and the problem will come up if one adopts a transformation that turns a Hamiltonian unbounded from above (bounded from below) into a Hamiltonian unbounded from below (bounded from above).

clear that the argument hinges upon the relation between a meaningful notion of time reversal, the spectrum condition, and the Wigner's general criterion.

The connection is relatively straightforward. If Hamiltonians must be *always* bounded from below, the second time translation with *t* decreasing must be generated by a bounded-from-below Hamiltonian. Otherwise, the time translation would be "physically meaningless" since it would involve eigenstates of the Hamiltonian whose spectrum is unbounded from below. Putting it drastically, unbounded-from-below Hamiltonians must *not* even be considered as *quantum-mechanics* systems. Therefore, the implementation of time reversal is demanded not only to generate a second time evolution with *t* decreasing, but also to generate a *quantum mechanical* second time evolution, which would be generated by a time-reversed H. Otherwise, Wigner's general criterion could not be applied since the second time translation could never be brought about. It is worth noting how this argument strengthens not only **A1** in the mathematical tailoring, but also **A2** and **A3** –the Θ-transformation could never transform solutions of the Schrödinger equation into time-reversed solutions, and it will thereby systematically fail to generate motion reversal. If we implement time reversal through some operator like Θ, it follows that the time-reversal transformation does not make physical sense This just stresses the necessity of relying on an anti-unitary representation of time reversal.

Roberts (2017) also offers a well-grounded argument for the requirement that the Hamiltonian's spectrum must remain invariant under time reversal. He begins by claiming that "all known Hamiltonians describing realistic quantum systems are bounded from below, which we will express by choosing a lower bound of $0 \leq (\psi, H\psi)$" (2017: 326). This can be empirically justified, at least partially: negative energies would turn matter unstable, but as matter looks reasonably stable, we would have good reasons to suppose that either negative energies do not exist (at least, within QM), or they remain undetected. This fact seems to be promoted to a general condition that a time-reversal operator must meet for its acceptability, meaning that $\langle\psi, H\psi\rangle$ and $\langle T\psi, TH\psi\rangle$ must be both non-negatives, as I specified above. Next, Roberts demands that there is at least "one realistic dynamical system" that satisfies time-reversal invariance in the sense that satisfies $Te^{itH}\psi = e^{-itH}T\psi$. Roberts makes the point that the time-reversal operator is demanded to be anti-unitary in order to meet these requirements, so for reductio, he assumes that such an operator is unitary, that is, that $T = \Theta$. This leads to $itH = -itTHT^{-1}$ and thus to $T\mathbf{H}T^{-1} = -H$. What we finally get is $0 \leq \langle\psi, H\psi\rangle = -\langle T\psi, TH\psi\rangle \leq 0$, which forces us to either accept that the Hamiltonian is unbounded from below (what he had previously ruled out) or that the Hamiltonian is the operator zero, which renders triviality. Therefore, by reductio, the time-reversal operator cannot be unitary but anti-unitary, $T = \mathcal{T}$.

4. The Philosophical Background of the Orthodoxy

Let me briefly summarize what I have shown so far. First, I outlined the mathematical tailoring of time reversal in QM according to the orthodoxy. What we learnt from it is that time reversal must be given by an anti-unitary implementation. Even though it differs from how time reversal is thought of in Hamiltonian classical mechanics, it captures a notion of time reversal that both transformations share –time reversal is implemented by tracing the state of a system back to the initial state. I claimed that the justification of the mathematical tailoring could not be purely formal, but it required some physical and conceptual background. This has been partially carried out in Section 3 by laying the physical foundations of time reversal as backtracking in QM. Regardless how much satisfactory this justification may be, a question remains: what entitles us to conceive of time reversal as backtracking in the first place? In other words, what is the philosophical background upholding the association between time reversal and backtracking? In this section I will address these philosophical aspects.

In my proposal, the philosophical background is a second step in the conceptual justification of the orthodoxy in QM. What I will argue is that this philosophical background mainly consists of two pillars:

- Temporal relationalism, which motivates a functional reductionist approach to time reversal in terms of motion reversal.
- The by-stipulation view on symmetries, which postulates that fundamental equations of motion must remain invariant under time reversal.

Before getting into the details, it is worth stressing the role that this philosophical background plays in the discussion about time reversal in QM as I framed it. My main claim is that the orthodoxy finds a friendly environment in such a philosophical background, which sensibly strengths its persuasive force when recognized. To put it differently, if temporal relationalism and the by-stipulation view on symmetries are adopted, then the orthodoxy comes out as a natural, and conceptually powerful, approach to time reversal in QM. The philosophical question that emerges from the physical justification is why we are entitled to call a specific piece of mathematics 'time reversal'. The mathematical tailoring, of course, does not provide such an answer and the physical justification simply assumes it by stating that time reversal must capture the idea of backtracking. The philosophical background provides the right sort of tools to answer this question in a conceptually clean and persuasive way.

