What is the value in an intrinsic formalism?

Eleanor March*

Abstract

I discuss the distinction between extrinsic and intrinsic approaches to reformulating a theory with symmetries, and offer an account of the special value of intrinsic formalisms, drawing on a distinction between which mathematical expressions are meaningful within an extrinsic formalism and which are not.

1 Introduction

Caspar Jacobs (2022) has recently considered afresh the motivation vs. interpretation debate about symmetries (originally due to Møller-Nielsen (2017)). Very roughly, the interpretationalist says that it is legitimate *ab initio* to interpret symmetry-related models (SRMs) of a theory as representing the same physical state of affairs, even in the absence of a metaphysically perspicuous characterisation of their common ontology, whereas the motivationalist denies this.[1] Jacobs takes up the question of what it means for a characterisation of the common ontology of SRMs to be 'metaphysically perspicuous.' He argues that this is captured by the demand for an intrinsic formalism, in the sense of Field (2016).[2] By contrast, extrinsic formalisms—in which the invariant content of SRMs is captured using structures which do not directly represent objects in the theory's ontology, e.g. by using equivalence classes of symmetry-variant quantities,[3] or more generally, by quotienting the theory's space of models under the symmetry group in question[4]—do not provide a similarly metaphysically perspicuous characterisation of a theory's ontological commitments. Instead, they

*Faculty of Philosophy, University of Oxford. eleanor.march@philosophy.ox.ac.uk

1. Though in light of recent work by Luc (2023), these positions are probably best understood as two extremes of a more-nuanced spectrum of motivationalist vs. interpretationalist views.

2. Note that some motivationalists may disagree here, i.e. on whether an intrinsic formalism is necessary to give a metaphysically perspicuous characterisation of the common ontology of SRMs. In particular, I would expect this to be the case for e.g. motivationalists who are fans of the Kleinian approach to geometry.

3. Following Wallace (2019) and Jacobs (2021b), I take using equivalence classes to subsume coordinate-based approaches.

4. Recall that given a space X on which a group G acts (from the left), one defines the quotient space X/G as the space of orbits of G, i.e. the space of equivalence classes $[x]$, $x \in [x]$ iff $gx \in [x]$ for all $g \in G$, which (providing the group action of G on X is suitably well-behaved, e.g. perhaps it is faithful) inherits the same kind of structure as the original space (i.e. if X is a smooth manifold, or Hilbert space, or whatever, then so is X/G).

only provide an effective decision procedure for determining whether the theory is committed to some piece of structure or other.

Or this is the story Jacobs tells. My aim in this essay will be to tell a rather different story about intrinsic formalisms, and why they are valuable. I have two motivations for this, neither of which I take to be dispositive, but each of which, to my mind at least, suggests the need for such a story. The first comes from a recent episode in the philosophy of spacetime physics literature, concerning how Maxwellian spacetime—which is supposed to be equipped with a standard of rotation, but not a standard of absolute acceleration—should best be characterized. Earman (1989) originally defined the rotation standard of Maxwellian spacetime extrinsically, as an equivalence class $[\nabla]$ of rotationally equivalent flat derivative operators, and Dewar (2018) also adopted this definition when writing down dynamics for the theory of Newtonian gravitation on Maxwellian spacetime (Maxwell gravitation). A number of authors (Weatherall 2018; Wallace 2020, 2019) then voiced concerns about the lack of an intrinsic formalism for Maxwell gravitation, which resulted in the development of an intrinsic characterization of the rotation standard (Weatherall 2018), and was then used to express the dynamics of Maxwell gravitation (Chen 2023; March 2023).

In light of Jacobs' discussion, one might expect the worries raised in this literature about Earman's definition of the rotation standard to concern, primarily, the lack of a 'metaphysically perspicuous' characterization of the ontology of Maxwellian spacetime. And so it is striking that they do not. Instead, what one finds are worries about the "mathematical impropriety" (Chen 2023, 23) of expressing the dynamics of Maxwell gravitation in terms of an arbitrary representative of this equivalence class; worries that the "minimalism" of Maxwellian spacetime ought to apply to the language used to express its dynamics as well as the structure in its models (22); worries about how to interpret intermediate terms in calculations on Maxwellian spacetime which do depend on the choice of representative of this equivalence class—that "one would like to be able to reason about quantities in Maxwellian spacetime without needing to introduce further structure" (Weatherall 2018, 34); and worries that the "awkwardness" of characterizing Maxwellian spacetime in terms of equivalence classes obscures the similarities between the dynamics of Maxwell gravitiation and those of Newton-Cartan theory (Wallace 2020, 28). Of course, these authors also discuss something like Jacobs' concern—e.g. Weatherall (2018), who suggests that using equivalence classes "obscures the intrinsic geometry of Maxwellian spacetime" (p. 35) and "makes reference to structure that one does not attribute to spacetime" (p. 34). But at the very least one might think, in light of this, that there is something missing from Jacobs' story about why intrinsic formalisms are valuable.

The second motivation begins from a somewhat different place. At a first pass, the concern is that it is not really clear that the Fieldian notion of intrinsicality is what is at issue in Jacobs' argument that one cannot 'read off' the ontology of a theory whose models are defined in terms of equivalence classes of symmetry-variant quantities. Here is just one reason to be worried about this. Jacobs (2022, 2) initially defines an intrinsic formalism as one in which the 'components' of the theory's models "directly correspond to its metaphysical posits." Later, he defines an intrinsic formalism as one "which is formulated in terms of mathematical entities that 'directly' represent physical fields" (Jacobs 2022, 5). But this is ambiguous, on two fronts.[5] First, it is ambiguous whether the criterion of 'directly representing the theory's metaphysical posits' is supposed to apply just at the level of objects in a theory's models, or also at the level of the objects in the dynamical equations used to pick out that class of models. Secondly, and relatedly, it is ambiguous what the objects, or 'components', of a theory's models are supposed to be—in particular, whether an equivalence class is supposed to count as a single 'component' or a collection thereof. This second ambiguity matters because of the first. In particular, if an equivalence class counts as a single component of a theory's models, and the criterion of 'directly representing the theory's metaphysical posits' is supposed to apply just at the level of objects in a theory's models, then it does not obviously follow that a theory whose models are defined in terms of equivalence classes of symmetry-variant quantities counts as extrinsic, on Jacobs' definition. That is, why not say that the equivalence class itself 'directly' represents one of the theory's metaphysical posits?

