



Degrees of freedom

Pieter Thyssen¹ · Sylvia Wenmackers¹

Received: 12 September 2019 / Accepted: 20 May 2020
© The Author(s) 2020

Abstract

Human freedom is in tension with nomological determinism and with statistical determinism. The goal of this paper is to answer both challenges. Four contributions are made to the free-will debate. First, we propose a classification of scientific theories based on how much freedom they allow. We take into account that indeterminism comes in different degrees and that both the laws and the auxiliary conditions can place constraints. A scientific worldview pulls towards one end of this classification, while libertarianism pulls towards the other end of the spectrum. Second, inspired by Hoefer, we argue that an interval of auxiliary conditions corresponds to a region in phase space, and to a bundle of possible block universes. We thus make room for a form of non-nomological indeterminism. Third, we combine crucial elements from the works of Hoefer and List; we attempt to give a libertarian reading of this combination. On our proposal, throughout spacetime, there is a certain amount of freedom (equivalent to setting the initial, intermediate, or final conditions) that can be interpreted as the result of agential choices. Fourth, we focus on the principle of alternative possibilities throughout and propose three ways of strengthening it.

Keywords Free will · Libertarianism · Physical theory · Auxiliary conditions · Principle of alternative possibilities · Nomological determinism · Statistical determinism · Indeterminism · Block universe · Phase space

1 Introduction

At least since the advent of Newtonian mechanics, an important aspect of the free-will debate has revolved around the tension between freedom of will and nomological determinism. Our focus in this article will be on the apparent irreconcilability

✉ Pieter Thyssen
pieterthyssen@gmail.com

Sylvia Wenmackers
sylvia.wenmackers@kuleuven.be

¹ Centre for Logic and Philosophy of Science, Institute of Philosophy, KU Leuven, Kardinaal Mercierplein 2, 3000 Leuven, Belgium

between libertarian free will,¹ which is often considered as demanding more of free will than compatibilism, and nomological determinism in the context of a block universe, with fixed events throughout the entire four-dimensional spacetime. Despite our focus on libertarianism, we assume the closure of physics throughout and remain firmly grounded in a naturalist outlook on the world.

In the remainder of this section, we state the principle of alternative possibilities and the challenges from determinism and indeterminism. In Sect. 2, we propose a classification of indeterminism in scientific theories. In Sect. 3, we take our cue from Carl Hoefer, who emphasized the importance of initial or intermediate conditions in the context of the free-will debate. This leads us to a more fine-grained classification that is based both on laws and on *auxiliary* conditions. In Sects. 4 and 5, we elaborate on the compatibilist account of Hoefer and show how it can be adapted into a libertarian account by combining it with ideas on level dependence from recent work by Christian List. We briefly conclude in Sect. 6.

1.1 Principle of alternative possibilities

One important requirement for libertarian free will is that an agent could have acted otherwise under exactly the same conditions.² Following Harry Frankfurt (1969), we call this ability to do otherwise the *principle of alternative possibilities*.^{3,4}

Definition 1 *Principle of alternative possibilities (PAP)* The action of an agent is free only if the agent could have acted otherwise under exactly the same conditions.

Libertarian free will is challenged in at least two ways:

1. Challenge from determinism;
2. Challenge from indeterminism.

¹ For an overview of the free-will debate, and the differences between compatibilism and libertarianism, see Kane (2011). The conversation between Kane, Fischer, Pereboom, and Vargas in *Four Views on Free Will* (Fischer et al. 2007) provides another introduction to the different views that are currently on the philosophical market.

² Not everyone agrees with this requirement. Source incompatibilists, for instance, argue that as far as moral responsibility is concerned, the ability to do otherwise is irrelevant (Frankfurt 1969; Vihvelin 2018). What *is* required for moral responsibility, in their view, is that an agent is the ultimate source of her actions. Leeway incompatibilists, in contrast, *do* take PAP to be crucial for free will (see also Pereboom 2003). In this paper, we do not aim to defend PAP as an important ingredient for free will: our analysis starts from the assumption that it is. For those who disagree, the subsequent analysis may be irrelevant.

³ In fact, the original term is ‘alternate’ possibilities. Although PAP was originally formulated in the context of discussions on moral responsibility, this will not be our focus here.

⁴ Even for philosophers who take PAP to be a *necessary* condition for libertarian free will, it is not a *sufficient* condition. Most libertarian authors include one or more additional principles as crucial for free will. List (2019), for instance, identifies three principles: (1) intentional agency, (2) alternative possibilities, and (3) causal control. Nonetheless, the sole focus of our discussion below will be PAP.

We look at both challenges in turn, with particular attention to the tension with PAP (and variants thereof) and gauge whether they can be dealt with in a satisfactory way. The ultimate goal of this paper is to develop an account of libertarian free will that answers both of these challenges.

1.2 Challenge from determinism

In what follows, we adopt John Earman's definition of determinism in terms of possible worlds (1986; see also Roberts 2006). Let \mathfrak{W} be the class of possible worlds and consider the subclass of physically possible worlds $W \subset \mathfrak{W}$. These are the worlds in which the same laws of physics apply as in our world.

Definition 2 *Determinism* The world $w \in W$ is deterministic if and only if for any $w^* \in W$ and any time t , if w and w^* agree on the complete physical state at t , then they agree on the complete physical state at all other times t' .⁵

Given a theory with deterministic laws, each possible history (*i.e.*, a path in the associated phase space) corresponds to a possible world. In the context of specific physical laws, such as those from the theories of relativity, possible worlds can be regarded as 'possible block universes'.

Note that Earman's definition does not say *why* the world is deterministic. It does not specify the mechanism by which determinism is supposed to work. In this paper, we consider a specific form of determinism: *nomological determinism*, which is the view that the state of the world at time t , along with the laws of nature, determines the state of the world at all other times t' .

The jury is still out on whether our world is nomologically deterministic or not. Observe that it is a *metaphysical* thesis rather than a *physical* one: it requires a comparison across (physically) possible worlds, whereas experiments only tell us results about the actual one.⁶

Assuming nomological determinism and a full description of the physical world at a given time t , possibly before you were born, it is fully determined what the state of the world will be at any other point in time t' , including those during and after you made your choices. So, PAP is incompatible with determinism. In summary, the challenge from determinism reads:

Challenge from determinism

(P1) Libertarian free will requires PAP;

⁵ If w and w^* only agree on the complete physical state at the times $t' > t$, then w is said to be *futuristically* deterministic. Likewise, if w and w^* only agree on the complete physical state at the times $t' < t$, then w is said to be *historically* deterministic.

⁶ Observe also that determinism is an *ontic* thesis, whereas predictability is an *epistemic* one. Determinism is therefore compatible with practical unpredictability: even in a deterministic world, we may not be able to practically carry out the retrodictions or predictions (for instance, because we do not know the laws of nature, or because we fail to attain complete knowledge of the state of the world at a certain time).

- (P2) Nomological determinism rules out PAP;
 (C) Free will and nomological determinism are incompatible.

The argument purports to show that there is no room for libertarian free will in a deterministic world.

1.3 Challenge from indeterminism

One common response to the challenge from determinism points to microphysical theories that are indeterministic, like quantum mechanics (but see Sect. 2). Another common response aims to argue that determinism at the microscopic physical level is compatible with indeterminism at the macroscopic agential level (see Sect. 5.1). Both strategies have in common that they regard indeterminism as indispensable for PAP and thus for free will. After all, in an indeterministic world, the state of the world at time t is compatible with multiple states at time t' . This seems to allow alternative possibilities, and hence to save PAP. But does it open up enough degrees of freedom to allow libertarian free will? Not all authors agree.

1.3.1 No half-way house

Hard incompatibilists, such as Derk Pereboom (2001, 2005), defend the view that libertarian free will is neither compatible with determinism *nor with indeterminism*. Indeterminism (or randomness), they claim, is not the same as deliberate choice. The problem, in the words of Arthur Eddington (1939), is that there is no ‘half-way house’ between determinism and randomness. Either the world is deterministic, in which case we are prisoners of a deterministic chain. Or the world is indeterministic, in which case everything depends on sheer chance and fluke events.

This traditional challenge from indeterminism goes back to David Hume (1888) who considered it “impossible to admit of any medium betwixt chance and absolute necessity”. Hume’s fork, however, targets the principle of agential causation. Even if the world is indeterministic, it is not clear how an agent can *use* this indeterminism to gain freedom. Although we do not dismiss this worry, it is not the target of our current discussion. Here, we want to focus on a different kind of *challenge from indeterminism*, or what we will call the *challenge from statistical determinism*, which challenges PAP but has received comparably little attention in the literature.

1.3.2 Challenge from statistical determinism

Indeterminism implies that there are alternatives, in the sense of multiple allowed future states consistent with a present state. Although this yields PAP, we claim that it is not yet sufficient for a strong sense of libertarian free will. To explain this, we propose to strengthen PAP as follows:

Definition 3 *Strengthened principle of alternative possibilities (PAP*)* The action of an agent is free only if the agent could have acted otherwise under exactly the same conditions, unbounded by probabilistic constraints.

