

PDF version of the entry
Symmetry and Symmetry Breaking
<https://plato.stanford.edu/archives/win2017/entries/symmetry-breaking/>
from the WINTER 2017 EDITION of the

STANFORD ENCYCLOPEDIA OF PHILOSOPHY



Edward N. Zalta Uri Nodelman Colin Allen R. Lanier Anderson
Principal Editor Senior Editor Associate Editor Faculty Sponsor

Editorial Board
<https://plato.stanford.edu/board.html>

Library of Congress Catalog Data
ISSN: 1095-5054

Notice: This PDF version was distributed by request to members of the Friends of the SEP Society and by courtesy to SEP content contributors. It is solely for their fair use. Unauthorized distribution is prohibited. To learn how to join the Friends of the SEP Society and obtain authorized PDF versions of SEP entries, please visit <https://leibniz.stanford.edu/friends/>.

Stanford Encyclopedia of Philosophy
Copyright © 2017 by the publisher
The Metaphysics Research Lab
Center for the Study of Language and Information
Stanford University, Stanford, CA 94305

Symmetry and Symmetry Breaking
Copyright © 2017 by the authors
Katherine Brading, Elena Castellani, and Nicholas Teh

All rights reserved.

Copyright policy: <https://leibniz.stanford.edu/friends/info/copyright/>

Symmetry and Symmetry Breaking

First published Thu Jul 24, 2003; substantive revision Thu Dec 14, 2017

Symmetry considerations dominate modern fundamental physics, both in quantum theory and in relativity. Philosophers are now beginning to devote increasing attention to such issues as the significance of gauge symmetry, quantum particle identity in the light of permutation symmetry, how to make sense of parity violation, the role of symmetry breaking, the empirical status of symmetry principles, and so forth. These issues relate directly to traditional problems in the philosophy of science, including the status of the laws of nature, the relationships between mathematics, physical theory, and the world, and the extent to which mathematics suggests new physics.

This entry begins with a brief description of the historical roots and emergence of the concept of symmetry that is at work in modern science. It then turns to the application of this concept to physics, distinguishing between two different uses of symmetry: symmetry principles versus symmetry arguments. It mentions the different varieties of physical symmetries, outlining the ways in which they were introduced into physics. Then, stepping back from the details of the various symmetries, it makes some remarks of a general nature concerning the status and significance of symmetries in physics.

- 1. The Concept of Symmetry
- 2. Symmetry Principles
 - 2.1 Relativity
 - 2.2 Symmetry and quantum mechanics
 - 2.3 Permutation symmetry
 - 2.4 C, P, T
 - 2.5 Gauge symmetry

- 2.6 Dualities
 - 3. Symmetry Arguments
 - 4. Symmetry Breaking
 - 4.1 Explicit symmetry breaking
 - 4.2 Spontaneous symmetry breaking
 - 5. General Philosophical Questions
 - Bibliography
 - Academic Tools
 - Other Internet Resources
 - Related Entries
-

1. The Concept of Symmetry

The term “symmetry” derives from the Greek words *sun* (meaning ‘with’ or ‘together’) and *metron* (‘measure’), yielding *summetria*, and originally indicated a relation of commensurability (such is the meaning codified in Euclid’s *Elements* for example). It quickly acquired a further, more general, meaning: that of a proportion relation, grounded on (integer) numbers, and with the function of harmonizing the *different* elements into a *unitary whole*. From the outset, then, symmetry was closely related to harmony, beauty, and unity, and this was to prove decisive for its role in theories of nature. In Plato’s *Timaeus*, for example, the regular polyhedra are afforded a central place in the doctrine of natural elements for the proportions they contain and the beauty of their forms: fire has the form of the regular tetrahedron, earth the form of the cube, air the form of the regular octahedron, water the form of the regular icosahedron, while the regular dodecahedron is used for the form of the entire universe. The history of science provides another paradigmatic example of the use of these figures as basic ingredients in physical description: Kepler’s 1596 *Mysterium Cosmographicum* presents a planetary architecture grounded on the five regular solids.

From a modern perspective, the regular figures used in Plato’s and Kepler’s physics for the mathematical proportions and harmonies they contain (and the related properties and beauty of their form) are symmetric in another sense that does not have to do with proportions. In the language of modern science, the symmetry of geometrical figures — such as the regular polygons and polyhedra — is defined in terms of their invariance under specified groups of rotations and reflections. Where does this definition stem from? In addition to the ancient notion of symmetry used by the Greeks and Romans (current until the end of the Renaissance), a different notion of symmetry emerged in the seventeenth century, grounded not on proportions but on an equality relation between elements that are opposed, such as the left and right parts of a figure. Crucially, the parts are *interchangeable with respect to the whole* — they can be exchanged with one another while preserving the original figure. This latter notion of symmetry developed, via several steps, into the concept found today in modern science. One crucial stage was the introduction of specific mathematical operations, such as reflections, rotations, and translations, that are used to describe with precision how the parts are to be exchanged. As a result, we arrive at a definition of the symmetry of a geometrical figure in terms of its *invariance* when equal component parts are exchanged according to one of the specified operations. Thus, when the two halves of a bilaterally symmetric figure are exchanged by reflection, we recover the original figure, and that figure is said to be invariant under left-right reflections. This is known as the “crystallographic notion of symmetry”, since it was in the context of early developments in crystallography that symmetry was first so defined and applied.^[1] The next key step was the generalization of this notion to the group-theoretic definition of symmetry, which arose following the nineteenth-century development of the algebraic concept of a group, and the fact that the symmetry operations of a figure were found to satisfy the conditions for forming a group.^[2] For example, reflection symmetry has

now a precise definition in terms of invariance under the group of reflections. Finally, we have the resulting close connection between the notion of symmetry, equivalence and group: a symmetry group induces a partition into equivalence classes. The elements that are exchanged with one another by the symmetry transformations of the figure (or whatever the “whole” considered is) are connected by an *equivalence relation*, thus forming an equivalence class.^[3]

The group-theoretic notion of symmetry is the one that has proven so successful in modern science. Note, however, that symmetry remains linked to beauty (regularity) and unity: by means of the symmetry transformations, distinct (but “equal” or, more generally, “equivalent”) elements are related to each other and to the whole, thus forming a regular “unity”. The way in which the regularity of the whole emerges is dictated by the nature of the specified transformation group. Summing up, a *unity of different and equal elements* is always associated with symmetry, in its ancient or modern sense; the way in which this unity is realized, on the one hand, and how the equal and different elements are chosen, on the other hand, determines the resulting symmetry and in what exactly it consists.

The definition of symmetry as “invariance under a specified group of transformations” allowed the concept to be applied much more widely, not only to spatial figures but also to abstract objects such as mathematical expressions — in particular, expressions of physical relevance such as dynamical equations. Moreover, the technical apparatus of group theory could then be transferred and used to great advantage within physical theories.^[4]

When considering the role of symmetry in physics from a historical point of view, it is worth keeping in mind two preliminary distinctions:

- The first is between implicit and explicit uses of the notion. Symmetry considerations have always been applied to the description of nature, but for a long time in an implicit way only. As we have seen, the scientific notion of symmetry (the one we are interested in here) is a recent one. If we speak about a role of this concept of symmetry in the ancient theories of nature, we must be clear that it was not used explicitly in this sense at that time.
- The second is between the two main ways of using symmetry. First, we may attribute specific symmetry properties to phenomena or to laws (*symmetry principles*). It is the application with respect to laws, rather than to objects or phenomena, that has become central to modern physics, as we will see. Second, we may derive specific consequences with regard to particular physical situations or phenomena on the basis of their symmetry properties (*symmetry arguments*).

2. Symmetry Principles

The first explicit study of the invariance properties of equations in physics is connected with the introduction, in the first half of the nineteenth century, of the transformational approach to the problem of motion in the framework of analytical mechanics. Using the formulation of the dynamical equations of mechanics due to W. R. Hamilton (known as the Hamiltonian or canonical formulation), C. G. Jacobi developed a procedure for arriving at the solution of the equations of motion based on the strategy of applying transformations of the variables that leave the Hamiltonian equations invariant, thereby transforming step by step the original problem into new ones that are simpler but perfectly equivalent (for further details see Lanczos 1949).^[5] Jacobi’s canonical transformation theory, although introduced for the “merely instrumental” purpose of solving dynamical problems, led to a very important line of research: the

general study of physical theories in terms of their transformation properties. Examples of this are the studies of invariants under canonical transformations, such as Poisson brackets or Poincaré’s integral invariants; the theory of continuous canonical transformations due to S. Lie; and, finally, the connection between the study of physical invariants and the algebraic and geometric theory of invariants that flourished in the second half of the nineteenth century, and which laid the foundation for the geometrical approach to dynamical problems. The use of the mathematics of group theory to study physical theories was central to the work, early in the twentieth century in Göttingen, of the group whose central figures were F. Klein (who earlier collaborated with Lie) and D. Hilbert, and which included H. Weyl and later E. Noether. We will return later in this section to Weyl (see Sections 2.1.2, 2.2, 2.5) and Noether (see Section 2.1.2). For more details on these developments see Brading and Castellani (2007).

On the above approach, the equations or expressions of physical interest are already given, and the strategy is to study their symmetry properties. There is, however, an alternative way of proceeding, namely the reverse one: start with specific symmetries and search for dynamical equations with such properties. In other words, we *postulate* that certain symmetries are physically significant, rather than deriving them from prior dynamical equations. The assumption of certain symmetries in nature is not, of course, a novelty. Although not explicitly expressed as symmetry principles, the homogeneity and isotropy of physical space, and the uniformity of time (forming together with the invariance under Galilean boosts “the older principles of invariance” — see Wigner 1967,^[6] pp. 4–5), have been assumed as prerequisites in the physical description of the world since the beginning of modern science. Perhaps the most famous early example of the deliberate use of this type of symmetry principle is Galileo’s discussion of whether the Earth moves in his *Dialogue concerning the two chief world systems* of 1632. Galileo sought to

neutralize the standard arguments purporting to show that, simply by looking around us at how things behave locally on Earth — how stones fall, how birds fly — we can conclude that the Earth is at rest rather than rotating, arguing instead that these observations do not enable us to determine the state of motion of the Earth. His approach was to use an analogy with a ship: he urges us to consider the behaviour of objects, both animate and inanimate, inside the cabin of a ship, and claims that no experiments carried out inside the cabin, without reference to anything outside the ship, would enable us to tell whether the ship is at rest or moving smoothly across the surface of the Earth. The *assumption* of a symmetry between rest and a certain kind of motion leads to the prediction of this result, without the need to know the details of the laws governing the experiments on the ship. The “Galilean principle of relativity” (according to which the laws of physics are invariant under Galilean boosts, where the states of motion considered are now those of uniform velocity) was quickly adopted as an axiom and widely used in the seventeenth century, notably by Huygens in his solution to the problem of colliding bodies and by Newton in his early work on motion. Huygens took the relativity principle as his 3rd hypothesis or axiom, but in Newton’s *Principia* it is demoted to a corollary to the laws of motion, its status in Newtonian physics therefore being that of a consequence of the laws, even though it remains, in fact, an independent assumption.