At this point, one might ask: so what? For even granting that Jacobs' definition is ambiguous in these ways, a charitable reconstruction of Jacobs would have it that he either takes an equivalence class to be a collection of 'components' of the models of a theory, or that the criterion of 'directly representing the theory's metaphysical posits' is supposed to apply also at the level of the objects in the theory's dynamical equations.[6] I agree. But this raises a deeper concern. To begin with, to suppose that anything significant—such as whether a theory counts as intrinsic or extrinsic—should turn on whether an equivalence class counts as a single 'component' of a theory's models or a collection thereof strikes me as a mistake. Indeed, Jacobs (rightly, to my mind) does not mention this question. So suppose we bracket that, and say that what is important for intrinsicality is that the criterion of 'directly representing the theory's metaphysical posits' applies at the level of the objects in the theory's dynamical equations.[7] (This, I take it, would also be more in accordance with Field's characterization of intrinsicality as the requirement that all the terms in a theory refer to physical quantities.) In that case, one might think that the problem

5. One might also worry that the distinction between what it is for a piece of mathematical structure to represent a piece of physical structure 'directly' vs. 'indirectly' is not sufficiently clear. I am sympathetic to this worry, but won't discuss it here.

6. I am setting aside, here, worries about how and whether this should exclude, e.g. numerical structure like the number 5, or mathematical operators like addition, etc.—which I take to partly explain Jacobs' decision to focus on objects in the models of a theory, rather than its dynamics, cf. (Jacobs 2022, 14). I will return to this in §2.

7. To preface the arguments of the next section somewhat, I do not think this strategy works particularly well either, and so will ultimately propose my own characterization of the distinction Jacobs has in mind. My point here is just that *if* one wants to do Jacobs exegesis, there is a mismatch between the kind of structures used to draw the intrinsic vs. extrinsic distinction, on the most charitable reading of Jacobs, and the kind of structures at issue in Jacobs' discussion of why that distinction is salient for the motivation vs. interpretation debate.

with formulating a theory extrinsically would itself have something to do with the dynamical equations of such theories—a point which is also suggested by my discussion of the first motivation above. But the dynamical equations of extrinsically formulated theories do not feature at all in Jacobs' discussion. Again, one is left with the sense that there is something more to be said.

Of course, these are just motivations for the story I am going to tell about why intrinsic formalisms are valuable, not justifications for it, nor outright criticisms of Jacobs' position (though I will raise some such criticisms in §3). My approach will be to draw out a contrast between the kinds of mathematical expressions that may be interpreted as physically meaningful in an intrinsic vs. an extrinsic formalism: in an intrinsic, but not an extrinsic formalism, are facts about whether some mathematical expression may be interpreted as physically meaningful essentially trivial—decidable by inspection, as a simple matter of the mathematical operations used to construct it themselves being well-defined. This means that intrinsic formalisms are better suited for getting a handle on what a theory says about the world, for two reasons. First, because to find out which equations may be interpreted as saying something physically meaningful within the theory, we need only take a cursory look at the form of those equations. Second, because one can more easily 'read off' what those (physically meaningful) equations say about the world—what relationships between physical quantities they express—without having to first identify which mathematical sub-expressions in those equations may be interpreted as representing physical quantities. Along the way, I will make several points which may be of independent philosophical interest, concerning the intrinsic vs. extrinsic distinction and its relationship to the reduction vs. internal sophistication vs. external sophistication distinction, and the kind of mathematical expressions which may be interpreted as physically meaningful within an extrinsic formalism.

In a little more detail, then, the structure of this article will be as follows. First, in §2, I introduce the distinction between intrinsic and extrinsic formalisms (in §2.1), and explain how this fits together with the distinction between reduction, internal sophistication, and external sophistication (in §2.2). In §3, I discuss Jacobs' arguments that extrinsic formalisms are not metaphysically perspicuous, and argue that most of them do not succeed. This takes us to §4, where I propose my alternative story about the special value of intrinsic formalisms, drawing on the idea that extrinsic formalisms come with a restriction on what mathematical expressions may be interpreted as physically meaningful. This will allow me to say why it is that intrinsic formalisms are more perspicuous than extrinsic formalisms. §5 concludes.

2 Prelude: what is an intrinsic formalism, anyway?

2.1 Intrinsic vs. extrinsic

I will begin by discussing the intrinsic vs. extrinsic distinction. As noted in §1, Jacobs (2022, 2) defines an intrinsic formalism as one in which the 'components' of the theory's models "directly correspond to its metaphysical posits," or one "which is formulated in terms of mathematical entities that 'directly' represent physical fields" (5). But, again as noted in §1, this is ambiguous—both because it is not clear what counts as a single 'component' of a theory's models, and because it is not clear whether the criterion of 'directly representing the theory's metaphysical posits' is supposed to apply just at the level of objects in a theory's models, or also at the level of the objects in the dynamical equations used to pick out that class of models.

Prima facie, one might think that addressing these worries would involve precisifying Jacobs' definition of an intrinsic formalism along one of these lines. I do not think this strategy is particularly attractive, for two reasons. First, suppose that the criterion of 'directly representing the theory's metaphysical posits' is supposed to apply just at the level of objects in a theory's models. Then, as mentioned in the introduction, whether or not e.g. Earman's Maxwellian spacetime—which, I take it, is supposed to be obviously extrinsic—counts as an extrinsic formalism on Jacobs' definition will depend on whether an equivalence class counts as a single 'component' of a theory's models or a collection thereof. But this last issue strikes me as mainly a terminological one, and so it would seem undesirable that something substantive, like the intrinsicality or otherwise of a theory, should turn on it.

Alternatively, suppose that the criterion of 'directly representing the theory's metaphysical posits' is supposed to apply at the level of objects in the theory's dynamical equations. In that case, one might worry about how and whether this is supposed to apply to numerical structure in the dynamical equations of a theory, like the number 5, which does not directly represent any *physical* quantity (though if one is a mathematical platonist, perhaps it does count as a 'metaphysical posit' of the theory). Now Jacobs (2022, 14) is clear that, unlike Field, he does not want the appeal to numerical structure like the number 5 to automatically classify a theory as extrinsic. But it is not immediately obvious how to spell this out in a general, non-question begging way, in terms of a restriction on which kinds of objects in the dynamical equations of a theory must directly represent the theory's metaphysical posits. At a more fundamental level, one might worry that even if Jacobs would be happy to countenance a commitment to mathematical platonism to get his definition off the ground, this seems to be something of a distraction. In particular, the various authors involved in the discussions of intrinsic vs. extrinsic formulations of Maxwell gravitation seem to have been able to carry out these discussions without needing to mention (much less make a commitment on) the nominalism vs. platonism debate (and likewise for Jacobs, in much of his discussion of the problems he sees with the extrinsic approach).

Let's take a step back. What kind of theories does Jacobs have in mind, in his discussion of the extrinsic approach? Jacobs (2022, pp. 7–8, fn. 11) gives a clear statement of what these theories are like: they are the kind of theories which are obtained by 'equivocating between' (in some sense) SRMs of some other theory—either by taking a formal quotient of the space of models under the symmetry group in question, or by 'adding arrows' between (non-isomorphic) SRMs of a theory understood category-theoretically.[8] The first case, in some detail, proceeds as follows. Recall that in general, a theory T has kinematically possible models (KPMs) of the form $\langle \mathcal{V}_1, \mathcal{V}_2, ..., Q_1, Q_2, ...\rangle$, where the $\mathcal{V}_i$ are a collection of (structured) value spaces with domains V_i, and the Q_i are a collection of quantities (defined as functions from some $\mathcal{V}_i$ into some $\mathcal{V}_j$). The dynamically possible models (DPMs) of T are those KPMs which satisfy the theory's equations of motion. A dynamical symmetry of T is a bijection on the KPMs induced by a collection of bijections $\chi_i : V_i \to V_i$ which preserves the space of DPMs.

