Gauge Symmetry Breaking in Gauge Theories—In Search of Clarification

Simon Friederich

friederich@uni-wuppertal.de

Universität Wuppertal, Fachbereich C – Mathematik und Naturwissenschaften, Gaußstr. 20, D-42119 Wuppertal, Germany

Abstract: The paper investigates the spontaneous breaking of gauge symmetries in gauge theories from a philosophical angle. Local gauge symmetry itself cannot break spontaneously in quantized gauge theories according to Elitzur's theorem—even though the notion of a spontaneously broken local gauge symmetry is widely employed in textbook expositions of the Higgs mechanism, the standard account of mass generation for the weak gauge bosons in the standard model. Nevertheless, gauge symmetry can be broken in gauge theories, namely, in the form of the breaking of remnant subgroups of the original local gauge group under which the theories typically remain invariant after gauge fixing. The paper discusses the relation between these instances of symmetry breaking and phase transitions and draws some more general conclusions for the philosophical interpretation of gauge symmetries and their breaking.

1 Introduction

The interpretation of symmetries and symmetry breaking has been recognized as a central topic in the philosophy of science in recent years. Gauge symmetries, in particular, have attracted a considerable amount of interest due to the central role they play in our most successful theories of the fundamental constituents of nature. The standard model of elementary particle physics, for instance, is formulated in terms of gauge symmetry, and so are its most discussed extensions. The defining characteristic of gauge symmetries is that they are not empirical but purely formal symmetries[1] in that

[1]See [Healey, 2007], Chapter 6, where the distinction between empirical and purely formal symmetries is spelled out in detail and the standard account of gauge symmetries as purely formal symmetries is defended with great care.

different configurations of the fields involved represent identical physical situations if they are related by gauge symmetry. For recent philosophical work on the interpretation of gauge symmetries see, for instance, [Redhead, 2002], [Brading and Castellani, 2003], [Lyre, 2004], [Healey, 2007].

The present paper focuses on a particular aspect of gauge symmetries, namely, the notion of a spontaneously broken gauge symmetry. This is a notion that may seem puzzling at first glance, for it seems natural to ask what it might mean to spontaneously break a purely formal symmetry that exists only on the level of our description of physical reality, not on the level of physical reality itself. The notion of a spontaneously broken gauge symmetry is not an exotic notion, however, for it is widely regarded as playing a crucial role in the generation of particle masses in the standard model of particle physics by the Higgs mechanism. Although it is almost universally accepted, the received view of the Higgs mechanism as a case of broken local gauge symmetry has been criticized by both physicists and philosophers of physics, see ['t Hooft, 2007], [Earman, 2004], [Healey, 2007], [Lyre, 2008]. 't Hooft, for instance, criticizing it from the point of view of a physicist, claims that the notion of a spontaneously broken local gauge symmetry is "something of a misnomer"[2], while Earman, from the point of view of a philosopher, expresses qualms concerning the Higgs mechanism as a spontaneously broken local gauge symmetry on grounds that "a genuine physical property like mass cannot be gained by eating descriptive fluff, which is just what gauge is."[3] Worries like these about the standard picture of the Higgs mechanism as a spontaneously broken local gauge symmetry are aggravated by a result known as Elitzur's theorem (see [Elitzur, 1975]), rigorously established for the context of lattice gauge theory, according to which local (gauge) symmetry cannot be spontaneously broken at all.

In order to develop an adequate perspective on the status of spontaneous symmetry breaking in quantized gauge theories, Earman proposes that the question be tackled by means of the constraint Hamiltonian formalism, an approach in which, in contrast to the standard approach discussed in Sections 5 and 6 of this paper, gauge orbits (that is, gauge field configurations related by gauge symmetry) are quotiented out before the resulting unconstrained system of variables is subjected to a quantization procedure (see [Earman, 2004]). An analysis of a particular gauge theory (the Abelian Higgs model) in terms of the constraint Hamiltonian approach has recently been proposed by [Struyve, 2011], but the generalization to the non-Abelian

[2]See (['t Hooft, 2007] p. 697).

[3]See ([Earman, 2004] p. 1239).

case and the quantization for the resulting unconstrained system remain to be done. Independently of the success of this enterprise, it seems reasonable to ask whether the puzzles surrounding the notion of spontaneous symmetry breaking in gauge theories might not be resolvable within the standard "Lagrangean" framework of quantum field theories, as it is actually used by those working in the field of high energy physics. My aim in the present paper will be to show that this can indeed be done. A proper assessment of the role of symmetry breaking in gauge theories that does not merely recite the standard narrative of the Higgs mechanism as a spontaneously broken local gauge symmetry, arguably, can be given on the basis of the conventional approach to quantum field theory alone.

The rest of this paper is organized as follows: Section 2 recalls some basic features of the concepts of (gauge) symmetry and (gauge) symmetry breaking, and Section 3 discusses the characterization of symmetry breaking as a "natural phenomenon" proposed by Liu and Emch[4] and considers in which sense it applies to cases where the broken symmetry is a *gauge* symmetry. Sections 4 and 5 assess the fate of the notion of *local* symmetry breaking in gauge theories, which, as argued in Section 4, makes sense at the classical level but is vacuous, as discussed in Section 5, in quantized gauge theories according to Elitzur's Theorem. Section 6 discusses the breaking of post-gauge fixing remnant *global* gauge symmetries and their relation to transitions between distinct physical phases. It is argued that there seems to be no fixed connection between these instances of symmetry breaking and phase transitions in that the distinction between broken and unbroken symmetries does not in general line up with a distinction between distinct physical phases. Section 7 turns to the more general philosophical relevance of these findings by considering their implications for claims brought forward in the literature on philosophical aspects of gauge symmetries and their breaking. The paper closes in Section 8 with a brief concluding remark.

2 Symmetries, gauge symmetries, and symmetry breaking

In this section, I give a brief review of the concepts in terms of which the questions discussed in this paper are formulated. The concepts are those of symmetry, gauge symmetry, symmetry breaking, and gauge symmetry breaking. Readers who are familiar with these notions can skip the section

[4] See ([Liu and Emch, 2005] p. 153).

without loss, perhaps apart from the last two paragraphs, which review the phenomenon of Bose-Einstein condensation in a free Bose gas in terms of broken gauge symmetry.

A *symmetry* α of a classical system is a transformation $\alpha : \gamma \mapsto \alpha\gamma$ of the (coordinate) variables in terms of which its configurations S_γ are individuated that induces an automorphism $\alpha : S_\gamma \mapsto \alpha S_\gamma \equiv S_{\alpha\gamma}$ which commutes with its time evolution. If the equations of motions for the system are derived from an action principle in the Lagrange formalism or as Hamilton's equations from a Hamiltonian, this translates into the statement that the Lagrangean or Hamiltonian from which they are derived is invariant under α. For a quantum system, a symmetry is an automorphism of the observables or canonical variables which preserves all algebraic relations. In analogy to the classical case, possible states of the system are individuated in terms of the expectation values they ascribe to these quantities. Time evolution, in the Heisenberg picture, counts as an algebraic relation among others, so the invariance of all algebraic relations under a symmetry in the Heisenberg picture implies that the symmetry commutes with the dynamics of the system.

