

On the meaning of symmetry breaking

ELENA CASTELLANI

To appear in K. Brading and E. Castellani, *Symmetries in physics: philosophical reflections*, Cambridge University Press 2003

Symmetries can be attributed to physical states or to physical laws. The focus of this volume is on the symmetries of physical laws, that is the physical symmetries postulated by means of invariance principles. Accordingly, the focus of this review paper is on symmetry breaking in the case of physical laws. In this case, there are two different forms – “explicit” and “spontaneous” – of symmetry breaking, the spontaneous symmetry breaking case being the more interesting from a physical as well as a philosophical point of view.

After some general preliminary remarks on symmetry breaking, we start by examining how symmetry breaking was first considered in the literature, and then turn to the main subject of this paper, that is the physical and philosophical meaning of symmetry breaking of the laws of nature.

1. Preliminaries - I

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.

Our concern here is the breaking of physical symmetries. In physics, symmetry properties may be attributed to physical laws (equations) or to physical objects/phenomena (solutions). As the contributions to this volume clearly show, the two cases must be distinguished when considering the meaning and functions of physical symmetries and, accordingly, of their breaking.¹

Another preliminary distinction is needed. The expression “symmetry breaking” is in fact ambiguously used. On the one hand, it is taken to indicate the process by means of which the considered symmetry is broken, and is therefore usually ascribed a “dynamic” character in the literature (in contrast with the “static” character attributed to a situation of symmetry).

¹Although there is of course a connection between the symmetries of physical laws and the symmetries of physical states, as will be seen in more detail later, especially when discussing the case of spontaneous symmetry breaking.

On the other hand, it is taken to indicate the result of a symmetry- breaking process, that is, a broken symmetry situation (or simply the fact that a symmetry is not there). Failure to keep these two meanings distinct can lead to confusion.²

In all cases, the general questions raised by symmetry breaking are of the following three types: a) why a symmetry breaks or is not there; b) how it breaks; c) what the effects or consequences of its breaking (or of its not being there) are.

Finally, let us emphasize once more the importance and usefulness of group theory. As is known, the symmetry of a “something” (a figure, an equation, ...) is defined in terms of its invariance with respect to a specified transformation group, its symmetry group. 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 situation where this symmetry is not broken. 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 transformations 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?”

2. Symmetry breaking and Curie’s analysis

Symmetry breaking was first explicitly studied in physics with respect to physical objects and phenomena. This follows naturally from the developments of the theory of symmetry, at the origin of which are the visible symmetry properties of familiar spatial figures and everyday objects.³ Moreover,

²Another source of confusion is due to unclear terminology in current literature, where usually no distinction is made between the result of a symmetry-breaking process (“broken symmetry”), the absence of one of the possible symmetries compatible with the situation considered (“non-symmetry”, or “dissymmetry” as it was termed in the nineteenth-century literature, notably by Louis Pasteur in his works on molecular dissymmetry), and the absence of all the possible symmetries compatible with the situation considered (“asymmetry”).

³See the introduction to this volume, sections 1 and 2.

examples of broken symmetries pervade the physical world, in particular the world of our everyday experience.

Any symmetry we can perceive (albeit in an approximate way) is indeed the result of a higher order symmetry being broken.⁴ This can actually be said of any symmetry which is not the “absolute” one (i.e. including all possible symmetry transformations). But we can say even more: in a situation characterized by an absolute symmetry, nothing definite could exist, since absolute symmetry means total lack of differentiation. For the presence of some structure, a lower symmetry than the absolute one is needed: in this sense, symmetry breaking is essential for the existence of structured “things”. Crystals, the natural objects “structured” *par excellence*, offer an exemplary illustration of this function of symmetry breaking, their many and striking symmetries being the result of the breaking of the symmetries of the initial medium from which they originated.⁵ More precisely, the case of crystals illustrates how, given a physical medium as much undifferentiated as physical constraints allow, structured patterns arise when some of the possible symmetries are broken for some physical causes (how this breaking actually occurs is another question, as we will see in the next section). The resulting pattern is thus characterized by the presence of some symmetries and the absence of others; or, in other words, by specified symmetries and specified “dissymmetries” (the term “dissymmetry” meaning that some of the possible symmetries compatible with the physical constraints are not present).⁶

The symmetries (and dissymmetries) of crystals, besides being at the origin of the systematic study and classification of the possible types of symmetric patterns, also represent the starting point for the first explicit analysis of the role of symmetry breaking in physics. This analysis is due to Pierre Curie, who towards the end of the nineteenth century devoted a series of works –

⁴The “order” of a symmetry being the order of the symmetry group – that is, the number of independent symmetry operations.

