40

SYMMETRY AND SUPERFLUOUS STRUCTURE: LESSONS FROM HISTORY AND TEMPERED ENTHUSIASM

Jenann Ismael

You know you've achieved perfection in design, not when you have nothing more to add, but when you have nothing more to take away. (De Saint-Exupery)

In this chapter, I will discuss a particularly powerful method for identifying and excising superfluous structure in a formalism. Superfluous structure is theoretical structure which plays no role in supporting dynamical behavior.

40.1 Mathematical Methods in Physics

In its most abstract form, an ontology is 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?" The rationalist metaphysician supposes that we have some form of rational insight into the nature of reality. The naturalistic metaphysician relies on observation and experiment. Her task is to infer ontology from data. Given an ontology and a set of laws, one can generate a range of possible behavior. (There are two ways of doing this. One can apply a combinatorial principle to the basic elements to obtain a set of possible configurations and then apply the laws to generate histories, or one can apply a combinatorial principle to generate all possible sequences of states and then use the laws to narrow those down to physically possible histories.) The naturalistic metaphysician faces an inverse problem: how does she infer backward *from* a range of observed behavior *to* underlying ontology?

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. In this setting, mathematics begins to emerge as an essential tool. For one thing, 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. For another thing, mathematics gives us new ways of recognizing and disclosing regularities in the data. From Galileo's discovery of the law of the pendulum (1602) to Kepler's discovery of the laws of planetary motion (1609, 1619), the most important discoveries turn out to be less a matter of finding new phenomena than of revealing hidden patterns in the data. And searching is not a matter of putting on boots and going out into the wild, but rather of sitting in the parlor in your slippers and looking through the data on paper 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~~ saw patterns everywhere. There was the famous scheme with the geometric solids that was supposed to reveal God's geometrical plan for the universe, and his harmonic theory, which linked geometry, cosmology~~,~~ and astrology through a series of musical intervals.[1] 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 to modern eyes, it has a nearly perfect correspondence to Brahe's measurements.~~

Once one has one's eyes on the right quantities, regularities in the phenomena emerge with clarity, but identifying those quantities takes special insight~~,~~ and can be a long and arduous process. Consider Galileo's discovery of the law of the pendulum. When he was 20 years old, 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 used his pulse to time large and small swings and noticed something that no one else had apparently realized: viz. that the period of each swing was the same. People had seen swinging lamps before, but they did not have their eyes on the right quantities and (therefore) had not seen the regularity ~~or the pattern~~.

What led Galileo to see the regularity? Not just his personal genius. Galileo was educated against the background of Aristotelian science, but there was a complex process that led his attention to the right quantities. Kuhn (1996) remarks on this process:

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 Neo-Platonism 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[2]

The (sometimes extended) process of zeroing in the invariant relationships can be hidden from view~~;~~ because once the invariant relationships have been identified they get embodied in the classifications we employ. (Invariant relationships are, in general, relative to a set of transformations. The general concept of invariance is a way of capturing what remains fixed when some specified set of other things are allowed to vary.) To call a system a pendulum, for example, 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 a part of the process of discovery. 'Pendulum' not a category 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,~~ and that the top half of ~~this scene~~ is the same as the bottom half, reflected through an axis.

~~These are quite simple patterns, but we can also see more complex ones, like that the pattern on the wall is the projection on a plane of the pattern on the lamp's surface.~~

In saying that we can *see* these patterns, I mean that the apprehension of the pattern is a perceptual phenomenon: it is a part of the processing of sensory information that requires no conscious effort or inference on our part. Recognizing first-order patterns (repetition, symmetry, regularity) in the visual~~, auditory, or tactual~~ field is something that the biological brain was made to do, and it does it very well. The regularity that Galileo discerned, ~~however,~~ and the sorts of regularities that get embodied in scientific laws quite generally, are not first-order patterns in an array of visually presented properties. They are typically higher-order invariant relationships between measurable quantities. They are not