Let T be a theory, as characterized above, and suppose that $\mathcal{V} = \langle V, \phi\rangle$ is the domain (or codomain) for just one quantity Q (the generalization to the case of multiple ϕ_i or Q_i is straightforward), and consider a group of G of dynamical symmetries $Q \to \chi^*Q$, $\chi : V \to V$. Then

1. We can construct a new space of models $\langle V, \phi, ..., [Q], ...\rangle$, $Q \in [Q]$ iff $\chi^*Q \in Q$ for all $\chi \in G$. Pairs of SRMs $\langle V, \phi, ..., Q, ...\rangle$, $\langle V, \phi, ..., \chi^*Q, ...\rangle$ now correspond to the same model $\langle V, \phi, ..., [Q], ...\rangle$.

2. When $\chi^*\phi \neq \phi$ for some $\chi \in G$, we can also construct a new space of models $\langle V, [\phi], ..., Q, ...\rangle$, $\phi \in [\phi]$ iff $\chi^*\phi \in [\phi]$ for all $\chi \in G$. Pairs of SRMs $\langle V, \phi, ..., Q, ...\rangle$, $\langle V, \phi, ..., \chi^*Q, ...\rangle$ now correspond to isomorphic pairs of models $\langle V, [\phi], ..., Q, ...\rangle$, $\langle V, [\phi], ..., \chi^*Q, ...\rangle$.

An example of 1 is the theory of electromagnetism on Minkowski spacetime, with models $\langle M, \eta_{ab}, [A_a], J^a\rangle$ where the $A_a \in [A_a]$ are all related by closed one-form shifts (i.e. Weatherall's (2016) $\mathbf{EM}_2'$). An example of 2 is Dewar's (2018) Maxwell gravitation, which has models $\langle M, t_a, h^{ab}, [\nabla], T^{ab}\rangle$ where $[\nabla]$ is an equivalence class of rotationally equivalent flat derivative operators.[9]

For the second case, one begins by associating T with a category $\mathbf{T}$, whose objects are the models of T, and whose arrows are isomorphisms of those models.[10] Again, let G be a group of dynamical symmetries of T, and suppose that

8. Another option along these lines, which I lack the space to discuss here, would be using 'stacky' constructions—on which, see Teh (2024). A thorough treatment of stacks would take us beyond the scope of this essay, so I will just note that one should feel free to add stacks as a fourth option to my characterization of extrinsic formalisms below.

9. Recall that two flat derivative operators ∇, ∇' on a spacetime $\langle M, t_a, h^{ab}\rangle$ are said to be rotationally equivalent iff for all unit timelike vector fields ξ^a on M, $\nabla^{[a}\xi^{b]} = \mathbf{0}$ iff $\nabla'^{[a}\xi^{b]} = \mathbf{0}$.

10. Or possibly a subclass thereof—see e.g. March, Read, and Chen (2025) for discussion of one such example, and Read (2025) for more general philosophical discussion—though I will not consider that case here. For general background on the category-theoretic approach, see e.g. Weatherall (2016).

these dynamical symmetries do not act as isomorphisms on the models of T (other than at the identity of G).[11] Then, very roughly

3. One constructs a new category $\overline{\mathbf{T}}$ whose objects are again the models of T, but whose arrows are now pairs (ψ, χ) consisting of an isomorphism ψ and an element ξ of G, which act on the models of T as that element ξ followed by that isomorphism ψ.

This provides an alternative way of formalizing the idea that SRMs of T are to be treated 'as if' they were isomorphic, for the purposes of understanding the structure of the theory. An example of 3 is the theory which Weatherall (2016) calls $\overline{\mathbf{EM}}_2$, whose objects are models of electromagnetism on Minkowski spacetime $\langle M, \eta_{ab}, A_a, J^a \rangle$, and whose arrows are pairs consisting of a diffeomorphism plus a closed one-form shift.

At this point, my suggestion is that we might as well just go ahead and define extrinsic formalisms as all and only those formalisms which are instances of the above constructions 1–3, and intrinsic formalisms as all the rest. First, because this precisely captures the class of theories which Jacobs takes to exemplify the extrinsic approach. Second, because it fits very closely with Jacobs' characterization of the intrinsic approach as requiring that one "lay down a set of relations, functions and operators which *explicitly* represent the world's physical structure" (Jacobs 2022, p. 13, emphasis mine)—i.e. essentially as implementing a ban on implicit definition in terms of symmetry-variant structures. Third, because it is a precise definition which avoids the aforementioned worries about 'components' of a theory's models, structures in the models of a theory vs. its dynamical equations, numerical structure, etc. (Anyone who is unhappy with my continuing to use Field's 'intrinsic' vs. 'extrinsic' terminology for this distinction should feel free to mentally replace the terms 'intrinsic' and 'extrinsic' with 'intrinsic*' and 'extrinsic*' going forward.)

My fourth reason for adopting this definition is that this way of articulating the intrinsic vs. extrinsic distinction also captures something like the idea that in an extrinsic formalism, not all the terms in the dynamical equations of the theory need 'directly represent the theory's metaphysical posits'—but whilst avoiding the worries about e.g. numerical structure raised above. This will take somewhat more work to articulate. As a way in, I want to begin with the following question: what kind of equations is it sensible to write down, when working with an extrinsic formalism? In particular: is it legitimate for these equations to make use of objects in the equivalence classes of symmetry-variant structures (or in the case of 3, structures in the models of T which are variant under the arrows of $\overline{\mathbf{T}}$)?

The answer which I want to suggest, and which I think captures the way in which extrinsic formalisms have previously been thought of, is 'yes'—so long as satisfaction of these equations is independent of the choice of representative of

11. The situation is a little more subtle in the case where some (but not all) of the non-trivial elements of G act as isomorphisms on the models of T—and will depend on the details of the group G in question—so I set aside this case for the sake of exposition.

the equivalence class. For example, when Earman defined Maxwellian spacetime as a structure $\langle M, t_a, h^{ab}, [\nabla]\rangle$, he wrote:

> Although questions about the acceleration of a body are not in general meaningful in this setting, it is, of course, meaningful to ask about the state of rotation of a fluid or an extended body. (Earman 1989, p. 32)

And when Dewar (2018) wrote down dynamics for Maxwell gravitation on Earman's Maxwellian spacetime, he went to some trouble to show that satisfaction of these equations was independent of the choice of $\nabla \in [\nabla]$. Or consider: it is often claimed that the only possible field equations for the theory of scalar electrodynamics with models $\langle M, \eta_{ab}, [\langle A_a, \psi\rangle]\rangle$ are ones which are $U(1)$ gauge invariant.[12]

So extrinsic formalisms don't just equivocate between symmetry-variant structures in defining the models of the theory; they also restrict the space of mathematical expression involving these structures which can be 'meaningfully' written down to ones which are invariant under the relevant class of symmetry transformations. Conversely, equations which are not so invariant—like the acceleration of a fluid relative to some $\nabla \in [\nabla]$ in Maxwellian spacetime—are not to be regarded as 'meaningful.' One way to spell this out is in terms of a supervaluationist semantics, in which the supervaluation is carried out over all objects in the relevant equivalence classes (see e.g. Dewar (2019) and Jacobs (2021a)), or in the case of the category-theoretic approach 3, over all objects in the isomorphism equivalence classes of $\overline{\mathbf{T}}$ (where 'isomorphism', here, means isomorphism under the arrows in $\overline{\mathbf{T}}$).