Although the spatial and temporal invariance of mechanical laws was known and used for a long time in physics, and the group of the global spacetime symmetries for electrodynamics was completely derived by H. Poincaré^[7] before Einstein’s famous 1905 paper setting out his special theory of relativity, it was not until this work by Einstein that the status of symmetries with respect to the laws was reversed. E. P. Wigner (1967, p. 5) writes that “the significance and general validity of these principles were recognized, however, only by Einstein”, and that Einstein’s work on special relativity marks “the reversal of a trend: until then, the principles

of invariance were derived from the laws of motion ... It is now natural for us to derive the laws of nature and to test their validity by means of the laws of invariance, rather than to derive the laws of invariance from what we believe to be the laws of nature". In postulating the universality of the global continuous spacetime symmetries, Einstein's construction of his special theory of relativity represents the first turning point in the application of symmetry to twentieth-century physics.^[8]

2.1 Relativity

2.1.1 The special theory of relativity

Einstein's special theory of relativity (STR) is constructed on the basis of two fundamental postulates. One is the light postulate (that the speed of light, in the "rest frame", is independent of the speed of the source), and the other is the principle of relativity. The latter was adopted by Einstein explicitly as a means of restricting the form of the laws, whatever their detailed structure might turn out to be. Thus, we have the difference between a "constructive" and a "principle" theory: in the former case we build our theory based on known facts about the constitution and behaviour of material bodies; in the latter case we start by restricting the possible form of such a theory by adopting certain principles.^[9]

The principle of relativity as adopted by Einstein (1905, p. 395 of the English translation) simply asserts that:

The laws by which the states of physical systems undergo changes are independent of whether these changes of states are referred to one or the other of two coordinate systems moving relatively to each other in uniform translational motion.

This principle, when combined with the light postulate (and certain other assumptions), leads to the Lorentz transformations, these being the

transformations between coordinate systems moving uniformly with respect to one another according to STR. According to STR the laws of physics are invariant under Lorentz transformations, and indeed under the full Poincaré group of transformations. These transformations differ from the Galilean transformations of Newtonian mechanics. H. Minkowski reformulated STR, showing that space and time are part of a single four-dimensional geometry, Minkowski spacetime. In this way, the Poincaré group of symmetry transformations is part of the structure of spacetime in STR, and for this reason these symmetries have been labelled "geometric symmetries" by Wigner (1967, especially pp. 15 and 17–19).

There is a debate in the literature concerning how the principle of relativity, and more generally the global space-time symmetries, should be understood. On one approach, the significance of space-time symmetries is captured by considering the structure of a theory through transformations on its models, those models consisting of differentiable manifolds endowed with various geometric objects and relations (see Anderson, 1967, and Norton, 1989). According to Brown and Sypel (1995) and Budden (1997), this approach fails to recognise the central importance of effectively isolated subsystems, the empirical significance of symmetries resting on the possibility of transforming such a subsystem (rather than applying the transformation to the entire universe). For further developments in this debate, including applications to local symmetries and to gauge theories, see Kosso (2000), Brading and Brown (2004), Healey (2007), Healey (2009), Greaves and Wallace (2014), Friederich (2015), Rovelli (2014) and Teh (2015, 2016).

The global spacetime invariance principles are intended to be valid for all the laws of nature, for all the processes that unfold in the spacetime. This universal character is not shared by the physical symmetries that were next introduced in physics. Most of these were of an entirely new kind, with no roots in the history of science, and in some cases expressly introduced to

describe specific forms of interactions — whence the name “dynamical symmetries” due to Wigner (1967, see especially pp. 15, 17–18, 22–27, 33).

2.1.2 The general theory of relativity

Einstein’s general theory of relativity (GTR) was also constructed using a symmetry principle at its heart: the principle of general covariance. Much ink has been spilled over the significance and role of general covariance in GTR, including by Einstein himself.^[10] For a long time he viewed the principle of general covariance as an extension of the principle of relativity found in both classical mechanics and STR, and this is a view that continues to provoke vigorous debate. Norton (2003) discusses the “Kretschmann objection” to the physical significance of general covariance. On invariance versus covariance, see Anderson (1967), Brown and Brading (2002), and Martin (2003, Section 2.2). What is clear is that the mere requirement that a theory be generally covariant represents no restriction on the form of the theory; further stipulations must be added, such as the requirement that there be no “absolute objects” (this itself being a problematic notion). Once some such further requirements are added, however, the principle of general covariance becomes a powerful tool. For a recent review and analysis of this debate, see Pitts (2006).

In Einstein’s hands the principle of general covariance was a crucial postulate in the development of GTR.^[11] The diffeomorphism freedom of GTR, i.e., the invariance of the form of the laws under transformations of the coordinates depending smoothly on arbitrary functions of space and time, is a “local” spacetime symmetry, in contrast to the “global” spacetime symmetries of STR (which depend instead on constant parameters). For a discussion of coordinate-based approaches to the diffeomorphism invariance of General Relativity, see Wallace (forthcoming), and for more on the physical interpretation of this

invariance, see Pooley (2017). Such local symmetries are “dynamical” symmetries in Wigner’s sense, since they describe a particular interaction, in this case gravity. As is well known, the spacetime metric in GTR is no longer a “background” field or an “absolute object”, but instead it is a dynamical player, the gravitational field manifesting itself as spacetime curvature.

The extension of the concept of continuous symmetry from “global” symmetries (such as the Galilean group of spacetime transformations) to “local” symmetries is one of the important developments in the concept of symmetry in physics that took place in the twentieth century. Prompted by GTR, Weyl’s 1918 “unified theory of gravitation and electromagnetism” extended the idea of local symmetries (see Ryckman, 2003, and Martin, 2003), and although this theory is generally deemed to have failed, the theory contains the seeds of later success in the context of quantum theory (see below, Section 2.5).

Meanwhile, Hilbert and Klein undertook detailed investigations concerning the role of general covariance in theories of gravitation, and enlisted the assistance of Noether in their debate over the status of energy conservation in such theories. This led to Noether’s famous 1918 paper containing two theorems, the first of which leads to a connection between global symmetries and conservation laws, and the second of which leads to a number of results associated with local symmetries, including a demonstration of the different status of the conservation laws when the global symmetry group is a subgroup of some local symmetry group of the theory in question (see Brading and Brown, 2003).

2.2 Symmetry and quantum mechanics

The application of the theory of groups and their representations for the exploitation of symmetries in the quantum mechanics of the 1920s

undoubtedly represents the second turning point in the twentieth-century history of physical symmetries. It is, in fact, in the quantum context that symmetry principles are at their most effective. Wigner and Weyl were among the first to recognize the great relevance of symmetry groups to quantum physics and the first to reflect on the meaning of this. As Wigner emphasized on many occasions, one essential reason for the “increased effectiveness of invariance principles in quantum theory” (Wigner, 1967, p. 47) is the linear nature of the state space of a quantum physical system, corresponding to the possibility of superposing quantum states. This gives rise to, among other things, the possibility of defining states with particularly simple transformation properties in the presence of symmetries.

In general, if G is a symmetry group of a theory describing a physical system (that is, the dynamical equations of the theory are invariant under the transformations of G), this means that the states of the system transform into each other according to some “representation” of the group G . In other words, the group transformations are mathematically represented in the state space by operations relating the states to each other. In quantum mechanics, these operations are implemented through the operators that act on the state space and correspond to the physical observables, and any state of a physical system can be described as a superposition of states of elementary systems, that is, of systems the states of which transform according to the “irreducible” representations of the symmetry group. Quantum mechanics thus offers a particularly favourable framework for the application of symmetry principles. The observables representing the action of the symmetries of the theory in the state space, and therefore commuting with the Hamiltonian of the system, play the role of the conserved quantities; furthermore, the eigenvalue spectra of the invariants of the symmetry group provide the labels for classifying the irreducible representations of the group: on this fact is grounded the possibility of associating the values of the invariant properties

characterizing physical systems with the labels of the irreducible representations of symmetry groups, i.e. of classifying elementary physical systems by studying the irreducible representations of the symmetry groups.

2.3 Permutation symmetry

The first non-spatiotemporal symmetry to be introduced into microphysics, and also the first symmetry to be treated with the techniques of group theory in the context of quantum mechanics, was *permutation symmetry* (or invariance under the transformations of the permutation group). This symmetry, “discovered” by W. Heisenberg in 1926 in relation to the indistinguishability of the “identical” electrons of an atomic system, [12] is the *discrete* symmetry (i.e. based upon groups with a discrete set of elements) at the core of the so-called quantum statistics (the Bose-Einstein and Fermi-Dirac statistics), governing the statistical behaviour of ensembles of certain types of indistinguishable quantum particles (e.g. bosons and fermions). The permutation symmetry principle states that if such an ensemble is invariant under a permutation of its constituent particles then one doesn’t count those permutations which merely exchange indistinguishable particles, that is the exchanged state is identified with the original state (see French and Rickles, 2003, Section 1).

Philosophically, permutation symmetry has given rise to two main sorts of questions. On the one side, seen as a condition of physical indistinguishability of identical particles (i.e. particles of the same kind in the same atomic system), it has motivated a rich debate about the significance of the notions of identity, individuality, and indistinguishability in the quantum domain. Does it mean that the quantum particles are not individuals? Does the existence of entities which are physically indistinguishable although “numerically distinct” (the so-called problem of identical particles) imply that the Leibniz’s Principle of the

Identity of Indiscernibles should be regarded as violated in quantum physics? On the other side, what is the theoretical and empirical status of this symmetry principle? Should it be considered as an axiom of quantum mechanics or should it be taken as justified empirically? It is currently taken to explain the nature of fermionic and bosonic quantum statistics, but why do there appear to be only bosons and fermions in the world when the permutation symmetry group allows the possibility of many more types? French and Rickles (2003) offer an overview of the above and related issues, and a new twist in the tale can be found in Saunders (2006). Saunders discusses permutation symmetry in classical physics, and argues for indistinguishable classical particles obeying classical statistics. He argues that the differences between quantum and classical statistics, for certain classes of particles, therefore cannot be accounted for solely in terms of indistinguishability. For further discussion and references see French and Krause (2006), Ladyman and Bigaj (2010), Caulton and Butterfield (2012), and the related SEP entry identity and individuality in quantum theory.

