Why Physical Symmetries?

Elena Castellani

January, 1998

Abstract

This paper is concerned with the meaning of "physical symmetries", i.e. the symmetries of the so-called laws of nature. The importance of such symmetries in nowadays science raises the question of understanding their very nature. After a brief review of the relevant physical symmetries, I first single out and discuss their most significant functions in today's physics. Then I explore the possible answers to the interpretation problem raised by the role of symmetries and I argue that investigating the real nature of physical symmetries implies, in some sense, discussing the meaning and methods of physics itself.

"Tell me, why should symmetry be of importance?"
(Mao Zedong to the physicist T. D. Lee, May 30, 1974.)

1. Introduction. It is well known that symmetry plays a very important role in today's physical science. A significant part of recent physical inquiry—especially the physics concerned with understanding the characteristics and behaviour of the fundamental building blocks of nature, the so-called elementary particles—is grounded on symmetry principles and their exploitation.

*Department of Philosophy, University of Florence, Italy.
†E-mail: castella@philos.unifi.it

The appropriate mathematical tool for investigating the consequences of symmetry principles is *group theory*: since the introduction in the 1920s of the theory of transformation groups in the quantum domain, mainly thanks to the seminal works of Hermann Weyl and Eugene Wigner, *symmetry groups* have become basic ingredients of theoretical physics. Fundamental physical theories can now be viewed as "symmetry theories", that is, theories centered on symmetry properties and formulated in terms of symmetry groups and their representations. The best example is the *standard model* for elementary particles developed around the 1970s, where the weak, electromagnetic and strong interactions are successfully described in terms of the symmetry groups SU(2), U(1) and SU(3).

The importance of symmetry in physics is thus a well-established fact. But why should symmetry play such a role in physics? The reasons for this effectiveness of symmetry are not at all evident. Is there a common origin of all physical symmetries? And what are indeed physical symmetries: useful conceptual tools introduced for understanding nature or properties existing in the physical world? Such are the kinds of questions constituting the "interpretation problem" raised by the role of symmetries in modern science.

The paper is devoted to the presentation and discussion of this problem. To present the question, I start with introducing the basic elements: first of all, what is exactly meant by "physical symmetries" and which are the main types of symmetries used in today's physics; secondly, which are the most important functions of physical symmetries in science. I then turn to the discussion of the interpretation problem raised by the use and functions of physical symmetries and explore the possible answers to the question of understanding their very nature. By enucleating the advantages and limits of the most significant positions, my main purpose is not so much to argue in favour of a given interpretation, but rather to emphasize that investigating the meaning of symmetry in physics inevitably implies a reflection on the meaning and methods of physical science itself.

2. Physical Symmetries and Symmetry Principles. *2.1 What is intended by "physical symmetries"?* As is known, the term "symmetry" can be understood in different ways. "Symmetry" may indicate, for example, a given "harmony of proportions" or a definite "relation of equality" between a number of elements. For a nowadays scientist, symmetry has the welldefined meaning of “invariance with respect to a group of transformations”. This group-theoretic definition is, in fact, the sense in which the concept of symmetry is currently applied in contemporary science.

In its group-theoretic version, the notion of symmetry can be referred to spatial configurations as well as to very abstract mathematical expressions. In particular, it can be referred to the fundamental equations of physical theories, the dynamical equations usually known as “laws of nature”. It is precisely as a property of physical laws that symmetry has become a basic concept in today’s theoretical physics. In this paper, the expression “physical symmetries” will be taken therefore to signify the invariance properties of physical laws with respect to given groups of transformations.

In physics, the symmetry properties of laws are postulated through principles, known as *symmetry principles* or *invariance principles*. The first principle to have been explicitly formulated as an invariance principle is the “principle of special relativity” by means of which Einstein, in 1905, established the invariance of physical laws with respect to the transformation group of inertial reference frames, the inhomogeneous Lorentz group.[1] Since Einstein’s first works on special relativity to nowadays quantum field theories, the history of the application of symmetry principles in physics is, in large part, the history of theoretical physics itself: think about the general theory of relativity, at the consequences of the introduction of group theory into quantum mechanics and at the development of the so-called gauge theories, the theories describing elementary particles and their fundamental interactions in terms of gauge symmetry groups.