The point I want to press here is that from the perspective of an extrinsic formalism, one does *not* need to appeal to the idea that SRMs represent the same physical state of affairs to justify the claim that equations which are not independent of the choice of $\phi \in [\phi]$ are physically meaningless. For suppose we take seriously the project of characterising structures in the models of a theory T extrinsically, e.g. via a preferred equivalence class of representations $[\phi]$. Then what it is for the models of T to have this kind of structure is *just is* to say: facts about the models of T can be represented equally (but redundantly) by any one of the $\phi \in [\phi]$. Since the models of T are supposed to represent physical states of affairs, it follows that equations which are not independent of the choice of $\phi \in [\phi]$ cannot be interpreted as saying something physically meaningful (since they cannot sensibly be thought of as *about* the models of T).[13] Conversely, insofar as it does make sense to talk about physical facts which are not so independent within the formalism of T, T simply fails to have the kind of structure we have defined it to have.[14] It is for this reason that

12. cf. also Mundy (1986).

13. Compare Lewis (1986): "A proposition [read: equation] is about a subject matter [...] if and only if that proposition holds at both or neither of any two worlds [read: mathematical structures] that match perfectly with respect to that subject matter."

14. I take this to be in the spirit of Belot's (2000) point that writing down equations which require for their formulation structure which the theory does not posit is "arrant knavery"—

equations which are not independent of the choice of $\phi \in [\phi]$ are physically meaningless.

That said, I do want to point out that just restricting to equations which are independent of the choice of representative of the equivalence class can't quite be the full story about which equations are meaningful in an extrinsic formalism, since it too easily falls prey to what one might think of as 'spurious' invariances. This is nicely illustrated with the example of Earman's Maxwellian spacetime. Since all the $\nabla \in [\nabla]$ are flat, the equation $R^a{}_{bcd} = 0$ is invariant between objects in the equivalence class, and will come out as being true. But this is, intuitively speaking, the wrong result: Maxwellian spacetime lacks full affine structure, and so it is simply not sensible to speak of it as flat or non-flat (though one can make sense of a weaker notion of rotational flatness, see March (2024)). Probably the right thing to do in this case is to note that one can also represent the rotation standard with a non-flat connection (any connection satisfying $R^{ab}{}_{cd} = 0$ will do), so that the equation $R^a{}_{bcd} = 0$ is not invariant between all the connections which can be used to represent the rotation standard. But it is not straightforward how to spell this out in general, at least without some intrinsic characterisation of the structure of interest already to hand.[15] In any case, I take the above arguments to show that being independent of the choice of representative of the equivalence class is a plausible minimal restriction on which equations are physically meaningful within an extrinsic formalism—so going ahead, we can adopt this restriction along with the proviso that it may need to be tightened up later.

2.2 Reduction vs. sophistication

I will now move on to discuss the distinction between reduction, internal sophistication, and external sophistication (see e.g. Dewar (2019) and Martens and Read (2021) for clear expositions of these views), and how this relates to the distinction between intrinsic and extrinsic formalisms. Reduction says that faced with SRMs of a theory T, one should reformulate T so that SRMs all map to the same model of the reduced theory. Importantly, this is the case even if the SRMs in question are isomorphic. By contrast, internal and external sophistication both say that if SRMs of T are isomorphic, one may interpret

i.e. certainly not sensible and maybe even incoherent; cf. also Wallace (2019) on coordinate-based approaches and Myrvold (2019) on Earman's SP2. I will discuss Earman's principles more in §3.

15. The other obvious thing to do here is to point out that the equation $R^a{}_{bcd} = 0$ coming out as true is not quite as bad as it sounds, since in Maxwellian spacetime, the left hand side of this equation lacks an interpretation in terms of parallel transport of timelike vectors along arbitrary (spacelike or timelike) curves. On this kind of view, the equation $R^a{}_{bcd} = 0$ might be true in Maxwellian spacetime, but it would not follow from this that Maxwellian spacetime is 'flat' in the usual sense of the word (flatness would require in addition e.g. a standard of parallel transport for timelike vectors along timelike curves). I think this kind of response can also probably be made to work, but again, it is not completely obvious how to spell this out in general.

them as physically equivalent via appeal to anti-haecceitism.[16] Where these two approaches differ is on their treatment of non-isomorphic SRMs. According to internal sophistication, one must first reformulate T so that SRMs map to isomorphic models of the sophisticated theory before one can appeal to anti-haecceitism to say that these models represent the same physical state of affairs, whereas according to external sophistication, one can 'stipulate' that SRMs are to 'count' as isomorphic, without reformulating the theory. External sophistication is thus very naturally articulated from the theories-as-categories standpoint, in which one can understand stipulating isomorphisms between non-isomorphic models as meaning that one is to add arrows between non-isomorphic models into one's category of models (e.g. as described in point 3 of §2.1).[17]

For what it's worth, I also think that external sophistication *should* be understood category-theoretically, if it is not either to be altogether mysterious or to collapse into a variant of reduction or internal sophistication (though one might also take this to mean that the distinction between external sophistication and internal sophistication or reduction was never the relevant one to start off with).[18] For example, Dewar characterizes external sophistication as "declaring, by fiat, that the symmetry transformations are now going to 'count' as isomorphisms" (Dewar 2019, pp. 502–3). But as Jacobs (2022) notes, this way of putting it is "somewhat puzzling" (Martens and Read (2021, p. 340) go further, saying that "[to] stipulate that qualitatively distinct, i.e. non-isomorphic models [...] are nevertheless isomorphic reads *prima facie* as nothing more than a flat-out contradiction"). It is in this vein that Jacobs (2022, 2021b) offers his own take on external sophistication, as meaning that the invariant content of SRMs should be captured by taking equivalence classes of symmetry-variant structures, i.e. via one of the constructions 1 or 2 mentioned in §2.1. But if this is what external sophistication is really about, then it becomes clear that it is really a form of reduction or internal sophistication, since it involves mathematically reformulating the models of the theory in such a way that SRMs will end up being either identical or isomorphic. Conversely, if we take Jacobs to mean that external sophistication should be understood as a commitment to characterizing structures in the models of a theory via their isomorphisms, but without reformulating the theory so that SRMs are in fact isomorphic, then we are back to the worry about how to make sense of 'declared isomorphisms' between non-isomorphic models, i.e. back to square one.