2.4 C, P, T

Because of the specific properties of the quantum description, the discrete symmetries of spatial reflection symmetry or *parity* (P) and *time reversal* (T) were “rediscovered” in the quantum context, taking on a new significance. Parity was introduced in quantum physics in 1927 in a paper by Wigner, where important spectroscopic results were explained for the first time on the basis of a group-theoretic treatment of permutation, rotation and reflection symmetries. Time reversal invariance appeared in the quantum context, again due to Wigner, in a 1932 paper.^[13] To these was added the new quantum particle-antiparticle symmetry or *charge conjugation* (C). Charge conjugation was introduced in Dirac’s famous 1931 paper “Quantized singularities in the electromagnetic field”.

The discrete symmetries C, P and T are connected by the so-called CPT theorem, first proved by Lüders and Pauli in the early 1950s, which states that the combination of C, P , and T is a general symmetry of physical laws. For philosophical reflections on the meaning and grounding of the CPT theorem see Wallace (2009) and Greaves (2010). A discussion of the proofs of the theorem, both within the standard quantum field theory framework and the axiomatic field theory framework, is provided in Greaves and Thomas (2014).

The laws governing gravity, electromagnetism, and the strong interaction are invariant with respect to C, P and T independently. However, in 1956 T. D. Lee and C. N. Yang pointed out that β -decay, governed by the weak interaction, had not yet been tested for invariance under P . Soon afterwards C. S. Wu and her colleagues performed an experiment showing that the weak interaction violates parity. Nevertheless, β -decay respects the combination of C and P as a symmetry. In 1964, however, CP was found to be violated in weak interaction by Cronin and Fitch in an experiment involving K-mesons. This implied, in virtue of the CPT theorem, the violation of T -symmetry as well; since then, there have been direct observations of T -symmetry violation, as reported by e.g., CPLEAR Collaboration (1998). For a careful analysis of the underlying assumptions working in the assessment of T -violation see Roberts (2015) and Ashtekar (2015). The philosophical puzzle is how we can come to know this time asymmetry, given that we cannot truly ‘reverse’ time. Roberts proposes three templates for explaining how this is possible, drawing in particular on a version of Curie’s principle. By contrast, Gołosz (2016) argues that this asymmetry should be understood as applying to a physical process, rather than to time itself. A yet more basic question is that of what time reversal symmetry actually means. On the philosophical debate on this point, see Roberts (2017) and references therein.

The existence of parity violation in our fundamental laws has led to a new chapter in an old philosophical debate concerning chiral or handed objects and the nature of space. A description of a left hand and one of a right hand will not differ so long as no appeal is made to anything beyond the relevant hand. Yet left and right hands do differ — a left-handed glove will not fit on a right hand. For a brief period, Kant saw in this reason to prefer a substantialist account of space over a relational one, the difference between left and right hands lying in their relation to absolute space. Regardless of whether this substantialist solution succeeds, there remains the challenge to the relationalist of accounting for the difference between what Kant called “incongruent counterparts” — objects which are the mirror-image of one another and yet cannot be made to coincide by any rigid motion. The relationalist may respond by denying that there is any *intrinsic* difference between a left and a right hand, and that the incongruence is to be accounted for in terms of the relations between the two hands (if a universe was created with only one hand in it, it would be neither left nor right, but the second hand to be created would be either incongruent or congruent with it). This response becomes problematic in the face of parity violation, where one possible experimental outcome is much more likely than its mirror-image. Since the two possible outcomes don’t differ intrinsically, how should we account for the imbalance? This issue continues to be discussed in the context of the substantialist-relationalist debate. For further details see Pooley (2003) and Saunders (2007).

2.5 Gauge symmetry

The starting point for the idea of *continuous internal symmetries* was the interpretation of the presence of particles with (approximately) the same value of mass as the components (*states*) of a single physical system, connected to each other by the transformations of an underlying symmetry group. This idea emerged by analogy with what happened in the case of

permutation symmetry, and was in fact due to Heisenberg (the discoverer of permutation symmetry), who in a 1932 paper introduced the $SU(2)$ symmetry connecting the proton and the neutron (interpreted as the two states of a single system). This symmetry was further studied by Wigner, who in 1937 introduced the term *isotopic spin* (later contracted to *isospin*). The various internal symmetries are invariances under phase transformations of the quantum states and are described in terms of the unitary groups $SU(N)$. The term “gauge” is sometimes used for all continuous internal symmetries, and is sometimes reserved for the *local* versions (these being at the core of the Standard Model for elementary particles).^[14]

The phase of the quantum wavefunction encodes internal degrees of freedom. With the requirement that a theory be invariant under *local gauge transformations* involving the phase of the wavefunction, Weyl’s ideas of 1918 found a successful home in quantum theory (see O’Raifeartaigh, 1997). Weyl’s new 1929 theory was a theory of electromagnetism coupled to matter. The history of gauge theory is surveyed briefly by Martin (2003), who highlights various issues surrounding gauge symmetry, in particular the status of the so-called “gauge principle”, first proposed by Weyl. The main steps in development of gauge theory are the Yang and Mills non-Abelian gauge theory of 1954, and the problems and solutions associated with the successful development of gauge theories for the short-range weak and strong interactions.

The main philosophical questions raised by gauge theory all hinge upon how we should understand the relationship between mathematics and physics. There are two broad categories of discussion. The first concerns the gauge principle, already mentioned, and the issue here is the extent to which the requirement that we write our theories in locally-symmetric form enables us to derive new physics. The analysis concerns listing what

premises constitute the gauge principle, examining the status of these premises and what motivation might be given for them, determining precisely what can be obtained on the basis of these premises, and what more needs to be added in order to arrive at a (successful) physical theory. For details see, for example, Teller (2000) and Martin (2003).

The second category concerns the question of which quantities in a gauge theory represent the “physically real” properties. This question arises acutely in gauge theories because of the apparent failure of determinism. The problem was first encountered in GTR (which in this respect is a gauge theory), and for further details the best place to begin is with the literature on Einstein’s “hole argument” (see Earman and Norton, 1987; Earman, 1989, Chapter 9; and more recently Norton, 1993; Rynasiewicz, 1999; Saunders, 2002; and the references therein). In practice, we find that only gauge-invariant quantities are observables, and this seems to rescue us. However, this is not the end of the story. The other canonical example is the Aharonov-Bohm effect, and we can use this to illustrate the interpretational problem associated with gauge theories, sometimes characterized as a dilemma: failure of determinism or action-at-a-distance (see Healey, 2001). Restoring determinism depends on only gauge-invariant quantities being taken as representing “physically real” quantities, but accepting this solution apparently leaves us with some form of non-locality between causes and effects.

Furthermore, we face the question of how to understand the role of the non-gauge-invariant quantities appearing in the theory, and the problem of how to interpret what M. Redhead calls “surplus structure” (see Redhead, 2003). For further detail on this topic, see Belot (1998), Nounou (2003), Weatherall (2016), and Nguyen, Teh, & Wells (forthcoming). For an approach to these questions using the theory of constrained Hamiltonian systems, see also Earman (2003b) and Castellani (2003, 2004). For an intuitive characterization of gauge symmetry, one that is more general than

the Lagrangian and Hamiltonian formulations of theories using which gauge symmetry is usually expressed, see Belot (2008). How best to interpret gauge theories is an open issue in the philosophy of physics. Healey (2007) discusses the conceptual foundations of gauge theories, arguing in favour of a non-separable holonomy interpretation of classical Yang-Mills gauge theories of fundamental interactions. Catren (2008) tackles the ontological implications of Yang-Mills theory by means of the fiber bundle formalism. Useful references are the Metascience review symposium on Healey (2007) (Rickles, Smeenk, Lyre and Healey, 2009), and the “Synopsis and Discussion” of the workshop “Philosophy of Gauge Theory,” Center for Philosophy of Science, University of Pittsburgh, 18–19 April 2009 (available online).

2.6 Dualities

Thus far, we have been discussing symmetries which act on the space of states of a physical theory. In recent years, much discussion within physics and philosophy has centered on certain kinds of symmetries that act on the space of theories. When such symmetries are interpreted as realizing an “equivalence” (which sense is itself something that requires philosophical work to explicate) between two theories, the theories are typically said to be related by a “duality symmetry” (and if we are speaking of “symmetry” in the strict sense of an automorphism, then such dualities are called “self-dualities”).

The dualities playing a central role in contemporary physics are of various types: dualities between quantum field theories (like the generalized electric-magnetic duality), between string theories (such as T-duality and S-duality), and between physical descriptions which are, respectively, a quantum field theory and a string theory, as in the case of gauge/gravity dualities (on the various types of dualities and their significance in physics, see Castellani and Rickles (eds.), 2017). Historically, the first

relevant dualities to be used in physics were electric-magnetic duality, momentum-position duality (via Fourier transform) – wave-particle duality in the QM context – and the Kramers-Wannier duality of the two-dimensional Ising model in statistical physics.

In general, because dualities are transformations between theories, their implications are more radical than those of symmetry transformations. While symmetries are mapping between the solutions of the same theory, different theoretical descriptions can have very different interpretations in terms of objects, properties, degrees of freedom, and spacetime frameworks. Thus, dualities naturally offer a new and interesting viewpoint on many traditional issues in the philosophy of science, such as a) reduction, emergence, and fundamentality; b) theoretical equivalence and underdetermination; and c) realism versus anti-realism.

Within the philosophical literature, work relating symmetries to dualities generally responds to one of the following three questions:

- What is an appropriate formal framework for understanding the inter-theoretic relationship realized by duality symmetries? Within the context of quantum gauge/gravity duality, a rudimentary framework for doing so was sketched in de Haro, Teh, and Butterfield (2017) and then further developed in de Haro and Butterfield (2017) for the case of bosonization dualities. And within the far simpler context of classical mechanics, Teh and Tsementzis (2017) explore the use of the quintessential symmetry of classical phase space (viz. symplectomorphisms) to relate Hamiltonian and Lagrangian theories, and proceed to discuss how can use this framework as a toy model for thinking about more sophisticated dualities.
- What is the relationship between the local symmetry of a theory and the symmetries of its dual theory? One reason that this question is pressing is that (as we mentioned in Section 2.5), local symmetries

are typically regard as “surplus” or “redundant” structure of a theory; thus one might hope to be able to construct a dual theory that does not contain such surplus structure. Indeed, it was conjectured in Polchinski and Horowitz (2009) that for gravity/gauge duality, the local symmetries (i.e. the diffeomorphisms) of the bulk theory will always be “invisible” from the perspective of its dual boundary theory. However, de Haro, Teh, and Butterfield (2017) argue that this conjecture is not true in full generality: there is a special class of diffeomorphisms which do not vanish in the boundary theory, but instead correspond to conformal transformations of the boundary CFT.

