Symmetry and Superfluous Structure: lessons from history and tempered enthusiasm[i]

Jenann Ismael

This is a draft prepared for a talk at a UC, Riverside workshop on Symmetries in Physics

At its most general level, we can think of the task of the naturalistic metaphysician as that of inferring ontology from data. Since ontology comes paired with laws the task is really to produce a pairing of ontology and laws that generates the data. In its most abstract form, an ontology is just an account of fundamental degrees of freedom in nature. The metaphysician asks, what are the independently varying components of nature, their internal degrees of freedom and the configurations they can assume? We can distinguish the rationalist metaphysician from the naturalistic metaphysician. The rationalist thinks that we have some special rational insight into the nature of reality, embodied in intuition or the clear light of reason, or some such thing to discern the fundamental structure of reality. The naturalistic metaphysician relies on observation and experiment. Given an ontology and a set of laws, we can generate a range of possible behavior (apply a combinatorial principle to the basic elements to obtain a set of possible configurations and then apply the laws to generate histories, or if you prefer apply a combinatorial principle to generate all possible sequences of states and then use the laws to narrow those down to physically possible histories), so the naturalistic metaphysician faces an inverse problem: how do we infer backwards *from* a range of observed behavior *to* underlying ontology? [ii]

Pre-scientific metaphysics was led by imagination. For the likes of Thales, Aristotle, Heraclitus, and Abelard, imagination was the primary tool in trying to develop insight into the nature of the world, but imagination has certain inherent limits. As complexity increases, visualizability decreases, and mathematics begins to emerge as an essential tool. First, the evidence becomes difficult to survey. Mathematics is needed to provide compact ways of representing the vast bodies of data we now bring to bear on theorizing. But mathematics also gives us new ways of recognizing and disclosing regularities in nature. From Galileo's discovery of the law of the pendulum, to Kepler's discovery of the laws of planetary motion, the most important discoveries turn out to be not a matter of finding new phenomena, but of revealing hidden patterns in the data. And searching is not a matter of putting on boots and going out into the wild, but sitting down in yours slippers with paper and looking through the data for hidden patterns. Brahe was a scientist in the first sense. He had his eyes trained on the stars. Kepler was a scientist in this second sense; he had his eyes trained on the numbers, and he famously saw patterns everywhere. There was the famous scheme with the geometric solids that was supposed to reveal God's geometrical plan for the universe. He did discover the laws of planetary motion, but he also proposed that the ellipticities of the planet orbits were determined by tunes they hummed as they circled. As crazy as it looks, it has nearly perfect correspondence to his fits to Brahe's measurements.

Once you have your eyes on the right quantities, regularities in the phenomena with clarity, but identifying those quantities takes special insights and can be a long and arduous process.[iii] Consider Galileo's discovery of the law of the pendulum: when he was 20 Galileo noticed a lamp swinging overhead while he was in a cathedral. Curious to find out how long it took the lamp to swing back and forth, he uses his pulse to time large and small swings and notices something that no one else had apparently realized: the period of each swing was the same.[iv] It is not that nobody had seen a swinging lamp or stone before: it is that they did not have their eyes on the right quantities and had not seen the regularity or the pattern. Kuhn's discussion gives some indication of the complex process that led to Galileo's attention to the right quantities, and I want to quote him, because I like the way he puts it. He was (of course) educated against the background of Aristotelian science

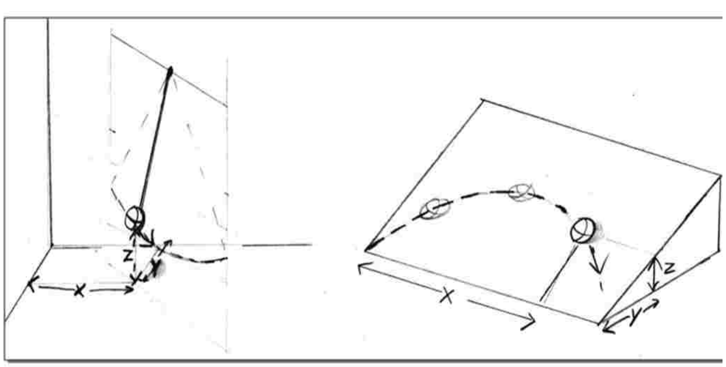
"Seeing constrained fall, the Aristotelian would measure… the weight of the stone, the vertical height to which it had been raised, and the time required for it to achieve rest. Together with the resistance of the medium, these were the conceptual categories deployed by Aristotelian science when dealing with a falling body…. Galileo saw the swinging stone quite differently. Archimedes' work on floating bodies made the medium non-essential; the impetus theory rendered the motion symmetrical and enduring; and Neoplatonism directed Galileo's attention to the motion's circular form. He therefore measured only weight, radius, angular displacement, and time per swing …given [these], pendulum-like regularities were very nearly accessible to inspection"[v]

This process of zeroing on the invariant relationships can be hidden from view, because once we've identified the invariant relationships they get embodied in the classifications we employ. To call a system a pendulum is already to assimilate it to a category of things that are characterized by these invariant relationships, but arriving at the *category* of a pendulum is not one that pre-Galilean philosophers employed.

