Advanced Philosophy of Physics: The Philosophy of Symmetries

James Read[1]

[1]Faculty of Philosophy, University of Oxford, UK, OX2 6GG

HT25-W1

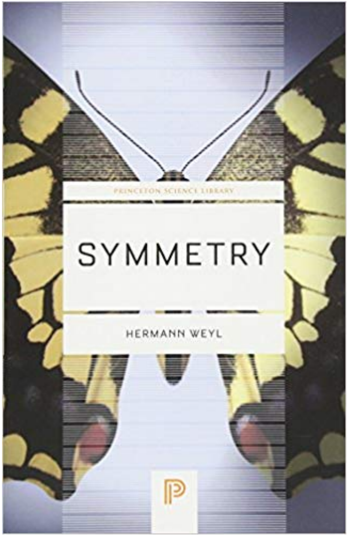
The plan

W1: The philosophy of symmetries

W2: The hole argument

W3: The Aharanov–Bohm effect

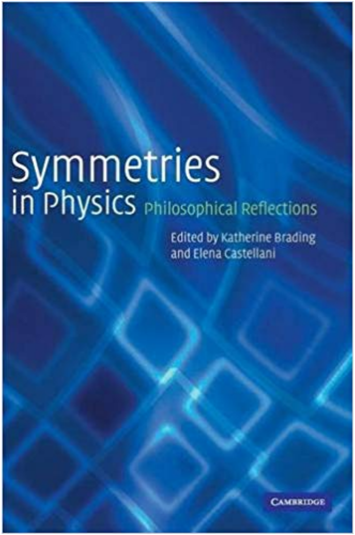
W4: The local validity of special relativity


PRINCETON SCIENCE LIBRARY
SYMMETRY
HERMANN WEYL

BAS C. VAN FRAASSEN

LAWS AND SYMMETRY

CLARENDON PAPERBACKS


Symmetries
in Physics
Philosophical Reflections
Edited by Katherine Brading
and Elena Castellani
CAMBRIDGE

“Symmetries in physics are a guide to reality.” (Dasgupta, 2016)

Today

Today

Kinematically possible models

- The *kinematically possible models* (KPMs) of a given theory specify the objects in terms of which the theory is defined.

Kinematically possible models

- The *kinematically possible models* (KPMs) of a given theory specify the objects in terms of which the theory is defined.
- For example, the KPMs of general relativity are triples $\langle M, g_{ab}, \Phi \rangle$, where M is a differentiable manifold, g_{ab} is a Lorentzian metric field on M, and Φ is a placeholder for matter fields.

Kinematically possible models

- The *kinematically possible models* (KPMs) of a given theory specify the objects in terms of which the theory is defined.
- For example, the KPMs of general relativity are triples $\langle M, g_{ab}, \Phi\rangle$, where M is a differentiable manifold, g_{ab} is a Lorentzian metric field on M, and Φ is a placeholder for matter fields.
- In Quine's terminology, KPMs pick out a theory's *ontology*.

Dynamically possible models

- The *dynamically possible models* (DPMs) of a given theory specify the dynamical equations satisfied by the objects in a theory's KPMs.

Dynamically possible models

- The *dynamically possible models* (DPMs) of a given theory specify the dynamical equations satisfied by the objects in a theory's KPMs.
- For example, the DPMs of general relativity are picked out by Einstein's equation, $G_{ab} = 8\pi T_{ab}$.

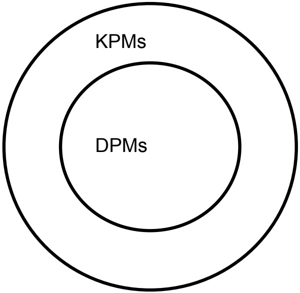
Dynamically possible models

- The *dynamically possible models* (DPMs) of a given theory specify the dynamical equations satisfied by the objects in a theory's KPMs.
- For example, the DPMs of general relativity are picked out by Einstein's equation, $G_{ab} = 8\pi T_{ab}$.
- In Quine's terminology, DPMs pick out a theory's *ideology*.

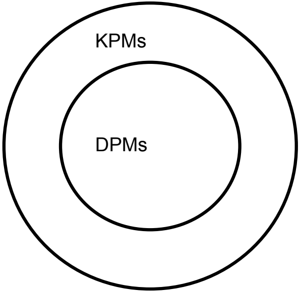
Dynamically possible models

- The *dynamically possible models* (DPMs) of a given theory specify the dynamical equations satisfied by the objects in a theory's KPMs.
- For example, the DPMs of general relativity are picked out by Einstein's equation, $G_{ab} = 8\pi T_{ab}$.
- In Quine's terminology, DPMs pick out a theory's *ideology*.
- A theory's DPMs therefore form a subclass of a theory's KPMs. Sometimes, a theory's DPMs are known as its *solutions*.

Kinematically and dynamically possible models



Kinematically and dynamically possible models



(For more on the kinematics/dynamics distinction, see Curiel (2016) and March (2024).)

Interpretation

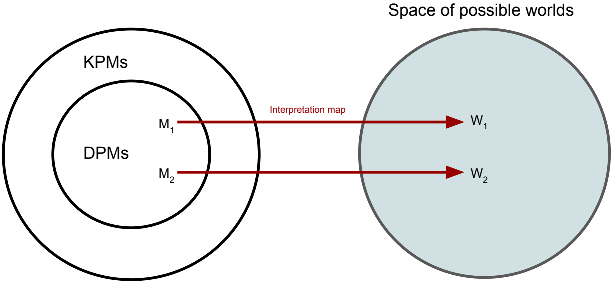
A theory's DPMs are *interpreted* as corresponding to possible worlds.

Interpretation

A theory's DPMs are *interpreted* as corresponding to possible worlds. (Or at least, this is a very common narrative which I won't question today.)