2.2 Listing symmetry principles. At the present status of physical knowledge, the symmetry principles playing a relevant role in fundamental physical theories concern properties of the following kinds:

a) “Space-time symmetries”: continuous space-time symmetries, such as the symmetries postulated through relativity principles, and discrete symmetries, such as space inversion and time reversal;

b) “Charge conjugation”: the discrete symmetry existing between particles and their antiparticles;

c) “Permutation symmetry”: the permutation symmetry between quantum particles of the same kind (the so-called identical particles), related to the

[1] On the significance of Einstein’s papers on special relativity as regards the role and history of symmetry principles, see in particular Wigner (1967, 5–6).

Bose-Einstein and Fermi-Dirac statistics;
d) "Unitary internal symmetries": internal symmetries described in terms of unitary groups SU(N), global or local (the local internal symmetries are commonly known as "gauge symmetries");
e) "Supersymmetry": the symmetry between fermions and bosons, whose local version is known as "supergravity".

All these symmetries have a precise significance and a distinct role in the theoretical contexts in which they are employed. And each kind of physical symmetry has its own peculiar history, motivated by some determinate developments and results in contemporary physical research. Continuous space-time symmetries, also known as "geometrical symmetries", are the "oldest" ones, with deep roots in the history of physics and intended to be valid for all the laws of nature. Such a universal character is not shared, on the contrary, by other types of symmetries like, for example, internal symmetries, which are properties of an entirely new kind, introduced in physics expressly to describe some specific forms of interactions.[2]

Such specificities of the various types of physical symmetries do not imply, however, that we cannot individuate some general features in the use of symmetry principles in physics. In particular, it is here maintained, it is possible to individuate some main functions characterizing *in general* the application of symmetries in contemporary science. The next section will be devoted to illustrating this point.

3. Main Functions of Physical Symmetries. In reviewing what can be singled out as the main functions of physical symmetries, let me begin with the most evident, general and familiar function of symmetry, that is the *classification function.* The possibility of classifying objects on the basis of their symmetry properties is indeed a most basic and constant feature in the history of the modern notion of symmetry, intimately related to the same essence of this concept: namely, the deep connection existing between the notions of "symmetry", "group" and "equivalence class".[3] This explains why it is actually in relation to the possibility symmetry offers of classifying objects that it has first been taken into consideration as a scientific notion: the systematic study of symmetry and its application to modern science started

[2] The distinction between "geometrical" and "non-geometrical" or "dynamical" symmetry principles is due to Wigner. See Wigner (1967, 17–27, 31–36, and 42–45).

[3] A discussion of this connection can be found, for example, in van Fraassen (1989, 243).

with the classification of crystals, the natural objects with the most apparent properties of symmetry.[4] In nowadays physical theories, the most striking example of this function of symmetry is undoubtly the classification of elementary particles by means of the irreducible representations of the fundamental physical symmetry groups. In fact, as first showed by Wigner (1939), elementary particles can be associated with the irreducible representations of the symmetry groups of their fundamental equations, with the resulting possibility of deriving, in this way, the essential (intrinsic) properties such as mass, spin, etc., characterizing each class of particles.

In the case of symmetry classifications including all the essential properties necessary for characterizing a given type of physical object (for example, all the quantum numbers necessary for characterizing a given type of particle), we can speak also of a *definition function* of physical symmetries. This is in substance nothing other than the possibility of defining entities on the basis of their transformation properties, a quite usual procedure in most mathematical domains.[5] This definition function, let us note, reveals to be particularly useful in the case of objects such as the particles of microphysics, by offering a criterium for determining entities which are far remote from our everyday experience and to which the "classical conception" of a physical object does no more apply.[6]

Physical symmetries are most frequently used as "constraints" in the physical theories in which they are applied. This *normative function*, as we shall call it, is actually very familiar to physicists. It is well known that the requirement of invariance with respect to a given transformation group imposes severe restrictions on the form that a theory may assume, limiting the possible types of quantities appearing in the theory as well as the possible forms of its fundamental equations. Under some conditions, it is even possible to derive the form of a physical law on the unique basis of such symmetry requirements. The most famous case is Einstein's derivation of the gravitational law, "the first attempt to derive a law of nature by selecting

[4]The best illustration of all the possible classifications of configurations with symmetry properties—the subject matter of the so-called theory of symmetry—is offered by Shubnikov and Koptsik (1974).

[5]Take, for example, the definition of tensors, which is based on the specification of the laws governing the transformation of the tensorial components.