Now, one might worry that this threatens to make external sophistication redundant, given that it is somewhat trivial to internally sophisticate or reduce a theory if one is satisfied with an extrinsic formalism (one simply takes appropriate equivalence classes). For my own part, I do not think that we should be

16. Or its analogue for quantities, anti-quidditism—though the difference does not really matter, since anti-quidditism about (the determinate magnitudes of) some physical quantity amounts to anti-haecceitism about points in the value space of that quantity.

17. cf. Dewar (2019, 2022)

18. This is how I am sometimes inclined to read Jacobs (2021b), when he instead distinguishes structure-first and symmetry-first approaches to capturing the invariant content of SRMs.

worried by this. The distinction between reduction vs. internal sophistication vs. external sophistication is about formal properties of a mapping between the KPMs of two (not necessarily distinct) theories. It is *not* about the extent to which it is non-trivial to arrive at such a theory. Whilst it is true that authors such as Martens and Read (2021) suggest that in cases of successful reduction or internal sophistication "there was no guarantee that [a reduced or internally sophisticated formalism] was even possible", I think we should take this to be a slip. What is (apparently, or at least much more plausibly) true is that there is no guarantee that an *intrinsic* reduced or internally sophisticated theory is always possible (especially in cases of reduction for theories with isomorphic SRMs)—but this just goes to show that the relevant distinction for these authors was never the reduction vs. internal sophistication vs. external sophistication distinction to begin with. And as for the worry that this leaves no interesting role for external sophistication, understood category-theoretically, I think this is simply false—as witness the whole host of cases (Galilean gravitation (Weatherall 2016), electromagnetism (Weatherall 2016; Nguyen, Teh, and Wells 2020), Newton-Cartan theory (March 2024), teleparallel gravity (March, Read, and Chen 2025; Weatherall and Meskhidze 2025), hidden symmetries (Read 2025)) where philosophers of physics have discussed (and said, in my view, interesting things about) precisely this strategy.

Having said this, it should be clear what I want to say about how the intrinsic vs. extrinsic formalisms distinction relates to the reduction vs. internal sophistication vs. external sophistication distinction, which is that the two are basically orthogonal. To be sure, one area of overlap remains, which is that externally sophisticated theories will *usually* be extrinsic (except in special cases where the arrows of one's original category of models for a theory do not include all the isomorphisms of those models, see again e.g. March, Read, and Chen (2025) and Read (2025)). But one can externally sophisticate an intrinsic or extrinsic formalism; likewise, one can reduce a theory by moving to an intrinsic or extrinsic formalism, and one can also internally sophisticate by moving to an intrinsic or extrinsic formalism.

To make this point absolutely clear, it is helpful to consider an example. Take the theory of Newtonian point-particle mechanics with models $\langle \mathcal{M}, \phi, \mathcal{B}, \gamma(i), \mathbb{R}^+, m(i) \rangle$, where $\mathcal{M} = \langle M, t_a, h^{ab}, \nabla, \xi^a \rangle$, ϕ is a scalar field which represents the gravitational potential, $\mathcal{B}$ is a (structured) domain of particles, the $\gamma(i) : \mathcal{B} \times \mathbb{R} \to \mathcal{M}$, are a collection of (smooth, future-directed) timelike curves which represent particle worldlines, and $m(i) : \mathcal{B} \to \mathbb{R}^+$ is an assignment of mass values to each particle. Suppose that uniform mass scalings—transformations of the form $m(i) \to \psi \circ m(i)$, where ψ is a bijection on the domain R^+ of $\mathbb{R}^+$ which preserves the relation $\leq$ and the operation $+$[19]—are dynamical symmetries of this theory. Table 1 outlines four possible approaches to reformulating the models of the theory, in light of this symmetry.

The first row of table 1 is fairly straightforward. In the intrinsic reduc-

19. Note that uniform mass scalings won't preserve the relation $\times$ on R^+, i.e. they are not automorphisms of the mass value space $\mathbb{R}^+ = \langle R^+, \leq, +, \times \rangle$. In particular, this means that models related by a uniform mass scaling are not isomorphic.

	Intrinsic	Extrinsic
Reduction	$\langle \mathcal{M}, \phi, \mathcal{B}, \gamma(i), \mathbb{R}^+, m(i,j) \rangle$	$\langle \mathcal{M}, \phi, \mathcal{B}, \gamma(i), \mathbb{R}^+, [m(i)] \rangle$
Internal sophistication	$\langle \mathcal{M}, \phi, \mathcal{B}, \gamma(i), D_m, \leq, \circ, m(i) \rangle$	$\langle \mathcal{M}, \phi, \mathcal{B}, \gamma(i), D_m, [f], \mathbb{R}^+, m(i) \rangle$

Table 1: Different reformulations of Newtonian point-particle mechanics, organised by where they fall with respect to the intrinsic vs. extrinsic formalisms and reduction vs. internal sophistication distinctions.

tion corner, we replace the assignment of mass values with an assignment of mass ratios $m(i,j) : \mathcal{B} \times \mathcal{B} \to \mathbb{R}^+$ to each pair of particles, subject to the constraints $m(i,i) = 1$ and $m(i,j)m(j,k) = m(i,k)$. In the extrinsic reduction corner, we replace $m(i)$ with an equivalence class of mass value assignments, i.e. $m(i) \in [m(i)]$ iff $\psi \circ m(i) \in [m(i)]$ for all uniform mass scalings ψ. Both of these count as reduced theories: the $m(i,j)$ are invariant under uniform mass scalings, and by construction any pair of models $\langle \mathcal{M}, \phi, \mathcal{B}, \gamma(i), \mathbb{R}^+, m(i) \rangle$, $\langle \mathcal{M}, \phi, \mathcal{B}, \gamma(i), \mathbb{R}^+, \psi \circ m(i) \rangle$ related by a uniform mass scaling map to the same model $\langle \mathcal{M}, \phi, \mathcal{B}, \gamma(i), \mathbb{R}^+, [m(i)] \rangle$.