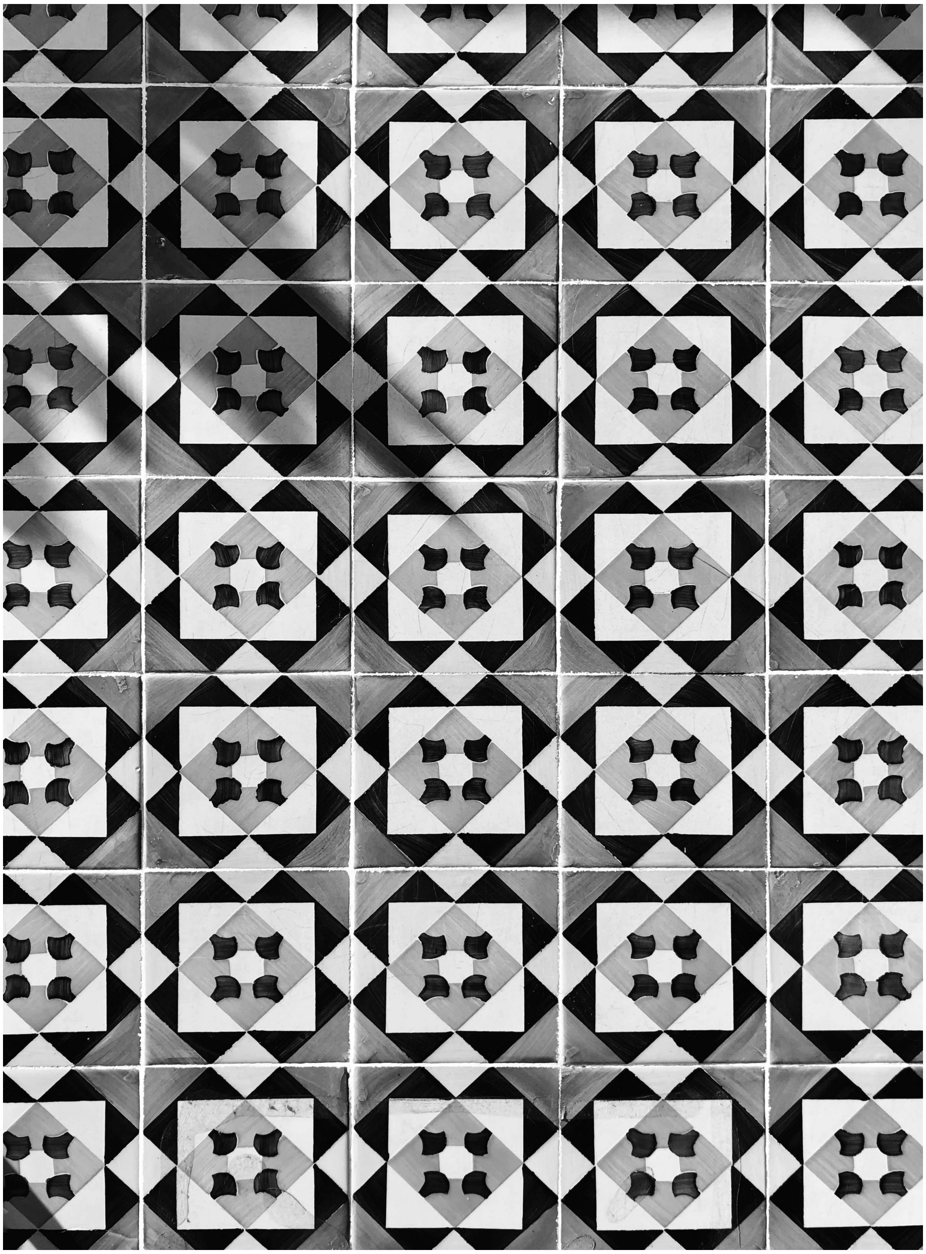

Figure 40.1 Wall tiling with repeated pattern.

AU: Please provide the figures citation.

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 – discloses regularities hidden in the phenomena. It is unavoidable in the search for dynamical regularities since in that case the items whose relations are being examined are not presented simultaneously in perception. We need to write down a time-line, draw a trajectory, or otherwise represent a system's

Figure 40.2 A reflected scene in water.

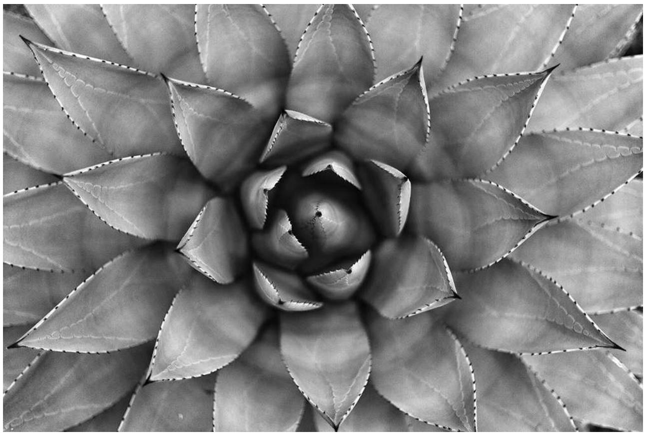

Figure 40.3 Circular patterns of cactus leaves.