[6]This is the basic idea motivating the "group-theoretic approach" to the problem of determining physical objects. For a presentation of this approach, see Castellani (1998).

the simplest invariant equation" (Wigner 1967, 6).[7]

Many physical phenomena can be explained on the basis of the assumption of symmetry principles. Just think, for example, at the symmetry constraints mentioned above, functioning as selection rules and thus providing a reason for the possibility or not of the occurrence of given events. The most important result at the ground of such an *explicative function* of physical symmetries is the connection between the symmetries of physical laws and the laws of conservation, first formulated in a general way by the mathematician Emmy Noether in 1918. This result, together with more recent theoretical achievements such as the "gauge principle", the mechanism of "spontaneous symmetry breaking" and the techniques developed for unifying symmetry groups, offers a very important basis for explaining natural phenomena: the periodic table, the structures of atomic and nuclear spectra, the same interaction forms of matter can all be understood, in some sense, as consequences of physical symmetries.

The group-theoretic treatment of physical symmetries, with the resulting possibility of unifying different types of symmetries by means of the unification of the corresponding transformation groups, has provided a decisive technical support to a function of symmetry constantly recurring in the history of the notion, that is the *unification function*. This function, generally also connected with a function of *simplification*, operates at the level of the formal elaboration of physical theories as well as at the level of the "influencing metaphysics". This is best illustrated by what is a dominant research program in today's theoretical physics, namely the program of unification of all fundamental forces of nature.[8] The ancient idea of a "Unity of Nature" is here reformulated in terms of a "Theory of Everything", to be obtained on the basis of symmetry principles and their group-theoretic exploitation.

All the precedent aspects of the role of physical symmetries may assume, in many cases, an important *heuristic function*. When used as constraints, for example, symmetry principles can serve as a guide both in the formulation of the laws of nature and in the prediction of the occurence of given

[7]That is, by selecting the most natural equation compatible with the principle of general relativity. It is not important, for the purpose of this paper, to enter into the details of the question regarding the very nature of this principle.

[8]The program is essentially grounded on the possibility of a common interpretation of the four fundamental interactions—gravitational, weak, electromagnetic and strong—in terms of underlying gauge symmetry groups.

phenomena.[9] An important prediction basis for the existence of new particles, moreover, is offered by the classification schemes grounded on physical symmetries,[10] as well as by the various conceptual developments connected with the exploitation of symmetry principles, such as in particular the interpretation of interaction theories as gauge theories and the consequent unification program. Let us just cite, as a paradigmatic example, the 1967 prediction of the particles W and Z^0 (experimentally "discovered" in 1983) in the context of the Weinberg-Salam gauge theory for the unification of the weak and electromagnetic interactions.

I conclude this review by signalling what can be called an *epistemological function* of physical symmetries. For the relevant position they have acquired in nowadays investigation of nature, symmetry principles furnish the basis for a new viewpoint on many themes typically discussed in the philosophy of science. Questions concerning the structure of physical theories, the relations between successive theories (inter-theory relations) and the scientific progress can be reformulated by taking into account symmetry groups and their interrelationships.[11] On the ground of the role of symmetry principles, group-theoretic approaches can be developed to address traditional issues such as the question of the nature of physical objects.[12] Finally, as we'll see below, symmetries have to do with basic epistemological subjects such as the methodology and objectivity of physical knowledge.

4. How to Interpret Physical Symmetries?

4.1 The basic question. The many important functions of symmetry in contemporary physics raise some natural questions. Why should symmetry be of such relevance to physics? Given the undisputable effectiveness of group-theoretic methods, are there other reasons for the success of the symmetry

---

[9] An important role in this respect is played by indifference arguments of the kind of the so-called *Curie's Principle* stating that the symmetries of the causes are to be found in the effects. For an analysis of this sort of arguments, see for example van Fraassen (1989, Ch. 10).

[10] Think at the case of the prediction in 1962 of the existence of the particle Ω^- ("discovered" in 1964) on the simple basis of the hadronic classification scheme known as the *Eightfold Way*.

[11] In regard, see in particular Post (1971) and Redhead (1975). In French (forthcoming) the results by Post and Redhead are reconsidered from the point of view of the "model-theoretic approach".

[12] See Castellani (1998).

approach in physics? In other words, which is the real nature of physical symmetries?

In the literature, we find two main groups of answers to these questions: a) answers of a mathematical-physical nature, such as those motivating the success of the symmetry approach simply on the ground of some mathematical result[13] or those explaining the presence of physical symmetries as a consequence of other features of physical laws;[14] and b) answers of a more general nature, taking into account also the philosophical aspects of the interpretation problem.