Naturally, this does not entail that anyone supporting the orthodoxy ought to embrace temporal relationalism or the by-stipulation view. Neither does it mean that the orthodoxy necessarily requires temporal relationalism or a by-stipulation view. This would be stronger than what I will hold here. To see this more clearly, take for instance temporal relationalism, which defends the reduction of time reversal to motion reversal. My claim is that there is a conceptually straightforward way to go from temporal relationalism to the justification of time reversal as backtracking, and from here to the orthodoxy's physical justification. To the contrary, I find that there is not the same conceptually straightforward way to make the route from temporal substantivalism, because it is not prima facie obvious that time reversal reduces to motion reversal in this framework to begin. This, of course, does not mean that temporal substantivalism ought to reject the orthodoxy, because it can always find the way to relate time reversal to motion reversal, even though this relation might not be reductive[9]. In any case, if I am right on this, the burden is now on the temporal substantivalist: she should provide an account that can make sense of the orthodoxy within a temporal substantivalist framework, by showing which specific relations connect time reversal with motion reversal. In the end, this might redound to imposing a more complex structure that can eventually favor temporal relationalism for simplicity.

Then, the philosophical background could well play a twofold role. First, it benefits the orthodoxy since it provides the right conceptual framework to justify many of its assumptions. Second, it can be seen as an argument in favor of relationalism and of the by-stipulation view, since it naturally articulates with the mathematics and the physics that the orthodoxy develops.

4.1 Leibniz meets time reversal at the Plank scale.

My first thesis is that temporal relationalism lays the conceptual groundwork for a straightforward philosophical justification of the orthodoxy, since it easily connects time reversal with motion reversal (i.e., with the notion of backtracking). The key here to achieve this connection is the functional reduction of time reversal in terms of motion reversal.

Relationalism was famously championed by Leibniz. In his third letter to Samuel Clarke (dated February 25, 1716), he claimed that:

> "what that argument really proves is that times, considered without the things or events, are nothing at all, and that they consist only in the successive order of things and events"

Our philosophical understanding of time plays a role in our conceptual understanding of time reversal –If we are said to invert the direction *of time*, it seems at least reasonable to suppose that our course of actions will be different depending on what we understand by 'time'. And in this sense our metaphysics of time comes first: It determines not only what *time* reversal *is* but also *upon what* it is meant to act. The relationalist, hence, is committed to understanding time reversal in a particular way according to her metaphysical principles.

There are many different types of relationalist-like views in metaphysics and in philosophy of physics that, in general, share the idea that time is nothing over and above temporal relations among events and things (Benovsky 2010: 492), though they can greatly vary on which it is considered as objective and fundamental in the physical world (see Sklar

[9] I thank an anonymous reviewer for this observation

1974, Earman 1989, Pooley 2013, for comprehensive overviews of the different kinds of relationalisms). To keep things simple, I will consider temporal relationalism as holding two tenets:

R1 **A monist ontology**. There are only events or physical bodies in the world and their temporal relations.

R2 **A reductionist attitude**. Time is nothing but change. The sort of relation between the physical world and the concept of 'time' is that of a *Leibnizian representation* or a *Machian abstraction*: time is an ideal, unreal entity parasitic on events-things' changing.

According to these tenets, the variable t occurring in most physical theories (setting aside general relativity) is merely an external and unreal parameter, which should not be taken as representing anything with physical meaning. In this sense, temporal relationalism implies some reductionist attitude toward temporal predicates. For instance, any reference to the 'directionality of time' should not be taken literally as if there were some primitive entity exemplifying the property of having a directionality. Rather, it should be taken metaphorically –the 'directionality of time' boils down to the directionality of the change of a series of temporal relations held by their relata.

One of the lessons we can take from these tenets is that time reversal should not be taken literally, as if it were a transformation of time itself (whatever it might mean in physical terms). In fact, the parlance of time reversal in physics and philosophy of physics is mostly metaphorical (see, for instance, Wigner 1932: 325, Gibson and Pollard 1976: 177, Ballentine 1998: 377, among many others). The task for the philosophical reflection on time and time reversal is thus to conceptually articulate the underlying notions and elements converging into the idea of time reversal as orthodoxly understood. So, relationalism makes coherent a series of assumptions and elements that build up the orthodoxy.