history in order to get a look at the relationships between its parts. We use diagrams, graphs, and equations ~~in presentations to~~ a set of relationships that we want to call attention to. In the mathematical investigation that precedes discovery, different notations are used in an exploratory way. The physicist introduces notations that highlight some relationships and suppress others. When she hits on a good representation, patterns emerge and regularities disclose themselves ~~on~~ inspection. One way of describing what she is doing is that she is using mathematical representation to transform a higher-order pattern recognition problem into a first-order problem. She plays around with different ways of representing the phenomena until she finds one in which previously hidden regularities are *rendered* visible. So she is using symbolic tools to transform a problem into one that the biological brain can handle. The idea that our ability to use symbolic tools in this way may be the cognitive innovation that underwrites a cascade of characteristically human abilities was originally developed by Karmiloff-Smith (Karmiloff-Smith, A. (1979). *A Functional Approach to Child Language*. Cambridge University Press, and (1986). "From Meta-Process to Conscious Access. *Cognition*, 23, 95–147").

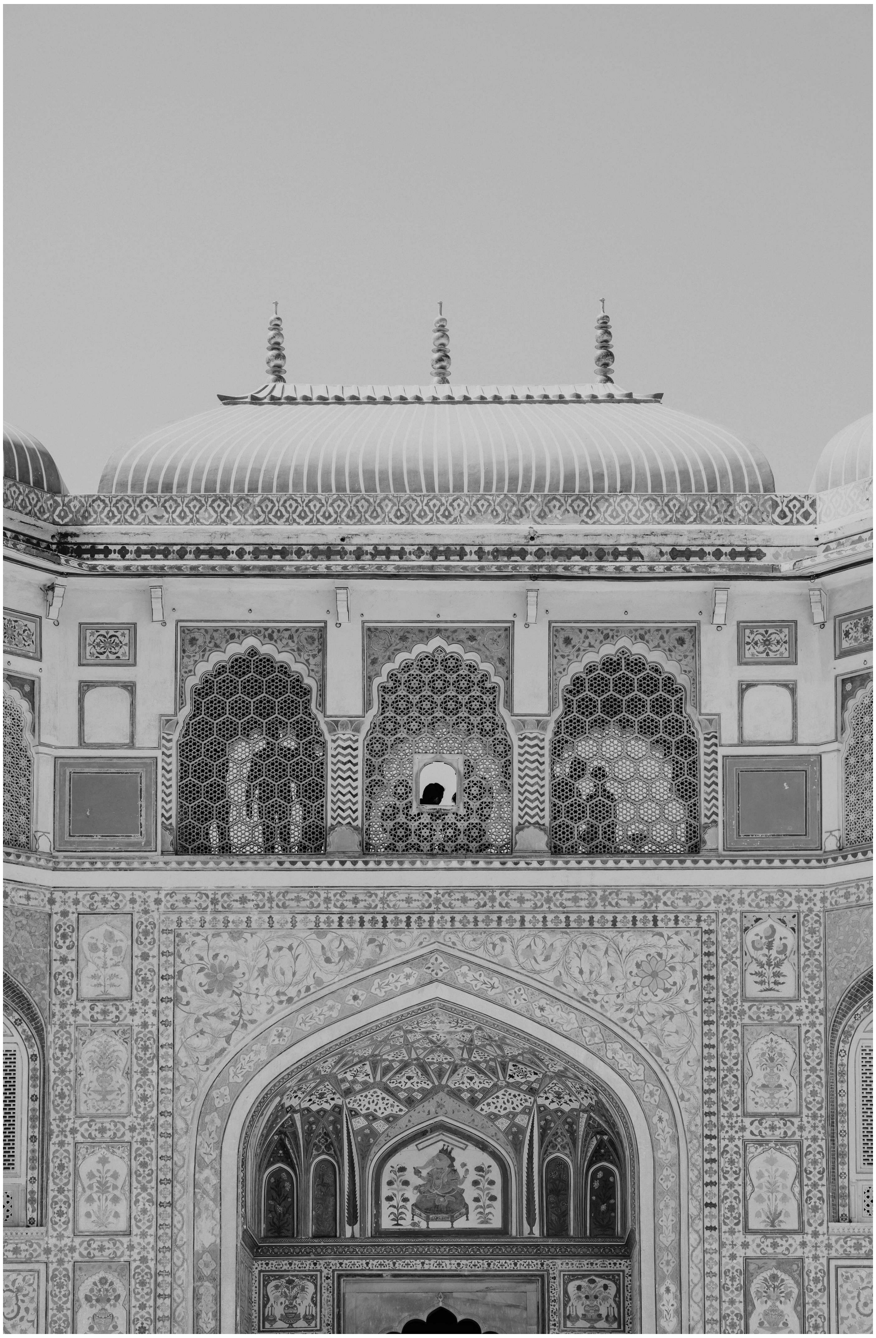

Figure 40.4 Nested patterns of tiles and architectural details in a mosque.