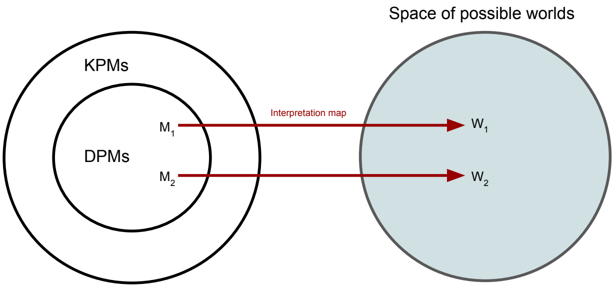
Interpretation

A theory's DPMs are *interpreted* as corresponding to possible worlds. (Or at least, this is a very common narrative which I won't question today.)



Interpretation

A theory's DPMs are *interpreted* as corresponding to possible worlds. (Or at least, this is a very common narrative which I won't question today.)



(Note: here, I am eliding certain important points about 'representational contexts', on which see Fletcher (2020) and Pooley & Read (2025).)

Symmetry and interpretation

Sometimes, multiple solutions of a given theory are interpreted as corresponding to the same possible world.

Symmetry and interpretation

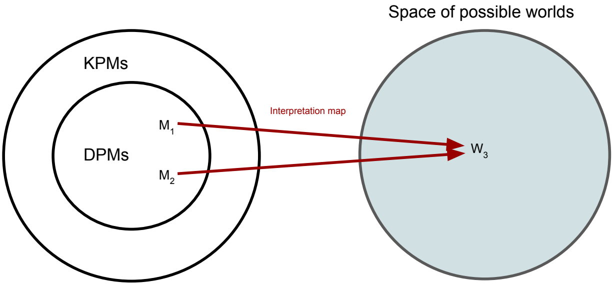
Sometimes, multiple solutions of a given theory are interpreted as corresponding to the same possible world.

In that case, one might say that there is ‘representational redundancy’ or ‘gauge redundancy’.

Symmetry and interpretation

Sometimes, multiple solutions of a given theory are interpreted as corresponding to the same possible world.

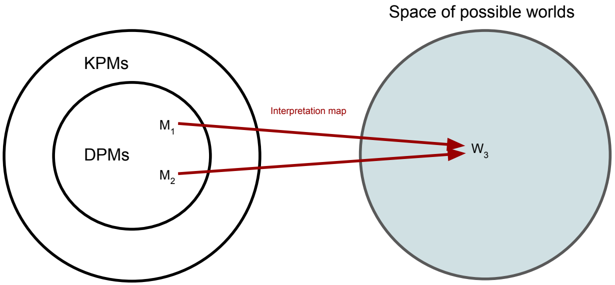
In that case, one might say that there is ‘representational redundancy’ or ‘gauge redundancy’.



Symmetry and interpretation

Sometimes, multiple solutions of a given theory are interpreted as corresponding to the same possible world.

In that case, one might say that there is ‘representational redundancy’ or ‘gauge redundancy’.

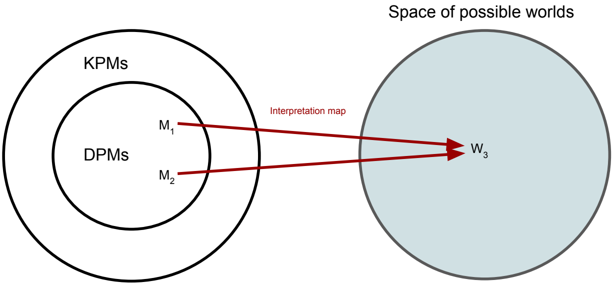


- Question: When is this the case?

Symmetry and interpretation

Sometimes, multiple solutions of a given theory are interpreted as corresponding to the same possible world.

In that case, one might say that there is ‘representational redundancy’ or ‘gauge redundancy’.



- Question: When is this the case?
- Tentative answer: In the presence of *symmetries*.

Symmetry and interpretation

If this tentative answer is to be substantiated, a number of questions must be addressed:

Symmetry and interpretation

If this tentative answer is to be substantiated, a number of questions must be addressed:

1. What *is* a symmetry transformation?
2. Should symmetry-related solutions of a given theory *invariably* be interpreted as representing the same physical state of affairs?
3. *How* do we articulate that putative common ontology of symmetry-related models?

Belot's catastrophic conclusion

Belot (2013) notes that the following two answers to, respectively, the first and second questions, are relatively widespread in the philosophical literature:

Belot's catastrophic conclusion

Belot (2013) notes that the following two answers to, respectively, the first and second questions, are relatively widespread in the philosophical literature:

- D1: The symmetries of a theory are those transformations that map solutions of the theory to solutions.
- D2: Two solutions of a theory are related by a symmetry transformation only if they are physically equivalent.

Belot's catastrophic conclusion

Belot (2013) notes that the following two answers to, respectively, the first and second questions, are relatively widespread in the philosophical literature:

- D1: The symmetries of a theory are those transformations that map solutions of the theory to solutions.
- D2: Two solutions of a theory are related by a symmetry transformation only if they are physically equivalent.

But the combination of D1 and D2 leads to catastrophe!

Belot's catastrophic conclusion

Belot (2013) notes that the following two answers to, respectively, the first and second questions, are relatively widespread in the philosophical literature:

- D1: The symmetries of a theory are those transformations that map solutions of the theory to solutions.
- D2: Two solutions of a theory are related by a symmetry transformation only if they are physically equivalent.

But the combination of D1 and D2 leads to catastrophe!

...So something has gone wrong with the orthodoxy.

Today's project

1. What *is* a symmetry transformation?
2. Should symmetry-related solutions of a given theory *invariably* be interpreted as representing the same physical state of affairs?
3. *How* do we articulate that putative common ontology of symmetry-related models?

We're going to work though the above questions in turn, and try to give more nuanced answers, which (hopefully!) avoid Belot's catastrophe.

Today

Two phases of the interpretative process

Caulton (2015) has presented a congenial framework, according to which there are two phases of the interpretative process:

Two phases of the interpretative process

Caulton (2015) has presented a congenial framework, according to which there are two phases of the interpretative process:

- In the **first phase** of the interpretative process, a model's *empirical* content is fixed: we establish how a world would look, according to that model.