By focusing on the second tenet, we see that the $t \rightarrow -t$ transformation (one of the, intuitively, most salient features of time reversal) must not be taken too seriously. It would be naïve, according to temporal relationalism, to take $t \rightarrow -t$ as performing a physically relevant action upon dynamical equations. What is really substantive in the understanding of time reversal is not the transformation of t, but the transformation of *change*. This suggests that time reversal should be considered simply as a "shortcut" standing for dynamically relevant transformations related to the change (or motion) of a system. To put it in a slogan, when it comes to time reversal, temporal relationalism holds that time reversal is *nothing but* change (or motion) reversal. This is the overarching concept grounding the physical justification and guiding to a good extent the mathematical tailoring: the formal representation ultimately seeks to capture the idea of reversing the change. The mathematical tailoring's task is then to identify those elements that represent change within each theory and to transform them in the right way.

So, we can postulate as a general scheme the following properties of a relational view on time reversal

$\boldsymbol{T}_{\mathbf{Rel}}$ (a) A mere re-parametrization of t by $T: t \to -t$, for any general time reversal.

(b) A change of all dynamically relevant magnitudes so as to generate a backward evolution, which is expressed by extensionally specifying the dynamically relevant transformations to take a system back to its original state.

The physically substantive part is given by the property (b), which genuinely generates the symmetry transformation. And it is important to highlight that the two physical justifications I laid out in Section 3 are philosophically based on this property: it motivates why certain observable has to transform in a specific way to generate the relevant symmetry transformation, namely, time reversal as motion reversal.

The second property can be clarified by proposing a sort of functionalist reduction of time reversal to motion reversal. Such an analysis will show how motion reversal realizes time reversal. The literature on functionalism (and, particularly, functional reduction) is abundant, so I will not get into details here. In general, it has mainly focused on either the relation between the mental and the physical in philosophy of mind, or the relations between high-order properties used in special sciences and low-order properties more frequent in physics. What I propose here is not a strict functional reduction as the one discussed in philosophy of mind or general philosophy of science, but a style of reasoning which can quite well capture the metaphysical and epistemic relations holding between time reversal and motion reversal in physics (for a functionalist approach in philosophy of physics, see for instance, Knox 2018).

The overall idea is that motion reversal (and those properties attached to it) *realizes* time reversal (and thus all those subsidiary properties attached to it). In particular, the notion of time reversal is functionally reduced to the idea of 'backtracking', in the sense that 'time reversal' refers to dynamical *realizers* that play the role of retracing a system's state to its origin. So, we can rephrase this by saying that if the state of a system has been "time-reversed", or that the history of a system has been "time-reversed", we have to find the *realizers* of such a state and such a history in terms of the dynamical operations that effectively generate a backtracking process. From an abstract perspective, the notion of time reversal is simply a placeholder, whose occupants will be those realizers playing the role of retracing a system's state to its origin. The problem of working out the right form of the time-reversal transformation is that of working out the right realizers within a specific physical theory.

The idea could be developed even further by offering the Ramsey sentence of time reversal, which has the structure

$$\exists X, \exists Y, \exists Z\ (\ldots X \ldots Y \ldots Z)$$

What we know is that the notion of backtracking or (change) motion reversal realizes the notion of time reversal. What we have to do now is to supply the roles of each key players (X, Y and Z) and to identify which specific transformations within a particular physical theory play each role. This realizes Ramsey sentence consequently. Then, in order to apply time reversal properly, we have to provide the ordered *n*-tuple of realizers within a physical theory that satisfies the Ramsey sentence for time reversal.

Under this framework, we could either adopt an eliminativist or conservative attitude. The eliminativist will be prone to simply eradicating any temporal predicate and structure in favor of predicates and structures exclusively referring to change. This would automatically remove temporal predicates from the physical picture when we want to be rigorous about what we are really doing when time reversing an equation of motion or a physical system. The term 'time reversal' is just *flatus vocis* (see Rovelli 2004 for such a radical attitude with respect to time). The conservative reductionist will be rather prone to preserving some temporal predicates and structures, though acknowledging their actual realizers relate to change. This view better preserves the classical Leibnizian-Machian framework, where durations are relative, time is essentially change, but other traditional temporal structures remain absolute (see Gryb and Thébault 2016 for a defense of this more conservative relationalism in quantum gravity).

When we center in the physical justification of the orthodoxy, we see how the relational view on time-reversal can explain in a natural way why we should understand time reversal as a two-time evolution, basing Wigner's general criterion. If time reversal is functionally realized by motion reversal, then any fair formal implementation of it has to pose the enough structure to represent the reversion of motion. In Wigner's general criterion, this is provided by guaranteeing that the time(-motion)-reversal transformation generates a second time translation that takes the evolution of the state back to its origin. Otherwise, the implementation of time(-motion) reversal fails to genuinely capture the idea of backtracking. But what entitles us to metaphysically relate time reversal to motion reversal is temporal relationalism. This is done by providing the adequate conceptual framework to articulate the mathematics and the physics in a coherent approach to time reversal. It is worth emphasizing that the persuasive force of the orthodoxy does not hinge only upon the mathematical tailoring (as often argued when discarding alternative implementations of time reversal), but also upon one's underlying metaphysics –it is this final step in the justification which entitles us to draw the right sort of conceptual connections.