The idea that symbolic tools allow us to represent knowledge in different formats allowing new kinds of cognitive operation and access has been further developed by a number of people.[3]

The process is not just useful for capturing patterns in motion over time. The recognition 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 very much in evidence in Galileo's notebooks. 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 that could be measured.

Metaphysics in the old days was a largely solitary project that made use of sparse, mostly qualitative data. Natural philosophers relied on first-order regularities in the behavior of visibly similar things (fire rises, massive objects fall), and a single powerful imagination produces a vision of the world that reproduces those regularities. Metaphysics in the days of mathematical physics, by contrast, uses large bodies of precise, quantitative data. Mathematical methods are employed to identify 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 took Newton's physical insight to see beyond the quantities that the impetus theory directed attention to, and fix on force, mass and acceleration as exhibiting the kinds of invariant relationships that could be captured in laws. 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 theoretical physicists quite generally.

40.2 Symmetries as Clues to Superfluous Structure

The first task for the physicist is to capture the regularities. But once we have a precise quantitative description of the regularities in dynamical behavior in hand, another role emerges 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 come into its own 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 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 catch up. 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*.

The first domain of application of these methods was spacetime physics. Here the focus is entirely on motion. We have formalism in hand that captures the laws of motion and the goal is (i) to excise physically insignificant structure in the underlying space, and (ii) to do it cleanly, leaving in place structure that is needed to support the dynamics. Let us begin by looking at dynamical symmetries i.e., symmetries that preserve a theory's laws. (This contrasts with the symmetries of individual solutions. There are further distinctions to be made, 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.) Let us also focus on dynamical symmetries that map solutions onto empirically indistinguishable counterparts. Now, we can 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. It generates no measurable effect, plays no role in the law-governed, observable dynamics of bodies, measuring instruments, etc. (This is true only if we demand that the equations be generally covariant. That requirement forces us to make explicit the features of a system that are crucial to its dynamical behavior.)[4] If, on the other hand, we identify the

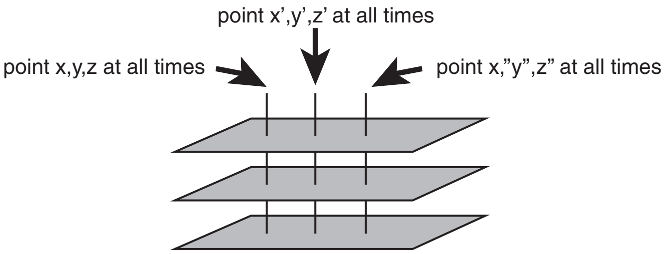

Figure 40.5 Newtonian spacetime.

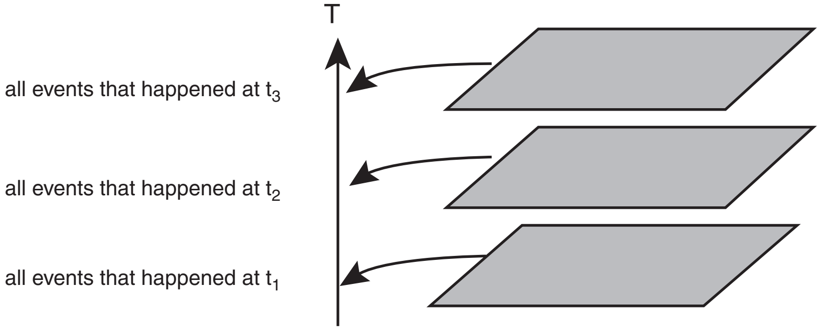

Figure 40.6 Galilean spacetime.

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 physical differences between systems that don't produce differences in the observable, law-governed behavior of the system.