Gauge theories are defined in terms of an action S that is invariant under transformations corresponding to an infinite dimensional Lie group and depending on a finite number of arbitrary functions. As follows from Noether's second theorem (see [Noether, 1918]), the equations of motion apparently fail to be deterministic in this case in that they involve arbitrary functions of space-time. This means, in particular, that any configuration of the coordinate variables at a given initial time t_0 does not uniquely determine the configuration of variables at a later time t_1. In the *physical* interpretation of gauge theories, however, determinism can be restored by assuming that variable configurations that are related by the symmetry represent identical physical situations. The symmetry is referred to as a *gauge symmetry* in this case. Classical electromagnetism is a paradigm example of a gauge theory in that (assuming the relativistic formulation in terms of four-vector fields) its action is invariant under local gauge transformations of the four-vector potential $A_\mu(x)$ having the form

$$A_\mu(x) \mapsto A_\mu(x) - \frac{1}{e}\partial_\mu \alpha(x)\,. \tag{1}$$

Only functions of the gauge fields that are invariant under gauge transformations of the form (1) correspond to physical quantities. The inertial frame-dependent electric and magnetic fields $\mathbf{E}(x)$ and $\mathbf{B}(x)$, obtained from $A_\mu(x)$ by taking certain derivatives, are examples of such quantities, and only these, not the gauge fields themselves nor any other gauge-dependent quantities, are observable. Since there are various different approaches to the quantization of classical theories that are gauge theories in the sense just discussed but no all-encompassing systematic account of what counts as a "quantization procedure", the notion of a "quantum gauge theory" does not seem to have a precise definition at present. For the discussion of gauge symmetry breaking in the quantum context in Sections 5 and 6 I shall introduce and rely on the functional integral approach to the quantization of gauge theories, which provides the basis for the vast majority of physicists' studies of quantum gauge theories.

Symmetries that are gauge symmetries in the sense just discussed are often referred to as *local* symmetries, alluding to the fact that the parameters of symmetry transformations can be chosen "locally", that is, independently of each other for distinct space-time regions (see, for instance, the freedom in the choice of $\alpha(x)$ in Eq. (1)). However, the idea that variable configurations related by symmetry correspond to identical physical situations applies also in contexts where the symmetry transformations depend on only finitely many parameters, and one speaks of "global gauge symmetries" with regard to these cases, in contrast to the "local" gauge symmetries discussed before. Only theories that are formulated in terms of *local* gauge symmetry are commonly referred to as gauge theories, however. The present paper adopts this standard use of terminology, understanding "gauge symmetry" to refer to both local and global gauge symmetries and "gauge theory" to refer to theories formulated in terms of *local* gauge symmetry only.

Having reviewed the notions of symmetry in general and of gauge symmetry in particular, I now turn to the notion of *spontaneous symmetry breaking* ("SSB" in what follows). The basic idea behind this concept is that the mapping of the state space onto itself which is induced by a symmetry transformation does not in general map each single state onto itself. Put differently, this means that the state of a physical system need not have all the symmetries which the laws of motions governing its behaviour have. States for which this is the case are candidates for exhibiting SSB, and for the purposes of the present paper, where the focus is on ground states and thermal states of theories with infinitely many degrees of freedom, one may simply identify them with the spontaneously symmetry breaking ones.[5]

[5]See [Strocchi, 2008] for a rigorous textbook account of SSB that avoids both unnecessary technicalities and misleading simplifications. Roughly speaking, for a state to exhibit SSB in the rigorous sense specified in that work, it needs to take an infinite amount of energy to transform the system from one asymmetric state into another. This is the reason why realistic systems (that is, systems without any potential barriers of infinite height)

For quantum theories, the basic idea behind the concept of SSB just sketched can be turned into a precise criterion using the language of the algebraic approach to quantum theories. One defines that for a symmetry α of the algebra of observables of a system to be spontaneously broken by a state ω, the GNS representations associated with the states ω and $\alpha^*\omega$ (defined by $\alpha^*\omega(A) = \omega(\alpha(A))$) have to be unitarily inequivalent.[6] Intuitively, this means that the states ω and $\alpha^*\omega$ cannot be written in the form of density matrices in one and the same Hilbert space $\mathcal{H}$. An expectation value $\omega(A)$ of an observable A for which

$$\omega(A) \neq \alpha^*\omega(A) \tag{2}$$

is called a symmetry breaking order parameter for the symmetry α in the state ω. Situations where the symmetry α is broken are characterized by the fact that this quantity is nonzero, whereas it vanishes for states that are symmetric with respect to α. Symmetry breaking order parameters in the sense of Eq. 2 can be used to *define* SSB in contexts where the algebraic criterion does not apply in that the quantum theory at issue is not formulated in terms of the algebraic approach. This holds, for instance, for the application of the concept of SSB in the framework of the functional integral formulation of quantum theories in which the quantization of gauge theories is most commonly carried out (see Section 5 of this paper for more details).

The notion of a spontaneously broken *gauge* symmetry may seem puzzling at first sight. As formulated by Smeenk, "[i]f gauge symmetry merely indicates descriptive redundancy in the mathematical formalism, it is not clear how spontaneously breaking a gauge symmetry could have any physical consequences, desirable or not."[7] Part of the aim of the present paper is to remove the puzzlement expressed in Smeenk's statement and to elucidate the physical significance of SSB for gauge symmetries. For the purposes of the present section it suffices to clarify the notion of a spontaneously broken

---

need to have infinitely many degrees of freedom to exhibit SSB. Furthermore, it does not suffice for a non-symmetric state to differ only *slightly* from a symmetric one (in the sense in which, say, a single particle state differs from a zero-particle, fully symmetric vacuum state) to qualify as symmetry breaking. Cases like these are automatically excluded by the criterion in terms of the algebraic approach to quantum theories stated in the next paragraph.

[6] For accessible introductions to the notions of algebras of observables, their representations, the unitary (in-) equivalence of representations, and a state's GNS representation, see, for instance, ([Ruetsche, 20011] Chap. 13) and [Strocchi, 2008].

[7] See ([Smeenk, 2006] p. 488). See ([Earman, 2004] Section 9) for a similar way of putting the challenge.

gauge symmetry at a technical level, and to do so, the account just given for SSB on the level of observables must be generalized by extending the algebra of observables to an algebra of canonical variables that are not themselves physical observables. A simple example of a quantum theory with a spontaneously broken gauge symmetry is that of Bose-Einstein condensation of a non-relativistic free Bose gas in the thermodynamic limit at zero temperature.[8] Since this theory is formulated in terms of *global*, rather than local, gauge symmetry, it does not qualify as a gauge theory, but the spontaneous breaking of a gauge symmetry can nevertheless nicely be illustrated with it. In this case, the canonical variables are (quantum) fields $\psi(x)$, and the system has infinitely many pure ground states Ω_θ, labelled by different values of an angle variable θ, all of which assign a nonzero expectation value to the (improper) field operator $\psi(x)$:

$$\Omega_\theta(\psi) = \sqrt{n}e^{i\theta}, \qquad \theta \in [0, 2\pi)\,, \tag{3}$$

where n is the average density $n = |\Omega_\theta(\psi)|^2$.