Two phases of the interpretative process

Caulton (2015) has presented a congenial framework, according to which there are two phases of the interpretative process:

- In the **first phase** of the interpretative process, a model's *empirical* content is fixed: we establish how a world would look, according to that model.
- In the **second phase** of the interpretative process, a model's *physical* content is fixed: we establish the ontology of the world, according to that model.

Symmetry-to-reality based reasoning

- We've already seen the (widespread) claim that "symmetries are a guide to reality".

Symmetry-to-reality based reasoning

- We've already seen the (widespread) claim that "symmetries are a guide to reality".
- Regardless of how one *defines* symmetries (more on which later), it is the symmetries which relate *empirically equivalent* solutions of a given theory which are involved in symmetry-to-reality based reasoning.

Symmetry-to-reality based reasoning

- We've already seen the (widespread) claim that "symmetries are a guide to reality".
- Regardless of how one *defines* symmetries (more on which later), it is the symmetries which relate *empirically equivalent* solutions of a given theory which are involved in symmetry-to-reality based reasoning.
- Any two empirically equivalent solutions should have the structure which varies between them excised—since this structure is *ex hypothesi* undetectable, and should therefore, given Occam's razor, be eliminated. (See Dasgupta (2016).)

Examples of symmetries

Everyone since Leibniz and Clarke agrees that the following are examples of symmetries:

Examples of symmetries

Everyone since Leibniz and Clarke agrees that the following are examples of symmetries:

1. Static Leibniz shifts: Shift the entire material content of the universe five metres to the North.

Examples of symmetries

Everyone since Leibniz and Clarke agrees that the following are examples of symmetries:

1. Static Leibniz shifts: Shift the entire material content of the universe five metres to the North.
2. Kinematic Leibniz shifts: Boost the entire material content of the universe, such that all bodies now possess an extra velocity component of five metres per second in the North direction.

Examples of symmetries

Everyone since Leibniz and Clarke agrees that the following are examples of symmetries:

1. Static Leibniz shifts: Shift the entire material content of the universe five metres to the North.
2. Kinematic Leibniz shifts: Boost the entire material content of the universe, such that all bodies now possess an extra velocity component of five metres per second in the North direction.
3. Constant shifts of the gravitational/electrostatic potential.

Today

Theories and interpretation

Further preliminaries

What is a symmetry?

The normative import of symmetries

Articulating common ontology

Conclusions

Three approaches to symmetries

Dasgupta (2016) distinguishes three possible approaches to the definition of symmetry transformations: *formal*, *ontic*, and *epistemic*:

Three approaches to symmetries

Dasgupta (2016) distinguishes three possible approaches to the definition of symmetry transformations: *formal*, *ontic*, and *epistemic*:

- **Formal approaches** “define [the notion of symmetry] in purely formal, set-theoretic terms” (Dasgupta, p. 861).

Three approaches to symmetries

Dasgupta (2016) distinguishes three possible approaches to the definition of symmetry transformations: *formal*, *ontic*, and *epistemic*:

- **Formal approaches** “define [the notion of symmetry] in purely formal, set-theoretic terms” (Dasgupta, p. 861).
- **Ontic approaches** “define a symmetry of a law to be a function that preserves the law and also preserves ... [salient physical] features F” (Dasgupta, p. 862).

Three approaches to symmetries

Dasgupta (2016) distinguishes three possible approaches to the definition of symmetry transformations: *formal*, *ontic*, and *epistemic*:

- **Formal approaches** "define [the notion of symmetry] in purely formal, set-theoretic terms" (Dasgupta, p. 861).
- **Ontic approaches** "define a symmetry of a law to be a function that preserves the law and also preserves ... [salient physical] features F" (Dasgupta, p. 862).
- **Epistemic approaches** define a symmetry such that "given any set of laws, any two situations related by a symmetry of those laws are observationally equivalent" (Dasgupta, p. 866).

Dasgupta on formal definitions of symmetries

The central problem raised by Dasgupta against formal approaches to symmetries is as follows. Any such definition

> *... must imply that given any set of laws, any two systems related by a symmetry of those laws will be observationally equivalent. And it is (to put it mildly) extremely hard to see how any purely formal definition could have this consequence.* (Dasgupta, p. 861)

Responses

1. Why assume that all symmetry-related models must be empirically equivalent? In other words: why assume that all symmetries must be involved in symmetry-to-reality based reasoning?

Responses

1. Why assume that all symmetry-related models must be empirically equivalent? In other words: why assume that all symmetries must be involved in symmetry-to-reality based reasoning?
2. Who said that establishing the observational equivalence of models was easy?

Ontic definitions

- According to ontic definitions, symmetry transformations preserve some salient physical quantities.

Ontic definitions

- According to ontic definitions, symmetry transformations preserve some salient physical quantities.
- Examples include dynamical symmetries, Lagrangian symmetries, etc.

Dasgupta on ontic definitions of symmetries

Dasgupta's central problem with ontic definitions is what he dubs the 'problem of inferential circularity':

> *The objection is that they get the order of justification backwards: we often use premises about symmetries in order to work out which physical features fix the data, so we cannot at the same time define symmetries to be those operations that preserve features that fix the data.* (Dasgupta 2016, p. 865)

Responses

1. Not all ontic definitions need be involved in symmetry-to-reality based reasoning (only the ones which preserve the observational data).
2. We don't fix *all* the physical quantities when we offer an ontic definition—just some. So it's not clear that the problem of inferential circularity is damning.

Responses

1. Not all ontic definitions need be involved in symmetry-to-reality based reasoning (only the ones which preserve the observational data).
2. We don't fix *all* the physical quantities when we offer an ontic definition—just some. So it's not clear that the problem of inferential circularity is damning.

(For more on these responses, see Read & Møller-Nielsen (2020b).)

Dasgupta on epistemic definitions of symmetries

In light of the (apparent) problems for formal and ontic definitions, Dasgupta proposes that we should embrace an *epistemic* approach to symmetries: symmetries *just are* transformations between empirically equivalent solutions.