- Should duality symmetries themselves be understood as a certain kind of “gauge symmetry”, thus explaining why we should treat dual theories as “physically equivalent”? On this question, Read (2016) undertakes the task of using the “hole argument” as a lens with which to compare string-theoretic dualities with (intra-theory) gauge symmetries.

3. Symmetry Arguments

Consider the following cases.

- Buridan’s ass: situated between what are, for him, two completely equivalent bundles of hay, he has no reason to choose the one located to his left over the one located to his right, and so he is not able to choose and dies of starvation.
- Archimedes’s equilibrium law for the balance: if equal weights are hung at equal distances along the arms of a balance, then it will remain in equilibrium since there is no reason for it to rotate one way or the other about the balance point.
- Anaximander’s argument for the immobility of the Earth as reported by Aristotle: the Earth remains at rest since, being at the centre of the

spherical cosmos (and in the same relation to the boundary of the cosmos in every direction), there is no reason why it should move in one direction rather than another.

What do they have in common?

First, these can all be understood as examples of the application of the Leibnizean Principle of Sufficient Reason (PSR): if there is no sufficient reason for one thing to happen instead of another, the principle says that nothing happens (the initial situation does not change). But there is something more that the above cases have in common: in each of them PSR is applied on the grounds that the initial situation has a given symmetry: in the first two cases, bilateral symmetry; in the third, rotational symmetry. The symmetry of the initial situation implies the complete equivalence between the existing alternatives (the left bundle of hay with respect to the right one, and so on). If the alternatives are completely equivalent, then there is no sufficient reason for choosing between them and the initial situation remains unchanged.

Arguments of the above kind — that is, arguments leading to definite conclusions on the basis of an initial symmetry of the situation plus PSR — have been used in science since antiquity (as Anaximander’s argument testifies). The form they most frequently take is the following: a situation with a certain symmetry evolves in such a way that, in the absence of an asymmetric cause, the initial symmetry is preserved. In other words, a breaking of the initial symmetry cannot happen without a reason, or *an asymmetry cannot originate spontaneously*. Van Fraassen (1989) devotes a chapter to considering the way these kinds of symmetry arguments can be used in general problem-solving.

Historically, the first explicit formulation of this kind of argument in terms of symmetry is due to the physicist Pierre Curie towards the end of

nineteenth century. Curie was led to reflect on the question of the relationship between *physical properties* and *symmetry properties* of a physical system by his studies on the thermal, electric and magnetic properties of crystals, these properties being directly related to the structure, and hence the symmetry, of the crystals studied. More precisely, the question he addressed was the following: in a given physical medium (for example, a crystalline medium) having specified symmetry properties, which physical phenomena (for example, which electric and magnetic phenomena) are allowed to happen? His conclusions, systematically presented in his 1894 work “*Sur la symétrie dans les phénomènes physiques*”, can be synthesized as follows:

- a. A phenomenon can exist in a medium possessing its characteristic symmetry or that of one of its subgroups. What is needed for its occurrence (i.e. for something rather than nothing to happen) is not the presence, but rather the absence, of certain symmetries: “Asymmetry is what creates a phenomenon”.
- b. The symmetry elements of the causes must be found in their effects, but the converse is not true; that is, the effects can be more symmetric than the causes.

Conclusion (a) clearly indicates that Curie recognized the important function played by the concept of *symmetry breaking* in physics (he was indeed one of the first to recognize it). Conclusion (b) is what is usually called “Curie’s principle” in the literature, although notice that (a) and (b) are not independent of one another.

In order for Curie’s principle to be applicable, various conditions need to be satisfied: the causal connection must be valid, the cause and effect must be well-defined, and the symmetries of both the cause and the effect must also be well-defined (this involves both the physical and the geometrical properties of the physical systems considered). Curie’s principle then

furnishes a *necessary condition* for given phenomena to happen: only those phenomena can happen that are compatible with the symmetry conditions established by the principle.

Curie's principle has thus an important methodological function: on the one side, it furnishes a kind of selection rule (given an initial situation with a specified symmetry, only certain phenomena are allowed to happen); on the other side, it offers a falsification criterion for physical theories (a violation of Curie's principle may indicate that something is wrong in the physical description).^[15]

Such applications of Curie's principle depend, of course, on our accepting its validity, and this is something that has been questioned in the literature, especially in relation to spontaneous symmetry breaking (see below, next section). Different proposals have been offered for justifying the principle. We have presented it here as an example of symmetry considerations based on Leibniz's PSR, while Curie himself seems to have regarded it as a form of the causality principle. In current literature, it has become standard to understand the principle as following from the invariance properties of deterministic physical laws. According to this "received view", as first formulated in Chalmers (1970) and then developed in more recent literature (Ismael 1997, Belot 2003, Earman 2002), Curie's principle is expressed in terms of the relationship between the symmetries of earlier and later states of a system, and the laws connecting these states. In fact, this is a misrepresentation of Curie's original principle: Curie's focus was on the case of co-existing, functionally related features of a system's state, rather than temporally ordered cause and effect pairs. For a discussion on such questions as to whether there is more than one "Curie's principle" and, independently of how it is formulated, what are the aspects that make it so scientifically fruitful, see Castellani and Ismael, 2016. As regards the status of the principle, Norton (2016) argues that it is a "truism", on the grounds that whether the principle succeeds or fails

depends on how one chooses to attach causal labels to the scientific description. For Roberts (2013), Curie's principle fails when the symmetry is time reversal. Roberts (2016) claims that the truth of the principle is contingent on special physical facts, and attributes to its failure an important role in the detection of parity violation and CP violation (on this point, see also Roberts, 2015, and Ashtekar, 2015).

4. Symmetry Breaking

A symmetry can be exact, approximate, or broken. Exact means unconditionally valid; approximate means valid under certain conditions; broken can mean different things, depending on the object considered and its context.

The study of symmetry breaking also goes back to Pierre Curie. According to Curie, symmetry breaking has the following role: for the occurrence of a phenomenon in a medium, the original symmetry group of the medium must be lowered (broken, in today's terminology) to the symmetry group of the phenomenon (or to a subgroup of the phenomenon's symmetry group) by the action of some cause. In this sense symmetry breaking is what "creates the phenomenon". Generally, the breaking of a certain symmetry does not imply that no symmetry is present, but rather that the situation where this symmetry is broken is characterized by a lower symmetry than the original one. In group-theoretic terms, this means that the initial symmetry group is broken to one of its subgroups. It is therefore possible to describe symmetry breaking in terms of relations between transformation groups, in particular between a group (the unbroken symmetry group) and its subgroup(s). As is clearly illustrated in the 1992 volume by I. Stewart and M. Golubitsky, starting from this point of view a general theory of symmetry breaking can be developed by tackling such questions as "which subgroups can occur?", "when does a given subgroup occur?"

Symmetry breaking was first explicitly studied in physics with respect to physical objects and phenomena. This is not surprising, since the theory of symmetry originated with the visible symmetry properties of familiar spatial figures and every day objects. However, it is with respect to the laws that symmetry breaking has acquired special significance in physics. There are two different types of symmetry breaking of the laws: “explicit” and “spontaneous”, the case of spontaneous symmetry breaking being the more interesting from a physical as well as a philosophical point of view.

4.1 Explicit symmetry breaking

Explicit symmetry breaking indicates a situation where the dynamical equations are not manifestly invariant under the symmetry group considered. This means, in the Lagrangian (Hamiltonian) formulation, that the Lagrangian (Hamiltonian) of the system contains one or more terms explicitly breaking the symmetry. Such terms can have different origins:

(a) Symmetry-breaking terms may be introduced into the theory by hand on the basis of theoretical/experimental results, as in the case of the quantum field theory of the weak interactions, which is expressly constructed in a way that manifestly violates mirror symmetry or parity. The underlying result, in this case, is parity non-conservation in the case of the weak interaction, first predicted in the famous (Nobel-prize winning) 1956 paper by T. D. Lee and C.N. Yang.

(b) Symmetry-breaking terms may appear in the theory because of quantum-mechanical effects. One reason for the presence of such terms — known as “anomalies” — is that in passing from the classical to the quantum level, because of possible operator ordering ambiguities for composite quantities such as Noether charges and currents, it may be that the classical symmetry algebra (generated through the Poisson bracket structure) is no longer realized in terms of the commutation relations of

the Noether charges. Moreover, the use of a “regulator” (or “cut-off”) required in the renormalization procedure to achieve actual calculations may itself be a source of anomalies. It may violate a symmetry of the theory, and traces of this symmetry breaking may remain even after the regulator is removed at the end of the calculations. Historically, the first example of an anomaly arising from renormalization is the so-called chiral anomaly, that is the anomaly violating the chiral symmetry of the strong interaction (see Weinberg, 1996, Chapter 22).

(c) Finally, symmetry-breaking terms may appear because of non-renormalizable effects. Physicists now have good reasons for viewing current renormalizable field theories as *effective field theories*, that is low-energy approximations to a deeper theory (each effective theory explicitly referring only to those particles that are of importance at the range of energies considered). The effects of non-renormalizable interactions (due to the heavy particles not included in the theory) are small and can therefore be ignored at the low-energy regime. It may then happen that the coarse-grained description thus obtained possesses more symmetries than the deeper theory. That is, the effective Lagrangian obeys symmetries that are not symmetries of the underlying theory. These “accidental” symmetries, as Weinberg has called them, may then be violated by the non-renormalizable terms arising from higher mass scales and suppressed in the effective Lagrangian (see Weinberg, 1995, pp. 529–531).

4.2 Spontaneous symmetry breaking

Spontaneous symmetry breaking (SSB) occurs in a situation where, given a symmetry of the equations of motion, solutions exist which are not invariant under the action of this symmetry without any explicit asymmetric input (whence the attribute “spontaneous”).^[16] A situation of this type can be first illustrated by means of simple cases taken from classical physics. Consider for example the case of a linear vertical stick

with a compression force applied on the top and directed along its axis. The physical description is obviously invariant for all rotations around this axis. As long as the applied force is mild enough, the stick does not bend and the equilibrium configuration (the lowest energy configuration) is invariant under this symmetry. When the force reaches a critical value, the symmetric equilibrium configuration becomes unstable and an infinite number of equivalent lowest energy stable states appear, which are no longer rotationally symmetric but are related to each other by a rotation. The actual breaking of the symmetry may then easily occur by effect of a (however small) external asymmetric cause, and the stick bends until it reaches one of the infinite possible stable asymmetric equilibrium configurations.^[17] In substance, what happens in the above kind of situation is the following: when some parameter reaches a critical value, the lowest energy solution respecting the symmetry of the theory ceases to be stable under small perturbations and new asymmetric (but stable) lowest energy solutions appear. The new lowest energy solutions are asymmetric but are all related through the action of the symmetry transformations. In other words, there is a degeneracy (infinite or finite depending on whether the symmetry is continuous or discrete) of distinct asymmetric solutions of identical (lowest) energy, the whole set of which maintains the symmetry of the theory.

