A Dynamical Perspective on the Arrow of Time

Kian Salimkhani

Revised and extended manuscript of a paper presented at the PTS meeting at the Eastern APA in January 2021.

Abstract

It is standardly believed that the generally time-reversal symmetric fundamental laws of physics themselves cannot explain the apparent asymmetry of time. In particular, it is believed that CP violation is of no help. In this paper, I want to push back against a quick dismissal of CP violation as a potential source for the arrow of time and argue that it should be taken more seriously for conceptualising time in physics. I first recall that CP violation is a key feature of our best physical theory which also has large-scale explanatory import regarding the matter–antimatter asymmetry of the universe. I then investigate how CP violation may help to explain the directionality of time. I argue that accounts à la Maudlin that posit an intrinsic fundamental direction of time are not convincing and instead propose to utilise recent results from work on the dynamical approach to relativity theory.

1 Introduction

The problem of the arrow of time basically arises from the following two observations: on the one hand, time and the ordering of events in time seem to be directed. Typically, reversing the ordering of events in time yields either a very different process, or a process that is not observed to happen at all. Learning is a very different process than forgetting. Recovering from a disease is a very different process than contracting a disease. An egg falls to the ground and smashes, but a smashed egg does not spontaneously jump up to reassemble in my trembling hand. Such examples demonstrate why we belief that directionality is essential to time.

On the other hand, however, it is standardly argued that this directionality cannot be explained by fundamental physics, since at the level of fundamental physics all processes are supposedly *un*directed, they do not distinguish between a future direction and a past direction, but are time-reversal invariant. So if a process is allowed by some fundamental law – that is the process is a solution of the respective equations – the time-reversed process is as well.

In other words, fundamental physics seems to suggest that there is no fundamental difference between temporal and spatial dimensions – between the reordering of events from left to right and from the past to the future. But if there is no directedness of time at the level of fundamental physics, why should there be one in those parts of physics that are allegedly derivative on these fundamental laws?

> So we have a puzzle at least, a paradox at worst: if there is no fundamental directedness in fundamental physics, how does it come to be present in other areas of physics? (Wallace, 2011, 262)

Now, it is well-known that there are exceptions: some processes do exhibit some form of time-reversal asymmetry. Maudlin therefore attacks the very starting point of the debate:

> the laws of physics as we have them . . . are not Time Reversal Invariant. The discovery that physical processes are not . . . indifferent to the direction of time is important and well known: it is the discovery of the violation of so-called CP invariance, as observed in the decay of the neutral K meson. . . . In short, the fundamental laws of physics, as we have them, do require a temporal orientation on the space-time manifold. So the argument . . . collapses at the first step. (Maudlin, 2002, 266–267)

Maudlin's subsequent argument for a fundamental arrow of time does not directly draw on this observation, but the existence of CP-violating laws is arguably meant to motivate his account further.

According to the standard view, however, such CP-violating processes are regarded as of no use for explaining the directedness of time (e.g., North (2011); Loew (2018)). In particular, many do acknowledge that there is some empirical evidence for a time asymmetry in certain physical processes, but then argue that this is a macroscopically irrelevant effect of certain "exotic . . . particles" (Loew, 2018), namely neutral K mesons, that are no constituents of ordinary material objects. Hence, the verdict is that CP violation is irrelevant for the problem of time's directedness.

In this paper I shall push back against such quick dismissals of CP violation as a potential source of explanation for the arrow of time and argue that CP violation should be taken more seriously for conceptualising time in physics. Here is why in a nutshell. There are essentially two arguments for the irrelevance claim that CP violation does not play any role in explaining the directedness of time: (1) as a tiny effect at the level of subatomic particles it is unclear how CP violation can be responsible for large scale macroscopic behaviour, and (2) CP violation does only affect a few specific subatomic particles that do not constitute ordinary matter (i.e., they are "exotic").

In its generality, the first argument, which I dub the *scale objection*, is hardly convincing without further elaboration. It is fair to assume that it is entirely uncontroversial that there are plenty of large-scale effects that can be traced back to the behaviour of subatomic particles. However, there are versions of this argument that have more force.

The central idea of the second argument, which I dub the *universality objection*, seems convincing at first: if CP violation is an effect of very specific exotic particles only, how should it affect ordinary material objects and their processes (in time) at all? Accordingly, why should CP violation have anything to do with time? My reply, in short, is that CP violation *does not* only affect exotic particles. CP violating processes are the result of a central aspect of our best theory of fundamental particle physics that ultimately affects all ordinary matter, as I shall argue. Thus, the second argument is based on a false premise.

Here is the plan of this paper in more detail. First, I argue that accounts à la Maudlin (2007) that posit an intrinsic fundamental direction of time are not convincing and instead propose to utilise recent results from work on the dynamical approach to relativity theory. I then briefly review some basics of fundamental particle physics and give five reasons why CP violation is not just some minor exotic effect, but a fundamental aspect of our best physical theory that deserves more attention. I continue that it is especially a refined understanding of Brown and Pooley's dynamical approach which makes more explicit why CP violation is relevant for the problem of time's directedness.

2 Four options

Prima facie, there are essentially four options for giving an account of time's directionality: one may try to draw on (1) asymmetric laws, (2) asymmetric boundary conditions, (3) asymmetric features of spacetime, and (4) metaphysics.

Options (2), (3), and (4) may be viewed as some kind of *ad hoc* explanation, in the sense that they are not accompanied by a genuine physical explanation as in option (1). So one might say that option (1) is generally preferable. When trying to explain some fundamental structure of the physical world, we typically inform our metaphysical inferences by consulting our best theories of fundamental physics. We seek a metaphysical account that is *based on* a physical explanation – tacitly presuming that a genuine physical explanation will typically be a dynamical explanation that refers to the dynamical evolution of the system captured in some law of nature.