Our brains are attuned to recognizing patterns. We can see that this rug has a repeated pattern, that this is the same pattern in different colors, that top half of this scene is the same pattern as the bottom half, reflected through an axis, even more complex ones, like the pattern on the wall is the projection on a plane of the pattern on the lamp's surface. And by 'see', I mean that literally, without inference or training, recognizing first order patterns (repetition, symmetry, regularity) in the visual, auditory, or tactual field is something that the biological brain does very well. But the sorts of patterns that Galileo discerned are not first-order patterns in the array of visually presented properties. They are typically higher-order invariant relationships between measurable quantities.[vi] They are not the kind of thing one sees in the way one sees the regularity in the pattern of a rug, or the symmetry in the photo of the reflected landscape.

Mathematical representation – whether in symbolic, geometric, or graphical form – provides ways of ways of representing motion (or more generally, dynamical change) that reveals the abstract, often higher order regularities hidden in the phenomena. If we represent the motion in the right way, the higher order patterns will reveal themselves on inspection. One way of describing what is going on here is that mathematical representation lets us transform a higher –order pattern recognition problem into a first order problem. We play around with different ways of representing the same phenomena, and when we find the right representation, the regularity will jump out at you. So we are using symbolic tools to transform a problem into one that the biological brain can handle.[vii]

And it is not just capturing patterns in motion over time; the ability to recognize that two apparently very different kinds of motion are the same is also a matter of recognizing an abstract, higher-order similarity. This kind of thing is also very much in evidence in Galileo. Here he is comparing the 3d motion of a pendulum to the motion of objects moving on an inclined plane. He didn't have stopwatches and modern instruments to measure how the motion of a stone thrown in the air changes speed with height, but he realized that by using gently sloped ramps or boards he could effectively slow the motion down to something he could measure.



Galileo was essentially comparing the 3d motion of pendulums (X,Y,Z) to the 3d motion of objects moving on an inclined plane. Doing this in full X,Y,Z detail for both objects would work perfectly but produce unnecessarily complicated equations.

Metaphysics in the old days: sparse, mostly qualitative data. Natural philosophers relied on first order regularities in the behavior of visibly similar things (fire rises, matter moves towards the center of the earth,), and a single powerful imagination produces a vision of the world that reproduces those regularities.

Metaphysics in the days of mathematical physics: large bodies of precise, quantitative data. Mathematical methods are used to reveal higher order patterns hidden in the data.

It is really with Newton that this method receives purest expression. It is obvious in retrospect, once you have your eyes on the right quantities where the regularities lie. But it was an aha moment of the highest order when Newton assimilated the motion of the falling apple to the movements of the planets around the earth. Whitehead said of Newton that what was so characteristic of him as a natural philosopher was that he had 'an eye attuned to the mathematicizable substructure of reality.' The same could be said for the theoretical physicist quite generally.

The first task for the physicist is to capture those invariant relationships. But once we have a precise quantitative description of the regularities in dynamical behavior in hand, we see yet another role for mathematics. This one is less obvious and less often talked about in philosophical reflections on science. It is much more recent: it is really only in the 20th century that it has emerged as an important technique in physical theorizing. It has less to do with the discovery of laws, and more to do directly with discerning the underlying ontology. This is where symmetries play an important role. What symmetries of the right kind do is help us identify and excise structure (either structure in the formalism or structure that we've attributed to the object) that is playing no dynamical role. It turns out that here as well, mathematical methods prove much more powerful than the imagination by itself. Indeed, the imagination is playing a catch up game. We use these methods to arrive at a new formal structure (which we now think captures the intrinsic structure of the physical reality we are trying to represent), and we are left to try to accommodate our imagination. It is really with these methods that physics comes into its own not just as a way of discovering laws, but as a very distinctive way of doing *metaphysics*.

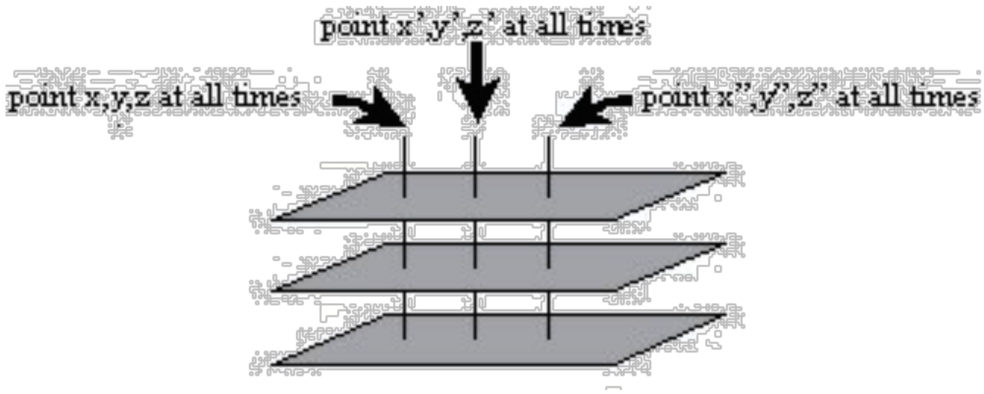
Space-time physics

The first domain of application of these methods was space-time physics. This is a very familiar story.