Responses

1. Isn’t this defining ourselves out of a problem?

Responses

1. Isn't this defining ourselves out of a problem?
2. The approach is not consonant with physical practice.

Responses

1. Isn’t this defining ourselves out of a problem?
2. The approach is not consonant with physical practice.
3. Why insist that *all* symmetries have to be involved in symmetry-to-reality based reasoning?

Responses

1. Isn't this defining ourselves out of a problem?
2. The approach is not consonant with physical practice.
3. Why insist that *all* symmetries have to be involved in symmetry-to-reality based reasoning?
4. Doesn't this make the notion of a symmetry redundant as a tool for metaphysical theorising about scientific theories, in the sense that the approach reduces the notion of a symmetry to an Occamist norm?

Responses

1. Isn't this defining ourselves out of a problem?
2. The approach is not consonant with physical practice.
3. Why insist that *all* symmetries have to be involved in symmetry-to-reality based reasoning?
4. Doesn't this make the notion of a symmetry redundant as a tool for metaphysical theorising about scientific theories, in the sense that the approach reduces the notion of a symmetry to an Occamist norm?

(See Read & Møller-Nielsen (2020b).)

Today

A normative question

Focus in the remainder on symmetry-related solutions which *are* interpreted as being empirically equivalent (whether by definition or otherwise).

A normative question

Focus in the remainder on symmetry-related solutions which *are* interpreted as being empirically equivalent (whether by definition or otherwise).

Should symmetry-related solutions be interpreted *ab initio* as being physically equivalent? (Møller-Nielsen 2017)

Of course, this isn't to say that there is no value to reformulating a theory's formalism in such a way that the surplus structure is made manifest, so that we can move to a formalism in which it is expunged entirely. Such a presentation lets us see what it is we are committed to by our (qualified) realism about the theory; if we want to know the answers to specific questions about the nature of a theory's ontology and ideology, then this is invaluable. ... But if we lack the mathematical tools to do so, then I maintain that there is nothing wrong with recognising that one's realism will only extend to structures that are invariant under the symmetries—whatever those may turn out to be. (Dewar 2015, pp. 326-7)

Of course, this isn't to say that there is no value to reformulating a theory's formalism in such a way that the surplus structure is made manifest, so that we can move to a formalism in which it is expunged entirely. Such a presentation lets us see what it is we are committed to by our (qualified) realism about the theory; if we want to know the answers to specific questions about the nature of a theory's ontology and ideology, then this is invaluable. ... But if we lack the mathematical tools to do so, then I maintain that there is nothing wrong with recognising that one's realism will only extend to structures that are invariant under the symmetries—whatever those may turn out to be. (Dewar 2015, pp. 326-7)

I argue, contrary to current orthodoxy, that the variance of a quantity under a theory's symmetries is not a sufficient basis for interpreting that theory as being uncommitted to the reality of that quantity. Rather, I argue, the variance of a quantity under symmetries only ever serves as a motivation to refrain from any commitment to the quantity in question. (Møller-Nielsen 2017, p. 1253)

Interpretation and motivation

- On the **interpretational approach** (Dewar), one may regard symmetry-related models *ab initio* as being physically equivalent.

Interpretation and motivation

- On the **interpretational approach** (Dewar), one may regard symmetry-related models *ab initio* as being physically equivalent.
- On the **motivational approach** (Møller-Nielsen), symmetry-related models may *only* be regarded as being physically equivalent once one has to hand a perspicuous explication of their common ontology.

Perspicuous explication

- In the case of the Leibnizian static shift, the perspicuous explication of the common ontology of the symmetry-related models is Newtonian gravity, married with *anti-haecceitism* about spacetime points.

Perspicuous explication

- In the case of the Leibnizian static shift, the perspicuous explication of the common ontology of the symmetry-related models is Newtonian gravity, married with *anti-haecceitism* about spacetime points.
- In the case of the Leibnizian kinematic shift, the perspicuous explication of the common ontology of the symmetry-related models is Newtonian gravity, expunged of the persisting points of absolute space (so absolute velocity not meaningful), while retaining affine structure (so absolute accelerations still meaningful).

Perspicuous explication

- In the case of the Leibnizian static shift, the perspicuous explication of the common ontology of the symmetry-related models is Newtonian gravity, married with *anti-haecceitism* about spacetime points.
- In the case of the Leibnizian kinematic shift, the perspicuous explication of the common ontology of the symmetry-related models is Newtonian gravity, expunged of the persisting points of absolute space (so absolute velocity not meaningful), while retaining affine structure (so absolute accelerations still meaningful).

Møller-Nielsen: We can only declare these shifted models to represent the same physical state of affairs once we've done the metaphysical hard graft of figuring out what that physical state of affairs *actually is*.

Interpretation/motivation and shifts

> *... [W]e can draw the conclusion of the inference [viz., that shifted solutions should be regarded as physically equivalent] only when we have the alternative theory in hand and have shown that all else is equal. This explains why it was rational for Newton to believe in absolute velocity even though he knew that it was variant ... and undetectable. The reason this was rational for him was that he had no good alternative theory to hand. He had good reason (his bucket argument) to think that relationalism was not empirically adequate. And relationalism was the only alternative view he knew of (he was not aware of Galilean space-time structures in which there is a well-defined feature of absolute acceleration ... but no absolute velocity). So for Newton, all else was not equal and he was rational to believe in absolute velocity.* (Dasgupta 2016, p. 854)

Two challenges for the interpretational approach

1. How are we to identify the common structure associated with symmetry-related models—and have we any reason to think that such structure is always there to be found?
2. Even supposing that such structure can be found, does it invariably admit of a coherent physical interpretation?

More nuanced positions in the debate

Luc (2023) identifies three salient questions in the interpretationalism/motivationalism debate:

More nuanced positions in the debate

Luc (2023) identifies three salient questions in the interpretationalism/motivationalism debate:

1. What should our *initial* reaction towards symmetry-related models of T be—should we regard them as physically equivalent or as physically inequivalent?
2. Should we look for a perspicuous account of the ontology shared by symmetry-related models of T?
3. How should we update our interpretation of symmetry-related models of T depending on the outcomes of the research mentioned in question (2)?