In the case of the apparent directionality of time, however, the received view takes it that fundamental physics can*not* provide an explanation, because *all* fundamental laws of physics are argued to be time reversal invariant. Accordingly, it is supposed to follow that the fundamental physical laws simply do not distinguish the future direction from the past direction in the dynamics.

Similarly, option (3) is dismissed since relativity theory itself does not introduce any distinction between future light-cones and past light-cones – such a distinction may only be introduced by adding some matter field (e.g., Castagnino et al. (2003); Castagnino and Lombardi (2009); Bartels and Wohlfarth (2014)).

Accordingly, the standard view argues that positing any asymmetry in time, i.e., any distinction between the future direction and the past direction can appeal to neither fundamental laws, nor the intrinsic nature of spacetime. Instead the directionality of time is taken to be a result of contingent facts – contingent asymmetric boundary conditions – about how matter is distributed in spacetime. This has been dubbed the *past hypothesis*. By help of the past hypothesis, we can then say, for example, that the directionality of time is nothing but the direction in which entropy increases.[1]

Is this convincing? In the literature we do find proponents of alternative views. For example, Maudlin (2007) – in line with option (3) – posits a fundamental time direction, conceptualised as engraved into a fundamental structure of physics, namely spacetime. Essentially, Maudlin offers an anti-formalistical argument: the fact that an entity or property (here, time's direction) is not present in the standard formalism does not imply that it is not part of the physical world. Inspired by Maudlin, Loew (2018) also posits a fundamental and primitive time direction, but – in line with option (4) – spelt out in metaphysical terms.

Against these non-reductive accounts, the aim of this paper is to develop a reductive explanation of the directionality of time – in line with option (1) – that resists simply positing a time direction, but more closely connects to physics. But let me first briefly discuss the accounts by Maudlin and Loew.

3 Grounding time's arrow

Tim Maudlin is among the most outspoken critics of the standard view. Against the standard view, Maudlin (2007) proposes to posit a primitive, fundamental and intrinsic time direction. Maudlin is after an explanation for why entropy was lower in the past. Why is it, Maudlin asks, that typicality arguments only work correctly in one direction of time? The idea then is, essentially, to introduce a *production relation*. According to Maudlin, the atypicality of the backward evolution of microstates is a result of there being a "fact about which states produce which" (Maudlin, 2007, 134). The assumption of an intrinsic direction of time – which is mathematically represented by an orientation of the spacetime manifold – is supposed to give rise to these facts about which states produce which: "earlier states produce later states" (Maudlin, 2007, 134). In turn, the production relation is supposed to explain the thermodynamic asymmetry.

However, Maudlin does not really spell out his essentially anti-formalistical proposal (which posits an entity despite its apparent absence from the standard formalism) in detail. In particular, it remains unclear how the intrinsic direction

---

[1] To be a bit more specific: the standard explanation of the thermodynamic asymmetry – that is, entropy tends to increase towards the future but not towards the past – appeals to typicality: entropy increases towards the future because given a macrostate most of the compatible microstates evolve towards higher entropy. However, typicality alone is not sufficient to explain the thermodynamic asymmetry. Typicality arguments are independent of the temporal direction: "having evolved from a state of higher entropy in the past . . . is . . . as typical for a system as evolving into a state of higher entropy in the future" (Loew, 2018, 486). To explain the thermodynamic asymmetry, we need to either posit that or explain why entropy was lower in the past. This is done by the asymmetric boundary condition, i.e., the past hypothesis.

of time gives rise to such a production relation and how the production relation helps to explain, for example, the thermodynamic asymmetry (Loew, 2018, 487).

In a recent paper Loew (2018) presents a *metaphysical* solution to these problems. Amongst others, Loew proposes how we may understand the posit of an intrinsic directionality of time such that it connects to the production relation. Loew argues that the intrinsic directionality of time is best understood in analogy with grounding – i.e., an asymmetric, extra-modal type of determination relation that gives rise to relative fundamentality (Loew, 2018, 488). The idea is, roughly, that as an asymmetric grounding-type determination, the intrinsic directionality of time trumps the potentially symmetric necessitation relation of the laws. This entails that the earlier states are metaphysically more fundamental than the later states (that exist in virtue of them), and underwrites a notion of production. In this way, we can think of Loew's proposal as offering a *metaphysical* account of Maudlin's production relation and, hence, the thermodynamic asymmetry. Note, however, that Loew, unlike Maudlin, does not openly oppose the standard explanation of the thermodynamic asymmetry. Loew's metaphysical strategy is intended to *improve* the standard posit of the past hypothesis, not replace it.

Now, some may not be satisfied. In particular, one might wonder whether there is any further sense in which physics could underwrite the apparent directionality of time. Otherwise, one might worry that Loew's metaphysics-heavy explanation, which is arguably based on our everyday conception of time only, is poorly justified or even in danger of contradicting physics. Also, how does the proposed relative fundamentality of earlier states (or the absolute fundamentality of the initial states) fare with other notions of fundamentality? For example, one might contest that the producer is metaphysically more fundamental than the product. Essentially, the question is whether it is possible to be metaphysically more neutral, i.e., to shift the explanatory burden from metaphysics back to physics. Or, to put it less critically: whether there are other (additional) ways to justify such a proposal and to explain the directionality of time.