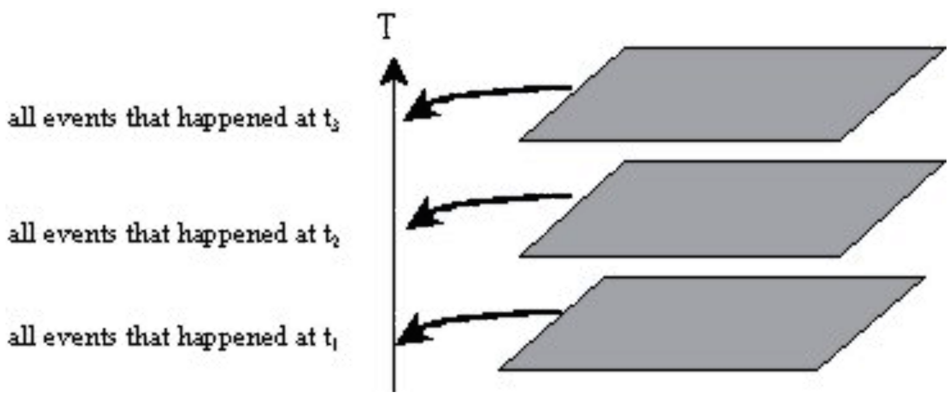
Here the focus is entirely on motion. We have formalism in hand that captures the laws of motion and the goal is to excise physically insignificant structure in the underlying space, and to do it cleanly, leaving in place structure that is needed to support the dynamics. We start with dynamical symmetries: symmetries that preserve a theory's laws.[viii] We focus on dynamical symmetries that map solutions onto empirically indistinguishable counterparts. And now we notice that if we regard those solutions as representing physically distinct situations, we are recognizing structure in the situations depicted by those solutions that is not just unobservable, but indiscernible in a much stronger sense, i.e., generates no measurable effect, plays no role in the law-governed, observable dynamics of bodies, measuring instruments, etc. [ix] If, on the other hand, we identify the physical situations represented by those solutions, effectively treating the structure that distinguishes them as a merely notational difference, we eliminate the dynamically irrelevant structure, and we are left with an account of the underlying structure that doesn't recognize hysical differences between systems that doesn't produce differences in the observable, law-governed behavior of the system.

It is now fairly customary to reconstruct the history of space-time physics as one in which the theoretical progression is driven by this kind of process. We start with a dynamics formulated in a space that is quite rich in structure, and then we use this kind of reasoning to whittle away structure that is playing no dynamical role. Newton formulated his dynamics against the background of absolute space and time to which he attributed a very rich geometrical structure: space has the structure of E3 and that time has a one-dimensional metrical structure and the individual parts of absolute space persist through time. We find that there are symmetries of the laws – Galilean boosts – that map solutions onto empirically indistinguishable solutions. There is the identification of structure that is not playing a dynamical role. We want to eliminate all and only structure not invariant under Galilean transformations. We find a way to do that by eliminating identity of points over time, retaining the structure to distinguish absolute acceleration but eliminating absolute velocity (retaining a distinction between inertial and non-inertial motion, but elimination a distinction between absolute rest and absolute velocity). There is the excision of structure that is not playing a dynamical role. The result is Galilean space-time.

The contrast between Newtonian space-time and Galilean space-time

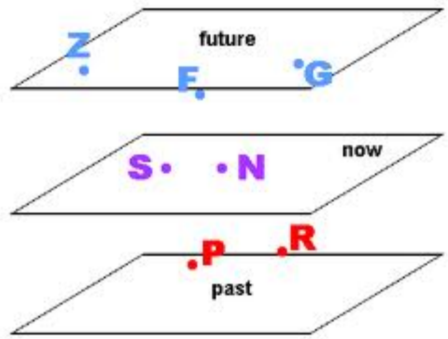


Galilean space-time absolute time, with spatial geometry on every spatial hyper surface is E3, but no identity of points over time.

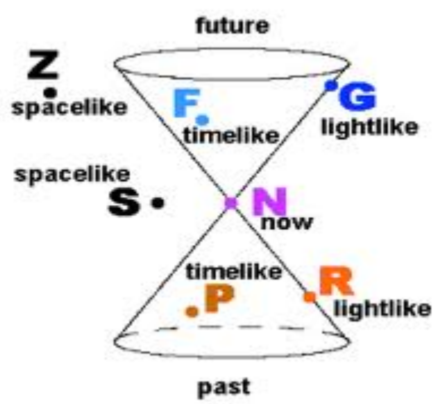


When phenomena proved resistant to Newtonian treatment, Newtonian physics was replaced by Special Relativity. Since we have new laws, we have to repeat the process. If we want to know the intrinsic structure of the space that supports STR, we take the laws of STR and do the same. STR laws are invariant under Poincare transformations and Poincare transformations map solutions onto empirically indiscernible counterparts. We eliminate geometric structure not invariant under those transformations, and we have Minkowski space-time.

Here's the contrast between Galilean space-time and Minkowski space-time (we're dealing now with light, so we can have color. In Galilean space-time, we don't have identity of points over time, but we have planes of absolute simultaneity.



Those have been eliminated in Minkowski space-time. In its place we have only the light-cone structure.