⁵A hot gas of identical atoms has a very high symmetry, the equations describing it being invariant under all rigid motions as well as under all permutations of the atoms. As the gas cools down, the only mathematical solution maintaining these symmetries, which is the state where all the atoms are in the same place, is ruled out by the physics. In fact, the initial symmetry breaks down and the physical system takes up a stable state with less symmetry, that is the crystal lattice. On this point, and on the general role of symmetry breaking in the formation of nature’s patterns from the smallest scales to the largest, see in particular Stewart and Golubitsky (1992), chapter 3.

⁶See n. 2.

in particular his famous 1894 paper, extracts from which are reprinted in this volume – to examining the role of symmetry and symmetry breaking in physical phenomena, especially electric and magnetic.

As mentioned in the introduction to this volume, Curie’s analysis was centred on the following question: which phenomena are allowed to occur in a given physical medium having specified symmetry properties? His studies of such properties as the pyro- and piezo-electricity of crystals had in fact persuaded him of the importance of the relationships between the symmetry of a physical medium (for example, a crystal) and the symmetry of the phenomenon occurring in it. By applying the methods and results of the theory of symmetry groups used in the crystallography of his times to the study of a number of physical phenomena, he arrived at definite conclusions known in the literature as “Curie’s principles”. Actually these principles amount to just one, Curie’s conclusions clearly not being independent of one another. In this section we focus on the following group of propositions (called by some “Curie’s first principle”):

- (a) “The characteristic symmetry of a phenomenon is the maximum symmetry compatible with the existence of the phenomenon.”
- (b) “A phenomenon may exist in a medium having the same characteristic symmetry or the symmetry of a subgroup of its characteristic symmetry.”
- (c) “In other words, certain elements of symmetry can coexist with certain phenomena, but they are not necessary. What is necessary, is that certain elements of symmetry do not exist. Dissymmetry is what creates the phenomenon.” (Curie, 1894, this volume.)

Curie makes very clear what he means with (a), (b) and (c) by discussing in his paper a number of concrete examples. These range from the determination of the characteristic symmetries of such phenomena as the Newtonian gravitation field, the electric field and the magnetic field (together with the symmetries of the media in which they originate), to the explanation – on the grounds of the symmetry groups implied and their relationships – of the possibility of such physical effects as the “Wiedemann effect” and the “Hall effect”.

In the case of the Wiedemann effect (the longitudinal magnetization of an iron cylindrical wire when a longitudinal electric field and an asymmetrical

torque are applied), for example, the original symmetry group of the wire before the application of the electric field and the torque is the symmetry group of a cylinder. For the magnetization of the wire (the “effect”) to occur, some elements of this group must be absent, that is, the group must be broken to a lower order group (which is what happens under the combined action of the electric field and the torque).⁷

According to Curie, symmetry breaking thus 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 (the electric field and the torque in the above example). In this sense symmetry breaking is what “creates the phenomenon”.

3. Preliminaries - II

A physical system is described in terms of its states, its observables and its equation of motion (the states being solutions to this equation). All these components together form the physical theory of the system.⁸

The symmetries playing a central role in today’s physics are the symmetries of physical theories. By this expression physicists mean the transformation groups leaving invariant (in form) the dynamical equations, the so-called physical laws (these invariance properties having then crucial implications for the physical states and observables).⁹

In the previous section we discussed symmetry breaking in the case of physical states. We now want to consider the meaning of symmetry breaking in the more general case of physical laws.

A preliminary question is the following: what is the relationship between the symmetry/asymmetry properties of states and the symmetry/asymmetry properties of laws? Since physical states are solutions to dynamical equations, it is natural to expect a relationship of some sort between the symmetry breaking of a state and the symmetry properties of the equation to

⁷For details and comments on the discussion of Wiedemann effect by Curie, see for example Radicati (1987), pp. 199-200.

⁸The terminology used in this and the following section is, for brevity’s sake, admittedly rough.

⁹See, for example, the Introduction and Martin, this volume.

which it is a solution. If we focus only on symmetry breaking of particular solutions (as in the previous section), nothing definite can be said about the symmetry properties of the equations. Depending on the case considered, asymmetric states can be solutions to symmetric as well as to asymmetric dynamical equations. In fact, the symmetry of a dynamical equation is not necessarily the symmetry of the individual solutions, but rather the symmetry of the whole set of solutions (in the sense that the symmetry of the dynamical equation transforms a given solution into another solution). Note that this explains, in particular, the role of asymmetric initial conditions in symmetry breaking of physical states (often emphasized in the literature): even if the considered equation is symmetric, an asymmetric solution at an initial time (i.e. asymmetric initial conditions) implies asymmetric solutions at later times. There remains the question of the origin of the asymmetric initial state (how do the laws give rise to asymmetric solutions?), which will be the subject of the next section.