It is now fairly customary to reconstruct the history of spacetime physics as one in which the theoretical progression is driven by a progressive elimination of excess structure, using symmetries as a guide.[5] We start with a dynamics formulated in a space that is quite rich in structure, and then we use the dynamical symmetries to help us identify, and whittle away, structure that is playing no dynamical role.[6] Newton formulated his dynamics against the background of an absolute space and time with a very rich geometrical structure. Space was attributed the structure of E3. Time was attributed the structure of the real line, and the individual parts of space were thought to persist through time. It was observed that there are symmetries of the laws – Galilean boosts – that map solutions onto empirically indistinguishable solutions that were nevertheless regarded as physically distinct. That meant that Newtonian Mechanics with absolute space and time recognizes physical structure that plays no 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 we have the excision of structure that is not playing a dynamical role. The result is Galilean spacetime.

Here is Newtonian spacetime, with absolute space and time.

Here is Galilean spacetime. As in Newtonian spacetime, time is absolute and the spatial geometry on every spatial hypersurface is E3. The difference is that there is no identity of points over time.

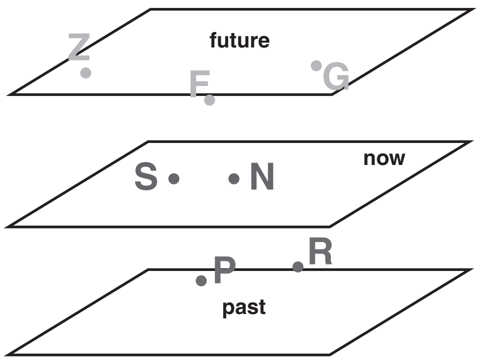

Figure 40.7 Planes of absolute simultaneity in Galilean spacetime.

When phenomena proved resistant to Newtonian treatment, Newtonian physics was replaced by Special Relativity (STR). 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 look at their symmetries. STR laws are invariant under Poincare transformations and Poincare transformations map solutions onto empirically indiscernible counterparts, so we eliminate geometric structure not invariant under those transformations. The result is Minkowski spacetime.

Here's the contrast between Galilean spacetime and Minkowski spacetime. In Galilean spacetime, we don't have identity of points over time, but we have planes of absolute simultaneity.

Those have been eliminated in Minkowski spacetime. In its place, we have only the light cone structure.

The schema for arriving at a spacetime that excises structure not invariant under symmetries goes like this: Start with the space of solutions to the laws. 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 regarded as 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 state-space obtained in this way, the reduced state-space.[7] I described this deliberately in the most general way. The only notions I used were (i) a very general notion of symmetry (the mappings of the solution space onto itself), and (ii) empirical indistinguishability. We can build more structure into the state-and require that the transformations that are candidates for interpretation as equivalence relations preserve that structure. So, for example, we might require that transformations that are candidates for interpretation as equivalence relations preserve differences that produce measurable effects when coupled with neighboring theories, or we might have other reasons for thinking that differences between models related by a given transformation are physically real. Whatever additional requirements we impose, however, would have to be motivated in physical terms.

What is characteristic of the 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 Mechanics 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 the linear motion of the universe's center of mass have been eliminated. Acceleration, but not position or velocity, remains absolute.

The inference from "S and S* are related by a spacetime symmetry" to "S and S* are physically equivalent" has been called "Leibniz equivalence," because it is a modern way of motivating Leibnizian conclusions about the identity of spacetimes that exhibit the right kind of indiscernibility.[8] We can apply Leibniz Equivalence to General Relativity (GR). Unlike the cases we just considered, the dynamics of GR isn't defined against the fixed background of a single spacetime. Instead, many distinct spacetimes (with their own symmetries) are solutions of the same theory. Since we are interested in symmetries of the *laws* that map solutions onto *empirically indistinguishable* solutions, in GR we look to the diffeomorphisms. Applying Leibniz equivalence, a physical spacetime corresponds to a diffeomorphism equivalence class of mathematical spacetimes. The structure transformed by diffeomorphisms is then regarded as physically insignificant.

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 spacetimes. (If one does assume, contrary to the present approach, that there are distinct physical possibilities related by diffeomorphisms, a kind of unobservable indeterminism arises in GR. This is because a diffeomorphism can sometimes leave a particular surface of simultaneity unchanged while shuffling 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.)[9] But the appeal to Leibniz equivalence also leads to new questions. So, for example, if we excise a structure that is not diffeomorphism invariant, then it would seem that the only quantities that remain are constant along gauge orbits. 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. This initiates a lot of discussion of what that could mean. Maudlin has argued that this is an absurd result. Healey has argued 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 (either as excess structure attributed to the world, or redundancy in the formalism), but they are *only* clues, part of the naturalistic metaphysician'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 judgment and care.