4 Engraving topology

As mentioned, Maudlin (2007) does offer some remarks on what positing a direction of time could mean precisely, namely positing that the fundamental manifold has an orientation (Maudlin, 2007, 118). However, Loew (2018) is right to call for a more explicit proposal that is linked to the production relation. This can be found in Maudlin's subsequent work on topology (see Maudlin (2014, 2015a,b)). Maudlin can be read as backing up his account of the direction of time by a particular understanding of topology. Maudlin's taking seriously the possibility of an intrinsic directionality of time may then be understood best in light of his 'line-based topology'. This is a non-standard axiomatisation of topology based on fundamental line elements instead of the standard topology's fundamental open sets. At first, this is just a reconceptualisation of standard topology. The interesting feature, however, is that the fundamental line elements have to be conceptualised as fundamentally *directed* line elements ('one-way streets') to be able to reproduce all possible topologies of standard open-set topology one by one. In this way, Maudlin argues, directed line-based topology reveals that most standard topologies actually have an instrinsic directionality that is hidden in the standard formulation (Maudlin, 2015b). Obviously, it is here where the link to time as a directed path between spatiotemporal events appears.

Theories like special relativity (SR) and general relativity (GR) which are often said to have 'spatialised' time – time is just another dimension – start from some ready-made topology that manifestly treats time as just another dimension. But why, asks Maudlin (2015b), should a four-dimensional manifold that locally looks like Euclidean space be the correct space to represent spacetime? Maudlin argues that his line-based topology makes obvious why this is misleading: the standard four-dimensional manifold allows for many lines that do not correspond to anything physical. The geometry of this four-dimensional space is too rich, most of it is not needed to represent physically possible worldlines – in fact, physics practice excludes such lines (like space-like worldlines) as 'unphysical' (Maudlin, 2015a).

So, essentially, what Maudlin does is engraving a directionality into the topology represented by a manifold and identifying this topological directionality with time. Since we typically do not try to further explain topological structure, but just start from some fundamental manifold, this is a forceful option. Indeed, what physical structure should ever be able to explain or ground manifold structure? The manifold is usually viewed to be a good candidate for a fundamental entity.[2]

Here is Maudlin's (2015a) conclusion: The essence of a line is a linear order among spacetime points or events at spacetime points. The question is then, what physical structure could generate a directed linear order among events. Maudlin (2015a) takes it that it is the relation of 'before and after' that should be identified as conveying this linear order.

It is not obvious, however, that this non-reductive account of the directionality of time is without problems. In particular, it must be explained in more detail how its mathematical and metaphysical motivation connect to physics. Arguably, one also needs to investigate how a topological orientation translates to picking out a preferred time direction.[3] It seems that Maudlin would still need something like Loew's proposal.

And, anyways, why is the engraving topology account by Maudlin better than simply excluding certain mathematical excess structure as unphysical after the fact – as the standard view has it – or putting in a direction via introducing a vector field, for example? Ultimately, this is just as well inscribing the desired feature into the fundamental ontology without further arguments from physics. One could argue that this is completely fine as long as such posits are consistent with science. Still, one might want to pursue a less prescriptive and more

---

[2] This connects to a recent debate on how to dynamically account for topological structure (see Norton (2008) and, against this, Menon (2019) and Linnemann and Salimkhani (2021)).

[3] This is similar to Loew's first worry regarding Maudlin (2007).

natural explanation based on the dynamical laws of physics (and a dynamical understanding of them).

5 No time feelers either

More importantly, however, we might contest that what Maudlin offers is actually providing an (appropriate) explanation in the first place. Essentially, the worry is that an intrinsic directionality of spacetime is isolated from having any bearing on the material world. Maudlin's proposal may seem appropriate and handy for our (mathematical) *description* of directed processes, but the processes themselves do not seem to have any access to the posited directionality: matter particles do not have time feelers. To solve this problem, the asymmetry in (space)time needs to be accompanied by or grounded in an asymmetry in the matter content.

Famously, Harvey Brown (2005) draws attention to a similar question. Concerning the so-called geometrical–dynamical debate in the philosophy of spacetime he asks: "[w]hat is geometry doing here – codifying the behaviour of free bodies in elegant mathematical language or actually explaining it" (Brown, 2005, 23–24)? The proponent of the geometrical approach opts for the latter. Spacetime is taken as fundamental and as explanatory of free body motion. Taking spacetime as explanatory in this sense seems to amount to taking the geodesics as "ruts or grooves in space-time which somehow guide the free particles along their way" (Brown, 2005, 24). This is arguably reminiscent of Hermann Weyl's *Führungsfeld* (guiding field).[4] Advocating a similar view as Weyl, Nerlich (1976) argues that particle motion is explained by the "shape" of spacetime, because a particle has no other way of 'knowing' how to move, "[i]t has no antennae to tell it where other objects are" (Nerlich, 1976, 264; as cited in Brown (2005, 24)). Brown concedes that "there is a prima facie mystery as to why objects with no antennae should move in an orchestrated fashion. That is precisely the pre-established harmony, or miracle" (Brown, 2005, 24). Still, Brown, pushes back:

> it is a spurious notion of explanation that is being offered here. If free particles have no antennae, then they have no space-time feelers either. How are we to understand the coupling between the particles and the postulated geometrical space-time structure? (Brown, 2005, 24)

In the case of SR, Brown concludes that spacetime geometry is merely a *codification* of free body motion – essentially, because he deems the geometrical explanation not only obscure, but also redundant:

> At the heart of the whole business is the question whether the spacetime explanation of inertia is not an exercise in redundancy. ... It is non-trivial of course that inertia can be given a geometrical

---

[4] See Weyl (1970, VI).