In the second row of table 1, we have modified the definition of the mass value space. In the intrinsic internal sophistication corner, we have replaced $\mathbb{R}^+$, with an additive extensive structure $\langle D_m, \leq, \circ \rangle$, where D_m is a domain of cardinality $2^{\aleph_0}$, $\leq$ is a total order on D_m, and $\circ$ is an associative binary operation (representing addition of mass values), subject to certain axioms.[20] Since uniform mass scalings are automorphisms of $\langle D_m, \leq, \circ \rangle$, models related by a uniform mass scaling are now isomorphic. In the extrinsic internal sophistication corner, the mass value space $\langle D_m, [f], \mathbb{R}^+ \rangle$ is again an additive extensive structure, but this time we have characterised it extrinsically rather than intrinsically. $[f]$ is an equivalence class of bijections $f : D_m \to \mathbb{R}^+$ defined as follows: $f \in [f]$ iff $f \circ \psi \in [f]$ for any bijection ψ on R^+ which preserves $\leq$ and $+$. To see that this is indeed an instance of internal sophistication, consider any bijection $\psi : D_m \to D_m$ which preserves $[f]$. It follows immediately from the definition of $[f]$ that these bijections are in one-to-one correspondence with our original uniform mass scalings. Now, in general, ψ is not an automorphism of $\langle \mathcal{M}, \phi, \mathcal{B}, \gamma(i), D_m, [f], \mathbb{R}^+, m(i) \rangle$, since $\psi \circ m(i) \neq m(i)$ unless $\psi = \mathrm{id}_{D_m}$, so this is not a reduced formalism. But it is an internally sophisticated formalism: ψ induces an isomorphism of $\langle \mathcal{M}, \phi, \mathcal{B}, \gamma(i), D_m, [f], \mathbb{R}^+, m(i) \rangle$, since it is an automorphism of $\langle D_m, [f], \mathbb{R}^+ \rangle$. This makes my point: that the intrinsic vs. extrinsic formalisms distinction is independent of the reduction vs. internal sophistication distinction.

20. See e.g. Hölder (1901) and Krantz et al. (1971) for details of these axioms. In effect, this amounts to 'forgetting' the multiplication operation $\times$ on R^+.

3 Metaphysical perspicuity

We can now return to Jacobs' argument that only intrinsic formalisms are metaphysically perspicuous. To reiterate, according to Jacobs, the ontological commitments of intrinsic formalisms can simply be 'read off' from the formalism, whereas extrinsic formalisms have only an 'effective decision procedure' for whether the theory is committed to some piece of structure or other, by determining whether that structure can be invariantly defined from objects in the equivalence classes.

The fact that the extrinsic approach has only an effective decision procedure for determining a theory's ontological commitments, Jacobs claims, means that extrinsic formalisms are not metaphysically perspicuous. This is for three reasons. First, extrinsic formalisms limit attempts at causal explanation, since only symmetry-invariant structures can be dynamically efficacious. Second, it is unclear what grounds (or explains, or justifies) the physical equivalence of SRMs from the perspective of an extrinsic formalism. This is supposed to be because the extrinsic approach reverses the natural order of explanation—the theory is committed to a certain ontology because SRMs represent the same physical state of affairs, rather than *vice versa*. Third, extrinsic formalisms appeal to physically irrelevant (i.e. symmetry-variant) quantities to characterise the structure of the theory's value spaces. This does not seem to tell us what these value spaces are really like, even if it fixes the correct structure (this is the constructivist complaint).[21]

However, I think that the extrinsic approach has the resources to resist these worries, at least when we take into account the fact that (a) the extrinsic approach comes with a restriction on what equations are physically meaningful, and (b) the extrinsic approach does not need to appeal to the idea that SRMs represent the same physical state of affairs to justify this restriction (recall §2). Indeed, I think this gives the extrinsic approach a very natural response to Jacobs' concern about what explains the physical equivalence of SRMs. That is, SRMs represent the same physical state of affairs because the only claims which would possibly allow one to distinguish between them (up to isomorphism) are physically meaningless, since they depend on the choice of representative in the equivalence classes. And the fact that the theory is committed to a certain ontology because certain distinctions are physically meaningless, and others physically meaningful, strikes me as the right result.[22] In fact, it seems to me precisely the kind of reasoning involved in Earman's (1989) famous symmetry

21. Note that for this kind of reason, extrinsic formalisms are likely to be repugnant to fans of Reichenbachian constructivism (see Linnemann and Read (2021) and Adlam, Linnemann, and Read (2022)), though here the issue is not so much that the quantities are 'physically irrelevant,' but rather a problem of not being able to help oneself to structure that one hasn't constructed yet.

22. Compare Leibniz's famous 'shift' argument against the reality of absolute space: absolute space is unreal because "[to] say that God can cause the whole universe to move forwards in a right line, or in any other line, without making otherwise any alteration in it; is another chimerical supposition. For two states indiscernible from each other, are the same state; and consequently, 'tis a change without any change." (Leibniz and Clarke 1998, p. 38)

principles:

SP1: Every dynamical symmetry of T is a spacetime symmetry of T.

SP2: Every spacetime symmetry of T is a dynamical symmetry of T.

In effect, SP2 says that T should posit enough spacetime structure that it can distinguish (in the sense that they are not isomorphic) between models which are not related by a dynamical symmetry. Conversely, SP1 says that T should not posit so much spacetime structure that it distinguishes (again, in the sense that they are not isomorphic) between models which are related by a dynamical symmetry. So providing we hold the dynamical symmetries of T fixed, Earman's principles suggest that it is entirely appropriate to let questions of what distinctions are physically meaningful underwrite questions of a theory's ontology in this way.

Turning now to the causal explanations worry, I think this is misplaced. Granted, the causal explanations which one can read off from an extrinsic formalism will often involve making reference to symmetry-variant quantities. But so long as the explanations themselves are symmetry-invariant,[23] it is not clear why this should hamper attempts at causal explanation. For example, Jacobs claims that the proponent of the extrinsic approach cannot without further argumentation explain the Aharanov–Bohm effect, in which a charged particle in the vicinity of an impenetrable solenoid picks up a phase proportional to the flux through that solenoid. This is supposed to be because the causal story involved in the explanation of the Aharanov–Bohm effect must appeal to invariant structures, such as the holonomies of the electromagnetic one-form. But here is an explanation to which the extrinsicalist can perfectly well appeal: the charged particle picks up a phase because up to $U(1)$ gauge symmetry, the electromagentic one-form can be represented as having a value such-and-such in the region surrounding the solenoid, and the phase difference picked up by the particle follows from this plus the dynamics of the theory. To the worry that such an explanation is not appropriately *causal* I say: whatever one's favourite account of causation is, either this explanation counts as appropriately causal (as for e.g. counterfactual, interventionist, or productive accounts),[24] or this is a

23. Note that we do also appear to give symmetry-variant explanations in physics, e.g. the appeal to the rest frame of the rocket in the explanation of Bell's rockets thought experiment, though one might dispute whether this really counts as a symmetry-variant explanation, given that the rest frame of the rocket can be defined invariantly from its worldline. However, this will depend on whether or not one has an operational understanding of the coordinate systems in question.