In *quantum physics* SSB actually does not occur in the case of finite systems: tunnelling takes place between the various degenerate states, and the true lowest energy state or “ground state” turns out to be a unique linear superposition of the degenerate states. In fact, SSB is applicable only to infinite systems — many-body systems (such as ferromagnets, superfluids and superconductors) and fields — the alternative degenerate ground states being all orthogonal to each other in the infinite volume limit and therefore separated by a “superselection rule” (see for example Weinberg, 1996, pp. 164–165).

Historically, the concept of SSB first emerged in condensed matter physics. The prototype case is the 1928 Heisenberg theory of the ferromagnet as an infinite array of spin 1/2 magnetic dipoles, with spin-spin interactions between nearest neighbours such that neighbouring dipoles tend to align. Although the theory is rotationally invariant, below the critical Curie temperature T_c the actual ground state of the ferromagnet has the spin all aligned in some particular direction (i.e. a magnetization pointing in that direction), thus not respecting the rotational symmetry. What happens is that below T_c there exists an infinitely degenerate set of ground states, in each of which the spins are all aligned in a given direction. A complete set of quantum states can be built upon each ground state. We thus have many different “possible worlds” (sets of solutions to the same equations), each one built on one of the possible orthogonal (in the infinite volume limit) ground states. To use a famous image by S. Coleman, a little man living inside one of these possible asymmetric worlds would have a hard time detecting the rotational symmetry of the laws of nature (all his experiments being under the effect of the background magnetic field). The symmetry is still there — the Hamiltonian being rotationally invariant — but “hidden” to the little man. Besides, there would be no way for the little man to detect directly that the ground state of his world is part of an infinitely degenerate multiplet. To go from one ground state of the infinite ferromagnet to another would require changing the directions of an infinite number of dipoles, an impossible task for the finite little man (Coleman, 1975, pp. 141–142). As said, in the infinite volume limit all ground states are separated by a superselection rule. Ruetsche (2006) discusses symmetry breaking and ferromagnetism from the algebraic perspective. Liu and Emch (2005) address the interpretative problems of explaining SSB in nonrelativistic quantum statistical mechanics. Fraser (2016) discusses SBB in finite systems, arguing against the indispensability of the thermodynamic limit in the characterization of SSB in statistical mechanics.

The same picture can be generalized to *quantum field theory* (QFT), the ground state becoming the *vacuum state*, and the role of the little man being played by ourselves. This means that there may exist symmetries of the laws of nature which are not manifest to us because the physical world in which we live is built on a vacuum state which is not invariant under them. In other words, the physical world of our experience can appear to us very asymmetric, but this does not necessarily mean that this asymmetry belongs to the fundamental laws of nature. SSB offers a key for understanding (and utilizing) this physical possibility.

The concept of SSB was transferred from condensed matter physics to QFT in the early 1960s, thanks especially to works by Y. Nambu and G. Jona-Lasinio. Jona-Lasinio (2003) offers a first-hand account of how the idea of SSB was introduced and formalized in particle physics on the grounds of an analogy with the breaking of (electromagnetic) gauge symmetry in the 1957 theory of superconductivity by J. Bardeen, L. N. Cooper and J. R. Schrieffer (the so-called BCS theory). The application of SSB to particle physics in the 1960s and successive years led to profound physical consequences and played a fundamental role in the edification of the current Standard Model of elementary particles. In particular, let us mention the following main results that obtain in the case of the spontaneous breaking of a continuous internal symmetry in QFT.

Goldstone theorem. In the case of a global continuous symmetry, massless bosons (known as “Goldstone bosons”) appear with the spontaneous breakdown of the symmetry according to a theorem first stated by J. Goldstone in 1960. The presence of these massless bosons, first seen as a serious problem since no particles of the sort had been observed in the context considered, was in fact the basis for the solution — by means of the so-called Higgs mechanism (see the next point) — of another similar problem, that is the fact that the 1954 Yang-Mills theory of non-Abelian gauge fields predicted unobservable massless particles, the gauge bosons.

Higgs mechanism. According to a “mechanism” established in a general way in 1964 independently by (i) P. Higgs, (ii) R. Brout and F. Englert, and (iii) G. S. Guralnik, C. R. Hagen and T. W. B. Kibble, in the case that the internal symmetry is promoted to a local one, the Goldstone bosons “disappear” and the gauge bosons acquire a mass. The Goldstone bosons are “eaten up” to give mass to the gauge bosons, and this happens without (explicitly) breaking the gauge invariance of the theory. Note that this mechanism for the mass generation for the gauge fields is also what ensures the *renormalizability* of theories involving massive gauge fields (such as the Glashow-Weinberg-Salam electroweak theory developed in the second half of the 1960s), as first generally demonstrated by M. Veltman and G. ’t Hooft in the early 1970s. (The Higgs mechanism is at the center of a lively debate among philosophers of physics: see, for example, Smeenk, 2006; Lyre, 2008; Struyve, 2011; Friederich, 2013. For a historical-philosophical analysis, see also Borrelli, 2012.)

Dynamical symmetry breaking (DSB). In such theories as the unified model of electroweak interactions, the SSB responsible (via the Higgs mechanism) for the masses of the gauge vector bosons is because of the symmetry-violating vacuum expectation values of scalar fields (the so-called Higgs fields) introduced *ad hoc* in the theory. For different reasons — first of all, the initially *ad hoc* character of these scalar fields for which there was no experimental evidence until the results obtained in July 2012 at the LHC — some attention has been drawn to the possibility that the Higgs fields could be phenomenological rather than fundamental, that is bound states resulting from a specified dynamical mechanism. SSB realized in this way has been called “DSB”.^[18]

Symmetry breaking raises a number of *philosophical* issues. Some of them relate only to the breaking of specific types of symmetries, such as the issue of the significance of parity violation for the problem of the nature of space (see Section 2.4, above). Others, for example the connection

between symmetry breaking and observability, are particular aspects of the general issue concerning the status and significance of physical symmetries, but in the case of SSB they take on a stronger force: what is the epistemological status of a theory based on “hidden” symmetries and SSB? Given that what we directly observe — the physical situation, the phenomenon — is asymmetric, what is the evidence for the “underlying” symmetry? On this point, see for example Morrison (2003) and Kosso (2000). In the absence of direct empirical evidence, the above question then becomes whether and how far the predictive and explanatory power of theories based on SSB provides good reasons for believing in the existence of the hidden symmetries. Finally, there are issues raised by the motivation for, and role of, SSB. See for example Earman (2003a), using the algebraic formulation of QFT to explain SSB; for further philosophical discussions on SSB in QFT in the algebraic approach, see Ruetsche (2011), Fraser (2012), and references therein. Landsman (2013) discusses the issue whether SSB in infinite quantum systems can be seen as an example of asymptotic emergence in physics.

SSB allows symmetric theories to describe asymmetric reality. In short, SSB provides a way of understanding the complexity of nature without renouncing fundamental symmetries. But why should we prefer symmetric to asymmetric fundamental laws? In other words, why assume that an observed asymmetry requires a cause, which can be an explicit breaking of the symmetry of the laws, asymmetric initial conditions, or SSB? Note that this assumption is very similar to the one expressed by Curie in his famous 1894 paper. Curie’s principle (the symmetries of the causes must be found in the effects; or, equivalently, the asymmetries of the effects must be found in the causes), when extended to include the case of SSB, is equivalent to a methodological principle according to which an asymmetry of the phenomena must come from the breaking (explicit or spontaneous) of the symmetry of the fundamental laws. What the real nature of this

principle is remains an open issue, at the centre of a developing debate (see Section 3, above).

Finally, let us mention the argument that is sometimes made in the literature that SSB implies that Curie’s principle is violated because a symmetry is broken “spontaneously”, that is without the presence of any asymmetric cause. Now it is true that SSB indicates a situation where solutions exist that are not invariant under the symmetry of the law (dynamical equation) without any explicit breaking of this symmetry. But, as we have seen, the symmetry of the “cause” is not lost, it is conserved in the ensemble of the solutions (the whole “effect”).^[19]

5. General Philosophical Questions

Much of the recent philosophical literature on symmetries in physics discusses specific symmetries and the interpretational questions they lead to. The rich variety of symmetries in modern physics means that questions concerning the status and significance of symmetries in physics in general are not easily addressed. However, something can be said in more general terms and we offer a few remarks in that direction here, starting with the main roles that symmetry plays in physics.

One of the most important roles played by symmetry is that of *classification* — for example, the classification of crystals using their remarkable and varied symmetry properties. In contemporary physics, the best example of this role of symmetry is the classification of elementary particles by means of the irreducible representations of the fundamental physical symmetry groups, a result first obtained by Wigner in his famous paper of 1939 on the unitary representations of the inhomogeneous Lorentz group. When a symmetry classification includes all the necessary properties for characterizing a given type of physical object (for example, all necessary quantum numbers for characterizing a given type of particle),

we have the possibility of defining types of entities on the basis of their transformation properties. This has led philosophers of science to explore a structuralist approach to the entities of modern physics, in particular a group-theoretical account of objects (see the contributions in Castellani, 1998, Part II). The relation between symmetry, group and structure and its exploitation in structuralist approaches, especially in the framework of structural realism, is a much debated issue. See, for example, Roberts (2011), French (2014) and Caulton (2015), and the SEP entry Structural Realism.

Symmetries also have a *normative* role, being used as constraints on physical theories. The requirement of invariance with respect to a transformation group imposes severe restrictions on the form that a theory may take, limiting the types of quantities that may appear in the theory as well as the form of its fundamental equations. A famous case is Einstein's use of general covariance when searching for his gravitational equations.

The group-theoretical treatment of physical symmetries, with the resulting possibility of unifying different types of symmetries by means of a unification of the corresponding transformation groups, has provided the technical resources for symmetry to play a powerful role in theoretical *unification*. This is best illustrated by the current dominant research programme in theoretical physics aimed at arriving at a unified description of all the fundamental forces of nature (gravitational, weak, electromagnetic and strong) in terms of underlying local symmetry groups.