> description .... But what is at issue is the arrow of explanation. The notion of explanation that Nerlich offers is like introducing two cogs into a machine which only engage with each other. It is simply more natural and economical – better philosophy, in short – to consider absolute space-time structure as a codification of certain key aspects of the behaviour of particles (and/or fields). (Brown, 2005, 24–25)

As already put forward above, this issue reappears with respect to the directionality of time. Even if there is an intrinsic directionality of spacetime, how is it supposed to be explanatory of the temporal behaviour of material bodies? How would they 'know' about the direction, given that they certainly do not have time feelers either? It seems that any intrinsic temporal asymmetry of spacetime needs to be either accompanied by or grounded in an appropriate asymmetry in the dynamical behaviour of matter fields.

Indeed, Maudlin himself did make use of this interdependence of spacetime and matter fields to boost the plausibility of his own account: for him, the existence of CP-violating phenomena (which originate in time-asymmetric dynamical laws) implies that spacetime requires an orientation to be able to support such time-asymmetric laws.

This draws attention to the fact that besides Loew's and Maudlin's accounts there is still option (1) left for further exploration. In the following, I intend to exploit that option and show – *pace* the received view – that CP violation in fundamental particle physics is a wrongfully neglected candidate for explaining the arrow of time.

6 CP violation

To get started, recall that in particle physics C, P, and T are discrete symmetry transformations. C denotes charge conjugation, which essentially means that a particle is transformed to its antiparticle. More specifically, an incoming particle X with 3-momentum $\vec{p}$ and spin vector $\vec{S}$, transforms under C to its complex conjugate $\bar{\mathrm{X}}$ as follows: $\mathrm{C}\ |\mathrm{X}; \vec{p}, \vec{S}\rangle_{\mathrm{in}} = |\bar{\mathrm{X}}; \vec{p}, \vec{S}\rangle_{\mathrm{in}}$. The parity transformation P flips the orientation of the three spatial coordinates, such that $\mathrm{P}\ |\mathrm{X}; \vec{p}, \vec{S}\rangle_{\mathrm{in}} = |\mathrm{X}; -\vec{p}, \vec{S}\rangle_{\mathrm{in}}$. And time reversal transformation T can be understood as flipping the time direction. There is a long debate on how to conceptualize time reversal transformations precisely (see Albert (2000), Earman (2002), Peterson (2015), and Roberts (2017)). In this paper I refer to the standard conception of particle physics where a state transforms under T as follows: $\mathrm{T}\ |\mathrm{X}; \vec{p}, \vec{S}\rangle_{\mathrm{in}} = |\mathrm{X}; -\vec{p}, -\vec{S}\rangle_{\mathrm{out}}$. Note that T also changes an incoming to an outgoing state. One can also combine the transformations, for example to charge conjugation parity transformation, short: CP transformation, which yields $\mathrm{CP}\ |\mathrm{X}; \vec{p}, \vec{S}\rangle_{\mathrm{in}} = |\bar{\mathrm{X}}; -\vec{p}, \vec{S}\rangle_{\mathrm{in}}$. In all standard quantum field theories the combined transformation CPT, which yields $\mathrm{CPT}\ |\mathrm{X}; \vec{p}, \vec{S}\rangle_{\mathrm{in}} = |\bar{\mathrm{X}}; \vec{p}, -\vec{S}\rangle_{\mathrm{out}}$, is an invariant symmetry transformation, i.e., the result is identified with the original state: $|\bar{\mathrm{X}}; \vec{p}, -\vec{S}\rangle_{\mathrm{out}} \equiv |\mathrm{X}; \vec{p}, \vec{S}\rangle_{\mathrm{in}}$. Notably, applying each transformation twice to some particle state does not change the state, i.e., $\mathrm{C}^2 = \mathrm{P}^2 = \mathrm{T}^2 = 1$.

Now, the question is whether physical processes are invariant under these transformations. For example, if all physical processes were invariant under parity, then physics would not distinguish between left- and right-handedness. While most processes are invariant under these transformations, some involving the weak force are not. For example, a positively-charged pion – which is composed of an up quark and an anti-down quark ($\mathrm{u\bar{d}}$) – always decays to an anti-muon and a left-handed neutrino ($\pi^+ \to \mu^+ + \nu_{\mathrm{L}}$).[5] This violates parity symmetry. So there are physical processes that distinguish between left- and right-handedness.

Similarly, neutral K-meson decay is observed to violate CP symmetry.[6] The decay rates of the CP-transformed and the original process are slightly different. Notably, via the CPT theorem, which essentially follows from Lorentz invariance and roughly states that any standard quantum field theory preserves CPT symmetry, CP violation implies T violation. So, on the face of it, there are physical processes that distinguish between orientations in time.

Typically, however, the relevance of CP violation in kaon decays for the directionality of time is questioned. This is for two reasons: (i) the effect is tiny – call this the scale objection – and (ii) apparently a feature of specific particles only, namely neutral K-mesons, that do not constitute ordinary material objects (e.g., Loew (2018)) – call this the universality objection. Essentially, the reasoning is that to explain a phenomenon as large and universal as the asymmetry of time, the explanans has to be large and universal, as well.

In the following, I intend to push back against such verdicts. Here are five reasons why CP violation cannot be so quickly dismissed as a curious and minuscule empirical observation that has no bearing on the arrow of time.

First, addressing the universality objection, CP violation is not observed only for neutral kaons. Since 2001 at the latest, it has been known that CP violation is not merely an effect in the Kaon sector. CP violation has also been observed in several B-meson decays and other processes involving the weak force. Moreover, direct T violation was observed independently of the assumption of the CPT theorem in 2012 by the BABAR Collaboration.