The same tenets also ground the Hamilton's spectrum argument I presented in Sub-section 3.2. Consider the following counterargument, based on Callender (2000)'s argument: The Hamiltonian is a first-time derivative magnitude, so it is natural to expect the Hamiltonian to change its sign under time reversal, which would lead to transforming a bounded-from-below Hamiltonian into an *unbounded*-from-below Hamiltonian (see fn. 2 for concerns). However, from a relational viewpoint the demand is excessive, even if formal and physical considerations are put momently aside. Whether physical magnitudes are canonically defined as first-time derivative does not play any substantive role in defining time reversal, because we are not *conceptually* interested in *t*. We should instead focus on elucidating what role such physical magnitudes play in the evolution of the state and what role they should play if the evolution were reversed. The conceptualization of time reversal at a physical level precisely attempt to work that out. From a conceptual viewpoint, the real issue is not whether or not it makes sense that the Hamiltonian changes its sign under time reversal, but whether such a transformation plays any role in formally implementing time reversal as backtracking. Clearly, it does not. This explains why the unitary transformation must be discarded, even under the implausible assumption that they might make mathematical and physical sense.

To emphasize my point. Temporal relationalism offers us a straightforward way to understand time reversal as conceived and formally implemented by the orthodoxy –it underpins, in a simple way, the underlying assumption that time reversal is just to track a system back to its initial state. For the sake of the argument, let us briefly consider the matter from the opposite view. Suppose now that time is a primitive substance independent of motion, *a lá Newton*. Time reversal should hence amount to a transformation of the intrinsic direction of such a substance. Within this view, there are no prima facie metaphysical reasons to identify motion reversal with time reversal, because they *are* different kinds of things. Therefore, there are no prima facie conceptual reasons to formally implement time reversal as it was recommended by Wigner's criterion. Whoever wants to hold the orthodoxy approach to time reversal in QM *and* temporal substantivalism should then provide us an account that shows how the justification of the orthodoxy can be achieved. This, I guess, can be done in different ways, but it would demand imposing further conceptual structure to get the connection between time reversal and motion reversal properly justified. What we miss in changing the metaphysical framework is the straightforward connection between time reversal and motion reversal we get from temporal relationalism. Without such a connection, the persuasive force of the orthodoxy sensibly diminishes.

Another virtue of relationalism when it comes to justifying the orthodoxy relates to the bridge that it builds with the empirical work. The substantivalist could insist on pointing out that we are still not allowed to call a piece of mathematics, as $\mathcal{T}$, 'time reversal', because it is implementing a different sort of transformation, namely, motion reversal. In the end, all we were just confused about names and concepts all along –the empirical information that physicists have been gathering so far in terms of time reversal has just been information about motion reversal A reply to the substantivalist could go in the following line. Even though from a strict substantivalist framework such a situation is possible, she has to accept that when we test time reversal, we always test motion reversal. And now the substantivalist faces an uncomfortable dilemma: either she gives us the way to test time reversal independently from motion reversal, or she declares time reversal untestable. The first option puts the burden on her, and we can just wait the answer. The second one forces her to provide further reasons of why an untestable symmetry transformation (i.e., time reversal as independent from motion reversal) should be preserved in our physical theories. In fact, it is similar to the case when we find redundant structure in our physical theory. By epistemic reasons (i.e., parsimony), we could just eliminate the redundant structure. In the end, the defense of the independency of time reversal would be self-defeating.

Whether the orthodoxy can be justified in a non-relational framework deserves, of course, much more work. My aim here is not to be exhaustive about the possible connection, but to show that temporal relationalism can make a case for the orthodoxy in a natural way. In addition, this also suggest that the orthodoxy might have been guided by relational intuitions when developing the mathematical tailoring and the physical justification. The contrast in the previous paragraph just shows some of the difficulties that an alternative philosophical background could face. If for any reason we reject any of the relationalist tenets, the physical justification of time reversal loses much of its persuasive force, dragging naturally down the anti-unitary representation of time reversal. This emphasizes the relevance of the metaphysical background for the orthodoxy –it is not merely an uncommitted defense of a particular mathematical tailoring, but a well-articulated general view on time reversal. Temporal relationalism is one of its pillars since it provides the adequate framework to build a robust and powerful conceptual justification of why time reversal should be thought of as backtracking.