There is an abstract general method for arriving at a space-time that excises structure not invariant under symmetries (described in more detail in the setting of classical mechanics by Belot, with indications about how to generalize).[x] Start with the space of solutions to the laws; now consider the symmetries of the laws that map solutions onto empirically indiscernible solutions. Those symmetries define an equivalence class of solutions that we now regard as physically equivalent. Structure not invariant under the symmetries (i.e., structure that distinguishes solutions in the same equivalence class) is interpreted as physically *in*significant. Structure that is invariant (i.e., structure shared by solutions in the same equivalence class) is physically significant. Of course this is a rather indirect method of representation, and it is better when we can replace the equivalence classes with an intrinsic representation of that geometrical structure but it serves well as a way of capturing the logic of the reasoning. Belot calls the space obtained in this way, the reduced state space. I described this deliberately in the most general way. The only notions I used were a very general notion of symmetry (the mappings of the solution space onto itself and empirical indistinguishability. We can build more structure into the state-space - as we typically do when we formulate a physical theory - and require that the transformations that are candidates for interpretation as equivalence relations preserve that structure. But there are various candidates for this additional structure and whatever additional requirements we impose would have to be motivated in physical terms.

What is characteristic of this move to a reduced state space is a simplification of the ontology in the sense that there is a reduction in degrees of freedom that are recognized in the state of a classical system. Belot carries out the construction and shows how degrees of freedom get eliminated in detail for the case of classical mechanics. The state space Newton used for his mechanics of N particles is 6N-dimensional. In constructing the reduced space of the Newtonian state space by identifying states related by Galilean symmetries, we reduce the theory's degrees of freedom by 10 (so the reduced state-space has 6N-10 degrees of freedom). Quantities like the position and linear motion of the universe's center of mass have been eliminated. Acceleration, but not position or velocity, is absolute.

The inference from space-time symmetry to physical equivalence has sometimes been called "Leibniz equivalence," because it is a modern way of motivating Leibnizian conclusions about the identity of space-times that exhibit the right kind of indiscernibility. We can apply Leibniz Equivalence to general relativity (GR). Unlike the cases we just considered, GR doesn't occur against the fixed background of a single space-time. Instead, many distinct space-times (with distinct symmetries) are solutions of the same theory. We are interested again in symmetries of the *laws* that map solutions onto *empirically indistinguishable* solutions. So we look to the diffeomorphisms. Applying Leibniz equivalence, a physical space-time corresponds to a diffeomorphism equivalence class of mathematical space times. The structure transformed by diffeormorphisms is then regarded physically insignificant. And it turns out that this has one nice consequence. It permits the interpreter of GR to avoid the "hole argument", and the chanceless, unobservable indeterminism that the argument would entail if diffeomorphism-related solutions were regarded as represented physically distinct space-times.[xi] But it also leads to new questions. So, for example, if we excise structure that is not diffeomorphism invariant, then it would seem that the only quantities that remain are constant along gauge orbits, but it turns out (as Earman (2002) pointed out) that these quantities don't evolve, so it seems we are left with the result that there is no temporal change. And so there begins a lot of discussion of what that could mean. Maudlin has argued that this is an absurd result. Healey has argued very effectively that there's a mistake in Maudlin's reasoning. But in any case, that's the method. You apply it and see what you get. You use the symmetries as *clues* to the presence of excess structure (whether you think of that as excess structure attributed to the world, or redundancy in the formalism), but they are *only* clues, part of the ontologist's toolkit. One has to work through the results of applying the method to see whether it generates a physically sensible conclusion, and it all has to be done 'with discretion and good taste'.

Remarks

A few remarks are in order here.

First, this sort of reasoning has become enshrined as part of the standard way for presenting the theoretical progression in space-time physics. It goes along with the very big theoretical shift from thinking of space-time as the fixed background against which all of physics takes place and thinking of it as part of the subject matter of physics. It canonizes the accepted justification for rejecting Newtonian space-time in favor of Galilean, motivating Minkowski space-time in STR, and guiding the interpretation of GR. But it is not an actual, on-the-ground historical description of the messy set of facts that drive the opinion of the individual scientist or the community as a whole. It is not that there *wasn't* this kind of reasoning on the ground. It is that there was a lot of other stuff as well, and we didn't see it as clearly as we do now. (You'll notice, for example, that I didn't mention coordinate systems, which played a central, and mostly unfortunate role in the actual history). In representing the history of physics for 'textbook' purposes, we tend to bring into relief arguments that lead us to what we now think is the right conclusion, forgetting all of the false starts, and confusions, portraying the theoretical progression as a rational exchange of one theory for another.[xii] But that also shows that this kind of history has an important normative component that actual history does not. The importance of this reasoning has emerged through backward looking filters in which we know which avenues of development turned out to be fruitful, and its authority derives from its historical success in leading to those avenues.