More fine-grained options in the debate

A. Motivationalism
B. Interpretationalism without motivation
C. Steadfast interpretationalism with motivation
D. Concessive interpretationalism with motivation
E. Graded interpretationalism with motivation

A: Motivationalism

Motivationalism gives us determinate answers to each of the above three questions:

A: Motivationalism

Motivationalism gives us determinate answers to each of the above three questions:

1. We should initially interpret symmetry-related models of T as physically inequivalent.
2. Yes.
3. If we find such an account, we should change our initial interpretation and begin to regard symmetry-related models of T as physically equivalent; if we do not find such an account, we should retain our initial interpretation.

B: Interpretationalism without motivation

Interpretationalism without motivation gives us determinate answers to each of the above three questions:

B: Interpretationalism without motivation

Interpretationalism without motivation gives us determinate answers to each of the above three questions:

1. We should initially interpret symmetry-related models of T as physically equivalent.
2. There is no need to do this.
3. Irrespective of whether we find such an account or not, we should retain our initial interpretation.

C: Steadfast interpretationalism with motivation

Steadfast interpretationalism with motivation gives us determinate answers to each of the above three questions:

C: Steadfast interpretationalism with motivation

Steadfast interpretationalism with motivation gives us determinate answers to each of the above three questions:

1. We should initially interpret symmetry-related models of T as physically equivalent.
2. Yes.
3. Irrespective of whether we find such an account or not, we should retain our initial interpretation.

D: Concessive interpretationalism with motivation

Concessive interpretationalism with motivation gives us determinate answers to each of the above three questions:

D: Concessive interpretationalism with motivation

Concessive interpretationalism with motivation gives us determinate answers to each of the above three questions:

1. We should initially interpret symmetry-related models of T as physically equivalent.
2. Yes.
3. If we find such an account, we should retain our initial interpretation, whereas if despite lots of effort we do not succeed in finding it, we should change our interpretation and begin to regard symmetry-related models of T as physically inequivalent.

E: Graded interpretationalism with motivation

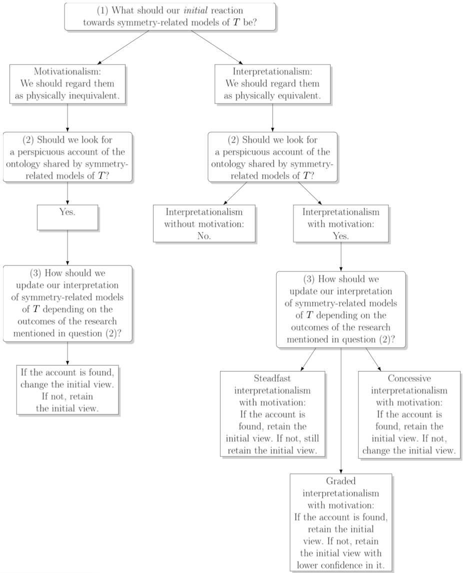
Graded interpretationalism with motivation gives us determinate answers to each of the above three questions:

E: Graded interpretationalism with motivation

Graded interpretationalism with motivation gives us determinate answers to each of the above three questions:

1. We should initially interpret symmetry-related models of T as physically equivalent.
2. Yes.
3. If we find such an account, we should retain our initial interpretation, whereas if despite lots of effort we do not succeed in finding it, we should still retain our initial interpretation as more plausible than its opposite, but we should significantly decrease our confidence about this interpretation.

A helpful chart



Rapprochement in the debate?

> *Therefore, concessive interpretationalism with motivation is as close to motivationalism as interpretationalism can be. It differs from motivationalism only in what it claims to be the most reasonable prima facie attitude towards models of T that we have recently discovered to be symmetry-related.* (Luc 2023, p. 20)

Rapprochement in the debate?

> *Therefore, concessive interpretationalism with motivation is as close to motivationalism as interpretationalism can be. It differs from motivationalism only in what it claims to be the most reasonable prima facie attitude towards models of T that we have recently discovered to be symmetry-related.* (Luc 2023, p. 20)

I, for one, would be happy to sign up to this—but you can make up your own minds!

Cases where motivation is lacking

- There are also cases of symmetry-related models where we simply lack motivation to regard them as being physically equivalent—even if they are empirically equivalent!

Cases where motivation is lacking

- There are also cases of symmetry-related models where we simply lack motivation to regard them as being physically equivalent—even if they are empirically equivalent!
- For example, Belot (2018) considers models of general relativity which differ by boundary diffeomorphisms: these are symmetry-related, but if we regard them as being physically equivalent then we lose the ability to define conserved quantities.

Cases where motivation is lacking

- There are also cases of symmetry-related models where we simply lack motivation to regard them as being physically equivalent—even if they are empirically equivalent!
- For example, Belot (2018) considers models of general relativity which differ by boundary diffeomorphisms: these are symmetry-related, but if we regard them as being physically equivalent then we lose the ability to define conserved quantities.
- (For more, see Luc (2022) and Read (2023, ch. 2).)

Today

An interpretative question

How are we to articulate the common ontology of symmetry-related solutions?

An interpretative question

How are we to articulate the common ontology of symmetry-related solutions?

(Note that this is distinct from our previous, normative question.)