4.2. Time-Reversal *Invariance*: by-stipulation or by-discovery

In the previous section, I focused on the metaphysical pillar of the orthodoxy. In this section, I will focus on the second pillar, which concerns epistemic and heuristic aspects of time-reversal symmetry in QM. Whereas the first pillar chiefly centered in the time-reversal *transformation* (what we metaphysically and physically mean by ‘time reversing’), the second pillar rather centers in the status of *symmetries* in physics. To be precise, it centers in the epistemic and heuristic aspects that connect the construction of a time-reversal transformation to the role that the time-reversal *symmetry* should play in a physical theory.

There are at least two opposing views on space-time symmetries in modern physics. One of them, which I will call *by-stipulation*, takes symmetries as postulated, being true independently of the details of the dynamics. The other, which I will call *by-discovery*, takes symmetries as a result of the details of the dynamics. In the former case, symmetries are principles that constrain the dynamics. In the latter, symmetries are derived from it. As time-reversal is prima facie a space-time symmetry, both views are also present in this case. What I will argue is that the orthodoxy can benefit from the by-stipulation view of time-reversal invariance, which offers, to a great extent, support to the physical justifications of an anti-unitary implementation.

What does justify the distinction between by-stipulation and by-discovery symmetries? Katherine Brading and Elena Castellani show that space-time symmetries are sometimes thought of as guides to theory construction. That is, principles that must be satisfied whatever the final details of the theory come to be. The mechanism whereby a symmetry is raised to a must-satisfied principle is that of *stipulation* –we postulate, independently of the details of a theory's dynamics, that a given symmetry holds, then the dynamics adapts to the symmetries' constraints. When laying the groundwork for Bohmian Mechanics, Dettlef Dürr and Stephan Teufel for instance write

> "A symmetry can be a priori, i.e., the physical law is built in such a way that it respects that particular symmetry by construction. This is exemplified by spacetime symmetries, because spacetime is the theater in which the physical law acts (as long as spacetime is not subject to a law itself, as in general relativity, which we exclude from our considerations here), and must therefore respect the rules of the theater". (2009: 43-44)

It is worth contrasting this quote to others we can find in the literature on symmetries. John Earman says

> "The received wisdom about the status of symmetry principles has it that one must confront a choice between the *a posteriori approach* (a.k.a. the bottom-up approach) versus the *a priori approach* (a.k.a. the top-down approach)". (2004: 1230)

Earman's distinction goes along with that of Brading and Castellani's (2007): whereas some take symmetries as postulated, guiding theory construction, others follow an opposite trend, according to which symmetries are a consequence of specific laws –like a *discovery* (2007: 1347). The idea of postulating a symmetry is *normative*, suggesting certain degree of necessity: a theory's dynamics must satisfy the symmetry principles, even though if the dynamics had been different. This gives symmetry principles certain modal robustness (or counterfactual robustness, see Lange 2009), entailed by its normative nature.

Who denies that symmetries have such normative nature is prone to regard symmetries as a property of dynamics. In this line, Earman says:

> "it would seem that the symmetry transformation could not fail to be a true symmetry of nature, contradicting the usual understanding that symmetry principles are contingent, that is, are true (or false) without being necessarily true (or false)" (1989: 121)

We hence come to know which symmetries a theory has by investigating the formal relations held by the elements in differential equations. Remarkably, this approach was followed by Isaac Newton in formulating classical mechanics in the *Principia* as the relativity principle appears as a corollary of the equations of motion (Corollary 5, see also Brading and Castellani

2003: 6) and by Joseph Lagrange (1811: 241). In this sense, a space-time symmetry plays a *descriptive* role, rather than a normative one.

To write all this out neatly, both approaches can be defined as follows. For a general space-time symmetry σ:

By-stipulation approach σ-symmetry plays a *normative role* in a theory's dynamics and it must thereby be regarded as a priori and necessary for a theory T.

By-discovery approach σ-symmetry plays a *descriptive role*, and it must thereby be regarded as a posteriori and contingent for a theory *T*.

It is worth remarking that the epistemic notions of "a priori/a posteriori" should be understood not in the traditional sense (independent or not of the experience), but in relation to a theory's dynamics: whether σ-symmetry is known independently of a theory's laws.