Physically, all states Ω_θ defined in Eq. (3) are equivalent to each other in that they (and their mixtures) yield the same expectation values for all observable quantities. Gauge symmetry comes into play in the form of an invariance of the dynamics under global gauge transformations of the form

$$\begin{aligned} \psi(x) &\mapsto \alpha^\lambda(\psi(x)) = e^{i\lambda}\psi(x)\,, \\ \psi^*(x) &\mapsto \alpha^\lambda(\psi^*(x)) = e^{-i\lambda}\psi^*(x)\,, \end{aligned} \tag{4}$$

where $\lambda \in [0, 2\pi)$.

The states Ω_θ are not invariant under these transformations in that

$$\Omega_\theta(\alpha^\lambda(\psi)) = \Omega_{\theta+\lambda}(\psi) \neq \Omega_\theta(\psi) \tag{5}$$

for $\lambda \neq 0$, so they break the gauge symmetry. The GNS representations associated with the states Ω_θ are unitarily inequivalent, so the symmetry α defined in Eq. (4) is spontaneously broken by each Ω_θ. Gauge symmetry breaking is an unavoidable feature of one's description if one wants to describe the free Bose gas in the thermodynamic limit (at zero temperature) in terms of gauge variables by means of a pure state, but the states Ω_θ, among which one can choose, are all *physically* equivalent in that they assign the same expectation values to all observables.[9]

[8] The following presentation relies on ([Strocchi, 2008] Chap. 7.2). See also Chapters 13.3 and 13.4 of [Strocchi, 2008] for further details.

[9] [Leggett, 2006] provides an illuminating discussion of Bose-Einstein condensation in

3 Symmetry breaking as a natural phenomenon

Spontaneous symmetry breaking, as explained in the previous section, is a feature of states that do not have the all the symmetries of the underlying laws of motions (in theories with infinitely many degrees of freedom). In order to *interpret* the notion thereby defined, let us first focus on cases where the broken symmetry is one of the algebra of observables (that is, not a gauge symmetry), so that its breaking manifests itself as an asymmetry on the level of observables. Having in mind these cases of SSB, Liu and Emch characterize symmetry breaking by means of the non-technical and intuitive notion of a "natural phenomenon"[10], contrasting it with "merely theoretical concepts" such as "renormalization, first- or second- quantization."[11] Whenever the state of a system as specified in terms of the expectation values of its observables spontaneously breaks some symmetry of the underlying laws of motion, the discrepancy between the symmetries of the state and those of the laws is an objective feature of the physical situation described by that state and not merely an artefact of our description. Liu's and Emch's characterization of SSB as a "natural phenomenon" therefore seems adequate for cases of SSB on the level of observables in that the breaking of these symmetries, whenever it occurs, is an objective matter and not merely a conventional or otherwise arbitrary feature of how we represent the physical situation.[12]

---

the absence of the thermodynamic limit that does not operate with the notion of a spontaneously broken gauge symmetry. The assumptions underlying Leggett's approach are more realistic than those of the discussion in the main text, since the number of atoms in physically realized examples of BEC is not exceedingly large (between roughly 10^3 and 10^5) and there are important inter-particles interactions. Leggett's "order parameter" (see [Leggett, 2006] Eq. (2.2.1)), in terms of which he defines Bose-Einstein condensation, is not a symmetry breaking order parameter in the sense of Eq. (orderparam).

[10] See ([Liu and Emch, 2005] p. 153). Liu and Emch focus on *quantum* spontaneous symmetry breaking, specifically, but the characterization of symmetry breaking as a natural phenomenon does no to seem to be based on any particular quantum (as opposed to classical) aspects.

[11] See ([Liu and Emch, 2005] p. 153, fn. 14).

[12] Note that to accept the characterization of SSB on the level of observables in quantum theories as a natural phenomenon, it does not seem necessary to endorse the standard *ontic* view of quantum states as states quantum systems "are in". The main reason for adopting the alternative, epistemic, conception of quantum states is that it elegantly avoids the paradoxes of measurement and nonlocality. (See [Friederich, 2011] for more details and an exploration of how the view might be spelled out in detail.) According to the epistemic conception of quantum states, quantum states reflect the state assigning agents' epistemic relations to these systems, so no such thing as the "true" quantum state of a quantum system is acknowledged, and SSB cannot be characterized in terms of

While SSB on the level of observables seems adequately characterized as a "natural phenomenon" in the sense just discussed, the status of SSB on the level of gauge variables seems less clear. The reason for this is that gauge symmetries, as explained in the first section, are purely formal symmetries that have no physical instantiations. Whenever we describe some physical situation in terms of broken gauge symmetry, there is thus no discrepancy between the *physical* symmetries of the situation and those of the underlying laws of motion. This can nicely be seen, for instance, in the case of Bose-Einstein condensation mentioned at the end of the previous section, where the gauge symmetry is broken by any of the states Ω_θ, but the physical properties of the system, i. e., the expectation values of observables, are exactly the same for all Ω_θ. There is in this case no asymmetry in the physical, gauge-invariant, properties of the system which the underlying laws of motion do not have. In just the same sense in which gauge symmetries contrast with empirical symmetries in that they have no physical instantiations gauge symmetry *breaking* seems to be rather an aspect of how we describe a physical situation than an objective feature of the situation itself.

One may feel, however, that to conclude from these considerations that gauge symmetry breaking does not deserve to be characterized as a "natural phenomenon" in any reasonable sense would be too rash. More specifically, one may feel that whether some physical system is described in terms of *broken* or *unbroken* gauge symmetry relates directly to objective features of that system. Even though SSB does not seem to be an *intrinsic physical* feature of systems described in terms of broken gauge symmetry in the same way as it is for systems described in terms of a broken symmetry on the level of observables, it might nevertheless be regarded as an *extrinsic physical* feature of these systems in the sense that their physical characteristics may strongly differ from those of systems described in terms of *unbroken* gauge symmetry. Instances of gauge symmetry breaking, one might want to say, deserve to be called "natural phenomena" if and only if situations described in terms of broken gauge symmetry are qualitatively different from those described in terms of unbroken gauge symmetry. However, since both

---

quantum systems' "being in" quantum states that break some symmetry of the algebra of observables. Nevertheless, proponents of the epistemic conception of states can hold that SSB is a natural phenomenon in that an observable called a "witness" of a symmetry of the observables may have a value that, if known, requires the assignment of a state that breaks that symmetry. (For an explanation of the notion of an observable being a "witness" for SSB, see ([Liu and Emch, 2005] p. 145).) The characterization of SSB in quantum theories as a natural phenomenon seems therefore independent of the question of whether quantum states are conceived of as ontic or non-ontic.

the notion of a natural phenomenon and that of a "qualitative difference" between physical situations are only intuitive notions, this idea is in need of further qualification and should be made more precise.