Dewar on this issue

> *It is often claimed that the symmetries of a theory reveal "surplus structure": structure which, in some sense, the theory could do without. For example, the boost symmetry of Newtonian mechanics indicates the superfluousness of absolute velocities; the gauge symmetry of electromagnetism reveals the superfluousness of absolute potentials; and so on and so forth. Moreover, it is widely held that if this is the case, then some modification of one's theory is appropriate, so as to make explicit what structure is not surplus (e.g. the replacement of Newtonian by Galilean spacetime, in response to the boost symmetry of Newtonian mechanics). [...] I compare and contrast two ways of making such a modification. The first is to replace the theory by (what I shall call) a reduced theory: a theory that deals only in quantities which are invariant under the relevant symmetry. The second is to replace the theory by (what I shall call) a sophisticated theory: a theory in which models related by a symmetry are isomorphic.* (Dewar 2019, pp. 485-6)

Dewar on this issue

It is often claimed that the symmetries of a theory reveal "surplus structure": structure which, in some sense, the theory could do without. For example, the boost symmetry of Newtonian mechanics indicates the superfluousness of absolute velocities; the gauge symmetry of electromagnetism reveals the superfluousness of absolute potentials; and so on and so forth. Moreover, it is widely held that if this is the case, then some modification of one's theory is appropriate, so as to make explicit what structure is not surplus (e.g. the replacement of Newtonian by Galilean spacetime, in response to the boost symmetry of Newtonian mechanics). [...] I compare and contrast two ways of making such a modification. The first is to replace the theory by (what I shall call) a reduced theory: a theory that deals only in quantities which are invariant under the relevant symmetry. The second is to replace the theory by (what I shall call) a sophisticated theory: a theory in which models related by a symmetry are isomorphic. (Dewar 2019, pp. 485-6)

Reduction and sophistication

- **Reduction**: “the idea is that we (i) identify some collection of invariants of the original theory; (ii) specify a theory in terms of those invariants; and (iii) show that the new theory captures all the symmetry-invariant content of the old theory.” (Dewar 2019, pp. 492-3)

Reduction and sophistication

- **Reduction**: “the idea is that we (i) identify some collection of invariants of the original theory; (ii) specify a theory in terms of those invariants; and (iii) show that the new theory captures all the symmetry-invariant content of the old theory.” (Dewar 2019, pp. 492-3)
- **Sophistication**: “the idea is that we need not insist on finding a theory whose models are invariant under the application of the symmetry transformation [...] the proposal is that we instead look for a theory such that [symmetry-related models of the original theory] *M* and *N* give rise to distinct but isomorphic models.” (Dewar 2019, p. 498)

A threefold distinction, summarised

- **Reduction**: In the presence of empirically equivalent symmetry-related models of some theory T, construct some new theory T' such that those symmetry-related models of T all map to some unique model of T'.
- **Internal sophistication**: In the presence of empirically equivalent symmetry-related models of some theory T, 'forget' structure while retaining the same number of models, in order to construct some new theory T' such that the symmetries of T under consideration (which needn't act as isomorphisms) act as isomorphisms of T'.
- **External sophistication**: In the presence of empirically equivalent symmetry-related models of some theory T, treat those models of T 'as if' they are isomorphic.

Two orthogonal debates

1. Interpretation + reduction (Caulton 2015)
2. Interpretation + sophistication (Dewar 2019)
3. Motivation + reduction (Dasgupta 2016)
4. Motivation + sophistication (Jacobs 2021)
5. Some more complicated cocktail (Møller-Nielsen 2017)

Reduction vs internal sophistication: illustration

- Let's illustrate the difference between reduction and internal sophistication with the simple example of electromagnetism.

Reduction vs internal sophistication: illustration

- Let's illustrate the difference between reduction and internal sophistication with the simple example of electromagnetism.
- Consider vector potential-based, source-free electromagnetism—KPMs are given by $\langle M, \eta_{ab}, A^a \rangle$; DPMs are given by the dynamical equations $d * dA = 0$ and $ddA = 0$ (the latter of which is in fact a mathematical identity by the Poincaré lemma).

Reduction vs internal sophistication: illustration

- Let's illustrate the difference between reduction and internal sophistication with the simple example of electromagnetism.
- Consider vector potential-based, source-free electromagnetism—KPMs are given by $\langle M, \eta_{ab}, A^a \rangle$; DPMs are given by the dynamical equations $d * dA = 0$ and $ddA = 0$ (the latter of which is in fact a mathematical identity by the Poincaré lemma).
- This theory has a *gauge symmetry*, $A \mapsto A + d\Lambda$; these symmetries do not relate isomorphic models.

Reduced version of electromagnetism

- Models of this version of electromagnetism which differ by a gauge transformation *agree* on the 'Faraday tensor' $F = dA$, which directly encodes the electric and magnetic fields.

Reduced version of electromagnetism

- Models of this version of electromagnetism which differ by a gauge transformation *agree* on the ‘Faraday tensor’ $F = dA$, which directly encodes the electric and magnetic fields.
- Thus, the reduced version of electromagnetism has KPMs $\langle M, \eta_{ab}, F_{ab} \rangle$, and DPMs given by $d * F = 0$ and $dF = 0$ (note that the latter is no longer a mathematical identity!).

Reduced version of electromagnetism

- Models of this version of electromagnetism which differ by a gauge transformation *agree* on the 'Faraday tensor' $F = dA$, which directly encodes the electric and magnetic fields.
- Thus, the reduced version of electromagnetism has KPMs $\langle M, \eta_{ab}, F_{ab} \rangle$, and DPMs given by $d * F = 0$ and $dF = 0$ (note that the latter is no longer a mathematical identity!).
- This is indeed a reduced theory, since gauge fields of A-EM map to unique models of F-EM.

Internally sophisticated version of electromagnetism

- One can also ‘forget’ about structure in the models of A-EM in order to construct a new version of the theory which has the same number of models, but where the symmetries now act as isomorphisms.

Internally sophisticated version of electromagnetism

- One can also ‘forget’ about structure in the models of *A*-EM in order to construct a new version of the theory which has the same number of models, but where the symmetries now act as isomorphisms.
- This will lead to a version of electromagnetism based upon *principal fibre bundles*. I will explain this in much more detail in Week 3!

Internally sophisticated version of electromagnetism

- One can also ‘forget’ about structure in the models of *A*-EM in order to construct a new version of the theory which has the same number of models, but where the symmetries now act as isomorphisms.
- This will lead to a version of electromagnetism based upon *principal fibre bundles*. I will explain this in much more detail in Week 3!
- In brief, the idea is that one moves to a theory with KPMs $\langle M, \eta_{ab}, P, \omega \rangle$, where P is a principal bundle, and ω is a connection on P. DPMs are given by the (abelian) Yang–Mills equations (I won’t write them).