It is often said that many physical phenomena can be explained as (more or less direct) consequences of symmetry principles or symmetry arguments. In the case of symmetry principles, the *explanatory* role of symmetries arises from their place in the hierarchy of the structure of physical theory, which in turn derives from their generality. As Wigner (1967, pp. 28ff) describes the hierarchy, symmetries are seen as properties

of the laws. Symmetries may be used to explain (i) the form of the laws, and (ii) the occurrence (or non-occurrence) of certain events (this latter in a manner analogous to the way in which the laws explain why certain events occur and not others). In the case of symmetry arguments, we may, for example, appeal to Curie's principle to explain the occurrence of certain phenomena on the basis of the symmetries (or asymmetries) of the situation, as discussed in section 3, above. Furthermore, insofar as explanatory power may be derived from unification, the unifying role of symmetries also results in an explanatory role.

From these different roles we can draw some preliminary conclusions about the status of symmetries. It is immediately apparent that symmetries have an important heuristic function, indicating a strong methodological status. Is this methodological power connected to an ontological or epistemological status for symmetries?

According to an *ontological* viewpoint, symmetries are seen as a substantial part of the physical world: the symmetries of theories represent properties existing in nature, or characterize the structure of the physical world. It might be claimed that the ontological status of symmetries provides the reason for the methodological success of symmetries in physics. A concrete example is the use of symmetries to predict the existence of new particles. This can happen via the classificatory role, on the grounds of vacant places in symmetry classification schemes, as in the famous case of the 1962 prediction of the particle Ω^- in the context of the hadronic classification scheme known as the "Eightfold Way". (See Bangu, 2008, for a critical analysis of the reasoning leading to this prediction.) Or, as in more recent cases, via the unificatory role: the paradigmatic example is the prediction of the W and Z particles (experimentally found in 1983) in the context of the Glashow-Weinberg-Salam gauge theory proposed in 1967 for the unification of the weak and electromagnetic interactions. These impressive cases of the prediction of

new phenomena might then be used to argue for an ontological status for symmetries, via an inference to the best explanation.

Another reason for attributing symmetries to nature is the so-called geometrical interpretation of spatiotemporal symmetries, according to which the spatiotemporal symmetries of physical laws are interpreted as symmetries of spacetime itself, the “geometrical structure” of the physical world. Moreover, this way of seeing symmetries can be extended to non-external symmetries, by considering them as properties of other kinds of spaces, usually known as “internal spaces”. The question of exactly what a realist would be committed to on such a view of internal spaces remains open, and an interesting topic for discussion.

One approach to investigating the limits of an ontological stance with respect to symmetries would be to investigate their empirical or observational status: can the symmetries in question be directly observed? We first have to address what it means for a symmetry to be observable, and indeed whether all symmetries have the same observational status. Kosso (2000) arrives at the conclusion that there are important differences in the empirical status of the different kinds of symmetries. In particular, while global continuous symmetries can be directly observed — via such experiments as the Galilean ship experiment — a local continuous symmetry can have only indirect empirical evidence. Brading and Brown (2004) argue for a different interpretation of Kosso’s examples,^[20] but agree with Kosso’s assessment that local symmetry transformations cannot exhibit an analogue of the Galileo’s Ship Experiment. Against this view, Greaves and Wallace (2014) and Teh (2016) have recently argued that when suitably understood, local symmetries can indeed have direct empirical significance (their argument turns on the possibility of using asymptotically non-trivial local symmetries to construct Galileo’s Ship scenarios for certain systems). In particular, Teh (2016) argues that this phenomenon is responsible for the empirically significant symmetries of

topological soliton solutions. On the other hand, Friederich (2015) contends that on one plausible axiomatization of the schema introduced by Greaves and Wallace, it is possible to deduce that local symmetries do not have direct empirical significance.

The direct observational status of the familiar global spacetime symmetries leads us to an *epistemological* aspect of symmetries. According to Wigner, the spatiotemporal invariance principles play the role of a prerequisite for the very possibility of discovering the laws of nature: “if the correlations between events changed from day to day, and would be different for different points of space, it would be impossible to discover them” (Wigner, 1967, p. 29). For Wigner, this conception of symmetry principles is essentially related to our ignorance (if we could directly know all the laws of nature, we would not need to use symmetry principles in our search for them). Others, on the contrary, have arrived at a view according to which symmetry principles function as “transcendental principles” in the Kantian sense (see for instance Mainzer, 1996). It should be noted in this regard that Wigner’s starting point, as quoted above, does not imply exact symmetries — all that is needed epistemologically is that the global symmetries hold approximately, for suitable spatiotemporal regions, such that there is sufficient stability and regularity in the events for the laws of nature to be discovered.

There is another reason why symmetries might be seen as being primarily epistemological. As we have mentioned, there is a close connection between the notions of symmetry and equivalence, and this leads also to a notion of irrelevance: the equivalence of space points (translational symmetry), for example, may be understood in the sense of the irrelevance of an absolute position to the physical description. There are two ways that one might interpret the epistemological significance of this: on the one hand, we might say that symmetries are associated with unavoidable redundancy in our *descriptions* of the world, while on the other hand we

might maintain that symmetries indicate a limitation of our epistemic access — there are certain properties of objects, such as their absolute positions, that are not observable. The view that symmetries are connected with the presence of non-observable quantities in the physical description, with the corollary that the empirical violation of a symmetry is intended in the sense that what was thought to be a non-observable turns out to be actually an observable, was particularly defended by one of the discover of parity violation, T.D. Lee (see section 2.4). According to Lee (1981), “the root of all symmetry principles lies in the assumption that it is impossible to observe certain basic quantities” (p. 178). See on this (and, more generally, on the relation between symmetry, equivalence and irrelevance) Castellani (2003). Dasgupta (2016) defends an epistemic interpretation of symmetry on a similar basis as Lee.

Finally, we would like to mention an aspect of symmetry that might very naturally be used to support either an ontological or an epistemological account. It is widely agreed that there is a close connection between symmetry and *objectivity*, the starting point once again being provided by spacetime symmetries: the laws by means of which we describe the evolution of physical systems have an objective validity because they are the same for all observers. The old and natural idea that what is objective should not depend upon the particular perspective under which it is taken into consideration is thus reformulated in the following group-theoretical terms: what is objective is what is invariant with respect to the transformation group of reference frames, or, quoting Hermann Weyl (1952, p. 132), “objectivity means invariance with respect to the group of automorphisms [of space-time]”.^[21] Debs and Redhead (2007) label as “invariantism” the view that “invariance under a specified group of automorphisms is both a necessary and sufficient condition for objectivity” (p. 60). They point out (p. 73, and see also p. 66) that there is a natural connection between “invariantism” and structural realism.

Growing interest, recently, in the metaphysics of physics includes interest in symmetries. Baker (2010) offers an accessible introduction, and Livanios (2010), connecting discussions of symmetries to dispositions and essences, is an example of this work.

To conclude: symmetries in physics offer many interpretational possibilities, and how to understand the status and significance of physical symmetries clearly presents a challenge to both physicists and philosophers.

Bibliography

- Anderson, J. L., 1967, *Principles of Relativity Physics*, New York: Academic Press.
- Ashtekar, A., 2015, “Response to Bryan Roberts: A new perspective on T violation”, *Studies in History and Philosophy of Modern Physics*, 52(A): 16–20.
- Baker, D., 2010, “Symmetry and the Metaphysics of Physics”, *Philosophy Compass*, 5(12): 157–1166.
- Bangu, S., 2008, “Reifying mathematics? Prediction and symmetry classification”, *Studies in History and Philosophy of Modern Physics*, 39: 239–258.
- Belot, G., 1998, “Understanding electromagnetism”, *British Journal for the Philosophy of Science*, 49: 531–555.
- , 2003, “Notes on symmetries”, in K. Brading and E. Castellani (eds.), *Symmetries in Physics: Philosophical Reflections*, Cambridge: Cambridge University Press, pp. 391–409.
- , 2008, “An Elementary Notion of Gauge Equivalence”, *General Relativity and Gravitation*, 40: 199–215.
- , 2012, “Symmetry and equivalence”, in R. Batterman (ed.), *The Oxford Handbook of Philosophy of Physics*, New York: Oxford University Press, Chap. 9.

- Born, M., 1953, “Physical Reality”, *Philosophical Quarterly*, 3: 139–149.
Reprinted in E. Castellani (ed.), *Interpreting Bodies: Classical and Quantum Objects in Modern Physics*, Princeton, NJ: Princeton University Press, 1998, pp. 155–167.
- Borrelli, A., 2012, “The case of the composite Higgs: The model as a “Rosetta stone” in contemporary high-energy physics”, *Studies in History and Philosophy of Modern Physics*, 43: 195–214.
- Brading, K. A., 2002, “Which symmetry? Noether, Weyl and conservation of electric charge”, *Studies in History and Philosophy of Modern Physics*, 33: 3–22.
- , 2010, “Mathematical and aesthetic aspects of symmetry” (rev. of Hon and Goldstein, 2008), *Metascience*, 19: 277–280.
- Brading, K. and Brown, H. R., 2003, “Symmetries and Noether’s theorems”, in K. Brading and E. Castellani (eds.), *Symmetries in Physics: Philosophical Reflections*, Cambridge: Cambridge University Press, pp. 89–109.
- , 2004, “Are gauge symmetry transformations observable?”, *British Journal for the Philosophy of Science*, 55: 645–665.
- Brading, K. and Castellani, E., 2007, “Symmetries and Invariances in Classical Physics”, in J. Butterfield and J. Earman (eds.), *Handbook of the Philosophy of Science. Philosophy of Physics*, Amsterdam: North Holland, Elsevier, 1331–1367.
- Brading, K. and Castellani, E. (eds.), 2003, *Symmetries in Physics: Philosophical Reflections*, Cambridge: Cambridge University Press.
- Brown, H. R. and Brading, K., 2002, “General covariance from the perspective of Noether’s theorems”, *Diálogos*, 79: 59–86.
- Castellani, E. (ed.), 1998, *Interpreting Bodies: Classical and Quantum Objects in Modern Physics*, Princeton, NJ: Princeton University Press.
- Castellani, E., 2000, *Simmetria e natura*, Roma-Bari: Laterza.
- , 2003, “Symmetry and equivalence”, in K. Brading and E. Castellani (eds.), *Symmetries in Physics: Philosophical Reflections*, Cambridge: Cambridge University Press, pp. 422–433.
- , 2004, “Dirac on Gauges and Constraints”, *International Journal of Theoretical Physics*, 43(5): 1503–1514.
- Castellani, E., and Ismael, J., 2016, “Which Curie’s Principle? ”, *Philosophy of Science*, 83(5): 1002–1013.
- Castellani, E. and Rickles, D. (eds.), 1998, *Dualities in Physics*, Special issue of *Studies in History and Philosophy of Modern Physics*, 59: 1–142
- Catren, G., 2008, “Geometric Foundations of Classical Yang-Mills Theory”, *Studies in History and Philosophy of Modern Physics*, 39: 511–531.
- Caulton, A., 2015, “The role of symmetry in the interpretation of physical theories”, *Studies in History and Philosophy of Modern Physics*, 52(B): 153–162.
- Caulton, A. and Butterfield, J. 2012, “On Kinds of Indiscernibility in Logic and Metaphysics”, *British Journal for the Philosophy of Science*, 63: 233–285.
- Chalmers, A. F., 1970, “Curie’s principle”, *British Journal for the Philosophy of Science*, 21: 133–148.
- Coleman, S., 1975, “Secret symmetry: an introduction to spontaneous symmetry breakdown and gauge fields”, in A. Zichichi (ed.), *Laws of hadronic matter*, New York: Academic Press, pp. 138–215.
- CPLEAR Collaboration, 1998, “First direct observation of time-reversal non-invariance in the neutral-kaon system”, *Physics Letters B*, 444: 43–5.
- Curie, P., 1894, “Sur la symétrie dans les phénomènes physiques. Symétrie d’ un champ électrique et d’ un champ magnétique”, *Journal de Physique*, 3rd series, vol.3, 393–417.
- Dasgupta, S., 2016, “Symmetry as an Epistemic Notion (Twice Over)”, *British Journal for the Philosophy of Science*, 67: 837–878.
- Debs, T. and Redhead, M., 2007, *Objectivity, Invariance, and Convention*:

- Symmetry in Physical Science*, Cambridge MA: Harvard University Press.
- De Haro, S., Teh, N., and Butterfield, J., 2017, “Comparing Dualities and Gauge Symmetries”, *Studies in History and Philosophy of Modern Physics*, 59: 66–80.
- Dirac, P. A. M., 1931, “Quantized Singularities in the Electromagnetic Field”, *Proceedings of the Royal Society*, A 133: 60–72.
- Earman, J., 1989, *World enough and spacetime*, Cambridge, MA: MIT Press.
- , 2002, “Laws, Symmetry, and Symmetry Breaking; Invariance, Conservation Principles, and Objectivity” (PSA 2002 Presidential Address), *Philosophy of Science*, 71 (2004): 1227–1241.
- , 2003a, “Rough guide to spontaneous symmetry breaking”, in K. Brading and E. Castellani (eds.), *Symmetries in Physics: Philosophical Reflections*, Cambridge: Cambridge University Press, pp. 334–345.
- , 2003b, “Tracking down gauge: an ode to the constrained Hamiltonian formalism”, in K. Brading and E. Castellani (eds.), *Symmetries in Physics: Philosophical Reflections*, Cambridge: Cambridge University Press, pp. 140–162.
- , 2004, “Curie’s Principle and Spontaneous Symmetry Breaking”, *International Studies in the Philosophy of Science*, 18: 173–198.
- Earman, J. and Norton, J., 1987, “What price spacetime substantivalism? The hole story”, *British Journal for the Philosophy of Science*, 38: 515–525.
- Fraser, D., 2012, “Spontaneous Symmetry Breaking: Quantum Statistical Mechanics versus Quantum Field Theory”, *Philosophy of Science*, 79(5): 905–916.
- Fraser, J. D., 2016, “Spontaneous Symmetry Breaking in Finite Systems”, *Philosophy of Science*, 83(4): 585–605.
- French, S. and Rickles, D., 2003, “Understanding permutation symmetry”,





- in K. Brading and E. Castellani (eds.), *Symmetries in Physics: Philosophical Reflections*, Cambridge: Cambridge University Press, pp. 212–138.
- French, S. and Krause, D., 2006, *Identity in Physics: A Formal, Historical and Philosophical Approach*, Oxford: Oxford University Press.
- French, S., 2014, *The Structure of the World*, Oxford: Oxford University Press.
- Friederich, S., 2013 “Gauge Symmetry Breaking in Gauge Theories—In Search of Clarification”, *European Journal for Philosophy of Science*, 3(2): 157–182.
- Friederich, F., 2015, “Symmetry, Empirical Equivalence, and Identity”, *British Journal for the Philosophy of Science*, 66(3): 537–559.
- Gołosz, J., 2016, “Weak Interactions: Asymmetry of Time or Asymmetry in Time?”, *Journal for General Philosophy of Science*, 48: 19–33.
- Greaves, H., 2010, “Towards a Geometrical Understanding of the CPT Theorem”, *British Journal for the Philosophy of Science*, 61: 27–50.
- Greaves, H. and Thomas, T., 2014, “On the CPT theorem”, *Studies in History and Philosophy of Modern Physics*, 45: 46–65.
- Greaves, H. and Wallace, D., 2016, “Empirical Consequences of Symmetries”, *British Journal for the Philosophy of Science* 65(1) : 59–89.
- Guay, A. and Hepburn, B., 2009, “Symmetry and Its Formalisms: Mathematical Aspects”, *Philosophy of Science*, 76: 160–178.
- Healey, R., 2001, “On the Reality of Gauge Potentials”, *Philosophy of Science*, 68: 432–455.
- , 2007, *Gauging What’s Real*, Oxford: Oxford University Press.
- , 2009, “Perfect symmetries”, *British Journal for the Philosophy of Science*, 60: 697–720.
- Heisenberg, W., 1926, “Mehrkörperprobleme und Resonanz in der Quantenmechanik”, *Zeitschrift für Physik*, 38: 411–426.
- , 1932, “Über den Bau der Atomkerne. I”, *Zeitschrift für Physik*, 77:

- 1–11.
- Hon, G. and Goldstein, B. R., 2008, *From Summetria to Symmetry: The Making of a Revolutionary Scientific Concept*, London: Springer.
- Ismael, J., 1997, “Curie’s principle”, *Synthese*, 110: 167–190.
- Jona-Lasinio, G., 2003, “Cross fertilization in theoretical physics”, in K. Brading and E. Castellani (eds.), *Symmetries in Physics: Philosophical Reflections*, Cambridge: Cambridge University Press, pp. 315–320.
- Kosso, P., 2000, “The empirical status of symmetries in physics”, *British Journal for the Philosophy of Science*, 51: 81–98.
- , 2003, “Symmetry, objectivity, and design”, in K. Brading and E. Castellani (eds.), *Symmetries in Physics: Philosophical Reflections*, Cambridge: Cambridge University Press, pp. 410–421.
- Ladyman, J. and Bigaj, T., 2010, “The Principle of the Identity of Indiscernibles and Quantum Mechanics”, *Philosophy of Science*, 77: 117–136.
- Lanczos, C., 1949, *The variational principles of mechanics*, Toronto: University of Toronto Press.
- Landsman, N. P., 2013, “Spontaneous Symmetry Breaking in Quantum Systems: Emergence or Reduction?”, *Studies in History and Philosophy of Modern Physics*, 44: 379–394.
- Lee, T. D. and Yang, C. N., 1956, “Questions of Parity Conservations in Weak Interactions”, *Physical Review*, 104: 254–258.
- Lee, T. D., 1981, *Particle Physics and Introduction to Field Theory*, New York: Harwood Academic Publishers.
- Liu, C., 2003, “Spontaneous symmetry breaking and chance in a classical world”, *Philosophy of Science*, 70: 590–608.
- Liu, C. and Emch, G. G., 2005, “Explaining quantum spontaneous symmetry breaking”, *Studies in History and Philosophy of Modern Physics*, 36: 137–163.
- Livanios, V., 2003, “Symmetries, dispositions and essences”, *Philosophical Studies*, 148: 295–305.
- Lyre, H., 2008, “Does the Higgs Mechanism Exist?”, *International Studies in the Philosophy of Science*, 22: 119–133.
- Mainzer, K., 1996, *Symmetries of nature*, Berlin: Walter de Gruyter.
- Martin, C., 2003, “On continuous symmetries and the foundations of modern physics”, in K. Brading and E. Castellani (eds.), *Symmetries in Physics: Philosophical Reflections*, Cambridge: Cambridge University Press, pp. 29–60.
- Morrison, M., 2003, “Spontaneous symmetry breaking: theoretical arguments and philosophical problems”, in K. Brading and E. Castellani (eds.), *Symmetries in Physics: Philosophical Reflections*, Cambridge: Cambridge University Press, pp. 346–362.
- Nguyen, J., Teh, N., and Wells, L., forthcoming, “Why surplus structure is not superfluous”, *British Journal for the Philosophy of Science*.
- Norton, J., 1989, “Coordinates and Covariance: Einstein’s view of space-time and the modern view”, *Foundations of Physics*, 19: 1215–1263.
- , 1993, “General covariance and the foundations of general relativity: eight decades of dispute”, *Rep. Prog. Phys.*, 56: 791–858.
- , 2003, “General covariance, gauge theories, and the Kretschmann objection”, in K. Brading and E. Castellani (eds.), *Symmetries in Physics: Philosophical Reflections*, Cambridge: Cambridge University Press, pp. 110–123.
- , 2016, “Curie’s Truism”, *Philosophy of Science*, 83(5): 1014–1026.
- Nounou, A., 2003, “A fourth way to the Aharonov-Bohm effect”, in K. Brading and E. Castellani (eds.), *Symmetries in Physics: Philosophical Reflections*, Cambridge: Cambridge University Press, pp. 174–200.
- Olver, P. J., 1995, *Equivalence, Invariants, and Symmetry*, Cambridge: Cambridge University Press.
- Pitts, B., 2006, “Absolute Objects and Counterexamples: Jones-Geroch Dust, Torretti Constant Curvature, Tetrad-Spinor, and Scalar