To motivate this move, suppose we accept, say, quantum mechanics as the source of indeterminism in the physical world. These forms of indeterminism carry a very limited sense of freedom: they come with a preset collection of alternatives associated with stable statistical properties (such as predictable measurement averages). In other words, they present us with a form of *statistical determinism*.

While PAP was challenged by determinism, PAP* is challenged by statistical determinism. In summary, the challenge from statistical determinism reads:

Challenge from statistical determinism

- (P1) Libertarian free will requires PAP*;
- (P2) Statistical determinism rules out PAP*;
- (C) Free will and statistical determinism are incompatible.

The argument purports to show that there is no room for libertarian free will in a probabilistic world. Libertarian free will and (probabilistic) indeterminism are incompatible notions.

1.3.3 Social, historical, and statistical determinism

The view that statistical regularities may threaten free will stems from the nineteenth century, although it has remained a minority position.⁷ As discussed by Ian Hacking (1983), the Belgian scientist and pioneer of statistical methods in the social domain, Adolphe Quetelet, already commented on the predictability of the numbers of yearly births and deaths and on the terrifying exactness of how crime rates are reproduced. This view extended to the moral domain, leading to a view of social determinism. Although Quetelet did not deny the existence of free will, other authors did. The English historian Henry Thomas Buckle (1865) defended historical determinism: he regarded the existence of stable statistics in the context of the social sciences as incompatible with human free will.

William James (1890), in his *Principles of Psychology*, came closest to the challenge from statistical determinism (although his comment is also related to the traditional challenge from indeterminism with respect to agential causation). As James observed, a brain exploiting some form of indeterminism in its decision-making process would be “like dice thrown forever on a table.”

Unless they be loaded, what chance is there that the highest number will turn up oftener than the lowest? [...] *Can consciousness increase its efficiency by loading its dice?* Such is the problem. Loading its dice would mean bringing

⁷ See Mueller (2017) for a recent review and Saka (manuscript, and references therein) for a contemporary defense.

a more or less constant pressure to bear in favor of those of its performances which make for the most permanent interests of the brain's owner; it would mean a constant inhibition of the tendencies to stray aside. (1890, 140, emphasis in original)

After the advent of quantum mechanics, some authors developed a similar view of statistical determinism in relation to this microphysical theory. For instance, Christof Koch (2012, p.104) wrote:

Given our current interpretation of QM [namely one that assumes collapses happen], a Popper-Eccles mind could exploit this idiosyncratic freedom. The mind would be powerless to change the probabilities, but it could decide what happens on any one trial. The mind's action would always remain covert, sub rosa, for if we considered many trials, nothing out of the ordinary would take place: only what is expected from natural law. Conscious will would act in the world within the straightjacket of physics. It would be indistinguishable from chance.

Given that both determinism and statistical determinism are incompatible with PAP*, it may seem that we have ruled out libertarian free will entirely, from the armchair. However, we have not yet explored all possible forms of indeterminism. So far, we have silently equated indeterminism with probabilism, as is customary in the literature, but this is not the only option. To broaden the discussion, we propose to consider not just two, but four classes of theories.

2 Classification of theories

We propose a classification of theories, as shown in Fig. 1.⁸ CLASS I theories are deterministic and CLASS II theories are probabilistic. Our main observation is that the latter only covers a subset of all indeterministic theories. In probabilistic theories, all possible outcomes are (assumed to be) known and all possibilities have specific probabilities associated with them. This specification is already suggestive of additional classes of theories, which allow more freedom than CLASS II. CLASS III theories allow for probability gaps (*i.e.*, possible outcomes *without* associated probabilities; Hájek 2003),⁹ but no possibility gaps (*i.e.*, cases in which even the list of possible outcomes is acknowledged to be incomplete), whereas CLASS IV theories also allow for the latter.

This fourfold classification resembles that of the 'Johari window' technique (Luft and Ingham 1955) with (I) known knowns, (II) unknown knowns, (III) known

⁸ For now, the classification we propose applies to candidate (micro-)physical theories, at the most fundamental level. We will broaden its application to other levels in Sect. 5.

⁹ The economist John Maynard Keynes, in his *Treatise on Probability* (1921), argued that numerical probabilities are the exception rather than the rule. In most cases, probabilities are non-numerical and non-quantifiable because they are either unmeasurable or incomparable.

unknowns, and (IV) unknown unknowns.¹⁰ To clarify our proposal, we give examples for each of the classes.

2.1 CLASS I theories

Newtonian physics is traditionally taken to be the paradigm example of a deterministic CLASS I theory. However, without additional conditions, classical mechanics is “not a paradise for determinism; in fact, Newtonian worlds provide environments that are quite hostile to determinism” (Earman 1986). If we understand Newtonian physics as including a condition or a postulate that ensures initial-value problems have unique solutions, then it does indeed serve as an excellent example for a CLASS I theory. (One such condition requires Lipschitz continuity: we return to this in the description of CLASS III.)

2.2 CLASS II theories

Quantum mechanics is often held as the paragon of a probabilistic CLASS II theory. This is only acceptable if we consider the Copenhagen interpretation or a spontaneous collapse interpretation. There also exist deterministic formalisms that are empirically equivalent with quantum mechanics, such as Bohmian mechanics, which belongs to CLASS I.

Also statistical mechanics fits in this category, at least when judged at the macro-level. Its probabilities, however, can be fully reduced to the deterministic micro-level. So, judged at that level, it is a CLASS I theory. (In Sect. 5.1, we will return to this level-dependence.)

2.3 CLASS III theories

A concrete example of a CLASS III theory is provided by non-Lipschitz mechanics: that is, Newtonian mechanics without the constraint that guarantees uniqueness of the solution to initial-value problems. An example of such an indeterministic system was rediscovered by John Norton (2003), who considered an idealized point mass initially at rest on top of a frictionless dome of a particular shape. The initial-value problem admits of a singular solution, where the point mass remains at the apex forever, as well as an infinite family of regular solutions: the mass spontaneously slides down the dome in an arbitrary direction after an arbitrary period of time. Non-Lipschitz mechanics does not supply probabilities for the two types of solutions, nor

¹⁰ CLASS I and CLASS II theories both spell out which possibilities exist (knowns) and, provided sufficient boundary conditions, they single out a single possibility (known) or provide probabilities over multiple possibilities (which of those will pertain is unknown), respectively. CLASS III and CLASS IV theories both deal with multiple possibilities (unknowns), but the possibilities themselves may all be known or not, respectively.

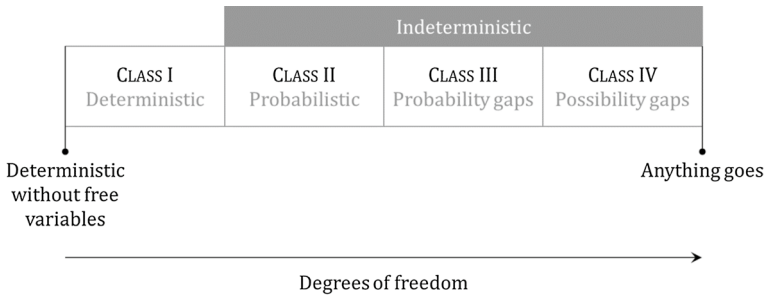


Fig. 1 Our proposed classification of physical theories

for the two variables in the family of regular solutions, thus providing us with a beautifully simple CLASS III example.¹¹

2.4 CLASS IV theories

CLASS IV theories have possibility gaps. They allow for radically new possibilities, not specified by the theory, to be realized. Such theories may specify some possible outcomes, and even some relative probabilities of a subset of possible events, but at least under some circumstances they allow for radical openness regarding possible outcomes. The most radical CLASS IV theory has the form: ‘anything can happen’. The other CLASS IV theories specify possible outcomes (with or without associated relative probabilities) for some but not all circumstances. They allow for possibility gaps, but only in specific cases, which the theory specifies.

Finding examples for CLASS IV is harder than for previous classes. Formalizing a theory that allows for radical uncertainty is a thorny issue: it does not seem to square well with the notion of a phase space (see “Appendix 1”), which has to be specified beforehand.¹² Another way of phrasing the difficulty is by observing that we are dealing with theories that aim to state their own incompleteness, which may well be impossible.¹³ As CLASS IV candidates, we may consider the most speculative theories from natural science (Kragh 2014) as well as metaphysical theories that allow for strong emergence. Strong emergence allows higher level properties to be incomputable from a full description in terms of lower level properties, and has been proposed for various phenomena, including chemical reactions, life, and consciousness

¹¹ Norton’s (forthcoming) infinite lottery logic model for pocket universes, which explicitly rejects the assignment of probabilities, also belongs to CLASS III.

¹² As such, the problem is similar to formalizing the notion of a catch-all hypothesis, which occurs in a probabilistic setting (CLASS II); see, e.g., Wenmackers and Romeijn (2016).

¹³ Indeed, this speculation pans out if we formalize it. Sentence φ is possible for theory T if and only if $\not\vdash_T \neg\varphi$ (with \vdash the deducibility relation). A theory with a possibility gap would simply be an incomplete theory and it is impossible for a theory to state its own incompleteness: it follows from Löb’s theorem that, if $\vdash_T \neg\text{Prov}_T \neg\varphi$ (with Prov_T a provability predicate for T), then $\vdash_T \varphi$ and, hence, T is not incomplete. We are grateful to Jan Heylen for suggesting this addition.