Internally sophisticated version of electromagnetism

- One can also ‘forget’ about structure in the models of A-EM in order to construct a new version of the theory which has the same number of models, but where the symmetries now act as isomorphisms.
- This will lead to a version of electromagnetism based upon *principal fibre bundles*. I will explain this in much more detail in Week 3!
- In brief, the idea is that one moves to a theory with KPMs $\langle M, \eta_{ab}, P, \omega \rangle$, where P is a principal bundle, and ω is a connection on P. DPMs are given by the (abelian) Yang–Mills equations (I won’t write them).
- Gauge symmetries now take $\langle M, \eta_{ab}, P, \omega \rangle$ to $\langle M, \eta_{ab}, P, d_* \omega \rangle$—so act as isomorphisms!

Internal sophistication and the need for anti-haecceitism/anti-quidditism

- When we internally sophisticate, we still have distinct models—we've just made them isomorphic.

Internal sophistication and the need for anti-haecceitism/anti-quidditism

- When we internally sophisticate, we still have distinct models—we've just made them isomorphic.
- So the case is akin to the Newtonian static shift scenario discussed previously.

Internal sophistication and the need for anti-haecceitism/anti-quidditism

- When we internally sophisticate, we still have distinct models—we've just made them isomorphic.
- So the case is akin to the Newtonian static shift scenario discussed previously.
- There, being an anti-haecceitist—i.e., denying that the physical spacetime points have primitive identities—then allowed us to say that those models represent the same world.

Internal sophistication and the need for anti-haecceitism/anti-quidditism

- When we internally sophisticate, we still have distinct models—we've just made them isomorphic.
- So the case is akin to the Newtonian static shift scenario discussed previously.
- There, being an anti-haecceitist—i.e., denying that the physical spacetime points have primitive identities—then allowed us to say that those models represent the same world.
- The situation when we move to fibre bundle EM is similar—but now the diffeomorphism d is on the value space, not spacetime!

Internal sophistication and the need for anti-haecceitism/anti-quidditism

- When we internally sophisticate, we still have distinct models—we've just made them isomorphic.
- So the case is akin to the Newtonian static shift scenario discussed previously.
- There, being an anti-haecceitist—i.e., denying that the physical spacetime points have primitive identities—then allowed us to say that those models represent the same world.
- The situation when we move to fibre bundle EM is similar—but now the diffeomorphism d is on the value space, not spacetime!
- To deny that points in value space have primitive identities is 'anti-quidditism'; embracing this is again sufficient to assert that these symmetry-related models represent the same possible world.

Externally sophisticated version of electromagnetism

- External sophistication would take the models of A-EM and treat them ‘as if’ they’re isomorphic, without doing the mathematical formulation which is associated with internal sophistication.

Externally sophisticated version of electromagnetism

- External sophistication would take the models of A-EM and treat them 'as if' they're isomorphic, without doing the mathematical formulation which is associated with internal sophistication.
- This has faced charges of metaphysical unperspicuity, especially from motivationalists—see e.g. Martens & Read (2020).

Externally sophisticated version of electromagnetism

- External sophistication would take the models of *A*-EM and treat them 'as if' they're isomorphic, without doing the mathematical formulation which is associated with internal sophistication.
- This has faced charges of metaphysical unperspicuity, especially from motivationalists—see e.g. Martens & Read (2020).

(To be clear: I like internal sophistication but not external sophistication!)

Merits of reduction vs internal sophistication

- Let's set aside external sophistication in what follows.

Merits of reduction vs internal sophistication

- Let's set aside external sophistication in what follows.
- There's an emerging consensus that internal sophistication is to be preferred over reduction.

Merits of reduction vs internal sophistication

- Let's set aside external sophistication in what follows.
- There's an emerging consensus that internal sophistication is to be preferred over reduction.
- E.g., for the following reasons:

Merits of reduction vs internal sophistication

- Let's set aside external sophistication in what follows.
- There's an emerging consensus that internal sophistication is to be preferred over reduction.
- E.g., for the following reasons:
  - Dynamical equations in reduced theory can become mathematical identities in internally sophisticated theories.

Merits of reduction vs internal sophistication

- Let's set aside external sophistication in what follows.
- There's an emerging consensus that internal sophistication is to be preferred over reduction.
- E.g., for the following reasons:
  - Dynamical equations in reduced theory can become mathematical identities in internally sophisticated theories.
  - Reduced theories sometimes don't admit of action principles, etc.

Merits of reduction vs internal sophistication

- Let's set aside external sophistication in what follows.
- There's an emerging consensus that internal sophistication is to be preferred over reduction.
- E.g., for the following reasons:
  - Dynamical equations in reduced theory can become mathematical identities in internally sophisticated theories.
  - Reduced theories sometimes don't admit of action principles, etc.
  - Easier to model subsystems with internally sophisticated theories (cf. Rovelli 2014) and other representational benefits (Bradley 2024).

Merits of reduction vs internal sophistication

- Let's set aside external sophistication in what follows.
- There's an emerging consensus that internal sophistication is to be preferred over reduction.
- E.g., for the following reasons:
  - Dynamical equations in reduced theory can become mathematical identities in internally sophisticated theories.
  - Reduced theories sometimes don't admit of action principles, etc.
  - Easier to model subsystems with internally sophisticated theories (cf. Rovelli 2014) and other representational benefits (Bradley 2024).
- For some pushback against this emerging consensus, see (Ehrenfels 2025).

Merits of reduction vs internal sophistication

- Let’s set aside external sophistication in what follows.
- There’s an emerging consensus that internal sophistication is to be preferred over reduction.
- E.g., for the following reasons:
  - Dynamical equations in reduced theory can become mathematical identities in internally sophisticated theories.
  - Reduced theories sometimes don’t admit of action principles, etc.
  - Easier to model subsystems with internally sophisticated theories (cf. Rovelli 2014) and other representational benefits (Bradley 2024).
- For some pushback against this emerging consensus, see (Ehrenfels 2025).

There’s very likely still more to be worked out here...