40.3 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 spacetime physics. It goes along with the very big theoretical shift from thinking of spacetime as the fixed background against which all of physics takes place to thinking of it as part of the subject matter of physics. It canonizes the accepted justification for rejecting Newtonian spacetime in favor of Galilean, motivating Minkowski spacetime 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. There was some of this kind of reasoning on the ground, but it was tangled up with less fruitful lines of reasoning, and it emerged with full clarity only in retrospect. (Anyone steeped in the actual history will notice that I didn't mention coordinate systems. These played a central, and mostly, unfortunate role in the actual history.)[10] 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 the false starts and confusions, portraying the theoretical progression as a rational exchange of one theory for another. 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. We now know which avenues of development turned out to be fruitful. The authority of this method of identifying and excising structure 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. 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.[11] 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. David Albert points out that we don't actually need to insist on a hard line between observable and unobservable quantities so long as our theories have a certain kind of causal integration (Albert, 2015). If we have a device D whose pointer observables are uncontroversially observable, then any physical quantity that can be made, by any physical process, no matter how indirect, to influence the pointer observable on D will then count as producing observable differences. The procedure will fail only for a theory that postulates equivocally observable quantities that can't be linked to any unequivocally observable quantity by any procedure.

There *are* more specific, rigorously definable notions of symmetry on offer in classical theories (e.g., the generalized, or Hamiltonian symmetries, or variational symmetries).[12] 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 (i) we don't want to tie the relevant notion of symmetry too tightly to any particular formalism, and (ii) they won't, in any case, do the work that we want them to do here.

Third, 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. But the clean separation of physically significant from physically insignificant structure is often very hard to achieve. It usually requires the development of new formalisms. And even when we can separate the structures, the interpretation of the results is difficult. The re-education of the imagination that was needed in the move from Newtonian to Galilean, and from Galilean to Minkowski spacetime, provides good examples. It is something that is even more amply illustrated in the case of modern gauge theories, so we turn now to those.

40.4 Gauge Theories

When we turn to non-spacetime theories such as electrostatics, electromagnetism, or quantum mechanics, we look to internal symmetries, i.e., transformations that alter the values of physical quantities that are supposed to represent the *internal degrees of freedom* pertaining to a system (or the matter in some volume of spacetime) rather than changing anything spatially or temporally external to it. We distinguish global internal symmetries from local symmetries. Global internal symmetries (such as the electrostatic potential, or phase transformations in quantum mechanics) act identically on the values of all points throughout spacetime. Local symmetries diverge from point to point. The best-known example of a theory with local gauge symmetry is classical electrodynamics. The vector potential is a four-dimensional vector; the time component is the familiar electric potential (V) while the space components collectively form the magnetic potential (A). Since physical predictions are fixed by the field tensor, which is left unchanged if we add the gradient of a scalar, if we follow the kind of reasoning we applied above, the most natural thing would be to regard structure not invariant under this transformation as physically unreal. The discovery and experimental confirmation of the Aharanov-Bohm effect, however, seemed to many to present an insuperable obstacle to that line of reasoning, because it showed that 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 (Vaidman, 2012). One *could* say that there is action at a distance from the field inside the solenoid, but if we want to avoid a new form of non-locality 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.

40.5 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 spacetime (Healey, 2007).[13] (Holonomies are images of oriented closed curves that can be constructed from line integrals of the gauge potential along closed curves.) On another view – discussed by Maudlin[14] – 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 spacetime 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. One of the defining principles 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 spacetime.

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 quantum mechanics 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 Lewis's Humean Supervenience program, 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.[15] 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.

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.

There is an important footnote to this debate. David Wallace has recently challenged the role that the Aharanov -Bohm effect has played in this discussion (Wallace MS). He argues that if we keep track of the complex scalar field to which the magnetic vector potential is coupled, we find gauge-invariant features of the scalar field and potential together that explain the effect and can be given a local characterization. If he is right, the interpretation of gauge –invariance as permuting

AU: Please check if we can change the spelling to "Aharonov" throughout the chapter.

physically insignificant structure be restored, without non-locality, and the more radical conclusions can be avoided. As can be seen from the ongoing nature of these debates, these issues are unsettled, and we can hope to gain some clarity in years to come.

40.6 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 a theory outside its primary domain of application. The Aharanov-Bohm effect provides a good example of this. The effect emerges only when an electromagnetic field is coupled to a complex vector field, and so it doesn't emerge until classical electromagnetism is combined with a theory that describes the behavior of quantum particles.

Second, as I've emphasized, 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 spacetime 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 the ongoing discussion of gauge theories.