A natural way of doing so is to say that gauge symmetry breaking qualifies as a "natural phenomenon" just in case the distinction between broken and unbroken gauge symmetry lines up completely with a distinction between two qualitatively different physical phases. Physical phases are regions in the space of parameters characterizing a theory (such as, for instance, particle masses, coupling constants, or temperature) in the interior of which the expectation values of macroscopic observables (derivatives of the Gibbs potential), written as functions of the parameters, vary only analytically. Boundaries between the different phases are called *phase transitions*.[13] Formulated in terms of this notion, the criterion for gauge symmetry breaking to qualify as a "natural phenomenon" stated above translates into the statement that it does so just in case the transition between broken and unbroken gauge symmetry coincides with a phase transition. Cases of SSB on the level of observables automatically count as natural phenomena in this sense, at least if there is a symmetry breaking order parameter in form of the expectation value of a macroscopic observable, which seems to be the case in all the typical cases of practical interest. In view of this tight connection between phase transitions and symmetry breaking it is not so surprising that the vocabulary of SSB is of crucial importance for our understanding and classification of phase transitions. An example of a phase transition that is accompanied by a change of a symmetry from broken to unbroken is the transition between a ferromagnetic and a paramagnetic phase of a magnetic material where the total magnetic moment of the system serves as an order parameter. This quantity is zero throughout the unbroken ("symmetric") regime but becomes nonzero in the broken regime and therefore must exhibit a non-analyticity (that is, a cusp or a jump) where the symmetry breaking occurs. For the breaking of a gauge symmetry, in contrast, it is not immediately clear on conceptual grounds whether it is necessarily accompanied by a non-analyticity on the level of observables, that is, by a phase transition. A more detailed investigation is required to decide whether specific instances

[13] Alternatively, one may reserve the notion of a phase transition for the physical *process* of crossing a phase boundary. This is the sense in which, for instance, cosmologists speak of phase transitions in the early universe. For a detailed and rigorous account of phase transitions in the sense of phase boundaries, see ([Sewell, 1986] Chapt. 4). Here I gloss over the difficulties of giving a rigorous account of thermodynamic notions such as the Gibbs potential in the relativistic, quantum field theoretical, context, assuming that at least for all practical purposes these difficulties can be met.

of broken gauge symmetries can count as natural phenomena in that sense.

For the case of Bose-Einstein condensation discussed in the previous section this question is settled rather easily. We saw that the ground states Ω_θ break the gauge symmetry α^λ for a free Bose gas at zero temperature. For temperatures T substantially higher than $T = 0$, however, the situation looks entirely different. Above a certain *critical temperature* T_c one finds that the expectation value $\Omega(\psi)$ vanishes, which may serve as a symmetry breaking order parameter, signalling that the gauge symmetry is unbroken above T_c. The most interesting question for present purposes is whether observable properties of the free Bose gas above the critical temperature T_c are qualitatively different from those below T_c. Clearly they are: Thermodynamic quantities such as the compressibility of the gas (which is infinite below T_c in the non-interacting case and nonzero yet finite above T_c) show qualitative differences below and above T_c, and the temperature dependence of its specific heat exhibits a cusp at T_c. Since for a free (i. e. non-interacting) system the many-particle states are just symmetrized products of the single-particle states, the microscopic origin of these features can be analysed in terms of occupation numbers of the single-particle states of the free bosons. For temperatures $T < T_c$ below the critical temperature the occupation number n_0 of the single-particle ground state (that is, the ground state for a single boson in the same volume) diverges, so that the ratio n_0/N remains finite even when the total number of particles N goes to infinity. At zero temperature, all particles have "condensed" into the single-particle ground state, so that $n_0 = N = n \cdot V$, where n is the particle density introduced in Eq. (3). For temperatures above the critical temperature T_c, in contrast, n_0/N goes to zero as N approaches infinity. The "condensation" of particles into the single-particle ground state vanishes together with the restoration of global gauge symmetry, as becomes manifest in the fact that n_0/N can be expressed in terms of the symmetry breaking order parameter. Therefore, in the case of Bose-Einstein condensation of a free Bose gas the distinction between broken and unbroken gauge symmetry corresponds exactly to a distinction on the level of macroscopic observables insofar as situations which are described in terms of broken gauge symmetry are separated by a phase transition from situations described in terms of unbroken gauge symmetry. In Section 6 of this paper I shall argue that this does not always hold for instances of symmetry breaking in gauge theories so that these do not (in general) qualify as natural phenomena in the (weak) sense introduced before in terms of phase transitions.

4 Local gauge symmetry breaking—the classical perspective

In this section, I briefly review the textbook account of the Higgs mechanism in classical terms as a spontaneously broken local gauge symmetry. To see the underlying idea, it suffices to consider, as an example, the Lagrangean of the Abelian Higgs model defined by

$$\mathcal{L} = D_\mu \phi^* D^\mu \phi - V(\phi) - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \,, \tag{6}$$

which exhibits a local $U(1)$ gauge symmetry in that it is invariant under gauge transformations of the form

$$\phi(x) \mapsto e^{i\alpha(x)}\phi(x), \qquad A_\mu(x) \mapsto A_\mu(x) - \frac{1}{e}\partial_\mu \alpha(x) \,. \tag{7}$$

The covariant derivative D_μ is defined as $D_\mu = \partial_\mu + ieA_\mu$, and the potential $V(\phi)$ in Eq. (6) is given by

$$V(\phi) = m_0^2 \phi^* \phi + \lambda_0 (\phi^* \phi)^2 \,. \tag{8}$$

If the coefficient of the term quadratic in the fields is taken to be negative, that is, if $m_0^2 < 0$, the potential V has a minimum at a nonzero value of the Higgs field ϕ, namely, $|\phi|^2 = -\frac{m_0^2}{2\lambda_0}$.

The classical ground states of the theory are configurations of the fields ϕ and A_μ of the form

$$\phi(x) = e^{i\theta(x)} v/\sqrt{2}, \qquad A_\mu(x) = -\frac{1}{e}\partial_\mu \theta(x), \tag{9}$$

where $\theta(x)$ is an arbitrary real-valued function of space and time and $v = \sqrt{-\frac{m_0^2}{\lambda_0}}$. For any two field configurations of the form Eq. (9) there exist gauge transformations of the form Eq. (7) that transform them into one another, so all these configurations are physically equivalent. However, since $v \neq 0$, the transformations (7) do not act trivially on these configurations, that is, none of the field configurations (9) is itself invariant under local gauge transformations. This means that local gauge symmetry is indeed spontaneously broken in any classical ground state of (6).