This paper will be concerned essentially with the second group of answers. From the more general philosophical point of view, the basic question to be faced is the following: do physical symmetries have the meaning of properties existing in the physical world or are they just part of the conceptual apparatus with the help of which we try to understand nature? In short, do symmetries belong to nature or do they belong to the theories? Accordingly, two main kinds of positions can be distinguished: the positions ascribing physical symmetries to nature, which I shall call *ontic interpretations*, and the positions attributing the symmetries to the theories, which I shall call *epistemic interpretations*.[15]

4.2 Ontic interpretations. According to these interpretations, physical symmetries pertain to the ontology of the physical world. They represent properties effectively existing in nature (or properties characterizing the structure of nature), and this is the deep reason for the success of the symmetry approach in physics.

The main motivation for such a position is the "geometric interpretation" of space-time symmetries, namely the possibility of interpreting the spatio-temporal symmetries of physical laws as symmetries of space-time itself, the "geometric structure" of the physical world. This "geometrization" of physical symmetries can be extended also to other types of symmetries,

[13]See, for example, Kreinovic and Longpré (1995).

[14]This is the position which can be found, for example, in Frogatt and Nielsen (1991).

[15]I am not taking into consideration, in this paper, the other sorts of interpretations which can be found in the literature and which are essentially concerned with symmetry as a property of natural forms. My discourse focuses exclusively on the question of the interpretation of the symmetries of fundamental physical laws, i.e. the "primary" natural symmetries.

by considering them as properties of other kinds of "spaces", usually known as "internal spaces". Whether the resulting picture is meant in a realistic sense, is not a matter of importance here. Let us just note that the ontological image of the physical world which is implied by considering all physical symmetries as pertaining to nature is of a very mathematical and abstract character.

If symmetries are parts of nature, is there a special reason for their existence and for the fact that we can arrive to know or "discover" them? Supporters of an "ontic" interpretation must deal with questions of this sort. The most common answers are based on biological and evolutionistic arguments: such as, for example, the argument (grounded on the much discussed "anthropic principle") according to which symmetries are present because they are necessary for the existence of life and observers, or the argument attributing the possibility (for us) of discovering the symmetries of nature to the evolution of our brain in a universe obeying the physical laws.[16]

4.3 Epistemic interpretations. According to these interpretations, symmetries have primarily to do with the modalities of physical knowledge: (a) as conditions on the nature (or even the possibility) of such knowledge, or (b) as expression of limits inherent in the nature of our description of the external world.

(a) A first and largely accepted interpretation of the epistemic kind is grounded on the connection which can be established between *symmetry* and *objectivity*. Starting point for this connection are again the space-time symmetries: the interpretation of the special theory of relativity as a theory of physical invariance with respect to the transformation group of reference frames (or "observers") provides the basis for a sort of *objectivity condition*—intending "objectivity" in the restricted sense of "intersubjectivity"—for our physical description. 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 common 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-theoretic terms: "objective" is what is invariant with respect to the transformation group of reference frames, or—quoting Hermann Weyl—"objectivity means invariance with respect to the

[16] Arguments of such a kind are discussed, for example, in Stewart and Golubitsky (1992).

group of automorphisms [of space-time]" (Weyl 1982, 132).

Space-time symmetries can be put into connection with an even stronger condition for what regards physical knowledge. According to Wigner, the spatio-temporal invariance principles are a sort of "prerequisites" for the 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, 29). Symmetry principles, being thus related to a *possibility condition* for describing nature in terms of laws, acquire the meaning of "super-laws", which "provide a structure or coherence to the laws of nature just as the laws of nature provide a structure and coherence to the set of events" (Wigner 1967, 17).[17]

Both the above positions are grounded on the equivalence of spatio-temporal reference frames for the formulation of the laws of nature. We can then ask: what about the symmetries which are not of a spatio-temporal nature? These symmetries cannot be related to the equivalence of reference frames.

(b) A unitary interpretation for all types of symmetries can be obtained, on the contrary, by considering the connection between symmetry and equivalence in its most general form. As already said, the two notions are intimately related: it is namely the essence of the notion of symmetry to express a situation of equivalence (indistinguibility, indifference) between a given number of elements.[18] In physics, it is quite usual to understand the presence of equivalent or indistinguishable elements in terms of *irrelevance*: take the most familiar case of the equivalence of space points (translational symmetry), understood in the sense of the irrelevance of the "absolute position" for the physical description, or, to mention a less familiar example, the equivalence of quantum particles of the same kind (permutation symmetry), which can be understood in the sense of the irrelevance of a distinction between "identical" particles in the context of quantum theory. On this viewpoint, physical

[17]While for Wigner this conception of symmetry principles as possibility conditions for a physical description in terms of laws is essentially related to our ignorance (if we could directly know all the laws of nature, we would not need some regulative principles in our search for physical laws), some have attributed a kind of "absolute" meaning to this status of symmetry principles, viewing them as "transcendental principles" in the Kantian sense. A position of this sort can be found, for example, in Mainzer (1996).