Third, applying the method is *never* free of interpretive assumptions. There is no *purely* formal way of recognizing a redundant structure. As I mentioned above, if we simply identify all solutions related by symmetries of the laws, we 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. The use of symmetries to identify structure that plays no dynamical role, however, need not be applied in this general setting. It can also be powerful when applied 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 sets 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.

Finally, it is a little misleading to say that structures related by a symmetry of the laws never have any physical interpretation at all, as we will see in the next section.

40.7 Justification

There are difficult, open questions about why the method works. There is no a priori *logical* principle that guarantees that this method for limning the structure of our theories won't lead us to eliminate a structure that represents something real. We can certainly imagine worlds in which that would be so. Consider, for example, Newton's world with absolute time and space or the non-localized gauge potential properties view considered by Healey are perfectly intelligible empirical possibilities.

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 counterexample 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 and S maps a onto empirically indistinguishable counterpart b," to "a=b," would fail.

There are, nevertheless, two broad routes to rationalizing the method that suggests 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. We start from the principle 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, unless differences between those elements produce empirically distinguishable effects. Then we point out that someone who recognizes structure that maps solutions of the laws onto empirically indistinguishable solutions can be faulted for recognizing structure that they can have no evidence for. That is a strong intuition, but making it more than an intuition would require showing that that given a body of data that respects 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 measure.

(ii) The second focuses on issues about reference and argues 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 observable (e.g., two points in absolute space at different times, or internal properties that don't make a discernibly different impact on measuring devices), we have not really managed to do so. We have not actually managed, that is to say, to get our referential hooks into two separate elements of reality. What we have, in that case, is actually two names (coined by rigidified definite descriptions, 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). This way of rationalizing the method 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. That is an issue that is very much contested, and very hard to see one's way through.

The question of why the method works is bound up with the theoretical preference for simpler theories. There is a temptation to try to justify the method by appeal to Ockham's razor. That temptation should be resisted.[16] Instead of looking to Ockham's razor for justification of the method, better to look at the logic behind the reasoning directly and see if it captures the valid heart of Ockham's razor. It might be that the preference for simpler theories is a generalized version of the method discussed here. The idea here would be that whenever you are choosing between competing, empirically adequate formalisms, one of which is simpler (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 theory has some redundancy in the formalism.

This might not be the right way to put it to the extent that it suggests we are identifying structures that have no physical interpretation at all, i.e., part of the mathematical machinery that is empty of any physical significance. But that is not quite right. Rovelli makes the point elegantly in "Why Gauge?" (http://arxiv.org/abs/1308.5599). 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). 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. 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 when we see differences in the values of those quantities, we are seeing differences in the perspective from which the object is viewed rather than differences intrinsic to the object. In a Minkowski spacetime, for example, observers traveling non-inertially relative to one another disagree on simultaneity in the way that observers in different locations disagree on what is nearby. Just so, Rovelli argues, 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. None of this means that there aren't some symmetries that permute physically empty features of notation. It is only to suggest that gauge invariance is not best understood that way.

The real authority of the method, in practice, rests 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.

40.8 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.

Acknowledgment

I would like to thank Jim Weatherall, Jeff Barrett, Richard Healey, Gordon Belot, Chris Smeenk, Giovanni Valente, Hans Halvorson, and other participants in the workshop on the Foundations of Gauge Theories. This paper exists because of that invitation, and I learned a great deal from the discussion there. Special thanks to Alastair Wilson for insightful comments on an earlier draft.

References

Albert, D. (2015). The technique of significables. In *After Physics*. Cambridge: Harvard University Press.

Baker, D.J. (2010). Symmetry and the metaphysics of physics. *Philosophy Compass*, 5(12): 1157–1166.

Barker, P. (2002). Constructing copernicus. *Perspectives on Science*, 10(2): 208–227.

Belot, G. (2001a). Symmetry and equivalence. In R. Batterman (ed.), *Oxford Handbook of Philosophy of Physics*. New York: Oxford University Press, pp. 318–339.

Belot, G. (2001b). The principle of sufficient reason. *Journal of Philosophy*, XCVIII: 55–74.

Brading, K.A. and Castellani, E. (2003). *Symmetries in Physics: Philosophical Reflections*. Cambridge: Cambridge University Press.

Cantwell, B. (2002). *Introduction to Symmetry Analysis*. Cambridge: Cambridge University Press.

Carroll, S. (1997). *Lecture notes on general relativity*. Available at: http://arxiv.org/abs/gr-qc/9712019.