In order to extract the physical content of the theory defined by the Lagrangean (6), it is useful to perform the θ-dependent local gauge transformation

$$\begin{aligned} \phi(x) = e^{i\theta(x)}\rho(x) \quad &\mapsto \quad \rho(x) \,, \\ A_\mu(x) \quad &\mapsto \quad A_\mu(x) + \frac{1}{e}\partial_\mu \theta(x) \equiv B_\mu(x) \,, \end{aligned} \tag{10}$$

which makes it possible to eliminate the θ-field from the Lagrangean, so that it becomes

$$\mathcal{L} = \partial_\mu \rho \partial^\mu \rho - V(\phi) + \frac{1}{2} e^2 \rho^2 B_\mu B^\mu - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \,. \tag{11}$$

Expanding the field ρ around its expectation value as $\rho = v/\sqrt{2} + \eta$ and neglecting terms which are of third or higher order in the fields η and B_μ one obtains

$$\mathcal{L}^{(2)} = \frac{1}{2} \left(\partial_\mu \eta \partial^\mu \eta + m_0^2 \eta^2 \right) + \frac{1}{2} e^2 v^2 B_\mu B^\mu - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \,. \tag{12}$$

The characteristic physical properties of the theory defined by this Lagrangean can easily be read off in that it describes a massive vector boson with a mass $M_B = ev$ and a massive scalar boson with mass $\sqrt{-2m_0^2}$. The real field θ, which would have played the role of a Goldstone boson in the case of an invariance under *global* gauge symmetries, has been eliminated, and this shows that there are no massless scalar particles in the theory. According to how this is often expressed, the Goldstone boson has been "eaten" by the gauge field. The Lagrangean (12) contains only gauge-invariant fields[14], and, from a classical point of view at least, one could have defined the theory directly in terms of these without introducing gauge symmetry at all.[15] Classically, as we see, the Higgs mechanism can be spelled out either in terms of broken local gauge symmetry or without introducing gauge symmetry in the first place. In the formulation using local gauge symmetry, as discussed before, it is broken in any classical ground state.

In the electroweak theory part of the standard model the implementation of the Higgs mechanism is slightly more complicated than in the case just discussed in that the broken local symmetry is a (non-Abelian) $SU(2)\times U(1)$ symmetry instead of the simpler (Abelian) $U(1)$ symmetry of our example. Moreover, the $SU(2) \times U(1)$ symmetry is not completely broken by the Higgs field, but only up to a residual $U(1)$ symmetry, which coincides with the gauge symmetry of electromagnetism. Despite these important conceptual differences, however, the conclusion just established that the Higgs mechanism can be described as a case of a spontaneously broken local gauge

[14] Equivalently, one could have arrived at a Lagrangean of exactly the same form by imposing the unitary gauge $\theta = 0$.

[15] See [Struyve, 2011] for a detailed discussion of these questions. If one chooses to use only gauge-invariant fields, however, one has to pay careful attention to the constraints for the variable η occurring in Eq. (12), see ([Struyve, 2011] Section 4) and ([Strocchi, 2008] p. 194). The analysis in terms of the constraint Hamiltonian approach given in ([Struyve, 2011] Section 7) avoids this problem.

symmetry is not affected and remains correct for the classical version of the electroweak theory. Leaving aside the classical context from now, I turn to the fate of spontaneously broken local gauge symmetry in *quantized* gauge theories.

5 Quantization without gauge fixing

At present we do not have any rigorous formulation of quantum gauge theories in the framework of the algebraic approach to quantum theories, so the status of symmetry breaking in quantized gauge theories has to be discussed within a different framework. Since functional integral quantization seems to be the most common and most convenient approach to the quantization of gauge theories, especially in the Non-Abelian case, it is taken as the basis for the discussion of symmetry breaking in quantized gauge theories in the present and following sections. The existence of non-vanishing symmetry breaking order parameter in the sense of Eq. (2) provides the criterion for SSB in this context.

In the functional integral formulation of quantum field theory, all expectation values of the observables and fields are obtained as derivatives of a generating functional $W[\eta_i]$ that depends on the so-called "source fields" η_i. In the case of a gauge theory with gauge field A_μ and scalar field ϕ this functional can be written as a functional integral (that is, as an integral over all possible field configurations) of the form

$$W[\eta, J] = N \int \mathcal{D}\phi \mathcal{D}A_\mu \, \exp\left(i \int \mathrm{d}^4 x (\mathcal{L} + \eta\phi + J_\mu A^\mu) \right) , \tag{13}$$

where $\mathcal{L}$ is the Lagrangean of the theory to be quantized and $S = \int \mathrm{d}^4 x \mathcal{L}$ the corresponding action. Correlation functions, which include the expectation values of gauge-dependent quantities that may serve as symmetry breaking order parameters (such as $\langle\phi\rangle$, where ϕ is the Higgs field), are obtained from $W[\eta, J]$ by taking derivatives with respect to the source fields η and J and then setting them to zero.

The expression (13) for $W[\eta, J]$ involves an integral over all possible configurations of the fields ϕ and A_μ, which means that each gauge-equivalent class of field configurations is integrated over infinitely often. As a result, the integral in Eq. (13) diverges in a "vicious way" in that the inverse free propagator, a function that is contained in the exponent of the integrand, cannot be inverted to obtain the free propagator itself that is required as a starting point for perturbative calculations.[16] Non-perturbative calculations that do not require the inverse free propagator in the exponent of the functional integral Eq. (13) to be invertible can be performed by starting from Eq. (13), but this requires the setting of lattice gauge theory, where the gauge theory to be quantized is not formulated on the space-time continuum but rather on a discrete lattice of space-time points.

There are at least two different possible reactions to this problem, which I shall consider in the present and following section, respectively. The first of these is to choose the non-perturbative route, which in practice means to quantize the theory on a lattice and to extrapolate the results to the continuum case by letting the lattice spacing go to zero; the second reaction, discussed in the following section, is gauge fixing—the insertion of terms in the functional integral that violate gauge invariance, but in such a way that correlation functions for gauge invariant quantities are independent of the choice of gauge fixing terms. Since local gauge symmetry is explicitly broken by gauge fixing terms, one has to consider the option without gauge fixing to assess the fate of local gauge symmetry breaking in quantized gauge theories. The next section will focus on the breaking of post-gauge fixing remnant *global* gauge symmetries that, depending on the choice of gauge, survive in the presence of gauge fixing terms.

In gauge theories that are quantized on a lattice, scalar and fermion fields are defined on a lattice representing discretized space-time, and the gauge fields are defined on the links connecting the lattice sites.[17] By considering finite lattices with periodic boundary conditions the functional integrals can be evaluated explicitly in a non-perturbative way, so that no expansion of expressions like Eq. (13) needs to be made that requires a free propagator for the gauge field. This means, in other words, that functional integral quantization can be carried out in the lattice setting without gauge fixing, so that gauge invariance is not violated during the quantization procedure. Since local gauge symmetry persists after quantization, it is possible to discuss whether *local* gauge symmetry can be broken in gauge theories that are

[16]The inverse free propagator of the gauge fields can be thought of as the coefficient of the part in the Lagrangean $\mathcal{L}$ which is quadratic in the gauge fields. For the Abelian case this part of the Lagrangean is given by $-\frac{1}{4}\left(\partial_m A_\nu - \partial_\nu A_\mu\right)^2$, and the resulting inverse free propagator for the gauge field is, in momentum representation, given by $\eta_{\mu\nu}k^2 - k^\mu k^\nu$ (where $\eta_{\mu\nu}$ corresponds to the matrix diag$[1,-1,-1,-1]$). As remarked in the main text, the operator $\eta_{\mu\nu}k^2 - k^\mu k^\nu$ is not invertible, which can be seen from the fact that it has k_ν as an eigenvector with eigenvalue zero.

[17]For the earliest introduction of lattice gauge models, see [Wegner, 1971]. Lattice gauge theory as sketched in this paragraph was essentially invented by Wilson, see [Wilson, 1974]. For a gentle modern introduction to lattice gauge theory, see Münster and Walzl [2000].

quantized in this way.