But the universality objection can be challenged further, which provides the second reason: CP violation is not only experimentally observed, but is sufficiently well understood theoretically. In particular, we know that it is a key feature of our best theory of fundamental particle physics, the Standard Model. Here is why. According to the Standard Model (and our experimental data, of course), there are three quark families. There being (at least) three quark families has an important implication for what is called quark mixing.[7] For three quark families, quark mixing is described by a unitary 3×3 matrix – the CKM

---

[5] The primary decay mode is a leptonic decay into an anti-muon and a muon neutrino, i.e., $\pi^+ \to \mu^+ + \nu_\mu$. The decay into the lighter positron and an electron neutrino $\pi^+ \to e^+ + \nu_e$ would offer a larger phase space, but is strongly helicity suppressed.

[6] A neutral K-meson is composed of a linear superposition of down and strange quarks, namely $d\bar{s} - \bar{d}s$.

[7] Quark mixing is essentially a result of the quarks having non-vanishing mass due to the Higgs mechanism, such that the quark mass eigenstates are generally different from quark interaction eigenstates.

matrix. The CKM matrix necessarily contains an irreducible complex phase. It is this complex phase which gives rise to CP-violating processes. In other words, CP violation is a consequence of the fact that there is an irreducible complex phase in the CKM matrix of the Standard Model. Thus, CP violation is a direct consequence of there being (at least) three interacting quark families with non-vanishing mass. If there were only two quark families, the corresponding quark mixing matrix would be real rather than complex: there would be no CP-violating complex phase. This shows that CP violation is a fundamental aspect of our world in a very robust sense.

Third, in its generality, the scale objection is hardly convincing without further elaboration. It is fair to assume that it is entirely uncontroversial that there are plenty of large-scale effects that can be traced back to the behaviour and properties of elementary and subatomic particles – take the temperature of a gas, for example. This is precisely what reductive explanations or part–whole explanations are usually about. After all, already the CP-violating meson decays are traced back to their constituent quarks.

However, there are versions of this argument that have more force. The phenomenon of decoherence, which screens off quantum effects from the classically describable world, or the robustness of higher-level physics against certain changes in the underlying micro-structure are indicative of how the relevance of CP violation might be confined to the level of subatomic particles. To avoid the problem of spelling out how CP violation cuts across the different levels from the scale of fundamental particles to the macroscopic scale, I shall therefore ultimately not propose to directly ground large-scale irreversible processes in CP violation, but argue how CP violation grounds a spatiotemporal orientation that can be used to align other, arguably independent arrows of time with it. This is also because CP and T violation do *not* yield an arrow of time in the sense of irreversibility. CP- and T-violating processes are *reversible*, it is only that the reversed process occurs at a different rate. So it is unclear how CP and T violation should be able to directly ground irreversible processes as in cases of smashed eggs.

Still, and here is the fourth reason, unlike parity violation, CP violation does have well-known large-scale explanatory significance: CP violation is crucial for explaining the matter–antimatter asymmetry of the universe (Sakharov, 1967). Arguably, this emphasizes the relevance of CP violation. In particular, it directly runs against the scale objection and can even be used to itself establish a global time direction for the universe. In this paper I shall not opt for this global time direction, since I take the evolution of the matter–antimatter abundance not as a fundamental arrow of time itself, but merely as an observable consequence that is derivative on the more fundamental CP-based arrow of time.

There remains an important and well-known caveat, though, that might reinforce some version of the scale objection: we know that the CP-violating phase in the CKM matrix of the Standard Model does not suffice to explain the observed matter–antimatter asymmetry (Sakharov, 1967). This is precisely why physicists are searching for additional CP-violating phases in extensions to the Standard Model. Theoretically, there are several candidate sectors where such additional CP-violating phases are expected: most importantly in the neutrino sector[8] and in supersymmetric extensions of the Standard Model.[9] So the fact that the CKM phase does not suffice to explain the matter–antimatter asymmetry should not be understood as a problem *per se* for CP-based explanations.

Fifth, the universality objection is severely contested by our understanding of vacuum fluctuations in quantum field theory and, more concretely, by our theory of hadron constitution: while a hadron's quantum numbers are determined by the so-called valence quarks, these valence quarks carry just one half of the hadron's total momentum. This is because every hadron also consist of, most importantly, gluons and so-called sea quarks. Sea quarks are virtual quark–antiquark pairs ($q\bar{q}$) that are continuously created and destroyed.[10] They are created when a gluon splits, and destroyed when two sea quarks annihilate to produce a gluon. Via this sea of quarks and gluons, essentially all ordinary matter participates at CP violation in the quark sector. Similarly, due to vacuum fluctuations even "empty" spacetime is not isolated from such processes.

To summarise, CP violation – and hence T violation – is not merely a peculiar feature of some exotic particles, but an integral and irreducible part of our best physical theory that affects, amongst others, (processes in) all baryonic, i.e., ordinary matter and is needed to explain the matter–antimatter asymmetry.

7 The dynamical–geometrical debate revisited

Nevertheless, the question remains how exactly CP and T violation become relevant to the directionality of time. In particular, it is important to recall that these symmetry violations do *not* yield an arrow of time in the sense that the respective processes are irreversible. Such processes are *reversible*, it is only that the reversed process occurs at a different rate. Accordingly, such asymmetries cannot directly ground irreversible processes as in cases of smashed eggs.[11]

Therefore, I shall propose a different route by connecting CP violation in the dynamics of fundamental matter fields directly to an orientation of spacetime itself.