Time-reversal invariance can be regarded from both approaches. My claim is that, if the by-stipulation view on symmetries is adopted, the orthodoxy is the natural approach to time reversal in QM. If we look at the orthodoxy closely, we can identify some assumptions giving support to the by-stipulation view since they play a normative role in the theory construction as well as in the mathematical tailoring of the time-reversal transformation. To illustrate this, Robert Sachs says:

> "In order to express explicitly the independence between the kinematics and the nature of the forces, we require that the transformations leave the equations of motion invariant when all forces or interactions vanish" (Sachs 1987: 7)

Time-reversal symmetry is required to hold by stipulation in the case of *the free Schrödinger* equation, that is, in the evolution of free-interaction quantum systems. This idea nicely comes along with the "theater picture" of Dürr and Teufel: the simplest systems' dynamics reflect genuinely the structure of the theater, both its asymmetries and symmetries. But such a structure is pre-existent and independent of the dynamics, playing the role of setting the (space)-time background for all models of the theory and of individuating the nature of forces, interactions, and the various structures (for instance, asymmetries) they generate.

To strength this point, let us briefly move to a different theory –classical electromagnetism. In discussing the mathematical form of the time-reversal transformation, Frank Arntzenius and Hillary Greaves (2009) claim that a widespread account, which they call 'the textbook's account', proceeds as follows (see also Peterson 2015):

> "Next let us consider the electric and magnetic fields. How do they transform under time reversal? Well, the standard procedure is simply *to assume* that classical electromagnetism is invariant under time reversal. *From this assumption* of time reversal invariance of the theory (…) it is inferred that the

> electric field E is invariant under time reversal (…)” (Arntzenius and Greaves 2009: 6. Italics mine)

The same mechanisms, mutatis mutandis, seems to be guiding the orthodoxy in QM. Indeed, the stipulation of time-reversal invariance is just the assumption **A2**, which easily justifies why the implementation of time reversal in QM must be anti-unitary. This is more evident when we contrast with the unitary implementation: If we previously presume that time-reversal invariance holds, then the formal implementation of time reversal cannot be one that make the free Schrödinger equation non-time-reversal invariant. Such a result would be in fact at odds with Sachs’ quote too: we, for instance, stipulate that the free Schrödinger equation is time reversal invariant *in order to* express the independence between dynamics and kinematics. Hence, its stipulation plays a heuristic role in our understanding of the theory, which we will be missed if the unitary transformation is rather adopted. So, everything converges at the same place: the anti-unitary operator emerges as the right implementation that carries out the sort of required transformations to keep the free Schrödinger equation invariant (**A2**), satisfying the epistemic and heuristic stipulation.

This by-stipulation view on symmetries and the implied justificatory mechanism can be also regarded from a different angle. The stipulation of time-reversal invariance also appears as a premise in Wigner’s definition of time reversal in Section 3. He invokes two explicit premises:

1. that a suitable time-reversal transformation must be able to restore “the system to its original state” (1932: 326),
2. and that time inversion must flip the direction of momentum to compensate for the twofold application of *T* in Wigner’s general criterion.

But there is also one fundamental implicit assumption:

3. for a time-reversal transformation to be well-defined (and to exist at all), the second time translation (from t_2 to t_1) must also be physically possible.

To see how this last assumption works let us suppose that a quantum state $|\psi\rangle$ evolves from t_1 to t_2, according to the Schrödinger equation (first time translation). At t_2, time reversal is applied upon the Schrödinger equation. *If* the time-reversal transformation is well-defined, *then* the time-reversed state $T|\psi\rangle$ should evolve from t_2 to t_1 also conforming with the Schrödinger equation (second time translation). And here the implicit assumption comes in. According to Wigner, the operation to be applied upon the state at t_2 must be of such a kind that yields a *quantum-mechanical* evolution –the transformation takes a solution of the free Schrödinger equation and transforms it into a solution of the free Schrödinger equation This is the standard definition of symmetry and the state that satisfies is $T|\psi\rangle = |\psi^*\rangle$, where $T = \mathcal{T}$.

To complete the argument, suppose now that at t_2, we apply a unitary time-reversal operator, $T = \Theta$. As remarked above, the free Schrödinger equation will not temporally translate the system back. But, even worse, the transformation will fail to generate the second time translation, turning a solution of the free Schrödinger equation into a non-solution. The metaphysical pillar of the orthodoxy shows that such a transformation is ill-conceived because it fails to represent time reversal as motion reversal. The by-stipulation symmetry view shows that such a transformation is ill-conceived because it makes the free Schrödinger equation non-time-reversal invariant, violating **A2**. As I commented before, whether we assume the by-stipulation or the by discovery view is motivated by various reasons, mainly concerning epistemic and heuristic features. Which view to take is not at issue here. What it is at issue is which one of these approaches can offer a straightforward, or more natural, justification of the orthodoxy. A commitment to the by-stipulation view imposes some constraints on the notion of symmetries that more directly leads us to the orthodoxy –the anti-unitary representation easily suits in such a framework along with the potential theoretical virtues attached to it.