(O'Connor and Wong 2015). Because of the radical openness needed for emergentism, this seems to require a CLASS IV theory.¹⁴

2.5 The need for a CLASS III theory

Hard incompatibilists have claimed that libertarian free will is incompatible with both determinism *and* indeterminism, as outlined in Sect. 1.3 (Pereboom 2005). We suggest it is more accurate to call this an incompatibility between free will on the one hand, and deterministic CLASS I and probabilistic CLASS II theories on the other hand. That is, rather than ruling out libertarian free will, the challenges from determinism and statistical determinism push libertarians towards theories with CLASS III or even CLASS IV indeterminacies. CLASS III theories introduce probability gaps; CLASS IV theories introduce, in addition, possibility gaps. The question therefore becomes what degree of freedom is minimally needed to realize libertarian free will.

Historically, CLASS III theories have often been invoked by libertarians who hoped to embed their view in a scientific worldview. In the nineteenth century, important contributions to this debate were made by James Clerk Maxwell, Balfour Stewart, Antoine Augustin Cournot, and Joseph Boussinesq, as reviewed by Marij van Strien (2015). Boussinesq (1879) is of particular interest here, since he was well-aware of the challenges from determinism and statistical determinism, and therefore focused on non-Lipschitz mechanics (with examples similar to that of Norton, 2013) as a possible source of human freedom.¹⁵ This is clearly better than quantum mechanics, which is only CLASS II and hence already in conflict with PAP*. However, even a CLASS III theory seems to limit the actions of free beings severely, since they can only choose from a predetermined menu of alternatives, as encoded by the phase space.

At first glance, this may seem insufficient for a strong sense of libertarian freedom which allows creativity and truly unforeseeable actions. For example, if you stand in a coffee bar, and you can only order what is on the menu, then this may make you doubt your free will. In order to be truly free, you may think, you should also be able to walk out, or to ask the barista to make you something that is not on

¹⁴ Depending on the interpretation of levels, however, this may yield a CLASS IV theory that does not apply to the microphysical level. We discuss this level-dependence in Sect. 5.1.

¹⁵ Besides advocating this CLASS III theory to beat the challenge from statistical determinism, Boussinesq also stipulated a non-physical, non-mechanical influence which he called the *principe directeur* and which could direct choices, thereby filling the gap between physical reality and the mathematical description of CLASS III systems (Mueller 2015; Bordoni 2017). The postulation of a *principe directeur* ran against the doctrines of materialism and positivism, and may be explained by Boussinesq's sympathy for the spiritualistic philosophy developed by Cousin, Vacherot, Caro and Janet, among others. According to the spiritualist doctrine, matter and motion are insufficient to provide a complete description of nature. Boussinesq thus attempted to save the freedom of will by postulating a spiritualistic metaphysics. The same has happened in more recent times. As Kane observed, many libertarians have posited "transempirical power centers, immaterial egos, noumenal selves outside of space of time, unmoved movers, uncaused causes and other unusual forms of agency or causation" in order to answer the challenge from physicalism, "thereby inviting charges of obscurity or mystery against their view" (Fischer et al. 2007, p. 9).

the menu (a filter coffee with added sugar, say). This thought might motivate a further strengthening of PAP*, which could only be fulfilled by a CLASS IV theory.

We argue, however, that this may be asking too much. After all, the fact that you cannot turn in a seahorse right now should not make you doubt your free will. So perhaps the requirement is not having unlimited but *sufficiently many* possibilities, which are not bounded by probabilities. We use this additional desideratum to strengthen the PAP* criterion further:

Definition 4 *Strong principle of alternative possibilities (PAP**)* The action of an agent is free only if the agent could have acted otherwise under exactly the same conditions, with sufficiently many alternatives unbounded by probabilistic constraints.

PAP** is satisfied in CLASS III theories. As such, our classification has helped us to clarify the traditional dynamics of the debate, with CLASS II as the gulf between determinists and libertarians.

However, we want to achieve more than mapping the terrain of old debates. One of our goals is to answer the challenge from determinism without giving up determinism or PAP**. To do this, we should return our focus to the case where the micro-level laws remain fully deterministic. But how can we get PAP** if the world is governed by CLASS I deterministic laws? To overcome this seemingly unbridgeable divide, we need a new map—one which adds an additional dimension to the debate. We offer a suggestion for such a map in the next section.

3 Classification of theories with auxiliary conditions

Our previous classification can be refined by taking into account an important insight from Hofer (2002).

3.1 The importance of auxiliary conditions

According to Eugene Wigner (1995, p. 699), Isaac Newton's greatest accomplishment was "the sharp distinction between initial conditions and laws of nature".¹⁶ Given that the initial conditions are often "quite arbitrary", the "prime focus" of physics has been the discovery of new laws of nature. But it is the laws of nature *together with* the initial conditions of the system which determine the behavior of that system. The laws alone are not sufficient. The initial conditions, after all, describe the state of the system at a definite time. Without a specification of the initial positions and velocities of the planets, for instance, Newton's laws are mute.

¹⁶ To the initial conditions and laws of nature, Wigner also added invariance principles, which he regarded as metalaws (laws which the laws of nature have to obey).

In Sect. 1.2, we have seen why determinism seems to preclude free actions. But Hoefer (2002) has reminded us that even deterministic laws, which leave no freedom for how one thing leads to another, do not constrain how things are at a given instant. After all, the laws do not determine their own auxiliary conditions—i.e., initial, intermediate or final conditions or boundary conditions.¹⁷ In that sense, they still allow a lot of freedom and the same laws of nature are compatible with multiple temporal evolutions. Hoefer, in other words, finds freedom not in the laws of nature, but in the auxiliary conditions. We use this idea here to refine our classification; we return to other aspects of Hoefer's (2002) work in Sect. 4.2.

3.2 Refined classification of theories

In Fig. 1, we classified theories into four classes. In the description, we focused on the laws. Now, taking into account Hoefer's (2002) message about auxiliary conditions, we present a refinement of our classification: as shown in Fig. 2, it still has four classes, but for each CLASS, we make an additional distinction according to the origin of the indeterminism. CLASSES Ia, IIa, IIIa, and IVa are classified as such due to indeterminism in the laws; CLASSES Ib, IIb, IIIb, and IVb are classified as such due to indeterminism in the initial values.¹⁸

For example, Bohmian mechanics is Class Ia and Class IIb, so Class II overall, which is the same as orthodox quantum mechanics (Class IIa). This example shows that only the classes in Fig. 1 are essential if we care only about distinguishing physical theories up to observational equivalence (*cf.* footnote 33).

As a second example, the proposal of Boussinesq (1879) had a law with probability gaps combined with fully specified initial conditions. Hence, it is a CLASS IIIa theory according to our refined classification. Following Hoefer (2002), we stick to deterministic laws but consider non-probabilistic constraints on the auxiliary conditions. In other words, we are looking for a CLASS IIIb theory.

4 The conditional interpretation of PAP

In this section and the next one, we introduce recent work by List (2014, 2019a, b), which we wish to combine in a novel way with further ideas of Hoefer (2002).

A first idea we wish to follow is List's (2019b) distinction between two different interpretations of PAP. In this section, we focus on the *conditional interpretation* of PAP. (We will consider the *modal interpretation* in Sect. 5.)

¹⁷ Hoefer published his ideas in a relatively obscure paper, which have therefore sadly been neglected in the free-will literature, despite their originality and novelty. In her book *How Physics Makes us Free*, Jenann Ismael (2016) has reached similar conclusions independently from Hoefer. The auxiliary conditions are usually taken to be initial conditions, but see Sect. 4.3.

¹⁸ For all $N, M \in \{I, II, III, IV\}$, a theory that is CLASS N a and M b is CLASS $\max\{N, M\}$ according to the previous classification.

Definition 5 *Conditional interpretation* If the agent had tried to do otherwise, then the agent would have succeeded.

4.1 Supervenience and multiple realizability

In statistical physics, macrostates *supervene* on microstates. Hence, whenever a macrostate obtains, there is an underlying microstate. Which microstate? This is hard to answer because macrostates are *multiply realizable*. (See the “Appendix 1” for a review of statistical mechanics.) Following List (2014), in the current context we may think of the macrostate as an *agential state*: the state of the agent as described by the relevant higher-level (macroscopic, psychological) theory. This stands in contrast with the microstate or *physical state* as described by the lower-level (microscopic, physical) theory.

Again following List (2014), the terms we just used can be defined as follows:

Definition 6 *Supervenience* No variation in the agential state is possible without a variation in the physical state.

Definition 7 *Multiple Realizability* Typically, more than one physical state corresponds to a particular agential state; hence, not every variation in the physical state will give rise to a variation in the agential state.

So, crucially, the agential state is (1) supervenient on and fully determined by the physical state and (2) multiply realizable and more coarse-grained than the physical state.

4.2 Apparent indeterminism due to coarse-graining

Now we return to Hofer (2002): he drew his inspiration from statistical mechanics, so he considered deterministic microphysical laws and, like List (2014), he acknowledged the important distinction between micro- and macrostates.

On Hofer’s (2002) view, exactly one microstate is realized at any given time. This implies that the laws are deterministic (CLASS Ia), and the initial conditions are fully specified as well (CLASS Ib). This makes the theory CLASS I overall. In such a theory, there is no room for PAP.