[18]Of course, the equivalence is to be intended as relative to the context considered.

symmetries are then related to the presence of irrelevant (or redundant) elements in our physical description: because of limits which are inherent in our way of "theorizing", in describing the physical world we inevitably introduce surplus theoretical structure. Quoting Dirac, the fundamental laws of nature "control a substratum of which we cannot form a mental picture without introducing irrelevancies. The formulation of these laws requires the use of the mathematics of transformation" (Dirac 1931, vii).

Another common way of interpreting the situation of equivalence corresponding to a symmetry is to relate it to the impossibility of observing a given quantity. On this view, for example, the equivalence of space points is to understood in the sense of the impossibility of observing the absolute position (a difference between the positions), the equivalence of "identical" quantum particles is to be understood in the sense of the impossibility of observing a difference between these particles, and so on. This interpretation in terms of *non-observability*, which is widely accepted among scientists although the problematic nature of the notion of observability, has been especially defended by the physicist T. D. Lee, according to whom "the root of all symmetry principles lies in the assumption that it is impossible to observe certain basic quantities" (Lee 1988, 178). A corollary of this view is the interpretation of "asymmetry" (more precisely: "symmetry breaking") in terms of "observability": the violation of a symmetry is to be intended in the sense that "what was thought to be a non-observable turns out to be actually an observable" (Lee 1988, 181).[19]

5. Conclusion. What are, at the end, physical symmetries? Properties of nature, conceptual tools, expression of surplus theorical structure or of non-observable features? The interpretation question cannot be easily answered.

As we have seen, discussing the role and meaning of symmetries in today's physical theories is something like discussing the methodologies and guiding strategies of physical science itself. On the one hand, physical symmetries play a crucial role in the progress of theoretical physics, they occupy a leading place in the elaboration and testing of fundamental theories and they offer a basis for proceeding towards theoretic unification. On the other hand, physical symmetries have a relevant epistemological role: they provide

[19]For a discussion, on this view, of some examples of symmetry violations (first of all the violation of parity) see Lee (1988, 181–188).

important new tools for discussing traditional issues in the philosophy of science, and they are found to be related to the nature of physical knowledge itself, as condition for the objectivity and even the possibility of a physical description in terms of laws, or as expression of limits inherent to our way of theorizing.

How to answer, then, the basic interpretation question? My suggestion is that what does really matter is the discussion itself, rather than a definite answer.

References

[1] Castellani, E. (1998), "Galilean Particles: An Example of Constitution of Objects", in E. Castellani, (ed.), *Interpreting Bodies. Classical and Quantum Objects in Modern Physics.* Princeton: Princeton University Press (forthcoming).

[2] Dirac, P. A. M. (1931), *The Principles of Quantum Mechanics.* Oxford: Oxford University Press.

[3] French, S. (forthcoming), "Models and Mathematics in Physics: the Role of Group Theory".

[4] Frogatt, C. D. and Nielsen, H. B. (1991), *Origin of Symmetry.* Singapore: World Scientific.

[5] Kreinovic, V. and Longpré, L. (1995), "Unreasonable Effectiveness of Symmetry in Physics", *International Journal of Theoretical Physics* 35, 7: 1549–1555.

[6] Lee, T. D. (1988), *Particle Physics and Introduction to Field Theory.* New York: Harwood Academic Publishers.

[7] Mainzer, K. (1996) *Symmetries of Nature. A Handbook.* New York: de Gruyter.

[8] Redhead, M. (1975), "Symmetry in Intertheory Relations", *Synthese* 32: 77–112.

[9] Post, H. R. (1971), "Correspondence, Invariance and Heuristics", *Studies in History and Philosophy of Science* 2: 213–255.

[10] Shubnikov, A. V. and Koptsik, V. A. (1974), *Symmetry in Science and Art.*, London: Plenum Press.

[11] Stewart,I. and Golubitsky, M. (1992), *Fearful Symmetry. Is God a Geometer?* London: Blacwell.

[12] van Fraassen, B. C. (1989), *Laws and Symmetry.* Oxford: Clarendon Press.

[13] Weyl, H. (1982), *Symmetry.* Princeton: Princeton University Press.

[14] Wigner, E. (1939), "On unitary representations of the inhomogeneous Lorentz group", *Annals of Mathematics* 40, 1: 149–204.

[15] Wigner, E. P. (1967), *Symmetries and Reflections.* Bloomington: Indiana University Press.