It might seem, at first glance, that the by-stipulation view arbitrarily declares that a given dynamics is invariant, without any further justification. This might then be regarded as a drawback of the by-stipulation view and that the by-discovery view then flats out win. And if this is the case, then the orthodoxy would have to revise some of its assumptions in the light of the by-discovery framework. Yet, I think it is not the case when the role of symmetries in the by-stipulation view is adequately addressed and considered from a broader perspective. I do not have enough room to develop the possible epistemic ramifications of adopting the by-stipulation view in detail, but here goes a hint of what, I think, is happening.

Time-reversal invariance is a property expressed by dynamical equations of motion, either a stipulated or discovered property. If we think of dynamical equations of motion as representing some primitive modality in the world, or at least some modally robust pattern, one is committed to squaring time-reversal invariance within such a framework. This could lead to consider time-reversal invariance as also expressing a (modal) property of the world. However, this might sound a bit odd: We come to know something substantial about the world by means of a stipulation. Further argumentation would then be required, which would place the by-discovery view in some advantage. However, this is not the only way to go. Neither is it the best way to go.

Instead of assuming that time-reversal invariance is a stipulated property of dynamical equations that express some laws of nature, we could hold a deflationary view on symmetries, conforming to which time-reversal invariance is a stipulated property of scientific laws understood as sentences in an axiomatic system. Therefore, symmetries are just theoretical postulates seeking for a better equilibrium between simplicity and informativeness. So, the motivation of stipulating a symmetry like time reversal is fundamentally representational and should be judged in such terms: We declare that general (or fundamental) dynamical equations are time-reversal invariant because it is a representational advantage to do it so –a time-reversal invariant dynamic just turns out to be simpler and more informative than non-time-reversal invariant ones. So, symmetries so understood square perfectly within the so-called Best System Approach (see Lewis 1973, Ramsey 1978, Loewer 1996 and Cohen and Callender 2009), which suggests us not to consider scientific laws, and symmetries I would add, as referring to some primitive modality in the world, but just playing a theoretical role striving for simplicity and informativeness. If symmetries are considered from this angle, the by-discovery view does not flat out win, but quite the opposite: the burden of the proof is on its side, since it has to show that symmetries have a more robust status in physical theories than the one given by a deflationary view.

5. Concluding remarks

In this paper I have offered novel insights to address the debate on time reversal in QM. I began by distinguishing three steps in construing the orthodoxy: the mathematical tailoring (Section 2), the physical justification (Section 3), and the philosophical background (Section 4). Each step was shown to be supported by an underlying one: the mathematical tailoring depends on its physical justification, which in turn relies on a philosophical background. The general aim was to bring to light the relevance of the philosophical background as a series of philosophical commitments from which the orthodoxy can sensibly benefit, once recognized, and made explicit. With respect to this, I have claimed that the orthodoxy is philosophically supported by two pillars:

- temporal relationalism, which promotes a functionalist reduction of time reversal in terms of motion reversal
- the by-stipulation view of time-reversal invariance

What this primarily shows is that the orthodoxy is not philosophically neutral, but it can be successfully articulated, and properly justified, when a series of metaphysical, epistemic and heuristic commitments are taken into account. These play a major justificatory role in the anti-unitary implementation of time reversal in QM. Contrarily to how the debate has developed thus far, the quid of the notion of time reversal in QM should not be primarily framed in terms of whether it ought to be anti-unitary or unitary, but if we have well-grounded reasons to call a piece of mathematics, as $\mathcal{T}$, 'time reversal'. This, when we looked at the physical justification, came down to the idea of thinking of time reversal in terms of motion reversal. Temporal relationalism and the by-stipulation view on symmetries came in to cement this connection, transmitting the justification all the way up in the chain.

So, what grounds the claim that a quantum system has been *genuinely* time reversed is a well-articulated view that is not simple, but quite complex, involving various mathematical, physical and, fundamentally, philosophical assumptions. The latter directly relate to big questions such as the nature of time and symmetries in physics in metaphysics, around which philosophers and scientist have long been gravitating. And it is such philosophical complexity what feeds the orthodoxy's persuasive force, rather than any seemingly obviousness or self-evident truth.

Acknowledgments. I am deeply grateful to Olimpia Lombardi, Michael Esfeld, Karim Thébault, Carl Hoefer, Bryan Roberts, Christian Sachse and María José Ferreira for so enriching and substantive discussions, criticisms, and suggestions on this manuscript and earlier versions. I also thank to Dustin Lazarovici, Frida Trotter, Andrea Oldofredi, Federico Benitez, Manuel Gadella, Federico Holik and Sebastián Fortín for suggestions and comments on some ideas expressed here. Finally, I would like to thank the anonymous reviewers: their reports significantly improved the original version. This work was supported by a FRS-FNRS (*Fonds de la Recherche Scientifique*) Postdoctoral Fellowship and made possible through the support of the grant #61785 from the John Templeton Foundation. The opinions expressed in this publication are those of the author and do not necessarily reflect the views of the John Templeton Foundation.