However, according to Hofer (2002), the agent cannot know the exact microstate, but only its coarse-grained, multiply realizable macrostate. The indeterminism which emerges at the coarse-grained macro-level is merely epistemic. Hence, for Hofer, we can only explain our *illusion* of freedom: as far as we know, i.e. up to the macrostate, the present is compatible with multiple futures. But in reality, only one microstate obtains at any given time, which is compatible with only one possible future history.

Regarding historical counterfactuals, for Hofer, the past has to remain the same *macroscopically*, and not “in all its gory *microphysical* detail” (2002, p. 215). Otherwise it would indeed take “a miracle to get the if-had-done-otherwise scenario

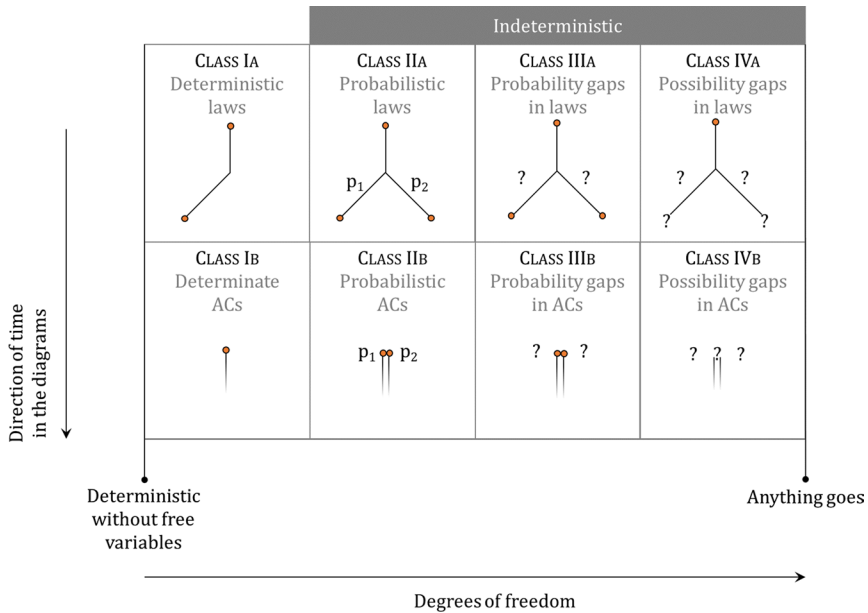


Fig. 2 Refinement of our proposed classification of theories, based on their laws and auxiliary conditions (ACs). The top row shows how different classes can be realized due to different kinds of laws. The bottom row shows how different classes can be realized by different constraints on the ACs. The diagrams show an example with two possible histories, with time running from top to bottom. In the diagrams, p_1 and p_2 indicate the presence of probabilities in the theories; their location indicates whether they pertain to initial conditions or subsequent branches. Question marks indicate the lack of probabilities or specification of states in the theories

started.” By merely keeping the macroscopic past fixed, you “don’t need miracles to postulate various different actions and their likely future consequences” (2002, p. 215). As a result, Hoefer focuses on the conditional interpretation of PAP, as defined at the start of Sect. 4.

The conditional interpretation relies on a counterfactual claim: *if* the agent had been in a different microstate compatible with (and for the agent indistinguishable from) the agent’s macrostate, *then* she would indeed have been successful in reaching a different outcome than the current one. In terms of possible worlds, it makes sense to consider an agent who was determined to type ‘s’ in the actual world, while there exists a nearest counterfactual possible world in which she tries to do otherwise and succeeds in typing ‘z’ instead.¹⁹ (See “Appendix 2” for more details on the typing example in terms of statistical mechanics.)

¹⁹ A complication here is that what the agent deems possible need not coincide with any physically possible world. But the program is to show the plausibility of alternative possibilities, not the accuracy of the agent’s representation of those possibilities.

4.3 Freedom in intermediate conditions

In Sect. 3, we already introduced Hoefer's (2002) proposal to consider the auxiliary conditions as an additional source of freedom, besides what is offered by the laws. We tend to think of the auxiliary conditions as fully specified initial conditions, but this leads to an additional argument in support of the challenge from determinism: the Consequence Argument.

4.3.1 The consequence argument

Clearly, there were no humans or other agents around at the time of the Big Bang to influence the initial conditions by their own free actions. This is also a key assumption in Peter van Inwagen's *Consequence Argument* (1975, 1983, 1989) against free will. In his *Essay on Free Will*, van Inwagen (1983, p. 16) formulated it as follows:

If determinism is true, then our acts are the consequences of the laws of nature *and events in the remote past*. But it is not up to us what went on before we were born; and neither is it up to us what the laws of nature are. Therefore, the consequences of these things (including our own acts) are not up to us (emphasis added).

Hoefer's reply to this argument was twofold: (1) not only do the laws fail to pin down their own auxiliary conditions, (2) they also fail to specify where to draw the boundary. So, instead of a past boundary condition, we can add the present condition to the dynamical rules and regard it as a constraint on the past and future. After all, we know most about the current situation, where we find ourselves—*in medias res*—choosing and acting. Why don't we take this direct observation more seriously? This is exactly what Hoefer (2002, p. 221) proposed: "the direction of determination (and, for most free actions, correct explanation) is from your choices to the ways the physical world can be—both toward the past and the future."

With his proposal, Hoefer denies a key premise of the Consequence Argument: our current actions do determine the conditions at the Big Bang—albeit very partially and indirectly. While this viewpoint may be unusual, common practice does not suffice to disallow it. As Hoefer (2002, p. 208) emphasizes, determinism gives us "logical relations of determination, not a unique *temporal* relation of determination" (emphasis in original). In that sense, one is free to choose where to draw the determining slice.²⁰ Hoefer argues that we might just as well draw the determining slice at the inside of the block universe, where we are now; hence his term: "freedom from the inside out".

²⁰ Or rather, *when* to draw it: the slice refers to a temporal slice or moment in time. In relativity theory, a moment in time is represented by a three-dimensional Cauchy hypersurface. Specifying the auxiliary conditions on this surface completely determines the entire past and future. Technically, a Cauchy hypersurface is an achronal set of spacelike separated events whose domain of influence (or Cauchy development) covers the entire spacetime manifold, both to the past and the future of this surface. See Hawking and Ellis (1973) for details.

However, this does not prevent someone else from drawing the determining slice at the Big Bang, and reasoning from then on forward in time. From that perspective, the initial conditions at the Big Bang predetermine everything, including all our thoughts, intentions, and actions. So, the challenge from determinism still applies.

4.3.2 Starry sky of free choices

Even if we would privilege the inside-out determination, this would not be enough to allow freedom at different points in time. We try to reconstruct Hoefer's view on this and elaborate on it. First, he remarks that relativity theory has taught us that local events do not influence spatially removed events instantaneously. Instead, there are finite domains of influence. This allows us to think of determining events (free actions) as spread out across the bulk of the block universe, rather than as being all located in one temporal slice. As such, it is consistent to associate at least part of the information in the auxiliary conditions with the free choices of agents that happen in the course of the dynamical evolution that is governed by deterministic laws.

On this view, freedom is on a strict budget: there is only a sparse amount of free choices available—the equivalent of one three-dimensional hypersurface—compared to the bulk of events in the four-dimensional volume of the block universe. This is unavoidable if all the freedom has to come from the auxiliary conditions. However, these free choices would be scattered within the block, yielding a starry sky of free choices in the middle of the block.

To conclude this section, recall that Hoefer defended a *compatibilist* account of free will, not a libertarian one. The aim of our paper, however, was to develop a libertarian account of free will that beats the challenge from determinism. Hoefer (2002) doesn't seem to think this is possible, but List clearly thinks otherwise, as we will see in the next section.

5 The modal interpretation of PAP

List (2019b) took the conditional interpretation of PAP to be insufficient for libertarian free will and argued that PAP requires a modal interpretation:

Definition 8 *Modal interpretation* It is *possible* for the agent to do otherwise (there are forks in the road).

Before offering our own proposal, we first review List's (2014, 2019a).

5.1 Emergent indeterminism

List (2014, 2019a) granted that the free-will sceptic might be right in concluding that, at the level of deterministic fundamental physics, there are no alternative possibilities. The mistake, according to List, is to claim that there are no alternative possibilities at all: free will is a higher-level concept, belonging to

the domain of the special sciences. As such, we should look for free will (and the ability to do otherwise) at the macroscopic agential level, not at the microscopic physical level.

5.1.1 Physical and agential levels

List (2019a) contended that the challenge from determinism, as outlined in Sect. 1.2, involves a category mistake due to a mixing of different levels. Premise (P2), after all, is a thesis about *physical* possibility, not *agential* possibility. It pertains to what is nomologically possible on the microscopic level. Yet, when discussing free will, we are interested in *agential* possibility or what is possible on the macroscopic level. The challenge from determinism should thus be modified to read:

Challenge from determinism*

- (P1) Libertarian free will requires PAP;
- (P2) Agential determinism rules out PAP;
- (C) Free will and agential determinism are incompatible.