Second, translating this reasoning into a technical setting is non-trivial. I used a very general notion of symmetry that is much wider than those that are used even in a classical setting when we talk of the symmetry group of a theory, and partly because of that, I needed to invoke a notion of empirical indistinguishability to narrow down the class of symmetries that are candidates for defining equivalence classes. Without that restriction, the directive to identify solutions related by symmetries of the laws would mean no theory would have more than one solution, so we would lose all of those physically real distinctions between solutions that represent *manifestly different* situations. The problem, however, is that empirical indistinguishability is a notion that we impose from the outside. It is not rigorously defined in a general way. There *are* more specific, rigorously definable notions of symmetry on offer in classical theories (e.g., the generalized, or Hamiltonian symmetries, or variational symmetries Belot discusses in his "Symmetry and Equivalence"). The reason for not employing those in this capacity (i.e., as identifying those symmetries that alert us to the presence of excess structure) is partly that we don't want to tie the relevant notion of symmetry too tightly to any particular formalism, but mostly because none of those will do the work that we want it to here. Gordon's paper makes that point about as forcefully as it can be made.[xiii]

Third, although there are few who would deny that (i) if you have a formalism that is empirically adequate, and you can identify structure that is playing no physically meaningful role generating testable predictions either with respect to the phenomena that are the primary domain of that theory *or* in coupling with other theories to generate that fall outside that domain, it should be excised and (ii) that symmetries of the laws that map solutions onto empirically indistinguishable counterparts can be a good guide at identifying structure that plays no physically meaningful role, it can be very difficult to implement this reasoning in the complex setting of a real theory. The clean separation of physically significant from physically insignificant structure is often very hard to do. It usually requires the development of new formalisms. And even when we can separate the structures, the interpretation of the results is difficult. Think of the re-education of the imagination that was needed in the move from Newtonian to Galilean, or Galilean to Minkowski space-time. But it is something that is even more amply illustrated in the case of modern gauge theories, so we turn now to those.

Gauge theories

When we turn to non-space-time theories like electrostatics, electromagnetism, or quantum mechanics we look to internal symmetries – transformations that alter the values of physical quantities that are supposed to represent the internal degrees of freedom for a system (or some region of space-time) rather than changing anything spatially or temporally external to it, and see where this kind of reasoning leads us. We distinguish global internal symmetries (like the electrostatic potential, or phase transformations in qm), which act identically on the values of all points throughout space (or space-time), from local symmetries, which diverge from point to point. The best-known example of a theory with local gauge symmetry is, of course, classical electrodynamics. The vector potential is a four-dimensional vector; the time component is the familiar electric potential V (x) while the space components collectively form the magnetic potential. Since physical predictions are fixed by the field tensor, which is left unchanged if we add the gradient of a scalar, the most natural thing following the kind of reasoning we applied above, would be to regard structure not invariant under this transformation as physically unreal. The problem with that of course is that it is undermined by Aharanov-Bohm effect: changes in the vector potential lead to a measurable difference in the behavior of quantum particles passing by solenoids despite the zero field in their vicinity. We *could* say that there is action at a distance from the field inside the solenoid, but if we don't want to do that, we have to work a little harder to separate physically significant from physically insignificant structure. So begins the search for a new ontology for classical electromagnetism. One is reminded here of the search for a space-time for Newtonian mechanics that eliminates absolute velocity but leaves in place absolute acceleration. Healey's *Gauging What's Real* defends a gauge invariant ontology for classical gauge fields (electromagnetism first, and then general Yang-Mills theories), that can explain Aharonov-Bohm without action at a distance.[xiv]

Where this reasoning leads

One of the most interesting aspects of this kind of reasoning from a philosophical point of view is that after getting rid of the surplus, we are left with quantities that aren't localized in point-sized regions of space. On Healey's view the fundamental quantities are the "holonomies," defined on closed loops in space-time (Healey, 2007). On another view – discussed by Maudlin[xv]- they are gauge-invariant values of the connection, which are (in Maudlin's parlance) "hyper-local," so that there is no determinate matter of fact about whether distant space-time points agree as to their value. Either way the clues we get from symmetry in this setting push us in new ontological directions. They push against common sense, and they push against programs in analytic metaphysics. The defining principle of David Lewis' metaphysical program – which occupies a very large share of research in contemporary analytic metaphysics – is that the fundamental building blocks of the universe are point-sized events in space-time.[xvi]

When we have this kind of conflict – i.e., when the methods described here which use purely mathematical clues to isolate the basic degrees of freedom in nature push against the programs in analytic metaphysics - it forces us to choose allegiances. Do we let our ideas about ontology be led entirely by physics, or is there some other form of reasoning embodied in the methods of analytic philosophy that has authority? A first step in the direction of providing an answer to that question is getting a clear idea of what is driving ontology in physics. I think that if there is a method that can be discerned in the bewildering detail and specificity of discussions in physics about fundamental ontology, this is a big part of it. We find laws that capture the invariant relationships, and then we use symmetries to identify and excise structure that is not playing a dynamical role.

Difficulties of application

Some general remarks about all of this worth emphasizing:

First, it is not a simple and straightforward judgment that some bit of structure is not playing an empirically meaningful role. There is no way around the case-by-case careful judgment about what is empirically significant and what is not. All of the phenomena have to be considered, including phenomena that are generated when a theory is coupled to another theory, outside its primary domain of application. The Aharanov-Bohm effect is the most obvious example. The effect emerges only when we couple an electromagnetic field to a complex vector field, and so it doesn't emerge until we combine classical electromagnetism with the behavior of quantum particles.

Second, even when we've become convinced that there is some structure that is not physically meaningful, excising that structure leaving in place all the physically meaningful structure is far from easy. The sanitized history of space-time theories made it look as though we identify empirically superfluous structure and go in to do a clean little surgery to remove the offending formation, leaving everything else in place. In retrospect, it looks clear as day. But while we are working in a formalism in which the offending structure is bound up with the living tissue, the discussion is complex and extremely hard to sort out. Hence the difficulty of ongoing discussion of gauge theories.