24. As an example, consider, e.g. a counterfactual account of causation. If the value of the electromagnetic one-form hadn't been such and such, up to $U(1)$ gauge symmetry, in the region surrounding the solenoid, then the phase picked up by the particle would have been different. Thus the value of the electromagnetic one-form in the region surrounding the solenoid, up to $U(1)$ gauge symmetry, counts as a cause of the phase shift picked up by the particle. A similar point goes for interventionist accounts. For anyone who is tempted to respond to this that for any segment of the particle's trajectory, one could choose a gauge in which it picks up zero phase over that segment, this is by-the-by, as far as the explanation I am suggesting is concerned—the point is that one cannot so choose a gauge throughout the entire region

problem for causal explanations of the Aharanov–Bohm effect in general, rather than the explanation which the extrinsic approach offers in particular (as for e.g. conserved quantity or causal mechanisms approaches), see Earman (2024).

As far as Jacobs' more general point goes—that extrinsic formalisms don't tell us what a theory's value spaces are really like—the extrinsic approach has an answer to this too. The answer is that those value spaces have just enough structure to capture all the invariant degrees of freedom of the symmetry-variant quantities that are used to define them, and no more. The proponent of the intrinsic approach will press the question: but just what structures are those? But here I think that the extrinsicalist can simply dig their heels in. Granted, the extrinsicalist cannot say without further argumentation just what the invariant structures in question are, but at least by their own standards, this seems like a perfectly acceptable answer.

With that said, I do think that Jacobs' insight about the fact that one cannot 'read off' the ontological commitments of extrinsic formalisms is basically on the right track. However, as I have argued, it it not clear why this fact by itself should act as a barrier to metaphysical perspicuity. So let us see what kind of consequences would act as such a barrier.

4 Triviality and non-triviality

Throughout this article, I have pressed the idea that within an extrinsic formalism, mathematical expressions which are not independent of the choice of representative of the equivalence classes of symmetry-variant structures should not be thought of as saying or representing something physically meaningful. I have also argued that from the perspective of an extrinsic formalism, this restriction is well-motivated, and perhaps even compulsory.

But it is also non-trivial. To see this, consider the kinds of equations that can be interpreted as physically meaningful in an intrinsic formalism. Since intrinsic formalisms do not equivocate between symmetry-variant structures in defining the models of the theory, any equation which is (i) constructed out of objects in the models of the theory, and (ii) is mathematically well-defined, expresses a statement which can be true or false of the models of the theory, and so can be interpreted as saying something physically meaningful. In other words, to check whether an equation is physically meaningful within an intrinsic formalism, one only needs to take a cursory look at the form of that equation.

Contrast this with the process of checking whether an equation is physically meaningful within an extrinsic formalism. In this case, not only does one need to verify that the equation is mathematically well-defined—in general, one also needs to verify that satisfaction of that equation is independent of the choice of representative of the equivalence class. And this second step is generally highly non-trivial (think e.g. of Dewar's (2018) dynamics for Maxwell gravitation, or of

surrounding the solenoid. Rather, it seems to me that this worry is not really to do with *causal* explanations *per se*, but local explanations, or separable explanations, or something along those lines. I am grateful to an anonymous referee for pressing this point.

the many pages of ink that physics undergraduates have spilled over the years verifying $U(1)$ gauge invariance of the Lagrangian of scalar electrodynamics).

A similar point applies at the level of interpreting those (physically meaningful) equations, once we have them, i.e. saying what relationships between physical quantities they express.[25] In order to interpret such an equation, in an extrinsic formalism, one first needs to identify which mathematical subexpressions of that equation have values which are independent of the choice of representative of the equivalence class, and so may be interpreted as representing physical quantities (I take it, this is precisely where Jacobs' insight that one cannot 'read off' the ontology of an extrinsic formalism comes in). And again, this process is non-trivial. By contrast, in an intrinsic formalism, one does not have to first parcel off which mathematical subexpressions in an equation have values which are independent of the choice of representative of an equivalence class in this way, and so can be interpreted as corresponding to physical quantities—the interpretation of these equations can simply proceed term by term.

But if this is right, we now have a handle on why it is that intrinsic formalisms are more perspicuous than intrinsic formalism. The issue is not just that one cannot read off the theory's ontology from the formalism—though that may come into it too—but that one also cannot read off which mathematical expressions that can be constructed in that formalism are candidates for being interpreted as saying or representing something physically meaningful. This means that intrinsic formalisms are better suited for getting a handle on what a theory says the world is like. Of course, those mathematical expressions will still need interpreting, but one does not need to do substantial mathematical heavy lifting in order to decide which bits of uninterpreted mathematics are candidates for being given a physical interpretation.

To illustrate this last point, I want to return to the example of Dewar's (2018) dynamics for Maxwell gravitation from §1. Dewar characterizes the models of Maxwell gravitation as follows: they are tuples these are tuples $\langle M, t_a, h^{ab}, [\nabla], T^{ab}\rangle$, where $\langle M, t_a, h^{ab}, [\nabla]\rangle$ is a Maxwellian spacetime, and T^{ab} the Newtonian mass-momentum tensor for whatever matter fields are present, such that at all points $p \in M$ where $\rho := T^{ab}t_a t_b \neq 0$, the following equations hold at p:

$$t_a \nabla_n T^{na} = \mathbf{0} \tag{1a}$$

$$\nabla_m(\rho^{-1}\nabla_n T^{nm}) = -4\pi\rho \tag{1b}$$

$$\nabla^c(\rho^{-1}\nabla_n T^{na}) - \nabla^a(\rho^{-1}\nabla_n T^{nc}) = \mathbf{0}, \tag{1c}$$

where ∇ is an arbitrary member of $[\nabla]$. To illustrate my first point—that it is non-trivial to see that the equations (1) are independent of the choice of $\nabla \in [\nabla]$—I will just note that this is the content of Dewar's (2018) proposition 2, which takes about a page of work to show. For my second point—that it is non-trivial to 'read off' the interpretation of these equations—note that the

25. I am grateful to an anonymous referee for suggesting this point.

value of the quantity $\nabla_n T^{na}$ is not independent of the choice of $\nabla \in [\nabla]$. So one cannot simply 'read off' e.g. the interpretation of (1a) as saying that the quantity $\nabla_n T^{na}$ is spacelike—for $\nabla_n T^{na}$ is not a candidate for representing any physical quantity at all. By contrast, in the intrinsic dynamics given for Maxwell gravitation by Chen (2023) and March (2023), (1a) is replaced by the continuity equation for the mass density ρ, $\pounds_\xi \rho - \frac{1}{2}\rho \hat{h}_{mn} \pounds_\xi h^{mn} = 0$, where ξ^a is a unit timelike vector field representing the net four-velocity of matter. Of course, this is also the conclusion about the interpretation of (1a) which Dewar (2018, 258) reaches; the point is not that one cannot arrive at this interpretation, but that because Dewar's presentation of Maxwell gravitation is extrinsic, it requires non-trivial mathematical work to massage (1a) into a form where one can see this. A similar point goes for (1b) and (1c).