In this modified form, the challenge from determinism shows that libertarian free will requires indeterminism at the agential level, but not necessarily at the physical level. This opens up a potential new avenue for libertarian free will.

For this, List had to show that determinism at the physical level is compatible with indeterminism at the agential level. We are convinced by his argument, the main points of which we summarized in “Appendix 3”, together with their connection to statistical mechanics, which helps to unify List’s proposal with Hoefer’s (2002) work.

List (2019a) called his account of free will *compatibilist libertarianism*. It is compatibilist because it takes nomological determinism to be compatible with free will and it is libertarian because free will requires agential indeterminism to allow for alternative possibilities—unlike other compatibilist accounts of free will. The source of List’s agential indeterminism is the multiple realizability of a macrostate by microstates: multiple physical microstates can realize one and the same psychological macrostate while leading to different alternative outcomes under the deterministic dynamics.

5.1.2 Two shortcomings and a way out

Although we find List’s work very inspiring, by itself it does not fully succeed. We see at least two shortcomings.

First, the agent must have alternative possibilities in the actual world, not merely in counterfactual situations. In order to have plenty of genuine alternatives, in the sense of PAP**, we need a CLASS III theory, as discussed above. However, because of supervenience, we need it not only at the macro-level (as List has successfully shown) but also at the micro-level. That is, we not only require PAP at the agential

level, but also at the physical level. Otherwise, emergent indeterminism is merely due to epistemic constraints.²¹

Secondly, List (2014, 2019a) did not explain *how* the agent can use this scope to act freely in a deterministic world. List (2019b) discussed agential causation, but he did not really specify how it is supposed to work.

To overcome these two shortcomings, we return to Hoefer (2002), who (1) emphasized the selection of auxiliary conditions by the agent and (2) can be reinterpreted in a way that may bring PAP at the physical level within reach.

5.1.3 A bold step: bundles of microstates

CLASS III theories have probability gaps, which leave room at the physical level for the agential level to act upon to have a real (as opposed to an epiphenomenal or illusory) effect on the microphysical state of affairs. This requires branching in at least some points in the microphysical histories (CLASS II or higher), as well as freedom from statistical regularities to allow these acts to be truly free (so, at least CLASS III, as already mentioned).

According to Hoefer (2002), such libertarian ‘freedom from the inside out’ would require that we could have done otherwise starting from the exact same initial microconditions. But this is impossible if we assume deterministic laws. Assuming CLASS Ia laws, however, we can still introduce CLASS III indeterminacy via the auxiliary conditions (CLASS IIIb).

Indeed, even if one requires PAP (or PAP**) at the physical level, Hoefer’s statistical mechanical account of compatibilist free will leaves room for a libertarian interpretation, if we take a bold step: let us, for a moment, assume that the present is not determined by a single microstate, but by a *bundle* of microstates. The suggestion in other words is to interpret the multiple realizability claim ontically, rather than epistemically. The idea is that a macrostate is not only multiply *realizable*, it is multiply *realized*. Let us now flesh out this proposal. We will do this in two steps. The first step (in Sect. 5.2) is merely a ladder, to be kicked away once the proposal is clear (in Sect. 5.4).

5.2 Pruning bundles

As remarked before, Hoefer takes there to be an actual microstate (unknown to the agent), whereas List (2019b, pp. 91–92) claims that there is no privileged microstate ‘within’ an agential state:

[A]t the psychological level, there is no fact of the matter as to which precise physical state obtains. [...] [T]he higher-level state, at any time, does not

²¹ After all, each microstate corresponds to a different auxiliary condition on which the deterministic laws can act. So, if we consider humans in combination with a fully specified micro-physical state of a supposedly deterministic universe, then no alternatives remain.

determine which underlying lower-level state obtains, and so we cannot treat any one of the lower-level states as the “true” one.

Although we do not attribute the following view to List, one could interpret this ontically by considering the macrostate as a common instant in an equivalence class of microhistories (where we use the term ‘history’ as defined in “Appendix 1”). In other words: we can think of a bundle of histories passing through the given macrostate at a given time. The alternative possibilities correspond to a partition into smaller bundles that pass from the current macrostate to multiple future macrostates. Effecting a choice then amounts to pruning all the other bundles but the chosen one. If the bundle contains infinitely many microhistories to begin with, this process may continue for countably many choices throughout history.

5.2.1 Typing example

To see how this works, consider again Hofer’s typing example mentioned in Sect. 4.2 (now using the notation from “Appendix 2”). At $t = 0$, you choose to type the letter ‘t’. You thereby select the macroregion Γ_{M_t} . Since this macrostate is multiply realized by all the microstates $x(0) \in \Gamma_{M_t}$, we have to keep track of their evolution through phase space. This bundle of trajectories starts off in Γ_{M_t} and ends up in the new region $\phi_1(\Gamma_{M_t})$ at $t = 1$. Due to Liouville’s theorem, $\mu(\Gamma_{M_t}) = \mu(\phi_1(\Gamma_{M_t}))$. The phase space region $\phi_1(\Gamma_{M_t})$ however will typically be *split* over two or more macrostates. Let us assume that it is spread over Γ_{M_s} and Γ_{M_z} .

The situation at $t = 1$ thus leaves you with two choices: typing ‘s’ or ‘z’. Suppose you freely choose to type the letter ‘s’. You thereby select the macroregion Γ_{M_s} , more specifically, the region $\Gamma_{M_s} \cap \phi_1(\Gamma_{M_t})$ corresponding to the overlap between Γ_{M_s} and $\phi_1(\Gamma_{M_t})$. It is as if your free choice has pruned part of the region $\phi_1(\Gamma_{M_t})$, leaving the subregion $\Gamma_{M_s} \cap \phi_1(\Gamma_{M_t})$.

5.2.2 Thinning bundles

The presentation in the previous paragraph was only forward-looking, but we may also reflect on the effect of pruning on the past bundle. If we take the past parts of non-selected histories as cut off as well, choices at an intermediate time thin out the bundle over all of time. This may eventually lead to a thin bundle or even a single microhistory throughout all of time. There are no branching points for this history, and it may be presented as ‘the’ block universe, although it was selected from branching points in bundles of histories—one among many possible block universes.

To get real freedom in this picture, we have to adopt an *ontic* view: the exact microstate is not fixed, only the coarse-grained macrostate is. That is, the initial conditions (at the time of the Big Bang) were only partially fixed, and therefore compatible with multiple futures. Whenever we make a free choice, we add a further constraint which refines the initial conditions.²² The evolution of our universe does

²² On an epistemic reading, we merely become aware of what we choose: this knowledge may also affect our knowledge of the past. On an ontic reading, affecting choices could be considered as a form of cau-

not correspond to a single trajectory in phase space, but to a bundle of trajectories, which is increasingly pruned every time a free choice is made. For a related view, see also Cristi Stoica (2012).²³

To be clear, this idea is *not* endorsed by Hoefer. In fact, this mode of presentation has fallen into the trap that Hoefer (2002, p. 206) called “the unholy marriage” of the eternalist, physical viewpoint, which presents events as a four-dimensional block universe, with our presentist, subjective experience of events, as they unfold right now. But this need not be a fatal objection, for we can rephrase the above proposal without reference to the dynamics of pruning on bundles of histories. Doing so effectively regains Hoefer’s idea of freedom from the inside out, but it now yields a libertarian reading (instead of Hoefer’s own compatibilist reading).

5.3 Principle of action uniqueness—against diverging worlds

Before we kick away our ladder, a warning is in order. Our bundle theory should be contrasted with an account of diverging worlds, in which each possibility gets realized in a different world (*i.e.*, a different block universe), leading to a block multiverse. The latter would be in tension with free will. The tension is with a principle that is rarely spelled out, but which we make explicit here:^{24,25}

Definition 9 *Principle of action uniqueness (PAU)* When an agent successfully carries out an action freely chosen out of a set of mutually incompatible alternatives, only the agent’s chosen action is realized.

Multiverse determinism says you could have done otherwise, *and did so* (in a parallel branch). So, all alternative possibilities were realized. PAU, on the other hand, requires that you could have acted otherwise, *but didn’t*. So only one alternative possibility was realized. We argue that PAU completes PAP, by stipulating that the agent takes one and only one action (in the relevant choice moment). The agent could have, but has not actually, acted otherwise (say, in a parallel world or a different part of the multiverse).

If we want an account of free will that satisfies both PAP and PAU, the above proposal for a diverging-worlds theory is ruled out. PAU requires determinateness of choice outcomes (which are intermediate conditions), not just relative to a branch (*i.e.*, a block universe), but absolutely and overall. The principle is reminiscent of

Footnote 22 (continued)

sation, in which case the proposal in the second paragraph would be bicausal. In other words, choices would have retrocausal effects on the past, as well as causal effects on the future.

²³ Although Stoica (2012) investigated this idea in the context of quantum mechanics, he did consider the deterministic case explicitly: “A deterministic universe can have incompletely determined initial conditions, which can be refined by ulterior choices.”

²⁴ Due to multiple realizability, the realization of a macroscopic alternative is compatible with a bundle of real histories rather than a unique real history. For simplicity, in what follows we assume that unique outcomes for all choices lead to a unique real history. In principle, this may still be a thin bundle instead.