Today

JR's package of views

- I'll close by putting my cards on the table and giving you my own package of views...

JR's package of views

- I'll close by putting my cards on the table and giving you my own package of views...
- ...you're welcome to adopt a different package!

1. **Definition of symmetries:** (J.R. & T.M-N., 2020b)
    - Symmetries should not be defined in epistemic terms—on pain of redundancy/lack of faithfulness to physics practice.
    - Those symmetries which are *discovered* to relate empirically equivalent solutions should be involved in symmetry-to-reality based reasoning (by Occam's razor).

1. **Definition of symmetries:** (J.R. & T.M-N., 2020b)
   - Symmetries should not be defined in epistemic terms—on pain of redundancy/lack of faithfulness to physics practice.
   - Those symmetries which are *discovered* to relate empirically equivalent solutions should be involved in symmetry-to-reality based reasoning (by Occam's razor).
2. **Interpretation/motivation:** (J.R. & T.M-N., 2020a)
   - Symmetry-related solutions should only be regarded as being physically equivalent when we have to hand a coherent explication of their common ontology.

1. **Definition of symmetries:** (J.R. & T.M-N., 2020b)
   - Symmetries should not be defined in epistemic terms—on pain of redundancy/lack of faithfulness to physics practice.
   - Those symmetries which are *discovered* to relate empirically equivalent solutions should be involved in symmetry-to-reality based reasoning (by Occam's razor).
2. **Interpretation/motivation:** (J.R. & T.M-N., 2020a)
   - Symmetry-related solutions should only be regarded as being physically equivalent when we have to hand a coherent explication of their common ontology.
3. **Reduction/sophistication:** (N.M. & J.R., 2020)
   - When symmetry-related solutions are isomorphic, we can appeal to anti-haecceitism/anti-quidditism to explicate the common ontology of those solutions.
   - When symmetry-related solutions are not isomorphic, we must appeal to reduction or internal sophistication to explicate the common ontology of those solutions.
   - As a means of explicating the ontology of symmetry-related solutions, external sophistication is unperspicuous.

References I

Gordon Belot, “Symmetry and Equivalence”, in R. Batterman (ed.), *The Oxford Handbook of Philosophy of Physics*, Oxford: Oxford University Press, pp. 318-339, 2013.

Gordon Belot, “Fifty Million Elvis Fans Can’t be Wrong”, Noûs 52(4), pp. 946–81, 2018.

Katherine Brading and Elena Castellani (eds.), *Symmetries in Physics: Philosophical Reflections*, Cambridge: Cambridge University Press, 2003.

Katherine Brading and Nicholas J. Teh, “Symmetry and Symmetry Breaking”, in E. N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, 2017.

Clara Bradley, “The Representational Role of Sophisticated Theories”, Philosophy of Science, 2024.

Adam Caulton, “The Role of Symmetry in the Interpretation of Physical Theories”, Studies in History and Philosophy of Modern Physics 52, pp. 153-162, 2015.

Erik Curiel, “Kinematics, Dynamics, and the Structure of Physical Theory”, 2016.

Shamik Dasgupta, “Symmetry as an Epistemic Notion (Twice Over)”, British Journal for the Philosophy of Science 67(3), pp. 837-878, 2016.

Neil Dewar, “Symmetries and the Philosophy of Language”, Studies in the History and Philosophy of Modern Physics 52, pp. 317-327, 2015.

Neil Dewar, “Sophistication About Symmetries”, British Journal for the Philosophy of Science 70(2), pp. 485-521, 2019.

Dominik Ehrenfels, “Reviving Reduction”, British Journal for the Philosophy of Science, forthcoming, 2025.

Samuel C. Fletcher, “On Representational Capacities, with an Application to General Relativity”, Foundations of Physics 50, pp. 228–249, 2020.

References II

- Jenann Ismael and Bas van Fraassen, “Symmetry as a Guide to Superfluous Theoretical Structure”, in K. Brading and E. Castellani (eds.), *Symmetries in Physics: Philosophical Reflections*, Cambridge: Cambridge University Press, pp. 371-392, 2003.
- Caspar Jacobs, *Symmetries as a Guide to the Structure of Physical Quantities*, D.Phil. thesis, University of Oxford, 2021.
- Joanna Luc, “Arguments from Scientific Practice in the Debate About the Physical Equivalence of Symmetry-related Models”, Synthese 200(72), 2022.
- Joanna Luc, “Motivationalism vs. Interpretationalism About Symmetries: Some Options Overlooked in the Debate About the Relationship Between Symmetries and Physical Equivalence”, European Journal for the Philosophy of Science 13(40), 2023.
- Eleanor March, “Categorical Equivalence and the Kinematics-Dynamics Distinction”, British Journal for the Philosophy of Science, 2024.
- Niels Martens and James Read, “Sophistry About Symmetries?”, Synthese, 2021.
- Thomas Møller-Nielsen, “Invariance, Interpretation, and Motivation”, Philosophy of Science 84, pp. 1253-1264, 2017.
- Robert Nozick, *Invariances: The Structure of the Objective World*, Cambridge, MA: Harvard University Press, 2001.
- Oliver Pooley and James Read, “On the Mathematics and Metaphysics of the Hole Argument”, British Journal for the Philosophy of Science, 2025.
- W. V. O. Quine, “Ontology and Ideology”, Philosophical Studies 2(1), pp. 11-15, 1951.

References III

James Read, *Background Independence in Classical and Quantum Gravity*, Oxford: Oxford University Press, 2023.

James Read and Thomas Møller-Nielsen, “Motivating Dualities”, Synthese 197, pp. 263-291, 2020.

James Read and Thomas Møller-Nielsen, “Redundant Epistemic Symmetries”, Studies in History and Philosophy of Modern Physics 70, pp. 88-97, 2020.

Carlo Rovelli, “Why Gauge?”, Foundations of Physics, 2014.

Theodore Sider, *The Tools of Metaphysics and the Metaphysics of Science*, Oxford: Oxford University Press, 2020.

Bas van Fraassen, *Laws and Symmetry*, Oxford: Clarendon Press, 1989.