Recall that also Maudlin claims that the existence of CP-violating phenomena implies that spacetime itself requires an orientation to be able to support such time-asymmetric laws. Such a claim can arguably draw on the well-known principles by John Earman which state that dynamical and spatiotemporal symmetries need to match in any given theory. Here are Earman's principles (see Earman (1989, 46)):

---

[8] The observation of neutrino oscillation indicates that neutrinos are massive, such that mixing occurs. As there are (at least) three neutrino families, the mixing matrix will generally contain a non-trivial complex phase as well.

[9] Theoretically one would also expect CP violation for the strong force, however, this has been experimentally excluded, which is known as the strong CP problem.

[10] Sea quarks may still hadronize to on-shell baryons or mesons.

[11] In fact, even if such processes were irreversible, it seems unclear how this would translate to paradigmatic cases of irreversible macroscopic processes, like the behaviour of an ideal gas, which seems to be independent of these properties of the constituents of the gas.

(E1) any dynamical symmetry of T is a spacetime symmetry of T, and

(E2) any spacetime symmetry of T is a dynamical symmetry of T.

Any violation of these principles would either result in spacetime's having undetectable structure (violation of (E1)),[12] or a situation where the dynamics can measure geometrical structure that does not exist (violation of (E2)).[13] Granted that violating these principles is problematic and that one should demand that they hold, the question is why they should hold – Maudlin does not provide an explanation, but merely posits that given certain time-asymmetric laws also spacetime needs to be oriented. Why (E1) and (E2) should hold is precisely what the recent geometrical–dynamical debate in the philosophy of spacetime is about. While the standard geometrical approach holds that the spacetime metric somehow explains the dynamical symmetries, the dynamical approach by Harvey Brown and Oliver Pooley opposes this view.

Centrally, Brown and Pooley[14] argue that the property of a field to act as the spacetime metric that is surveyed by material rods and clocks is not an intrinsic property of that field – as the standard geometrical view seems to claim. Rather, this chronogeometricity is said to be a property that a metric field only has in virtue of the dynamical symmetry properties of those matter fields that make up the material rods and clocks. Only if the dynamical symmetry properties of the matter fields coincide with the symmetry properties of a candidate metric field, does a metric field have chronogeometricity such that it is the metric field.

In the case of special relativity, the proponent of the dynamical approach therefore advocates a form of relationalism: the coincidence of the symmetry properties of the Minkowski metric and the matter field dynamics is explained by the fact that the Minkowski metric is ontologically reducible to the Lorentzian symmetry properties of the matter field dynamics.

If the metric field and the matter fields are to be regarded as ontologically independent entities, as the geometrical approach has it, it is unclear why the symmetry properties should coincide. Accordingly, this coincidence must be regarded as a "miracle", as an unexplained empirical fact: "As a matter of logic alone, if one postulates spacetime structure as a self-standing, autonomous element in one's theory, it need have no constraining role on the form of the laws governing the rest of the theory's models." (Brown and Pooley, 2006, 84). Without additional constraints, it is unclear why the dynamical symmetry properties of the matter fields should not be completely different than the symmetry properties of the metric field.

Accordingly, problem cases can be constructed (Read et al., 2018). Regarding general relativity, for example, the Einstein's field equations alone do not

---

[12] If E1 is violated, then the dynamics has certain symmetries that are not symmetries of spacetime as well. In other words, spacetime has *more structure* than the dynamics. This means that there are spacetime structures that are undetectable by the dynamics.

[13] If E2 is violated, then spacetime has certain symmetries that are not dynamical symmetries. In other words, the dynamics has *more structure* than spacetime. This means that the dynamics can measure spacetime structures that do not exist.

[14] See Brown and Pooley (2001; 2006) and Brown (2005).

put any constraints on the matter field dynamics, that is they admit not only Lorentzian but also, for example, Galilean matter fields.

However, the case of general relativity presents a challenge for the dynamical approach as well. This is because the proponent of the dynamical approach to general relativity agrees that the metric field g is a fundamental entity, as the geometrical view has it. Hence, also the proponent of the dynamical approach has to explain why the symmetry properties of the matter fields should coincide with those of g, such that g obtains its chronogeometricity. The proponent of the dynamical approach does so by referring to the strong equivalence principle, which states that all laws of physics are locally Lorentz-invariant. So g obtains its chronogeometricity due to the empirical fact that the strong equivalence principle holds: the strong equivalence principle fixes the dynamical symmetry properties of the matter fields so that they coincide with the local symmetry properties of g.

Due to this recourse to an unexplained empirical fact – the equivalence principle – the dynamical approach to GR is explanatorily weaker than in the special-relativistic case. In this sense, the dynamical approach to GR is less successful. In particular, the dynamical approach is no longer preferable to the geometrical approach in terms of explanatory strength. In fact, Read (2020) argues that the dynamical approach to GR is indistinguishable from any tenable geometrical approach.

Now, it has recently been argued that this shortcoming of the dynamical approach to GR can be fixed by the spin-2 theory of gravity, which yields an ontological reduction of g to matter field dynamics (see Salimkhani (2020b)). In brief, Salimkhani demonstrates that spin-2 theory provides us with a non-geometrical derivation of the Einstein field equations and can be understood as a fixed-field formulation of GR that reduces GR to a special-relativistic theory of an interacting massless spin-2 field. Put in terms of the famous God metaphor (e.g., Barnes (2013)): If God had created a Lorentzian spin-2 field and the other Lorentzian matter fields, she would have created a world described by general relativity with g as the effective metric field.