²⁵ Under the header “The irrelevance of forks”, Saka (manuscript) has reached a similar conclusion.

the requirement for definite outcomes in quantum mechanics, which is used by opponents of the many worlds approach.²⁶ Here, too, PAU serves as a uniqueness condition that rules out a multiverse or bundle interpretation.

In other words, PAU requires us to treat the full ensemble of block universes as merely hypothetical, not actual. Within this hypothetical bundle, there is only one real history (past and present); this is similar to the medieval concept of the actual history as a thin red line (Øhrstrøm and Hasle 2015).²⁷

5.4 Block universe regained

Now we are ready to show how Hoefer's freedom from the inside out may be combined with List's compatibilist libertarianism. Our proposal in Sect. 5.2 (the ladder) was CLASS Ia and IIIb at the microphysical level, so CLASS III overall at the microphysical level. Our final proposal is CLASS I at the microphysical level and CLASS III at the agential macro-level.

We can assume that the laws at the micro-level remain fully deterministic (CLASS Ia) and that there is a unique initial condition (CLASS Ib). Like Hoefer, however, we are free to interpret this initial condition as being the net result of all free choices over the course of history (recall Sect. 4.3). These choices cannot be 'seen' at the level of a single microphysical history, which contains no branching points (by the assumption of microphysical determinism). Instead, they are located at the higher level of macrostates of such microphysical histories: each macrostate corresponds to a region of the phase space, which we interpret as a bundle of hypothetical block universes, including all the options we do *not* choose. That is, the macro-, agential histories do have branching points and thus offer real alternatives.

As List and others have shown, higher-level probabilism (CLASS II) is indeed compatible with micro-level determinism (CLASS I). But List may have overlooked the fact that agential choice need not even be governed by chance: it may be as free as CLASS III. The fact that this is a form of non-probabilistic indeterminism becomes especially clear if we regard it, as Hoefer does, as stemming from a freedom equivalent to setting the initial conditions, which is unconstrained by the dynamics of the equations themselves. The freedom of agents thus relies in their ability to select auxiliary conditions at the macro-level. This selection is bounded by possibilities, but unbounded by probabilities. Their range is fixed due to the underlying phase space, where each macrostate is surrounded by only a finite number of other macrostates, each of which corresponds to a possible future action. But there is no additional probability distribution over this range. So, this is CLASS III at the agential level, but not beyond. Although this does not allow as much freedom as CLASS IV, we have argued that this is all we need to be truly free as suggested in our motivation for PAP**.

²⁶ For instance, the requirement is called 'determinate outcomes' by Maudlin (1995).

²⁷ Or, at most, a subset of the full bundle is real: *cf.* footnote 24.

In short, the upshot of this reading is that we have something akin to a CLASS III theory at the agential level, fully compatible with CLASS I at the microphysical level. Now that we have kicked away the ladder of an actual bundle theory, we are back at a unique block universe, just like Hofer. Yet, we have managed to give a libertarian account of free will within a single deterministic universe, just like List, by taking into account the hypothetical ensemble or bundle of block universes.

6 Conclusions

Let us now take stock. In this paper, we have made at least four contributions to the contemporary debate on the freedom of will.

6.1 Classification of theories

Firstly, we proposed a classification of theories (Fig. 1) based on how much freedom they allow. This classification makes explicit at least two points that are often left implicit. It shows that beyond the determinism–indeterminism dilemma, there is a wider range of options. Indeterminism comes in different degrees; we distinguished three. Moreover, our more fine-grained classification (Fig. 2) shows that both the laws and the auxiliary conditions can place constraints that allow more or less freedom. We hope that this classification will help to clarify the debate. In particular, we think it helps to analyze a central tension in the free-will debate: a scientific worldview pulls towards lower-number classes, while libertarianism pulls towards higher-number classes.

Usually the question about human freedom is asked in a categorical, all-or-nothing way. However, our analysis suggests, among other things, that freedom may be a gradable notion rather than a Boolean on/off switch. In our paper, we have looked at the spectrum of theoretical possibilities, hence the ‘degrees of freedom’ in the title of this article (with a nod to its technical meaning in physics). A complementary way to read our proposal is that it exposes a free-will spectrum of sorts: we hope this viewpoint will inspire future work.

While the classification that we developed deals with the amount of freedom that various theories admit, it might be interesting to explore the connection to a different debate: whether freedom applies in different degrees to various living beings. For instance, in his correspondence with Saint-Venant, Boussinesq defended the position that consciousness and free will may apply to animals in various degrees, depending on their complexity (see, for instance, Mueller 2015, p. 628).

6.2 Auxiliary conditions

Secondly, taking our cue from Hofer (2002), we have given a specific interpretation to the auxiliary conditions. In the context of statistical mechanics, an interval of auxiliary conditions corresponds to a region in phase space. We have made explicit that this also corresponds to a bundle of block universes. Whereas the laws specify

possible trajectories in phase space, or possible block universes, they do not determine their own auxiliary conditions. Hence, the auxiliary conditions leave freedom in a direction orthogonal to that of the laws, so to speak. This may leave some room for a form of non-nomological indeterminism within a generally physicalist approach.

6.3 Libertarian free will

Thirdly, we aimed to combine crucial elements from the work of Hoefer and List and attempted to give a libertarian reading of this combination. Our proposal for a combination is as follows: (1) We assume a deterministic microtheory and an indeterministic macrotheory, (2) like List and unlike Hoefer, we require PAP at the agential level, and (3) like Hoefer, we assume agents can use the macroscopic indeterminism. What we have called the ‘starry sky of free choices’ is crucial to make the third element work: throughout spacetime, there is a certain amount of freedom (equivalent to setting the initial, intermediate or final conditions) that can be interpreted as the result of agential choices. As such, the totality of all agential choices in spacetime determines (part of) the auxiliary conditions, from the inside out, as Hoefer would call it.

6.4 Varieties of PAP

Fourthly, our paper focused on one libertarian principle: the principle of alternative possibilities. In examining PAP throughout the paper, we have proposed three ways to strengthen it. In PAP*, we added that the alternatives should not be bound by probabilities, in order to escape statistical determinism. In PAP**, we added that there should be sufficiently many alternatives, such that the agent does not feel restricted by a limited menu. (In passing, we remarked that some libertarians may require even more: that there be no menu at all, allowing truly novel options to arise. We have not developed this here.)

Finally, we added the principle of action uniqueness to PAP (or a strengthened version thereof): PAU stipulates that exactly one of the alternatives is chosen, which is relevant in the context of multiverse or actual bundle theories. Combining PAP and PAU brings out a second central tension in the free-will discussion: the requirement of real alternatives, at a moment when an agent has a choice, of which only one is realized subsequently.

This tension has been part of the free-will debate since antiquity and is closely related to the issue of future contingents. Unsurprisingly, our proposal in the final section of our paper is also similar to a medieval suggestion: that of the actual history (past and future) as a thin red line. Within the formalism of statistical mechanics, this thin red line can be represented as a privileged trajectory in the sample space. And this can also be interpreted as one actual block universe in an infinite bundle of possible block universes.

Acknowledgements We are grateful to two anonymous reviewers for their constructive feedback on an earlier version of this article, which greatly helped us to improve the presentation, and to Stefaan Cuypers for helpful pointers to the literature. PT thanks Gerard Bodifée for inspiring discussions on the nature of free will. We gratefully acknowledge funding for this research project by FWO (Research Foundation—Flanders) through Grant Number G0B8616N.

Funding This work was supported by the Research Foundation Flanders (Fonds Wetenschappelijk Onderzoek, FWO), Grant Number G0B8616N “Time, Causality, and Probability in Quantum Mechanics: Assessing Retro-causal Explanations in Light of Recent Experiments”.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix 1: Review of statistical mechanics

There is no unified and generally accepted formalism of statistical mechanics, instead there are two: the Boltzmannian approach and the Gibbsian approach. For our purposes, the Boltzmannian approach is the most useful one.

Microstates

Consider a dynamic system S of N point-particles. The system can be in a number of states which evolve over time. The set of all possible states of S is called the *state space* of S . In the case of Boltzmannian statistical mechanics, the state of S is provided by the (fine-grained) *microstate* $x = (q, p)$, which specifies the positions $q = q_1, \dots, q_{3N}$ and momenta $p = p_1, \dots, p_{3N}$ of all the particles, and the state space is the $6N$ -dimensional *phase space* Γ .

The change of the system’s microstate over time t is governed by the Hamiltonian equations of motion:

$$\frac{dq_i}{dt} = \frac{\partial H}{\partial p_i}, \quad \frac{dp_i}{dt} = -\frac{\partial H}{\partial q_i},$$

with $H(q, p)$ the Hamiltonian of the system. This induces a *phase flow* $\phi_t : \Gamma \rightarrow \Gamma$ on the phase space, with ϕ_t a one-to-one mapping. For example, if the system starts out in the microstate $x(0) \in \Gamma$ at time $t = 0$, it will trace out a trajectory in phase space under the Hamiltonian dynamics, and be mapped to $\phi_{1-0}(x) = x(1) \in \Gamma$ at time $t = 1$. What is important for our purposes is that the Hamiltonian dynamics, which maps an initial state to its final state, is perfectly deterministic.