Third, applying the method is *never* free of interpretive assumptions. There is no *purely* formal way of recognizing redundant structure. As I mentioned above, if we simply identified all solutions related by symmetries of the laws, we would end up with a single solution to every set of laws. Before we have a structure to which we can apply these sorts of symmetry probing reasons without that kind of triviality, we need a set of differences that can function as sufficient conditions for distinctness. That has to come from outside the formalism we are working with. I have suggested empirical distinguishability as a generic criterion, because it captures the reasoning that we are trying to apply in the murkier, technically involved setting of modern theory. The reasoning is that ultimately we are only warranted in postulating unobservable structure that plays a role in supporting observable dynamical behavior, and so we should be getting rid of structure that doesn't play such a role.

But I want to point out that can also be powerful when not applied in that general setting, but in a setting in which we are willing to take certain structures for granted as representing real physical differences for purposes at hand, and we are interested in seeing whether some other set of structures play a physically meaningful role (the role of a difference-maker) in producing differences at that level.

The only way in which this reasoning can be applied in a *theory neutral* way is if there is some basic set of differences that are regarded *pre-theoretically* as sufficient conditions for distinctness.[xvii] Most people wrapped up in the thickets of everyday disputes about gauge fields are not working at such a high level and bristle at the imposition of an epistemic distinction like this in what ought to be straightforward physics reasoning, a distinction that comes from the outside that seems way too soft and human centered to be playing a role in discussion of fundamental ontology.[xviii] In practice it is enough if we have a set of differences that are agreed by participants in the discussion, for the purposes at hand, as sufficient conditions for distinctness.[xix]

Justification

There are difficult, open questions about why the method works. There is no a priori *logical* principle that guarantees the success of these methods, in one sense. We can certainly imagine worlds in which real structure is not invariant under symmetries of the laws that map solutions onto empirically indistinguishable counterpart: Newton's world with absolute time and space or the non-localized gauge potential properties view considered by Healey are perfectly intelligible physical possibilities.[xx]

There are, nevertheless, two broad routes to rationalizing the method (not necessarily just two, but these are the two that suggest themselves as promising)

(i) The first is the one that I adverted to above: there is a weaker form of non-logical rationality, which we might call 'evidential rationality', that shifts the burden of proof so that we aren't warranted in distinguishing elements related by certain kinds of symmetry that we know only indirectly, as elements in a structure interpolated behind experience. So that someone who recognizes structure that maps solutions of the laws onto empirically indistinguishable solutions is recognizing structure that they can have no evidence for.[xxi]

(ii) Alternatively, we can focus on issues about reference and argue that if we have a theory in which we purport to refer to distinct elements that are not distinguishable by any observation, or their relations to anything else, (e.g., two points in absolute space at different times, or internal properties that don't make discernibly different impact on measuring devices), we have not really managed to do so. We have not actually managed to get our referential hooks into two separate elements of reality; what we have is actually two names (coined by rigidified definite descriptions perhaps obtained from the Ramsey sentence of our theory) for the same thing (a coarse-grained object constituted by the pair or more generally, an equivalence class of weakly indiscernible objects). The idea this would have the advantage of reducing every case of recognizing too much structure in the world to a case of unacknowledged redundancy in the formalism. But it is a tough row to hoe, because it will have to rest on questions about how it is that terms in a theoretical formalism come to represent (or refer to) some particular bit of the natural world, and that is an issue that is very much contested, and very hard to see one's way through. (Healey has some very interesting related arguments, not quite as strong, to the effect that it is never justifiable to accept a theory that recognizes distinct elements with this kind of indistinguishability, but thinks that we can nevertheless coin terms that refer to individually by use of demonstratives.

The question of why the method works is bound up with the theoretical preference for simpler theories. It might be that the preference for simpler theories is a generalized version of the method here that eliminates physically insignificant (or redundant) structure in the formalism. The idea here would be that whenever you are choosing between competing, empirically adequate formalisms, one of which is simpler than the other (where simplicity is measured in the most obvious way, by counting up the number of basic degrees of freedom postulated), it is a good methodological bet to suppose that the more complex has some redundancy in the formalism.[xxii] But these are very much unsettled questions.

For practical purposes, however, I don't think the justification matters. The authority of the method doesn't rest on a priori rationalization, but on its historical success at leading to fruitful lines of development. It is one of the most distinctive of modern and the question of why it works is part of the more general puzzle of why physics works.

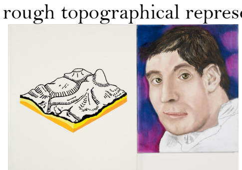
Conclusion

I have given the very broad-brush historical background and articulated the kind of reasoning behind the use of symmetries of the right kind to zero in on the appropriate ontology, mostly by way of providing some context. The discussion now passes to the highly technical and highly specific issues about how to separate the physically significant from the physically insignificant structure in classical and quantum gauge theories carried on by those in the trenches, so to speak. They are unavoidably complex, unavoidably empirical, and they demand looking at theories in full detail, one by one. [xxiii]

[i] Prepared for a workshop on Symmetries in Physics at UC, Riverside.
[ii] Since there will be any number of ontologies that are compatible with the observed behavior (regard it as within the range of possibility, allowing for errors of observation) the question for the naturalistic metaphysician is which of these have the best fit with the evidence, and so we unavoidably in practice appeal to criteria of fit that go beyond mere compatibility. It is in the criteria of fit that go beyond mere compatibility that we find a preference in practice for ontologies that reconstruct the data from a few basic ingredients exhibiting simple, and regular behavior.
[iii] The invariant relations aren't the surface regularities, pendula come in different colors, sizes, and material. They are higher-order regularities: relations between quantities that themselves take different values.
[iv]The period depends on the length of the pendulum and also on the amplitude of the oscillation, but, if the amplitude is small, the period is almost independent of the amplitude.
[v] Kuhn, *The Structure of Scientific Revolutions*, p. 124.