6. References

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

Albert, D. Z. (2000). *Time and Chance*. Cambridge, MA: Harvard University Press.

Arntzenius, F. and Greaves, H. (2009). "Time reversal in classical electromagnetism". *The British Journal for the Philosophy of Science*, 60: 557-584

Ballentine, L. (1998). *Quantum Mechanics. A modern Development*. Singapur: World Scientific.

Barbour, J., and Bertotti, B. (1982). Mach's principle and the structure of dynamical theories. *Proceedings of the Royal Society A, 382,* 295–306.

Benovsky, J. (2010). "The relationalist and substantivalist theories of time: foes or friends?" *European Journal of Philosophy*, 19 (4): 491-506.

Brading, K. and E. Castellani. (2007). "Symmetries and Invariances in Classical Physics". *Handbook of the Philosophy of Science, Philosophy of Physics, Part B*. Eds. J. Butterfield and J. Earman. The Netherlands: Elsevier, 1331–1367.

Callender, C. (2000). "Is time 'handed' in a quantum world?" *Proceedings of the Aristotelian Society,* 100: 247-269.

Chevallier, G. (2007). "Wigner's theorem and its generalizations". In K. Engesser, D. M. Gabbay and D. Lehmann (eds.), *Handbook of Quantum Logic and Quantum Structures,* New York: Elsevier Publisher, pp. *429-475*.

Cohen, J. and Callender, C. (2009). "A better best system account of lawhood". *Philosophical Studies* 145: 1-34.

Costa de Beauregard, O. (1980). "CPT invariance and interpretation of quantum mechanics", *Foundations of Physics*, 10: 513-530.

Dürr, D. and Teufel, S. (2009). *Bohmian Mechanics: The Physics and Mathematics of Quantum Theory,* Berlin: Springer-Verlag.

Earman, J. (1989). *World enough and space-time. Absolute versus relational theories of space-time*. Cambridge, Massachusetts: MIT Press.

Earman, J. (2002). “What time-reversal invariance is and why it matters”. *International Studies in the Philosophy of Science,* 16: 245-264.

Earman, J. (1989). *World enough and space-time. Absolute versus relational theories of space-time*. Cambridge, Massachusetts: MIT Press.

Gasiororowicz, S. (1966). *Elementary Particle Physics.* New York: John Wiley and Sons.

Gibson, W.M. and Pollard, B.R. (1976). *Symmetry Principles in elementary particle physics*. Cambrige: Cambridge University Press.

Gryb, S. and Thébault, K. (2016). “Time remains”. *British Journal for Philosophy of Science,* 67: 663-705.

Knox, E. (2018). “Physical relativity from a functionalist perspective”. *Studies in History and Philosophy of Modern Physics*, 67: 118-124.

Lagrange, J. (1811). *Mécanique Analytique.* Paris: xxx.

Lewis, D. (1973). *Counterfactuals*. Cambridge: Harvard University Press.

Loewer, B. (1996). “Humean Supervenience”. *Philosophical Topics*, 24: 101–126.

Lopez, C. (2019). “Roads to the past: how to go and *not* to go backward in time in quantum theories”. *European Journal for Philosophy of Science*, 9: 27.

Messiah, A. (1966). *Quantum Mechanics.* New York: John Wiley and Sons.

Peterson, D. (2015). “Prospect for a new account of time reversal”. *Studies in History and Philosophy of Modern Physics,* 49: 42-56.

Pooley, O. (2013). Substantivalist and relationalist approaches to spacetime, In Batterman, R. (Ed.) *The oxford handbook of philosophy of physics* (pp. 522-586). Oxford: Oxford University Press.

Ramsey, F. (1978). *Foundations*. London: Routledge and Kegan Paul.

Roberts, B. (2017). “Three myths about time reversal invariance”. *Philosophy of Science,* 84, 2: 315-334.

Roberts, B. (2018). “Time reversal” http://philsci-archive.pitt.edu/15033/1/Roberts2018-Time Reversal.pdf.

Rovelli, C. (2004). *Quantum gravity*. Cambridge: Cambridge University Press.

Sachs, R. (1987). *The Physics of Time Reversal*. London: University Chicago Press.

Sklar, L. (1974). *Space, Time and Spacetime.* Berkeley: University of California Press.

Uhlhorn, U. (1961). “Representation of Symmetry Transformations in Quantum Mechanics.” *Arkiv för Fysik* 23:307–40.

Wigner, E. (1932). *Group Theory and its Application to the Quantum Mechanics of Atomic Spectra.* New York: Academic Press (1959).