I also take it that this is very much in the spirit of the concerns raised by Weatherall (2018), Chen (2023), and Wallace (2020) in their discussions of the problems they see with Earman's extrinsic characterization of Maxwellian spacetime—especially Weatherall's concern about how to interpret terms in calculations on Maxwellian spacetime, and Wallace's concern that characterizing Maxwellian spacetime extrinsically obscures what its dynamics really look like. Neither of these authors' points, I take it, are that one cannot arrive at a clear interpretation of Maxwell gravitation and calculations therewith, characterized extrinsically, nor that one cannot say the same things about Maxwell gravitation characterized extrinsically as one can when it is characterized intrinsically—the point is that it requires non-trivial work to do so.

5 Close

In this article, my aim has been to get clear on the special value of intrinsic formalisms, in the context of reformulating a theory with SRMs. I have argued that this value comes from the fact that which mathematical expressions may be interpreted as physically meaningful in an intrinsic formalism is trivial in a way in which it is not for extrinsic formalisms. The special value of an intrinsic formalism is just that it allows us to 'read off' what these mathematical expressions are.

Of course, being able to 'read off' a theory's ontology is also valuable, if one thinks that considerations of ontology have a privileged role to play in theory interpretation. But my focus on being able to read off what mathematical expressions may be interpreted as physically meaningful means that my argument about the value of intrinsic formalisms is able to stand apart from this issue. Rather than just being the purview of a certain brand of philosopher of physics, intrinsic formalisms should be of value to everyone.

Acknowledgements

I would like to thank Dominik Ehrenfels, Caspar Jacobs, and James Read for helpful discussions and comments. I am grateful to Balliol College, Oxford, and the Faculty of Philosophy, University of Oxford, for financial support.

References

Adlam, Emily, Niels Linnemann, and James Read. 2022. *Constructive Axiomatics in Spacetime Physics Part II: Constructive Axiomatics in Context.* https://arxiv.org/abs/2211.05672.

Belot, Gordon. 2000. "Geometry and Motion." *The British journal for the philosophy of science* 51:561–595.

Chen, Elliott. 2023. "Newtonian gravitation on Maxwell spacetime." [Forthcoming]. http://philsci-archive.pitt.edu/id/eprint/22440.

Dewar, Neil. 2018. "Maxwell gravitation." *Philosophy of Science* 85 (2): 249–270.

Dewar, Neil. 2019. "Sophistication about Symmetries." *The British Journal for the Philosophy of Science* 70 (2): 485–521. https://doi.org/10.1093/bjps/axx021. https://doi.org/10.1093/bjps/axx021.

Dewar, Neil. 2022. *Structure and Equivalence.* United Kingdom: Cambridge University Press.

Earman, John. 2024. *Explaining the Aharonov-Bohm Effect.* https://philsci-archive.pitt.edu/22970/.

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

Field, Hartry. 2016. *Science without Numbers.* 2nd ed. Oxford: OUP.

Hölder, Otto. 1901. *Die Axiome der Quantität und die Lehre vom Mass.* Teubner.

Jacobs, Caspar. 2022. "Invariance, intrinsicality, and perspicuity." *Synthese* 200 (2).

Jacobs, Caspar. 2021a. "Invariance or equivalence: a tale of two principles." *Synthese* 199 (3-4): 9337–9357.

Jacobs, Caspar. 2021b. "Symmetries as a Guide to the Structure of Physical Quantities." DPhil thesis, University of Oxford.

Krantz, David, Robert Duncan Luce, Patrick Suppes, and Amos Tversky. 1971. *Foundations of Measurement.* Mineola, New York: Dover Publications.

Leibniz, Gottfried Wilhelm, and Samuel Clarke. 1998. *The Leibniz-Clarke correspondence.* Edited by Henry Gavin Alexander. Manchester: Manchester University Press.

Lewis, David. 1986. "A Subjectivist's Guide to Objective Chance." In *Philosophical Papers: Volume II,* 83–132. Oxford: OUP. https://doi.org/10.1093/0195036468.003.0004.

Linnemann, Niels, and James Read. 2021. *Constructive Axiomatics in Spacetime Physics Part I: Walkthrough to the Ehlers-Pirani-Schild Axiomatisation.* https://doi.org/10.48550/ARXIV.2112.14063. https://arxiv.org/abs/2112.14063.

Luc, Joanna. 2023. "Motivationalism vs. interpretationalism about symmetries: some options overlooked in the debate about the relationship between symmetries and physical equivalence." *European journal for philosophy of science* 13 (3).

March, Eleanor. 2024. "Are Maxwell Gravitation and Newton-Cartan Theory Theoretically Equivalent?" Forthcoming, *British Journal for the Philosophy of Science,* https://doi.org/10.1086/730863.

March, Eleanor. 2023. *Maxwell gravitation without reference to equivalence classes of derivative operators.* https://philsci-archive.pitt.edu/22404/.

March, Eleanor, James Read, and Lu Chen. 2025. "Equivalence, reduction, and sophistication in teleparallel gravity." *European Journal for Philosophy of Science* 15 (2): 32. https://doi.org/10.1007/s13194-025-00658-0.

Martens, Niels C. M., and James Read. 2021. "Sophistry about symmetries?" *Synthese* 199 (1): 315–344. https://doi.org/10.1007/s11229-020-02658-4. https://doi.org/10.1007/s11229-020-02658-4.

Møller-Nielsen, Thomas. 2017. "Invariance, Interpretation, and Motivation." *Philosophy of Science* 84 (5): 1253–1264.

Mundy, Brent. 1986. "On the General Theory of Meaningful Representation." *Synthese* 67 (3): 391–437.

Myrvold, Wayne. 2019. "How could relativity be anything other than physical?" *Studies in the History and Philosophy of Modern Physics* 67:137–143.

Nguyen, James, Nicholas Teh, and Laura Wells. 2020. "Why surplus structure is not superfluous." *British Journal for the Philosophy of Science* 71 (2): 665–695.

Read, James. 2025. *Good VIBES only.* https://philsci-archive.pitt.edu/25073/.

Teh, Nicholas J. 2024. *The Philosophy of Symmetry.* Elements in the Philosophy of Physics. Cambridge University Press.

Wallace, David. 2020. "Fundamental and emergent geometry in Newtonian physics." *British Journal for the Philosophy of Science* 71 (1): 1–32.

Wallace, David. 2019. “Who’s afraid of coordinate systems? An essay on the representation of spacetime structure.” *Studies in the History and Philosophy of Modern Physics* 67:125–136.

Weatherall, James Owen. 2018. “A brief comment on Maxwell(/Newton)[-Huygens] spacetime.” *Studies in the History and Philosophy of Modern Physics* 63:34–38.

Weatherall, James Owen. 2016. “Are Newtonian gravity and geometrised Newtonian gravity theoretically equivalent?” *Erkenntnis* 81 (5): 1073–1091.

Weatherall, James Owen, and Helen Meskhidze. 2025. “Are General Relativity and Teleparallel Gravity Theoretically Equivalent?” *Philosophy of Physics* 3:6. https://doi.org/10.31389/pop.152.