This ontological reduction of the metric field explains the coincidence of the symmetry properties of the matter fields and the metric field, and thus its chronogeometricity. In particular, the dynamical approach has to accept only one "miracle" as unexplained and is thus preferable to the geometrical view.

But precisely because metric fields like the Minkowski metric or g are ontologically reducible to matter field dynamics, it is not merely the universal dynamical symmetry properties of the matter fields which are inherited by the metric field. Rather, the matter field dynamics pass on *all* dynamical properties. In particular, the matter field dynamics equip the metric field with an orientation, *if at least some of the matter field dynamics violate CP symmetry.*

Already Brown has occasionally pointed out the following: the spacetime fundamentalist, who thinks that the spacetime metric determines the dynamical symmetry properties of the matter fields, should be surprised why some matter fields violate parity, for example, although the respective metric field does not. Take special relativity: according to the symmetry properties of the

Minkowski metric there should be no parity violating matter field dynamics. If the Minkowski metric is a fundamental entity that determines the dynamics of matter fields, then why should the Minkowski metric only transfer one of its symmetries, namely Lorentz symmetry, but not parity symmetry etc.? Turning around the ontological dependence relation between the metric and the matter fields solves this problem. Then the respective derivative metric field inherits *all* properties of the fundamental matter field dynamics. In other words, one only obtains the Minkowski metric, *if all matter fields have Lorentzian dynamics and do not violate parity symmetry etc.*

So, if all matter fields have Lorentzian symmetry properties and in addition are not parity violating etc. – for example in a world in which there are only electrons, positrons, and electromagnetic interactions, i.e., photons – the effective spacetime is flat Minkowski spacetime. If the world contains gravitons in addition, then spacetime is curved, as is described by GR. And if there are fields that violate CP (for example the fields associated to the weak force), this curved spacetime is equipped with an orientation.[15,16]

On this view, it is the contingent field content that decides whether a world has a spacetime that is time-symmetric or time-asymmetric – only spacetimes that do not exclusively ontologically depend on time-symmetric matter field dynamics exhibit an orientation. To be able to empirically access this orientation, we need to use the CP-violating dynamics.

Notably, obtaining a global orientation seems to work only for sufficiently well-behaved spacetimes or regions of spacetime, namely globally hyperbolic spacetimes. But this is in a sense automatically fulfilled, since spin-2 theory is only strictly equivalent to GR for globally hyperbolic spacetimes – all empirically relevant solutions to the Einstein Field Equations are globally hyperbolic (Wald, 1984, 202).

8 Conclusion

I have argued that CP violation is a fundamental aspect of our best physical theory that should not be dismissed when it comes to explanations of time's directedness. In particular, according to Earman, such dynamical symmetry properties need to be reflected in the symmetry properties of spacetime. For making this manifest, I have proposed to utilise recent results regarding the dynamical approach to relativity theory which suggest that the spacetime metric is ontologically reducible to the matter field dynamics. In turn, the symmetries match. Depending on the concrete matter field dynamics, the metric field is then equipped with an orientation, such that a direction of time can be conceptualised.

---

[15] As already argued in Salimkhani (2020a, 181, footnote 113).

[16] At first sight, this orientation may be conceptualised as a *local* orientation if the processes themselves are understood to ground it. However, since the properties of these processes are universally fixed, the local orientations imply a global orientation. Alternatively, we may as well understand the orientation as a *global* orientation from the start, if we take it to be grounded in the (plenist) fields.

I take it that the dynamical perspective thereby makes more precise and more pressing a conjecture that can already be formulated by spacetime substantivalists like Maudlin, who argues that since some laws of nature are time-reversal asymmetric (referring to CP violation), the laws require some intrinsic time asymmetry – and, hence, also spacetime itself requires some orientation in order to support such laws. The dynamical view is most apt for making this precise, because it dynamically explains this additional structure instead of just putting it in by hand – as Maudlin does it.[17] Furthermore, due to the ontological reduction we do not need to allude to an unexplained adaption between spacetime and laws and vague concepts of spacetime 'supporting' such laws. After all, this is at the heart of the debate on the dynamical approach. It is the ontological reduction that does away with certain otherwise unexplained miraculous facts.

Acknowledgements

I thank Emily Adlam, Andreas Bartels, Thomas Blanchard, Maren Bräutigam, Natalja Deng, Andreas Hüttemann, Niels Linnemann, and Martin Voggenauer for helpful discussions and comments on earlier versions of this paper. I also thank the audience of the Philosophy of Time Society Group Meeting at the Eastern Division Meeting of the APA "in" New York City (online) in January 2021 for valuable feedback. This work was supported by the German Research Foundation (DFG, grant number: FOR 2495).

References

David Albert. *Time and Chance.* Harvard University Press, Cambridge, MA, 2000.

Elizabeth Barnes. Emergence and fundamentality. *Mind*, 121(484):873–901, 2013. DOI: https://doi.org/10.1093/mind/fzt00110.1093/mind/fzt001.

Andreas Bartels and Daniel Wohlfarth. *The directedness of time in classical cosmology*, pages 1–15. 2014. DOI: https://doi.org/10.1142/9789814596640$_0$00110.1142/9789814596640$_0$001.

Harvey R. Brown. *Physical relativity: Space-time structure from a dynamical perspective.* Oxford University Press, Oxford, 2005.

---

[17] Similarly, the dynamical perspective might shed light on the issue that a mere direction of time does not suffice to establish a 'passing' of time (cf. Maudlin (2007)). Just because space may be equipped with a preferred direction, doesn't make space 'passing'. If 'passage' of time is understood in terms of some 'production' of 'new' states out of previous states, the dynamical picture might have something to contribute. After all, spacetime is conceived as derivative on matter field *dynamics*, which one could link to a process-based ontology.