Histories

The temporal path of the system through phase space is called its *history*. Formally, the history of S is a map h from the reals (time) into the system's phase space, assigning to each instant of time t a corresponding microstate $x(t)$ of S :

$$h : \mathbb{R} \rightarrow \Gamma; t \mapsto x(t).$$

If the system S starts out in a different microstate, it will trace out a different trajectory in Γ , yielding a different history h' for S .

Notice that the different possible trajectories in Γ represent different possible *block universes*—or different possible worlds in the terminology of our definition for determinism, satisfying the same Hamiltonian laws of evolution. Due to the Hamiltonian dynamics, each of these block universes is completely deterministic. Which one of these possible block universes gets actualized depends on the auxiliary conditions (*i.e.*, on which microstate is selected as initial—or final, or intermediate—state, see Sect. 4.3).

Liouville's theorem

The phase space Γ is further endowed with the *Lebesgue measure* μ . According to *Liouville's theorem*, the Lebesgue measure is preserved under the dynamics of the system. That is, the dynamic evolution of a system will preserve the volume of its initial phase space region:

$$\mu(R) = \mu(\phi_t(R)),$$

for all regions $R \subseteq \Gamma$. The *shape* of R , on the other hand, can change quite dramatically over time.²⁸ In summary, the mathematical framework of Boltzmannian statistical mechanics is the triple $\langle \Gamma, \phi_t, \mu \rangle$.

Macrostates

Now, let M denote the set of all (coarse-grained) *macrostates* M_i ($i = 1, \dots, m$) of the system. Each macrostate specifies the state of the system at a macroscopic level of grain, but not the microscopic configuration of all the point-particles. There is therefore a many-to-one mapping σ from Γ into M , such that for every microstate x

²⁸ This is the case in particular for systems with chaotic dynamics. Informally, deterministic chaos occurs for systems that have regions in their phase space such that nearby points evolve to points that are far removed. This is related to the property of sensitive dependence on initial conditions and the rate of separation is made numerically precise by the Lyapunov exponent. Werndl (2009b) has argued that chaos is best defined in terms of mixing on a subset of the phase space. (Here, mixing is a mathematical notion inspired by the physical notion: see Werndl's paper for the precise definition.) Werndl concludes that "for predicting any event at any level of precision $\varepsilon > 0$, all sufficiently past events are approximately probabilistically irrelevant."

in Γ there is exactly one corresponding macrostate $\sigma(x) = M(x)$ in M , but where a given macrostate M_i in M can be realized by more than one microstate x in Γ .

Every macrostate M_i picks out a particular *macroregion* Γ_{M_i} in phase space:

$$\Gamma_{M_i} := \{x \in \Gamma \mid M(x) = M_i\},$$

such that the different Γ_{M_i} form a partition (or coarse-graining) of Γ . In other words, the different M_i do not overlap and jointly cover Γ :

$$\Gamma_{M_i} \cap \Gamma_{M_j} = \emptyset;$$

$$\bigcup_{i=1}^m \Gamma_{M_i} = \Gamma.$$

With the help of the supervenience mapping σ from microstates to macrostates,²⁹ we can turn every fine-grained history h into a coarse-grained history \mathfrak{h} , with

$$\mathfrak{h} : \mathbb{R} \rightarrow M; t \mapsto M(t) = \sigma(x(t)).$$

Appendix 2: Counterfactuals and Hoefer’s typing example in terms of statistical mechanics

According to Hoefer (2002), when an agent makes a free choice—for instance, the choice to type the letter ‘t’ on her laptop—she thereby selects a particular macrostate, say M_t , corresponding to this coarse-grained action. There is usually an enormous number of microstates x corresponding to each macrostate (multiple realizability). So, which x will be realized by freely choosing M_t ? “When I freely choose to type the ‘t,’” says Hoefer (2002, p. 211), “I do not thereby choose to actualize a *particular* microstate!” (emphasis in original). The best we can say is that “some one of this enormous number of microstate-types shall be, and that is all” (2002, p. 210). This sounds reasonable. After all, when we say that we feel free, we do not mean to imply that we have the power to influence every single atom by our free actions. We merely intend to say that we have the power to shape the world in a very coarse-grained way.³⁰

²⁹ In statistical physics, the supervenience mapping from micro- to macrostates is also crucial to define entropy: the Boltzmann entropy of a microstate depends logarithmically on the number of microstates ‘in’ (that is, consistent with) the corresponding macrostate. To be specific, the Boltzmann entropy of a macrostate M is defined as $S_B(M) = k_B \log [\mu(\Gamma_M)]$, with k_B the Boltzmann constant.

³⁰ The mechanism by which one freely selects one or the other macrostate is left underspecified by Hoefer. Hoefer is skeptical about strong reductionism and does not subscribe to the idea of causal completeness with *upward causation* from the microphysical to the macrophysical. Instead, Hoefer endorses the perspective of *downward causation* from the macrophysical to the microphysical. Thus “my intention to type the letter ‘t’ causes the particular motions experienced by all the atoms in my left forefinger as I type it [...] rather than (for example) the immediately preceding motions of other nearby atoms, or any other such particle-level events” (2002, 201, emphasis added). That is, our choices, thoughts and intentions are

According to Hoefer, PAP is at “the heart of the issue” (2002, p. 214). What is really interesting about Hoefer’s account of free will is that it leaves room for *counterfactual beliefs*. That is the sense in which we take Hoefer’s account to be “much more robust” than other compatibilist accounts (2002, p. 203). “When I type the letter ‘s’ I may think that I could have chosen to type a ‘z’ instead [...]. And I think I could have done so, *with the past being, macroscopically, just the way I know it to be*” (2002, p. 214, emphasis in original).

Here is what we take Hoefer to mean. Consider Fig. 3 (which is explained in more general terms in “Appendix 3”). When you choose to type the letter ‘t’ at $t = 0$, you thereby selected the macroregion Γ_{M_t} . A particular microstate $x(0) \in \Gamma_{M_t}$ was thereby realized at random, or at least beyond your control. Now, under the Hamiltonian dynamics, $x(0)$ will deterministically evolve to $x(1)$ at $t = 1$. And as it turns out $x(1) \in \Gamma_{M_s}$, where Γ_{M_s} is the macroregion corresponding to you typing the letter ‘s’.

However, even though it was predetermined that you would type an ‘s’ after you typed the letter ‘t’, you still have the feeling that you could have typed a ‘z’ instead. What grounds this belief, according to Hoefer, is that there is indeed another microstate $x'(0) \in \Gamma_{M_t}$ which could have been realized instead of $x(0)$, and which evolves to $x'(1) \in \Gamma_{M_z}$ at $t = 1$ with Γ_{M_z} the macroregion corresponding to you typing the letter ‘z’.³¹

Hoefer (2002), like List (2014), thus finds freedom in the multitude of microstates realizing the same macroscopic present, but leading to different futures and world histories. Whereas List modified the interpretation of PAP, such that it can apply in a deterministic world, Hoefer explained why PAP appears to (but does not actually) apply in such a context.

Appendix 3: PAP from multiple realizability

List (2019a) argued that the challenge from determinism shows that libertarian free will requires indeterminism at the agential level, but not necessarily at the physical level. Here, the link with Boltzmannian microstates and macrostates, as introduced in “Appendix 1”, is easily made.³² To be precise, let the agential states correspond to the macrostates M_i and the physical states to the microstates $x = (q, p)$. (1) The macrostates in M supervene on the microstates in Γ . In other words, you cannot change M_i without changing x : this is supervenience (see Definition 6). And, (2) to each microstate corresponds exactly one macrostate, but many distinct microstates can correspond to the same macrostate: this is multiple realizability (see Definition 7).

Footnote 30 (continued)

primary explainers of our physical actions (2002, 207). It thus seems Hoefer is assuming a form of *non-reductive physicalism*.

³¹ Hoefer emphasized that “we are not concerned with the actual past history of the world in all its microscopic detail; that *does*, of course, determine the present” (2002, 216, emphasis in original).

³² Although List does not expound his theory in these terms, it can be easily adopted to a statistical mechanics framework. To the best of our knowledge, List only makes the link with statistical mechanics in the conclusions of his 2014 paper (174): “This echoes the way in which statistical mechanics accounts for the emergence of stochasticity in a deterministic Newtonian world.”.