4. Symmetry breaking of physical laws

Symmetries of physical laws can be broken in two ways: *explicitly* or *spontaneously*. We examine the two cases separately.

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, as the following cases illustrate.

(a) Symmetry-breaking terms may be introduced into the theory by hand on the basis of theoretical/experimental results. An illustration is offered by 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* (see Pooley, this volume). 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 (and then confirmed by others in the following year) considering the case of the

decays $K \rightarrow 3\pi(\tau)$ and $K \rightarrow 2\pi(\theta)$ of the kaon K , known at the time as the “ $\tau - \theta$ puzzle”).¹⁰

(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 the following: 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.¹¹

In practice it is not always possible to master the commutation rules of fields, and the alternative way is to check whether the Noether currents are conserved order by order in perturbation theory after an appropriate renormalization.¹² 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, “discovered” in connection with the problem of understanding the observed decay rate of the neutral pion (see Weinberg, 1996, chapter 22).

(c) Finally, symmetry-breaking terms may appear because of non-renormalizable effects. As mentioned also in other contributions to this volume (see Martin, section I, and Castellani, section IV), 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”

¹⁰For a detailed history of parity violation, see in particular Telegdi (1987).

¹¹See for example J. Govaerts, “The quantum geometer’s universe: particles, interactions and topology”, 2002, hep-th/0207276, section 4.2.

¹²For details see for example Itzykson and Zuber (1980), pp. 510-11.

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. 529-31).

4.2. Spontaneous symmetry breaking

Spontaneous symmetry breaking (SSB) indicates a situation where, given a symmetry of the equations of motion, solutions exist which are not invariant under the action of this symmetry *without the introduction of any term explicitly breaking the symmetry* (whence the attribute “spontaneous”).¹³ How can this happen?

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, discussed also by Jona-Lasinio in this volume. 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.¹⁴

¹³Historically, the name “spontaneous breakdown of symmetry” appeared first in Baker and Glashow (1962), to indicate the fact that starting with a model possessing the higher symmetry $SU(3)$, “without the introduction into the Lagrangian of any symmetry-breaking terms, solutions exist which have only the lower symmetries of isotopic spin and hypercharge”.

¹⁴Another 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. The theory has a discrete inversion symmetry which is respected by the lowest energy solution when the velocity is below a specified critical value; when the velocity reaches the critical value, the symmetric equilibrium configuration becomes unstable and two lowest energy stable configurations appear, each one being asymmetric but related to the other by the action of the symmetry of the theory. This case, typically used as a simple model for the bifurcation theory of relative equilibria and its connection with dynamic stability theory in textbooks on classical mechanics, has been recently discussed in detail by C. Liu, “The meaning of spontaneous symmetry breaking (I): From a simple classical model”, 2002, PITT-PHIL-SCI00000563.

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.

SSB occurs in classical and quantum physics. In what follows, we focus on the meaning of SSB in quantum physics, since it is in this context that it leads to the most important consequences.

The first distinction to be drawn is between *finite* and *infinite* physical systems. In the case of finite systems, SSB actually does not occur: 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”.¹⁵

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 (the spin-spin interaction as well as the rest of the Hamiltonian) 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.¹⁶

¹⁵For details see for example Weinberg (1996), pp. 164-5.

¹⁶For this image see Aitchison (1982), pp. 75-7.

To use a famous image by Sidney 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.¹⁷ As said, in the infinite volume limit all ground states are separated by a superselection rule.

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. As stated by Baker and Glashow (1962, p. 2463), “it is thus made plausible that the intricacies of the physical world are not reflected in an equally intricate fundamental theory”.

How the concept of SSB was in fact transferred from condensed matter physics to QFT in the early 1960s, thanks especially to works by Y. Nambu and G. Jona-Lasinio, is described in the other contributions to this section of the volume (see Jona-Lasinio, Earman, and Morrison). In particular, Jona-Lasinio’s paper 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. We just mention some of the main ideas here, referring to the other chapters of this section of the volume (and the references therein) for further details.

¹⁷See Coleman (1975), pp. 141-2.

In short, as it turned out, a number of field theories may display spontaneous breaking of internal symmetries. In the case of a *discrete symmetry*, this does not lead to especially interesting consequences. In the case of a *continuous symmetry*, what results from SSB can be synthesized in the following points.