The most important result in this context, mentioned already in the introductory section of this paper, is a theorem due to Elitzur[18], which states that local gauge symmetry cannot be spontaneously broken at all. Mathematically, what the theorem says is that for any local gauge transformation $\alpha : \phi \mapsto \alpha(\phi)$ the vacuum expectation value of the Higgs field ϕ is invariant under the gauge transformation α in the sense that

$$\langle\phi\rangle = \langle\alpha(\phi)\rangle\,, \tag{14}$$

from which it follows that $\langle\phi\rangle = 0$. Since the proof of the theorem can be generalized to show that the expectation values of *arbitrary* gauge-dependent combinations of fields must be zero, this means that there can be no spontaneous breaking of local gauge symmetry at all.[19] The proof of the theorem is crucially based on the fact that local gauge transformations depend on infinitely many parameters and does not extend to the case of *global* gauge symmetries, which depend on only finitely many parameters. Thus, the impossibility of breaking local gauge symmetries is not a direct consequence of the general unobservability of gauge transformations but has to do with the specific features of *local* symmetries.

A possible reaction to Elitzur's theorem, tempting perhaps for those who are used to think of what is usually called the Higgs mechanism as a spontaneously broken local gauge symmetry, is to regard the theorem as an embarrassment for lattice gauge theory rather than as a reductio of that way of conceiving the Higgs mechanism. The temptation to make this move, however, should be resisted on at least two grounds. The first is that if we want to assess the fate of the notion of a spontaneously broken local gauge symmetry in quantized gauge theories at all, we must do so in the context of an approach where local gauge symmetry is not explicitly broken from the start, that is, an approach to the quantization of gauge theories that does not rely on gauge fixing. The Wilsonian lattice formulation of gauge theory fulfills this requirement and provides the natural framework for investigating the status of (allegedly broken) local gauge symmetries, especially in the absence of a workable approach to quantization without gauge fixing for the continuum case. The second reason for not dismissing

[18] Elitzur proved the theorem for the case of a Higgs field with fixed modulus, see [Elitzur, 1975]. The result was generalized to the case of a Higgs field with variable modulus by de Angelis, de Falco and Guerra, see [De Angelis et al., 1978].

[19] See ([Strocchi, 1985], Chap. II 2.5) and ([Fröhlich et al., 1981], Section 3) for helpful sketches of the generalized version of the proof and ([Itzykson and Drouffe, 1989], Chap. 6.1.3) for a more rigorous textbook version applied to the special case of Eq. (14).

Elitzur's theorem as an oddity of the lattice formulation is that the crucial ingredient of its proof—the fact that local gauge transformations, in contrast to global ones, depend on an infinity of parameters labelled by the points of space-time—carries over to the continuum case, where the number of independent parameters characterizing a gauge transformation is even non-denumerably large. Another important aspect is that the proof does not seem to depend on any peculiarities of the lattice formulation of which one could be sure that they would not be valid in whatever precise formulation one might have for the continuum case without gauge fixing in the future.

Elitzur's theorem raises the question of whether the Higgs mechanism may perhaps not work as an account of mass generation in the standard model as it shows that the notion of a spontaneously broken local gauge symmetry is not sound, on which textbook expositions of the Higgs mechanism are commonly based. Fortunately, however, as demonstrated by Fröhlich, Morchio, and Strocchi[20], such fears are ungrounded, since the physical phenomena which are usually associated with the Higgs mechanism can be recovered in terms of an approach that uses only entirely gauge-invariant fields. They develop a recipe for constructing gauge-invariant combinations of the Higgs and gauge fields that allows to reformulate any gauge theory in terms of such gauge-invariant combinations. Observable quantities, such as the (Yukawa) couplings between the gauge bosons and fermions in the conventional formulation, are obtained as functions of expectation values of these gauge-invariant fields. In particular,Fröhlich et al. provide a list of gauge-invariant quantities that are non-vanishing and correspond directly to quantities identifiable with the particle masses in the conventional formulation using gauge dependent fields. So, mass generation through the Higgs mechanism can get along very well without assuming a nonzero expectation value for any gauge dependent combination of fields, in particular not for the Higgs field itself.

One may conclude from the fact that mass generation through the Higgs mechanism, as demonstrated by Fröhlich, Morchio and Strocchi, can be accounted for in terms of gauge-invariant fields that to characterize the Higgs mechanism as a spontaneously broken local gauge symmetry is, as Smeenk puts it, a "relatively benign case of abuse"[21] of terminology. An alternative conclusion to draw, however, would be that the abuse of terminology involved in characterizing the Higgs mechanism as a spontaneously broken local gauge symmetry is not so benign—after all, the notion is demonstrably

[20]See [Fröhlich et al., 1981].

[21]See ([Smeenk, 2006] p. 498).

vacuous—, but that despite its being vacuous the notion of a spontaneously broken local gauge symmetry has an important heuristic value, at least historically, and may still be useful for semi-classical calculations.

Another worry that might be brought up by Elitzur's theorem is that if we do not dispose of the notion of a spontaneously broken local gauge symmetry, we can no longer make the important distinction between cases where local gauge symmetry is broken and cases where it is unbroken ("restored"). This distinction, however, is apparently crucial to describe the *electroweak phase transition*, a phase transition between two different phases of the universe as described by the electroweak theory at different values of the fundamental parameters such as temperature and the Higgs boson mass. This transition is widely believed to have actually taken place as temperature decreased in the history of the early universe so that it supposedly evolved from a phase where the $SU(2) \times U(1)$ local gauge symmetry of electroweak theory is unbroken to the phase in which we now exist, where that symmetry is allegedly broken.[22] In the phase where the electroweak symmetry is said to be unbroken ("restored") the electron and the neutrino are not yet distinguishable in that they correspond to degenerate states of one and the same particle. At the present state of the universe, in contrast, there is obviously a substantive physical difference between the electron and the neutrino, so the supposed phase transition seems to have taken the universe from one phase to another, qualitatively very different, one. Do we have to conclude from Elitzur's theorem that the very idea of an electroweak phase transition rests on an error in that there cannot be a transition from a situation where electroweak symmetry is unbroken to a situation where it is broken since electroweak symmetry can never be broken at all?

Fortunately, this conclusion need not be drawn since the electroweak phase transition, just as the Higgs mechanism itself, can be described in purely gauge invariant terms. An example of an observable, that is, gauge-

[22]Detailed calculations (see, for instance, [Kajantie et al., 1996]) have shown that for values of the Higgs mass not excluded by experiment the electroweak phase transition is actually not a real phase transition (in the sense of an abrupt change in thermodynamic quantities) but rather a steep crossover between two qualitatively different regimes of electroweak theory, meaning that at least some expectation values of observables vary very strongly (yet analytically) from one regime to the other. In the context of the present paper, however, the question of whether, for realistic values of the Higgs mass, the electroweak phase transition is a genuine phase transitions or rather a continuous crossover is not important since we are concerned here with the more general questions of whether the notion of a spontaneously broken local gauge symmetry is needed to give meaning to the distinction between the high and low temperature phases of the electroweak theory, which are sharply separated for *some* values of the Higgs mass.

invariant quantity that may quite drastically change at the phase transition is the expectation value $\langle\phi^*\phi\rangle$ (where ϕ is the Higgs field), which, if displayed as a function of parameters such as temperature and the Higgs boson mass, exhibits a "jump" across the planes in the phase diagram where the electroweak phase transition occurs.[23] From the fact that phase transitions are often accompanied by the breaking (or restoration) of certain symmetries and the fact that the electroweak phase transition is often associated with "electroweak symmetry breaking" one might mistakenly conclude that there is an incompatibility between Elitzur's theorem and the electroweak phase transition. As we have just seen, however, this is not the case, for the distinction between the two different phases, one where electroweak symmetry is allegedly broken and one where it is allegedly unbroken, can be made in an entirely gauge invariant way so that the dubious notion of a spontaneously broken local gauge symmetry is altogether avoided. Phase transitions are indeed often accompanied by instances of symmetry breaking, but the definition of a phase transition in terms of non-analytic behaviour of observable quantities does not require symmetry breaking. The electroweak phase transition, as we see, is a case in point.