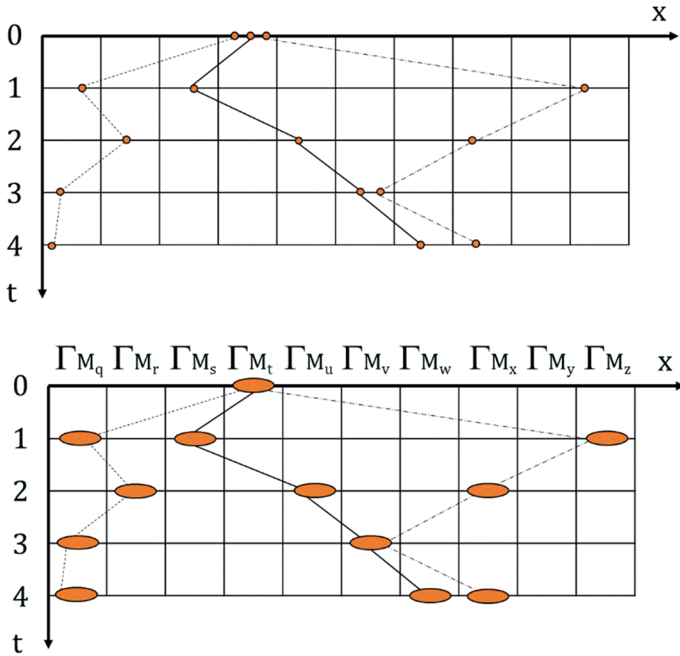


Fig. 3 In these two diagrams, microstate x is represented as a one-dimensional, continuous variable and time t (running from top to bottom) is taken to be discrete. Microstates are indicated by small circles, macrostates by larger ellipses. Histories are indicated by lines between instantaneous states. The diagram at the top shows micro-level determinism; the diagram at the bottom shows that this is compatible with macro-level indeterminism

As List (2014) observed, the multiple realizability of macrostates by microstates makes it possible for the supervenience mapping σ to be such that determinism at the micro-level is mapped to indeterminism at the macro-level. In the context of physics, List was not the first to consider this idea. Charlotte Werndl (2009a) has given examples of observationally equivalent systems that can be modelled by deterministic and stochastic equations.³³

The coexistence of micro-level determinism and macro-level indeterminism can be understood better with the help of Fig. 3 (which is inspired by Fig. 1 in List and Pivato 2015). In the top part of Fig. 3, every circle represents a possible microstate $x \in \Gamma$ of the system at a particular time t . At the start, $t = 0$, the system can be in one of multiple microstates: for definiteness, we have indicated three of them. Each of these possible microstates evolves in time, tracing out a trajectory through phase space, as indicated by the various lines connecting the dots. Each trajectory represents one of the three possible histories h_j ($j = 1 \rightarrow 3$). All histories are deterministic, since no branching occurs at the micro-level.

³³ See also Butterfield (2012) for a terminologically distinct account of how micro-level determinism can be compatible with macro-level indeterminism.

The macro-level supervenes on the micro-level. That is, to each microstate x corresponds a macrostate $M(x)$, as represented by the ellipses in the bottom diagram. Specifically, all microstates lying in the same x -interval are mapped to the same macrostate (multiple realizability), with the ellipses representing the different macroregions $\Gamma_{M(x)}$. To every history h_j at the physical micro-level corresponds a coarse-grained history $\mathbb{h}_j = \sigma(h_j)$ at the agential macro-level. In this case, however, branching points do occur, leading to indeterminism. For instance, although the three histories are the same initially, they diverge at $t = 1$. Clearly then, indeterminism at the macro-level is consistent with determinism at the micro-level, given supervenience and multiple realizability. Call this *emergent indeterminism*.

The reasoning here works in both directions (indeterminism at the macroscale can emerge from determinism at the microscale) and can be iterated: for example, it is consistent to consider determinism at the micro- and macroscale with indeterminism at the mesoscale. This can be iterated indefinitely (List and Pivato 2015, section 9). In addition, it is also possible to consider different classes of indeterminism at different levels. An idea of this kind was present in the work of Quetelet, who presumed that at the level of populations humans are more predictable than at the individual level, without being fully deterministic.

References

- Bordoni, S. (2017). *When historiography met epistemology: sophisticated histories and philosophies of science in French-speaking countries in the second half of the nineteenth century*. Leiden: Koninklijke Brill NV.
- Boussinesq, J. V. (1879). Conciliation du véritable déterminisme mécanique avec l'existence de la vie et de la liberté morale. *Mémoires de la société des sciences de l'agriculture et des arts de Lille*, 6, 25–256.
- Buckle, H. T. (1865). *Histoire de la Civilisation en Angleterre*. Paris: Librairie Internationale A. Lacroix.
- Butterfield, J. (2012). Laws, causation and dynamics at different levels. *Interface Focus*, 2, 101–114.
- Earman, J. (1986). A primer on determinism. In R. E. Butts (Ed.), *University of Western Ontario series in Philosophy of Science*. Dordrecht: Reidel.
- Eddington, A. S. (1939). *The philosophy of physical science*. Cambridge: Cambridge University Press.
- Fischer, J. M., Kane, R., Pereboom, D., & Vargas, M. (2007). *Four views on free will*. Oxford: Blackwell.
- Frankfurt, H. G. (1969). Alternate possibilities and moral responsibility. *The Journal of Philosophy*, 66(23), 829–839.
- Hacking, I. (1983). Nineteenth century cracks in the concept of determinism. *Journal of the History of Ideas*, 44, 455–475.
- Hájek, A. (2003). What conditional probability could not be. *Synthese*, 137, 273–323.
- Hawking, S. W., & Ellis, G. F. R. (1973). *The large scale structure of space-time*. Cambridge: Cambridge University Press.
- Hofer, C. (2002). Freedom from the inside out. In C. Callender (Ed.), *Time, reality & experience*. Cambridge: Cambridge University Press.
- Hume, D. (1888). *Treatise of human nature*. Oxford: Clarendon Press.
- Ismael, J. T. (2016). *How physics makes us free*. Oxford: Oxford University Press.
- James, W. (1890). *The principles of psychology*. New York: Henry Holt and Company.
- Kane, R. (Ed.). (2011). *The Oxford handbook of free will* (2nd ed.). Oxford: Oxford University Press.
- Keynes, J. M. (1921). *Treatise on probability*. London: Macmillan & Co.
- Koch, C. (2012). *Consciousness: confessions of a romantic reductionist*. Cambridge, Massachusetts: The MIT Press.

- Kragh, H. (2014). Testability and epistemic shifts in modern cosmology. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 46, 48–56.
- List, C. (2014). Free will, determinism, and the possibility of doing otherwise. *Noûs*, 48(1), 156–178.
- List, C. (2019a). XII—What’s wrong with the consequence argument: A compatibilist libertarian response. *Proceedings of the Aristotelian Society*, 119, 253–274.
- List, C. (2019b). *Why free will is real*. Cambridge, Massachusetts: Harvard University Press.
- List, C., & Pivato, M. (2015). Emergent chance. *Philosophical Review*, 124, 119–152.
- Luft, J., & Ingham, H. (1955). The Johari window, a graphic model of interpersonal awareness. In *Proceedings of the Western Training Laboratory in Group Development*, Los Angeles: University of California.
- Maudlin, T. (1995). Three measurement problems. *Topoi*, 14, 7–15.
- Mueller, T. M. (2015). The Boussinesq debate: reversibility, instability, and free will. *Science in Context*, 28, 613–635.
- Mueller, T. M. (2017). Can you put free will into an equation? The debate on determinism and mathematics at the end of the nineteenth century. *European Journal for the History of Economic Thought*, 24, 441–464.
- Norton, J. D. (2003). Causation as folk science. *Philosopher’s Imprint*, 3, 1–22.
- Norton, J. D. (forthcoming). Eternal inflation: When probabilities fail. *Synthese*. <https://doi.org/10.1007/s11229-018-1734-7>.
- O’Connor, T., & Wong, H. Y. (2015). Emergent properties. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/spr2020/entries/properties-emergent/>.
- Øhrstrøm, P., & Hasle, P. (2015). Future contingents. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/win2015/entries/future-contingents/>.
- Pereboom, D. (2001). *Living without free will*. Cambridge: Cambridge University Press.
- Pereboom, D. (2003). *Source incompatibilism and alternative possibilities* (pp. 184–199). Burlington: Ashgate.
- Pereboom, D. (2005). Defending hard incompatibilism. *Midwest Studies in Philosophy*, 29, 228–247.
- Roberts, J. T. (2006). Determinism. In S. Sarkar & J. Pfeifer (Eds.), *The philosophy of science: An encyclopedia*. New York: Routledge.
- Saka, P. (manuscript). Statistical determinism: The odds get odder and necessity gets even. <http://www.uh.edu/psaka/sylla/stet.htm>.
- Stoica, O. C. (2012). Modern physics, determinism, and free-will. *Noema*, 11, 431–456.
- van Inwagen, P. (1975). The incompatibility of free will and determinism. *Philosophical Studies*, 27, 185–199.
- van Inwagen, P. (1983). *An essay on free will*. Oxford: Clarendon Press.
- van Inwagen, P. (1989). When is the will free? *Philosophical Perspectives*, 3, 399–422.
- van Strien, M. (2015). Vital instability: Life and free will in physics and physiology, 1860–1880. *Annals of Science*, 72, 381–400.
- Vihvelin, K. (2018). *Arguments for incompatibilism*. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/fall2018/entries/incompatibilism-arguments/>.
- Wenmackers, S., & Romeijn, J.-W. (2016). New theory about old evidence. *Synthese*, 193, 1225–1250.
- Werndl, C. (2009a). Are deterministic descriptions and indeterministic descriptions observationally equivalent? *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 40, 232–242.
- Werndl, C. (2009b). What are the new implications of chaos for unpredictability? *British Journal for the Philosophy of Science*, 60, 195–220.
- Wigner, E. P. (1995). Events, laws of nature, and invariance principles. In J. Mehra (Ed.), *Philosophical reflections and syntheses. The collected works of Eugene Paul Wigner (Part B historical, philosophical, and socio-political papers)* (Vol. B/6). Berlin: Springer.