- Density”, *Studies in History and Philosophy of Modern Physics*, 37: 347–371.
- Pooley, O., 2003, “Handedness, parity violation, and the reality of space”, in K. Brading and E. Castellani (eds.), *Symmetries in Physics: Philosophical Reflections*, Cambridge: Cambridge University Press, pp. 250–280.
- Pooley, O., 2017, “Background Independence, Diffeomorphism Invariance, and the Meaning of Coordinates”, in D. Lehmkuhl, G. Schieman, and E. Scholz (eds.), *Towards a Theory of Spacetime Theories* (Einstein Studies: Volume 13), Birkhauser, pp. 105–143.
- O’Raifeartaigh, L., 1997, *The dawning of gauge theory*, Princeton, NJ: Princeton University Press.
- Read, J., 2016, “The Interpretation of Sting-theoretic Dualities”, *Foundations of Physics*, 46: 209–235.
- Redhead, M., 2003, “The interpretation of gauge symmetry”, in M. Kuhlmann, H. Lyre, and A. Wayne (eds.), *Ontological Aspects of Quantum Field Theory*, Singapore: World Scientific. Reprinted in K. Brading and E. Castellani (eds.), *Symmetries in Physics: Philosophical Reflections*, Cambridge: Cambridge University Press, pp. 124–139.
- Rickles, D., Smeenk, C., Lyre, H. and Healey, R., 2009, “Gauge Pressure” (rev. symposium of Healey, 2007), *Metascience*, 18: 5–41.
- Roberts, B. W., 2011, “Group Structure Realism”, *British Journal for the Philosophy of Science*, 62(1): 47–69.
- , 2013, “The Simple Failure of Curie’s Principle”, *Philosophy of Science*, 80: 579–592.
- , 2015, “Three merry roads to T violation”, *Studies in History and Philosophy of Modern Physics*, 52(A): 8–15.
- , 2016, “Curie’s hazard: From electromagnetism to symmetry violation”, *Erkenntnis*, 81(5): 1011–1029.
- , 2017, “Three myths about time reversal in quantum theory”, *Philosophy of Science*, 84(2): 315–334.
- Ryckman, T. A., 2003, “The philosophical roots of the gauge principle: Weyl and transcendental phenomenological idealism”, in K. Brading and E. Castellani (eds.), *Symmetries in Physics: Philosophical Reflections*, Cambridge: Cambridge University Press, pp. 61–88.
- Rynasiewicz, R., 1999, “Kretschmann’s analysis of covariance and relativity principles”, in H. Goenner, J. Renn, J. Ritter and T. Sauer (eds.), *The expanding worlds of general relativity (Einstein Studies 7)*, The Centre for Einstein Studies, Boston: Birkhauser, 431–462.
- Ruetsche, L., 2006, “Johnny’s So Long at the Ferromagnet”, *Philosophy of Science*, 73: 473–486.
- , 2011, *Interpreting quantum theories*, New York: Oxford University Press.
- Saunders, S., 2002, “Indiscernibles, general covariance, and other symmetries”, in A. Ashtekar, D. Howard, J. Renn, S. Sarkar, and A. Shimony (eds.), *Revisiting the Foundations of Relativistic Physics: Festschrift in Honour of John Stachel*, Dordrecht: Kluwer, pp. 151–173.
- , 2006, “Are quantum particles objects?”, *Analysis*, 66: 52–63.
- , 2007, “Mirroring as an A Priori Symmetry”, *Philosophy of Science*, 74: 452–480.
- Shubnikov, A. V. and Koptsik, V. A., 1974, *Symmetry in science and art*, London: Plenum Press.
- Smeenk, S., 2006, “The Elusive Higgs Mechanism”, *Philosophy of Science*, 73: pp. 487–499.
- Stewart, I. and Golubitsky, M., 1992, *Fearful symmetry. Is God a geometer?*, Oxford: Blackwell.
- Strocchi, F., 2008, *Symmetry Breaking*, Berlin-Heidelberg: Springer.
- , 2012, “Spontaneous Symmetry Breaking in Quantum Systems. A review for Scholarpedia,” arXiv: 1201. 5459v1 [physics.hist-ph].
- Struyve, W., 2011, “Gauge Invariant Accounts of the Higgs Mechanism”,

- Studies in History and Philosophy of Modern Physics*, 42: 226–236.
- Teh, N., 2015, “A Note on Rovelli’s ‘Why Gauge?’ ”, *European Journal for the Philosophy of Science*, 5(3): 339–348.
- 2016, “Galileo’s Gauge: Understanding the empirical significance of gauge symmetry”, *Philosophy of Science*, 83(1): 93–118.
- Teh, N. and Tsementzis, D., 2017, “Theoretical equivalence in classical mechanics and its relationship to duality”, *Studies in History and Philosophy of Modern Physics*, 59: 44–54.
- Teller, P., 2000, “The gauge argument”, *Philosophy of Science*, 67: S466–S481.
- ’t Hooft, G., 1980, “Gauge theories and the forces between elementary particles”, *Scientific American*, 242: June, 90–166.
- van Fraassen, B. C., 1989, *Laws and symmetry*, Oxford: Oxford University Press.
- Wallace, D., 2009, “QFT, Antimatter, and Symmetry ”, *Studies in the History and Philosophy of Modern Physics*, 40: 209–222.
- Wallace, D., forthcoming, “ Who’s Afraid of Coordinate Systems? An essay on the representation of spacetime structure”, *Studies in the History and Philosophy of Modern Physics*.
- Weatherall, J., 2016, “Understanding Gauge”, *Philosophy of Science*, 83(5): 1039–1049.
- Weyl, H., 1952, *Symmetry*, Princeton, NJ: Princeton University Press.
- Wigner, E. P., 1927, “Einige Folgerungen aus der Schrödingerschen Theorie für die Termstrukturen”, *Zeitschrift für Physik*, 43: 624–652.
- , 1937, “On the Consequences of the Symmetry of the Nuclear Hamiltonian on the Spectroscopy of Nuclei”, *Physical Review*, 51: 106–119.
- , 1939, “On Unitary Representations of the Inhomogeneous Lorentz Group”, *Annals of Mathematics*, 40: 149–204.
- , 1967, *Symmetries and reflections*, Bloomington, Indiana: Indiana University Press.

Academic Tools

-  How to cite this entry.
-  Preview the PDF version of this entry at the Friends of the SEP Society.
-  Look up this entry topic at the Indiana Philosophy Ontology Project (InPhO).
-  Enhanced bibliography for this entry at PhilPapers, with links to its database.

Other Internet Resources

[Please contact the author with suggestions.]

Related Entries

quantum mechanics | quantum theory: identity and individuality in | space and time: inertial frames | space and time: the hole argument | time: thermodynamic asymmetry in

Notes to Symmetry and Symmetry Breaking

1. Symmetry considerations were used by R.J. Haüy to characterize and classify crystal structure and formation (see his 1801 *Traité de minéralogie*, Volume 1), and with this, crystallography emerged as a discipline distinct from mineralogy.
2. A group is defined to be a set G , together with a product operation (\cdot) , such that: for any two elements g_1 and g_2 of G , $g_1 \cdot g_2$ is again an element of G ; the group operation is associative; the group contains the identity element; and for each element there exists an inverse.

3. An equivalence relation is a relation that is reflexive, symmetric and transitive. For further details on the connection between symmetry, equivalence and group in mathematics and physics, see Olver (1995). Castellani (2003) is devoted to exploring the relevance of these connections to the issue of the meaning of physical symmetries. On the connection between groups and equivalence classes, when generalized to groupoids, and, in general, for a discussion of the relation between the concept of symmetry and its formalisms, see Guay and Hepburn (2009). A thorough analysis of the relation between the notion of a symmetry of a physical theory and that of the physical equivalence of two solutions or models of such a theory is provided in Belot (2012).

4. A classic reference for the general theory of symmetric configurations is Shubnikov and Koptsik 1974. An overview is provided also in Mainzer 1996. Further details of the material in this section can be found in Castellani (2000, Chapters 1-3). Hon and Goldstein (2008) offers a detailed historical study of the term ‘symmetry’ and the concepts associated with it up to the early 1800s (on this book see Brading, 2010).

5. This is an example of a methodological use of symmetry properties: on the basis of the invariance properties of the situation under consideration (in this case, the dynamical problem in classical mechanics), a strategy is applied for deriving determinate consequences. The underlying principle is that equivalent problems have equivalent solutions. This type of symmetry argument (see Section 3, below) is discussed also by van Fraassen (1989, Chapter 10).

6. Wigner (1967) is a collection of twenty-four reprinted papers covering several decades.

7. Whence the name ‘Poincaré group’ introduced later by Wigner, whereas Poincaré himself named the group after Lorentz. Nowadays the name

‘Poincaré group’ is reserved for the inhomogeneous Lorentz group; that is, the Lorentz group plus spatial translations.

8. General relativity marks a further important stage in the development, as we will see below.

9. An example of such a principle is the non-existence of perpetual motion machines in thermodynamics.

10. For references see the literature on Einstein’s “hole argument”, detailed in Section 2.5, below.

11. For more on Einstein’s route to GTR and the changing fortunes of general covariance as one of his postulates, see Norton (1993). See also the references on Einstein’s “hole argument” in Section 2.5, below.

12. The problem which led Heisenberg to introduce this symmetry (and connect it with the statistical behaviour of quantum particles) in Heisenberg (1926) was to obtain a quantum description of atomic systems --- seen as ensembles of identical electrons subject to Coulomb interaction --- in agreement with the spectroscopies results of the time.

13. In that paper, entitled “Über die Operation der Zeitumkehr in der Quantenmechanik”, Wigner introduced time reversal invariance in order to reinterpret results previously obtained by H. Kramers.

14. Subsequently, new symmetries acquired relevance in theoretical physics, such as *supersymmetry* (the symmetry relating bosons and fermions and leading, when made local, to the theories of *supergravity*), and the various forms of *duality* used in today’s superstring theories.

15. See, for example, Mach’s discussion of the Oersted effect in his *Die Mechanik in ihrer Entwicklung historisch--kritisch dargestellt* of 1883.

16. Strocchi (2012) criticizes this too simplified description. See his paper for a precise formulation and Strocchi (2008) for a thorough and rigorous treatment of symmetry breaking in classical and quantum physics. Note that there are both analogies and disanalogies between the classical and quantum cases of SSB. These are the subject of discussion in the newly-emerging philosophical literature on SSB, and include the role of asymmetrical causes and the transition from a symmetrical to an asymmetrical situation.

17. Another example from classical physics which is often used in the literature to illustrate SSB is the case of a ball moving with no friction in a hoop constrained to rotate with a given angular velocity. This case is discussed in detail by Liu (2003).

18. Note that SSB was in fact first introduced in the DSB form. In the BCS theory of superconductivity, as well as in the 1961 theory of broken chiral symmetry by Nambu and Jona-Lasino, SSB is realized dynamically through a fermion condensate. In the BCS theory, for example, the gauge invariance of electromagnetism is spontaneously broken by pairs of electrons that condense -- forming a bound state -- in the ground state of a metal. Although DSB has not (so far) proved successful as an alternative route to the problem raised by the Higgs fields in the Standard Model, it has been applied with success to specific cases: for example, besides the already mentioned case of the BCS theory, the current quantum field theory of the strong interaction (quantum chromodynamics), in the approximation that quark masses are very small, possesses chiral symmetries that are spontaneously broken by a condensation of quark—antiquark pairs. For a historical—philosophical analysis of the notion that the Higgs boson is a composite particle, see Borrelli (2012).

19. Stewart and Golubitsky (1992), for example, speak of an “Extended Curie's principle” to indicate this situation.

20. Kosso's analysis begins from a set of examples offered by 't Hooft (1980).

21. The significance of the notion of invariance and its group-theoretic treatment for the issue of objectivity is explored in Born (1953), for example. For more recent discussions see Kosso (2003) and Earman ([2002] 2004, Sections 6 and 7).

Copyright © 2017 by the authors

Katherine Brading, Elena Castellani, and Nicholas Teh