The topic of phase transitions in gauge theories will concern us again in the following section while discussing the role of spontaneous symmetry breaking in the presence of gauge fixing terms.

6 Gauge fixing and symmetry breaking

Having discussed the quantization of gauge theories without gauge fixing in the lattice formulation of gauge theories, I now turn to their quantization by means of gauge fixing terms, which makes it possible to perform perturbative computations using the diagrammatic techniques invented by Feynman in the continuum as well as on the lattice. In classical gauge theories, gauge fixing amounts to the implementation of constraints for the Higgs and gauge fields such as, for instance, the unitary gauge mentioned in Section 4, which fixes the phase of the Higgs field at a constant value, say zero, at any space-time point. For the Higgs field in the Abelian Higgs model discussed in Section 4, which can be written as $\phi(x) = e^{i\theta(x)}\rho(x)$, this means setting $\theta(x) = 0$ for all x. Other choices of gauge fixings that tend to be better suited for practical calculations include the Coulomb gauge, defined by $\partial_i A^i = 0$ (where the summation is over spatial indices only), and the Lorenz gauge, defined by $\partial_\mu A^\mu = 0$.

[23] See, for instance, [Buchmüller et al., 1994] pp. 134-6.

In the functional integral formulation of quantum field theory, gauge fixing is implemented in the form of field-valued Dirac-δ-functions in the functional integral. The introduction of these δ-functions can be seen as part of a change of integration variables involving a Jacobi determinant, the so-called Faddeev-Popov determinant $\Delta(A)$, and it requires, at least in certain gauges, the introduction of additional, purely formal, fields as integrations variables. These are the so-called ghost fields, which do not correspond to any physical degrees of freedom.[24] The original gauge-invariant action S of the gauge theory to be quantized (corresponding to the integral in the exponent of Eq. 13) is replaced by an "effective" action S_{eff} of the form

$$S_{\mathit{eff}} = S + S_{gf} + S_{ghost} \,, \tag{15}$$

where S_{gf} implements the gauge fixing in that it contains the gauge fixing constraint and S_{ghost} is an additional term in the presence of ghost fields.

The gauge fixing term S_{gf} in the "effective" action S_{eff} explicitly violates local gauge invariance in that some gauge-dependent term is inserted in the functional integral. This is done in such a way that the physical content of the theory remains unchanged, so the gauge fixing does not have any physical significance whatsoever. However, the way in which local gauge invariance is violated by this procedure depends on the choice of gauge fixing made. One possibility is that the gauge freedom is *completely* eliminated by the gauge fixing in the sense that out of any class of gauge-equivalent field configurations exactly one is singled out by the gauge fixing constraint. This is the case for the unitary gauge, which, in the case of the locally $U(1)$-symmetric Abelian Higgs model discussed before, is given by $\theta(x) = 0$. Here, local gauge symmetry is eliminated completely (and explicitly) at the level of the "effective" action S_{eff}, so *spontaneous* symmetry breaking cannot occur any more, for there simply is no unbroken symmetry left to be broken.

For other choices of gauge fixing terms, however, the action S_{eff} can still be invariant under symmetry transformations corresponding to some finite-parameter subgroup of the original infinite-parameter local gauge group. In the presence of gauge fixing terms of this class, the action S_{eff} still exhibits certain *global* gauge symmetries, but no longer a local one. The spontaneous breaking of global symmetries is not forbidden by Elitzur's theorem, and indeed the breaking of these remnant global gauge symmetries is a common phenomenon in gauge theories in the presence of gauge fixing. In what follows, I will refer to it as the spontaneous breaking of "global subgroups" of

[24]This can be seen, for instance, from the fact that ghost fields formally correspond to spinless fermion fields the physical existence of which is excluded by the spin-statistics theorem.

the original, local, gauge group or just "remnant symmetry breaking". It can also be studied in the formulation without gauge fixing, discussed in the previous section, by introducing fields which depend on the spacetime variable x not only in an explicit manner, but also implicitly, via an additional dependence on the gauge fields. An example of such a field is[25]

$$\Phi(x; A) = g(x; A)\phi(x)\,, \tag{16}$$

where $\phi(x)$ is the Higgs field and $g(x; A)$ is a transformation that transforms it into a chosen gauge such as, say, the Coulomb or Landau gauge. The so defined $\Phi(x; A)$ has a nonzero expectation value just in case the Higgs field $\phi(x)$ itself has a nonzero expectation value for the respective choice of gauge fixing, that is, for the choices mentioned, in the Coulomb or Landau gauge.

Since local gauge symmetry cannot be spontaneously broken according to Elitzur's theorem, the breaking of these remnant global subgroups is the only way in which gauge symmetries can be broken in quantized gauge theories.[26] To answer the question of whether gauge symmetry breaking in quantized gauge theories can count as a natural phenomenon in the sense spelled out in Section 3 in terms of phase transitions, we therefore have to investigate whether the distinction between broken and unbroken remnant gauge symmetry always lines up with a contrast between distinct physical phases. We have to ask, in other words, whether the transition from unbroken to broken global subgroups is always accompanied by an abrupt change in the expectation values of some observables.

Even though there does not seem to be any rigorous statement about the relation between (remnant) symmetry breaking and the occurrence of phase transitions in gauge theories, there is strong evidence, based on a combination of exact and numerical results, that there is *no* rigid connection between the two and that, therefore, remnant gauge symmetry breaking does not in general qualify as a natural phenomenon in the sense specified in Section 3. A particularly illuminating discussion of the relation between the breaking of remnant subgroups and phase transitions is given by Caudy and Greensite in the context of a study of an $SU(2)$-symmetric lattice gauge model

[25] The example taken is Eq. (1.1) in [Caudy and Greensite, 2008].

[26] There are other, non-gauge, symmetries which can be broken is quantized gauge theories such as, for instance, chiral symmetry in QCD or centre symmetry in non-Abelian gauge theories (the centre of a group is the set of elements which commutes with all other elements), which seems to be linked to the confinement-deconfinement phase transition, see [Greensite, 2011]. The present paper is not concerned with the breaking of these symmetries but only with that of gauge symmetries.

with a fixed-modulus Higgs field.[27] For this model, there is robust numerical evidence that there exists, in a limited region of the phase diagram, a phase transition between a "Higgs phase", where the spectrum exhibits a gauge boson mass, and a "non-Higgs phase", where there is no such mass and the properties of the model are more similar to those of quantum chromodynamics (QCD) in the presence of confinement.[28] The main conclusion drawn by Caudy and Greensite from their results is that there is no *general* agreement between the two transition lines (that between the different phases and that between broken and unbroken remnant gauge symmetry), even though for *some* values of the parameters of the model the transition between the two phases does coincide with that between a regime where remnant symmetry is broken and one where it is unbroken.