[vi] Mathematics is essential to process because the imagination has certain kinds of inherent limits. We are very good at pattern recognition when it concerns relations between first order properties. We can recognize repetition, symmetry, and regularity in a visual array immediately and without inference. If I ask you whether these objects exhibit the same pattern; you would be immediately able to tell me

If I asked you whether the two images below exhibit the same pattern, you probably wouldn't be able to say, even though the first is a rough topographical representation of the face depicted on the right.



The similarity in the second case is a higher-order similiarity that requires looking not at the first-order properties, but the relations among those relations (judging sameness or difference between the relations defined over the first order properties).

[vii]Our biological brains are not that different from other primates, the basic cognitive mechanisms are innate capacities for pattern recognition. What makes us special is the use of symbolic representation to supplement the built-in capacities for pattern recognition. It transforms higher order pattern recognition problems into one that our brains can handle. This speculation is supported by famous experiments on chimps that show that language trained chimps can perform higher-order recognition tasks that non-language trained chimps cannot. Chimps can be easily be trained to do first order pattern recognition tasks – e.g., to tell whether a pair of objects (a pair of spoons, or a pair of cups) in an array are similar, but they can't be trained to do higher-level pattern recognition tasks until they are provide with the use of external tags of some kind. So for example, if they are sent into a room with a set of things already sorted into pairs, they can't be trained to separate the pairs of similar things (e.g., two spoons, two shoes, two cups) from pairs of different things (e.g., a spoon cup pair, a fork shoe pair) unless they are given the use of external tags of some kind. So, for example, they are given a marker, or colored tags, or (in the initial experiments. What they will do is first associate a tag or symbol with the pairs depending on whether they are similar – red tags with similar pairs, blue ones with different pairs - and then use the tags to sort into pairs of similar or different objects. In a background review of these experiments "Thompson and Oden (1996)… conclude that .. the use of simple, arbitrary external tags for independently identifiable relational properties [opens] up the more abstract space of knowledge about relations between relations." Andy Clark commenting on these results speculates that words and symbols perform the same function for us, rendering "[higher order] features of our world concrete and salient, and allow[ing] us to target our thoughts (and learning algorithms) on a new domain of basic objects. This new domain compresses what were previously complex and unruly sensory patterns into simple objects. These simple objects can then be attended to in ways that quickly reveal further (otherwise hidden) patterns, as in the case of relations between-relations." ("Magic Words")
[viii]This contrasts with the symmetries of individual solutions. We can make some further distinctions, between continuous and discrete symmetries. Continuous symmetries correspond to continuous changes in the geometry of the system, described by continuous or smooth functions. Discrete symmetries correspond to non-continuous changes in a system (e.g., reflections or rotations through a fixed degree, permutations or *interchanges*.)

[ix] This is true only if we demand that the equations be generally covariant (hold in all frames of reference rather than just some). That requirement forces us to make explicit the features of a system that are crucial to its dynamical behavior.
[x] Belot, Gordon (2001), \The Principle of Sufficient Reason," Journal of Philosophy 98:55{74.

[xi] If one does assume, contrary to the present approach, that there are distinct physical possibilities related by diffeomorphisms, a kind of unobservable indeterminism pops up in GR. This is because a diffeomorphism can sometimes leave a particular surface of simultaneity

unchanged while shufflng around what happens at which point in the future of that surface. Since the shuffling changes nothing invariant under GR's symmetries, the indeterminism disappears on the present approach, where only such invariants are regarded as physically real.

[xii] There's nothing wrong with that kind of reconstruction. It serves a pedagogical purpose, that doesn't depend on its being an actual, on-the-ground historical description of the messy set of facts that drive the opinion of the individual scientist or the community as a whole.
[xiii] See Belot, "Symmetry and Equivalence" for a detailed consideration.

[xiv] He considers three options:
• "The no gauge potential properties view": the ontology is given by the field
strength
• "The localized gauge potential properties view": the gauge potential makes up the ontology, or some other (yet to be determined) quantities that are associated to arbitrarily small neighbourhoods in space-time and that can be derived locally from the gauge potential.
• "The non-localized gauge potential properties view": the quantities that make up the ontology can be derived from the gauge potential, but are not necessarily associated to arbitrarily small neighbourhoods in space-time. Healey presents two arguments against the first view. First, this view implies action at a distance. This follows from the Aharonov-Bohm effect, where a magnetic field confined to a localized spatial region still has an effect on the interference pattern of a quantum particle that passes outside that region. Second, in the case of non-Abelian Yang-Mills theories, there exist gauge inequivalent gauge potentials that give rise to the same field strength in a simply connected region. So the field strength at that region would not completely capture the gauge invariant content associated to that region.
The second view is Rejected (in good part) on the ground of underdetermination: gauge equivalent potentials would yield different properties that can not be distinguished by observable facts.
After having rejected the first two possible views, Healey argues in favor of the third one. In particular, Healey defends a holonomy interpretation, whereby the ontology is given by the holonomies associated to loops (images of oriented closed curves) in space-time (these can be constructed from line integrals of the gauge potential along closed curves).