- *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 (massless because of the gauge invariance).
- *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, i.e. to a *gauge symmetry*, the Goldstone bosons “disappear” and the gauge bosons acquire a mass; or, in other words, the Goldstone bosons are “eaten up” to give mass to the (otherwise massless) gauge bosons, and this happens without (explicitly) breaking the gauge invariance of the theory.
- *Renormalizability.* The above 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.¹⁸

¹⁸Their Lagrangians in fact involve only massless gauge fields coupled minimally to conserved currents, the mass of the gauge vector bosons and the non-conservation of the currents (whence the apparent non-renormalizability) being the result of SSB. It is a general result that the divergence structure of a renormalizable quantum field theory is not affected by SSB.

Let us end this brief review of SSB by mentioning *dynamical symmetry breaking* (DSB). In such theories as the Glashow-Weinberg-Salam unified model of electroweak interactions, the SSB responsible for the masses of the gauge vector bosons is due to 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 *ad hoc* character of these scalar fields, for which there is no experimental evidence (no “Higgs particle” has been observed up to now) – increasing attention has been drawn to the possibility that the scalar 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”.

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.¹⁹

5. Symmetry breaking and philosophical questions

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 Pooley and Huggett, this volume, section II). 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, discussed in section IV of this volume (see also the introduction,

¹⁹In a similar way to the spontaneous breaking of the chiral symmetry of the strong interactions hypothesized in 1961 by Nambu and Jona-Lasino, although of course not in terms of quarks and antiquarks.

section 3). Finally, there are issues raised by the role and nature of symmetry breaking in general; these are the subject of this concluding section.

From what we have seen so far, the main points about the general role of symmetry breaking in physics can be summarized as follows:

- (1) Symmetry breaking is essential for the existence of something structured (for the presence of some structure, a lower symmetry than the absolute one is needed). In Curie's words, the absence of certain elements of symmetry or dissymmetry is what creates the phenomenon.
- (2) Asymmetric states can be solutions to symmetric as well as asymmetric equations (the symmetry of a dynamical equation not being necessarily a symmetry of the individual solutions but rather the symmetry of the whole set of solutions). In the case of symmetric equations, asymmetric initial conditions explain the presence of asymmetric solutions, but there remains the question of the origin of the asymmetric initial state.
- (3) Spontaneous symmetry breaking allows symmetric theories to describe asymmetric reality. The asymmetries of the physical world of our experience do not necessarily belong to the fundamental laws of nature: 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 short, SSB provides a way of understanding the complexity of nature without renouncing fundamental symmetries.

In the light of these points, the philosophical discussion has focused on methodological and epistemological aspects of spontaneous symmetry breaking, and in particular the following questions:

(A) If, according to point (1), the absence or breaking of some symmetry is not only common but even necessary for the existence of the phenomena, why search for a way (provided by the mechanism of SSB) of avoiding attributing these asymmetries to the fundamental laws?²⁰ The question is actually twofold:

- Why should we prefer symmetric to asymmetric fundamental laws? We refer to the other contributions to this volume for this point, the answer

²⁰On this sort of question see, for example, Earman, this section of the volume, and Kosso, 2000.

being provided by the many relevant functions and consequences of symmetry principles in contemporary physics.

• Why should we assume that an asymmetry needs an explanation? The underlying idea is that we can start with symmetries, but not asymmetries.²¹ In other words, it is assumed 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 idea is very similar to the one expressed by Curie in his famous 1894 paper. In section 2 we have mentioned one of the forms in which Curie expressed his conclusions about the relationships between the symmetries of a physical medium and the symmetries of a phenomenon occurring in it. Another equivalent formulation but in terms of “causes” and “effects” – usually known as “Curie’s principle” *tout court* in the literature – is the one stating that the symmetries of the causes must be found in the effects; or, equivalently, the asymmetries of the effects must be found in the causes (see Curie, 1894, this volume). This formulation, when extended to include the case of SSB, is well appropriated to express the methodological principle considered here: an asymmetry of the phenomena (“the effects”) must come from the breaking – “explicit” or “spontaneous” – of the symmetry of the fundamental laws (“the causes”). What the real nature of this principle is remains an open issue, at the centre of a developing debate (see the introduction to this volume, section 2).