Harvey R. Brown and Oliver Pooley. The origins of the spacetime metric: Bell's Lorentzian pedagogy and its significance in general relativity. In Craig Callender and Nick Huggett, editors, *Physics Meets Philosophy at the Plank Scale*, pages 256–272. Cambridge University Press, Cambridge, 2001.

Harvey R. Brown and Oliver Pooley. Minkowski space-time: A glorious non-entity. In Dennis Dieks, editor, *The Ontology of Spacetime*, volume 1 of *Philosophy and Foundations of Physics*, chapter 4, pages 67–89. Elsevier, Amsterdam, 2006.

Mario Castagnino and Olimpia Lombardi. The global non-entropic arrow of time: From global geometrical asymmetry to local energy flow. *Synthese*, 169(1):1–25, 2009. DOI: https://doi.org/10.1007/sll229-009-9495-y10.1007/sll229-009-9495-y.

Mario Castagnino, Olimpia Lombardi, and Luis Lara. The global arrow of time as a geometrical property of the universe. *Foundations of Physics*, 33(6): 877–912, 2003.

John Earman. *World Enough and Space-Time*. MIT Press, Cambridge, Massachusetts, 1989.

John Earman. What time reversal is and why it matters. *International Studies in Philosophy of Science*, 16(3):245–264, 2002.

Niels Linnemann and Kian Salimkhani. The constructivist's programme and the problem of pregeometry. PhilSci-Archive: http://philsci-archive.pitt.edu/id/eprint/2003520035, 2021. Preprint.

Christian Loew. Fundamentality and time's arrow. *Philosophy of Science*, 85: 483–500, 2018.

Tim Maudlin. Thoroughly muddled McTaggart: Or, how to abuse gauge freedom to generate metaphysical monstrosities. *Philosophers' Imprint*, 2(4):1–23, 2002. DOI: https://doi.org/2027/spo.3521354.0002.0042027/spo.3521354.0002.004.

Tim Maudlin. *The Metaphysics within Physics*. Oxford University Press, Oxford, 2007.

Tim Maudlin. *New Foundations for Physical Geometry. The Theory of Linear Structures*. Oxford University Press, Oxford, 2014.

Tim Maudlin. I—Tim Maudlin: Time, Topology and Physical Geometry. *Aristotelian Society Supplementary Volume*, 84(1):63–78, 2015a. ISSN 0309-7013. DOI: https://doi.org/10.1111/j.1467-8349.2010.00186.x10.1111/j.1467-8349.2010.00186.x.

Tim Maudlin. Geometrical structure and the direction of time. Yale University, Franke Program in Science and the Humanities, 21 April 2015, URL: `https://www.youtube.com/watch?v=O-ZU4NP1N_0`, 2015b. Online; accessed 19 March 2020.

Tushar Menon. Algebraic fields and the dynamical approach to physical geometry. *Philosophy of Science*, 86(5):1273–1283, 2019. DOI: https://doi.org/10.1086/70550810.1086/705508.

Graham Nerlich. *The Shape of Space*. Cambridge University Press, Cambridge, 1976.

Jill North. Time in thermodynamics. In Craig Callender, editor, *The Oxford Handbook of Philosophy of Time*, pages 312–350. Oxford University Press, Oxford, 2011.

John D. Norton. Why constructive relativity fails. *The British Journal for the Philosophy of Science*, 59(4):821–834, 2008. DOI: https://doi.org/10.1093/bjps/axn04610.1093/bjps/axn046.

D. Peterson. Prospects for a new account of time reversal. *Studies in History and Philosophy of Modern Physics*, 49:42–56, 2015.

James Read. Explanation, geometry, and conspiracy in relativity theory. In Claus Beisbart, Tilman Sauer, and Christian Wüthrich, editors, *Thinking About Space and Time: 100 Years of Applying and Interpreting General Relativity*, volume 15 of *Einstein Studies*, pages 173–205. Birkhäuser, Basel, 2020. PhilSci-Archive: http://philsci-archive.pitt.edu/id/eprint/1525315253.

James Read, Harvey R. Brown, and Dennis Lehmkuhl. Two miracles of general relativity. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 64:14–25, 2018. DOI: https://doi.org/10.1016/j.shpsb.2018.03.00110.1016/j.shpsb.2018.03.001.

Bryan W. Roberts. Three myths about time reversal in quantum theory. *Philosophy of Science*, 84:315–334, 2017.

A. D. Sakharov. Violation of CP-Invariance, C Asymmetry and Baryon Asymmetry of the Universe. *Zh. Eksp. Teor. Fiz. Pis'ma*, 5:32–35, 1967.

Kian Salimkhani. *On the Fundamentality of Spacetime in General Relativity and Approaches to Quantum Gravity*. PhD thesis, Rheinische Friedrich-Wilhelms-Universität Bonn, 2020a.

Kian Salimkhani. The dynamical approach to spin-2 gravity. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 2020b. DOI: https://doi.org/10.1016/j.shpsb.2020.05.00210.1016/j.shpsb.2020.05.002.

Robert M. Wald. *General Relativity.* The University of Chicago Press, Chicago, 1984.

David Wallace. The arrow of time in physics. In Heather Dyke and Adrian Bardon, editors, *A Companion to the Philosophy of Time*, pages 262–281. Whiley-Blackwell, West Sussex, 2011.

Hermann Weyl. *Raum, Zeit, Materie.* Springer-Verlag, Berlin/Heidelberg, 6 edition, 1970.