[xv] *The Metaphysics within Physics*, 2007, 78-103.

[xvi] More specifically, when we excise structure that is not invariant under local and global gauge transformations, we are left with fundamental quantities that aren't point-sized. Something similar happens in qm with respect to phase transformations. In this case, when we 'mod out' the structure not invariant under phase transformations, we are left with a failure of Humean Supervenience, because the state of a complex system of spatially separated parts does not supervene on the states of the regions in which the parts are located. All of these point to the failure of the idea that the world is composed of spatially localized degrees of freedom, which jointly determine the state of the world as a whole.
[xvii] Bas and I proposed that if observable differences, perhaps in position can play this role… we might have added observable differences of other kinds – e.g., in color or temperature - but this adds nothing so long as these kinds of differences can themselves be measured with spatial pointer observables. The hard cases would be unmeasurable auras that only some people are able to see and which can't be measured using any known physical process. There is no non-question-begging way of rejecting such things, but in practice, these kinds of cases don't seem to arise in a practically pressing way.sufficient conditions get turned into necessary and sufficient conditions for distinctness by saying that we have warrant for recognizing structure that is either itself observable or makes a difference to observable dynamical behavior .

[xviii] that resistance is ameliorated to some extent when combined with the epistemic justification for the reasoning.
[xix] Choose your sufficient condition for distinction and apply the reasoning to excise structure that doesn't produce differences by that criterion. You can make perverse choices: auditory or gustatory.

[xx] Given any rule that says 'whenever S is a symmetry of the laws and S maps a onto empirically indistinguishable counterpart b, then a=b, we can immediately construct a counterexamples in the form of a universe in which S permutes physically real elements but in which S is a contingent symmetry of our world. Indeed combinatorial principles will *entail* that for any account of what the basic elements are, there will be some metaphysically possible distinct situations a and b that are related by S. If S is reflection through a plane P, and the laws are symmetric with respect to S, there will be bilaterally symmetric solutions to the laws. So the inference from S is a symmetry of the laws which maps our world onto an empirically indistinguishable counterpart would fail.

[xxi]That is a strong intuition, but making it more than an intuition would require showing that that given a body of data that respect the symmetry is much *more likely* to have been produced by a situation in which a=b than a world in which a≠b. i.e., the probability of evidence obeying the symmetry being generated by an a=b world is higher than a a≠b world. The problem is that these kinds of probabilistic claims always depend on a measure over possible worlds, and can be reversed with different choices of meature.

[xxii]There is a temptation to justify the method by appeal to Ockham's razor. The problem with that is that there are as many different versions of Ockham's razor as there are notions of simplicity. No one of these has gained wide acceptance as a principle of theory choice, and the justification of the principle in any of its versions is not well understood. (See Elliot Sober, *Ockham's Razors: A User's Manual* (2015) for a terrific discussion) My view is that the kind of reasoning described here, which seeks to excise excess theoretical structure gives explicit content to the only version of Ockham's razor that has been uncontentiously vindicated by the history of science. Instead of looking to Ockham's razor for justification of the method, we should look at the logic behind the reasoning directly and see it as capturing the valid heart of Ockham's razor.

[xxiii] A final point that seems worth making, motivated by Rovelli's "Why Gauge?". The standard way of putting what the method described here is doing is misleading in some cases because it suggests we are identifying structures that have no physical interpretation at all, part of the mathematical machinery that is just empty. But that is not quite right. In the paradigm cases, the structures in question are not physically insignificant, but rather implicitly relativized to a reference frame that can be embodied in another physical object (like an observer or a measuring device), the coupling between the physical situation or state of an observer or measuring device is invariant, so that if the state of the observer or measuring device is held fixed except for, changes in the non-invariant quantity produce changes in the physical situation produce changes in the observer (or measuring device). Frame-dependent quantities like 'being simultaneous with', or 'being at rest' are real and measurable (indeed, in some cases, directly observable), but they are not *absolute*. So, the problem was not that we were regarding as real structures that are 'merely mathematical', but that we were regarding as *absolute* structures that are implicitly *relational* or perspectival. The relations are invariant, and when we couple to another object, we can 'measure' the non-invariant quantities directly (though we can't translate them into a description of the invariant quantities when we know our own situation, so to speak). So that when we see differences in the values of those quantities, we are seeing between differences in the perspective from which the object is viewed rather than differences intrinsic to the object. A single observer can distinguish what is nearby from what is far away, and observers travelling non-inertially relative to one another disagree on simultaneity in the way that observers in different locations disagree on what is nearby. And observers who see changes in the values of gauge-dependent quantities are seeing changes in their relations to the gauge-invariant quantities that characterize the intrinsic properties of a system. This is not to say that there aren't *some* symmetries that permute physically empty features of notation. Transforming the color of the pen in which one writes down the equations, for example, or the particular temperature one chooses to count as zero on a temperature scale. Gauge cases have been traditionally conceived in that way.