For the sake of completeness, let us mention an aspect of the debate on the validity of Curie’s principle that is in apparent contrast to the “extended” way in which it has been presented above. The point is sometimes made in the literature that SSB appears to violate Curie’s principle because of the absence of any asymmetric cause. Thus, when SSB occurs we have a situation where an asymmetry of the effect does not derive from an asymmetry of the cause; or, equivalently, a situation where the symmetry of the cause (the law) is not to be found in the effect (the phenomenon). That this is not the case is easily shown. 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 the symmetry of the “cause” is not lost, it is conserved in the ensemble of the solutions (the whole “effect”) (see section 4, and point (2), this section).²²

²¹Of course, there remains the fundamental issue of the origin of the symmetries, which is the subject matter of other papers in the volume (in particular, sections I and IV).

²²Stewart and Golubitsky (1992), for example, speak of an “Extended Curie’s principle”

(B) What is the epistemological *status* of a theory based on SSB? This general issue is currently addressed by tackling the following cluster of questions:

- What is the empirical stand of the “hidden” symmetries of the fundamental laws (given that, in the SSB case, what is directly observed – the physical situation, the phenomenon, ... – is asymmetric)? How can we know them, if they are hidden? The question is part of the general issue of the empirical evidence for the symmetries of physical laws, here addressed in the case where there is no direct correspondence between the symmetry of the laws and the symmetry of the phenomena (see for example Morrison, this volume, 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.

This is in part a specific issue:

(a) In each of the particular cases in which SSB is applied, is the existence of the hidden symmetries of the laws the best explanation for the behaviour of the physical world? How successful is the specific SSB case considered in producing observable consequences?

And in part a general issue:

(b) Can the predictive and explanatory power of these theories be a basis for scientific realism with respect to the assumed hidden symmetries, and, more generally, with respect to all the (directly) non-observable features introduced for the SSB mechanism to function? In this respect, for example, Morrison (this volume) discusses two particular issues arising in the specific but paradigmatic case of the Glashow-Weinberg-Salam theory of electroweak interactions: that is, the question of the epistemological *status* of the Higgs particles, and the issue of the nature of the physical assumptions about the quantum field theory vacuum, which play a crucial role in the history of SSB.

Conclusions. Concluding this final section, it is worth underlining that the philosophical reflection on symmetry breaking (in particular in its “spontaneous” form) is very recent in the literature. The discussion of the methodological and epistemological issues mentioned above, while important from both a physical and a philosophical point of view, has just started. Stimulating this discussion by offering an overview of its basic elements and possible developments is the aim of this review paper and the other contributions

to indicate this situation.

(Jona-Lasino, Earman and Morrison) to the section III of this volume.

Acknowledgements

Many thanks to Katherine Brading, Roberto Casalbuoni and Leonardo Castellani for helpful comments and suggestions.

References

Aitchison, I. J. R. (1982). *An informal introduction to gauge field theories*. Cambridge: Cambridge University Press.

Baker, M., and Glashow, S. L. (1962). “Spontaneous breakdown of elementary particles symmetries”. *Physical Review*, **128**, 2462-71.

Coleman, S. (1975). “Secret symmetry: an introduction to spontaneous symmetry breakdown and gauge fields”. In *Laws of hadronic matter*, ed. A. Zichichi, pp. 138-215. New York: Academic Press.

Curie, P. (1894). “Sur la symétrie dans les phénomènes physiques. Symétrie d’un champs électrique et magnétique”. *Journal de Physique*, 3rd series, vol. 3, 393-417.

Itzykson, C., and Zuber, J.-B. (1980). *Quantum field theory*. New York: McGraw-Hill.

Kosso, P. (2000). “The epistemology of spontaneously broken symmetries”. *Synthese*, **122**, 359-76.

Lee, T.D., and Yang, C.N. (1956). “Question of parity conservation in weak interactions”. *Physical Review*, **104**, 254-8.

Radicati, L.A. (1987). “Remarks on the early developments of the notion of symmetry breaking”. In *Symmetries in physics (1600-1980)*, ed. M.G. Doncel, A. Hermann, L. Michel, and A. Pais, pp. 195-206. Barcelona: Servei de Publicacions.

Stewart, I., and Golubitsky, M. (1992). *Fearful symmetry. Is God a geometer?* Oxford: Blackwell.

Telegdi, V.L. (1987). Parity violation. In *Symmetries in physics (1600-1980)*, ed. M.G. Doncel, A. Hermann, L. Michel, and A. Pais, pp. 433-49. Barcelona: Servei de Publicacions.

Weinberg, S. (1995). *The quantum theory of fields. I.* New York: Cambridge University Press.

Weinberg, S. (1996). *The quantum theory of fields. II.* New York: Cambridge University Press.