Castrillon, L. and Marsden, J.E. (2008). Covariant and dynamical reduction for principal bundle field theories. *Annals of Global Analysis and Geometry*, 34: 263–285.

Clark, A. (1993). *Associative Engines: Connectionism, Concepts and Representational Change*. Cambridge, MA: MIT Press.

Clark, A. (1998). Magic words: How language augments human computation. In P. Carruthers and J. Boucher (eds.), *Language and Thought: Interdisciplinary Themes*. Cambridge: Cambridge University Press, pp. 162–183.

Clark, A. and Karmiloff-Smith, A. (1993). The Cognizer's innards: A psychological & philosophical perspective on the development of thought. *Mind & Language*, 8(4): 487–519.

Dennett, D. (1994). Labeling and learning (Commentary on Clark & Karmiloff-Smith 1994). *Mind and Language*, 8: 540–548.

Earman. (2002).

Field, V. and Bowden, M.E. (1994). Kepler's geometrical cosmology. *Annals of Science*, 51(1): 95–97.

van Fraassen, B. and Ismael, J. (2003). Symmetry as a guide to superfluous theoretical structure. In E. Castellani and K. Brading (eds.), *Symmetries in Physics: Philosophical Reflections*. Cambridge: Cambridge University Press, pp. 371–392.

Friedman, M. (1983). *Foundations of Space-Time Theories: Relativistic Physics and Philosophy of Science*. Princeton, NJ: Princeton University Press.

Harman, P.M. (1990). Kepler's geometrical cosmology. *International Studies in Philosophy*, 22(3): 103.

Healey, R. (2007). *Gauging What's Real*. Cambridge: Cambridge University Press.

Ismael, J. (1997). Curie's principle. *Synthèse*, 110(2): 167–190.

Karmiloff-Smith, A. (1992). *Beyond Modularity: A Developmental Perspective on Cognitive Science*. Cambridge, MA: MIT Press/Bradford Books.

Kuhn, T. (1996). *The Structure of Scientific Revolutions* (3rd edition). Chicago, IL: University of Chicago Press.

Leeds, S. (1999). Gauges: Aharonov, Bohm, Yang, Healey. *Philosophy of Science*, 66: 606–627.

Maudlin, T. (1998). Healey on the Aharonov-Bohm effect. *Philosophy of Science*, 65: 361–368.

Maudlin, T. (2007). *The Metaphysics Within Physics*. Oxford: Oxford University Press.

Maudlin, T. (2012). *Philosophy of Physics: Space and Time*. Princeton, NJ: Princeton University Press.

Norton, J. (1993). General covariance and the foundations of general relativity: Eight decades of dispute. *Reports of Progress in Physics*, 56: 791–861.

Norton. (2011). Norton, J.D. (2015). The hole argument. In E.N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2015 Edition). Available at: https://plato.stanford.edu/archives/fall2015/entries/spacetime-holearg/.

Rovelli, C. (2014). Why gauge? *Foundations of Physics*, 44(1): 91–104.

Sober, E. (2015). *Ockham's Razors: A User's Manual*. Cambridge: Cambridge University Press.

Vaidman, L. (2012). On the role of potentials in the Aharonov-Bohm effect. *Physical Review A*, 86: 040101.

Wallace, D. (MS). *Deflating the Aharonov-Bohm effect*. Available at: https://arxiv.org/abs/1407.5073.

AU: Please provide complete details for "Earman (2002)".

AU: Please provide complete details for "Norton (2011)".

Notes

1 See Field and Bowden (1994), Harman (1990), Barker (2002).
2 Kuhn (1996, p. 124).
3 For a sampling, see Karmiloff-Smith (1992), Clark (1993), Clark and Karmiloff-Smith (1993), Clark (1998), Dennett (1994).
4 See Ismael (1997), Norton (1993).
5 Baker (2010), Brading and Castellani (2003), Maudlin (2012), Friedman (1983).
6 See Dasgupta (this volume).
7 Belot (2001a, 2001b).
8 Norton (2015), Carroll (1997). See also the chapter by Pooley (this volume).
9 Norton (2011), ibid.
10 Michael Friedman (1983) contains an historical overview.
11 Van Fraassen and Ismael (2003) make a similar suggestion.
12 Castrillon Lopez and Marsden (2008), Cantwell (2002).
13 Healey (2007); also Leeds (1999), Maudlin (1998).
14 Maudlin (2007, pp. 78–103).
15 See Healey (this volume) for further discussion.
16 There are as many different versions of Ockham's razor as there are notions of simplicity; Sober (2015) is an excellent discussion of different versions of the principle and its various applications.