This conclusion has two distinct interesting aspects the first of which is that, according to the results reported by Caudy and Greensite, both in Coulomb and Landau gauge part of the separation line between broken and unbroken gauge symmetry is found for parameters where the existence of an accompanying phase transition can be definitely excluded.[29] Remnant gauge symmetry breaking, thus, is not always linked to a transition between distinct physical phases as in the Bose-Einstein case discussed in Section 3 in that the transition between broken and unbroken remnant subgroups can occur in regimes where all observables vary analytically. This shows that remnant symmetry breaking is not in general a natural phenomenon in the sense specified in Section 3. A second interesting aspect of the conclusions

[27]See [Caudy and Greensite, 2008]. More precisely, their results are for a model with a fixed-modulus Higgs field in the fundamental colour representation. Their results clearly show that in a certain range of parameters the system exhibits the typical features of a "Higgs phase" such as, for instance, the appearance of a massive spectrum associated with the gauge field degrees of freedom, even though there is no "Mexican hat potential" (which makes sense only for a Higgs field with a variable modulus).

[28]See [Greensite, 2011] for an introduction to the problem of confinement that includes an in-depth discussion of how confinement should be defined in the first place.

[29]It is known from an exact result due to Fradkin and Shenker [1979] that in the model considered by Caudy and Greensite, for any two pairs of gauge and Higgs couplings β and γ, there exists a continuous path in the β-γ-plane along which the expectation values of all observables vary analytically. This implies that the phase boundary that separates the "Higgs phase" from the "confinement phase" cannot be such that it divides the phase diagram into a pair of half-planes, but that it rather must have an endpoint, just as the phase transition between the liquid and gaseous phases in the typical phase diagram of ordinary matter has a (critical) endpoint beyond which the distinction between liquid and gas is only gradual. Caudy and Greensite find the distinction between regimes with broken and unbroken remnant gauge symmetries to coincide *in part* with the phase transition between the Higgs and confinement regimes, but they also find it continuing beyond the endpoint of that transition line for parameters where all observables vary only analytically.

presented by Caudy and Greensite is that, according to their results, the values of the parameters (couplings of the theory) for which there is a transition between unbroken and broken gauge symmetry are dependent on the *choice* of remnant subgroup, that is, if gauge fixing is used, on the choice of gauge fixing terms. As Caudy and Greensite conclude, gauge symmetry breaking in gauge theories is "ambiguous" in the sense that whether or not remnant gauge symmetry is broken for a specific choice of parameters can depend on the (from a physical point of view) arbitrary choice of remnant subgroup. This observation illustrates further why remnant symmetry breaking does not deserve to be called a "natural phenomenon" in that whether or not it occurs for a given choice of parameters depends on the unphysical (gauge) freedom of description.

In the following, final, section of this paper, I consider some consequences of the considerations presented in this and the previous sections for philosophical debates about the interpretation of gauge symmetries and their breaking.

7 Philosophical implications

The considerations on gauge symmetry breaking presented in the previous sections have interesting philosophical ramifications. In particular, they imply that some interpretive claims about gauge symmetries and their breaking in the literature are misleading. I discuss three examples of such claims.

The first example is Peter Kosso's contention that broken gauge symmetries belong to the class of cases where "the relevant laws of nature are exactly symmetric, but the phenomena expressing these laws are not."[30] That this characterization cannot really be adequate follows already from the fact that gauge symmetries have no physical instantiations. If a theory such as that of the Bose gas discussed in Section 3 has ground states that break (global) gauge symmetry, all these ground states are still physically equivalent in that with respect to observable quantities they all assign the same expectation values. Kosso's question of why we should think that the fundamental interactions of nature are "gauge symmetric" even though the phenomena which we observe are not is misleading since there is no asymmetry in the phenomena that is not found in the basic laws due to the fact that gauge symmetries are purely formal and hence unobservable. The defence of the Higgs mechanism as an account of mass generation in the standard model may still raise interesting epistemological challenges, but

[30]See ([Kosso, 2000], p. 359).

this has nothing to do with the issue of conjecturing the fundamental laws to be symmetric in a way in which the phenomena we observe are not.

A number of claims on the nature and role of gauge symmetry breaking in gauge theories are based on failure to take into account Elitzur's theorem and the fact that whether or not the Higgs field has a nonzero expectation values depends on the choice of gauge fixing. Margaret Morrison, for instance, argues that the Higgs mechanism is "based on the idea that even the vacuum state can fail to exhibit the full symmetries of the laws of physics."[31] As a claim about ideas that have historically played a role in the development of the Higgs mechanism this statement may be true, but Morrison argues further that even from a methodological point of view "one needs the underlying vacuum assumptions regarding the plenum and degeneracy as part of the 'physical' picture."[32] An integral part of this picture, as she claims, is the thought that here "we are dealing with fields whose *average value* is non-zero, where the vacuum is said to have a non-zero expectation value."[33] This statement, as we have seen, is not correct in that, as we know from Elitzur's theorem, the vacuum expectation of the Higgs field *is* actually zero in the absence of any gauge fixing, whereas in the presence of gauge fixing it depends on the choice of gauge fixing which, practical considerations aside, is arbitrary from a physical point of view. Morrison's central conclusion that "it would be folly to accept a robust physical interpretation of the SSB story"[34] in the electroweak theory is quite plausible (depending on what exactly is meant by "robust physical interpretation"), but the reason she gives for drawing the conclusion, namely, "that the various vacuum hypotheses which provide the necessary theoretical foundations are essentially problematic, for both physical and philosophical reasons"[35], is not completely convincing. The problematic aspect of the notion of spontaneous symmetry breaking in the context of the $SU(2) \times U(1)$ symmetry of the electroweak theory is not that it is based on a questionable "theoretical story about the nature of the vacuum"[36], but that the $SU(2) \times U(1)$ *local* gauge symmetry is in fact unbroken, whereas the breaking of remnant subgroups depends on the gauge fixing.

Misunderstandings about the nature and significance of SSB in gauge theories can be found not only among philosophers but also among eminent

[31] See ([Morrison, 2003] p. 356).
[32] See ([Morrison, 2003] p. 357).
[33] See ([Morrison, 2003] p. 359).
[34] See ([Morrison, 2003] p. 361).
[35] See loc. cit.
[36] See loc. cit.

physicists. Steven Weinberg, for instance, argues in a ground-breaking paper on phase transitions in gauge theories that these phase transitions have the "philosophical implication" as regards the "reality" of gauge symmetries that "if a gauge symmetry becomes unbroken for sufficiently high temperature, it becomes difficult to doubt its reality."[37] Weinberg's reasoning here seems to be that if gauge symmetries exist in both broken and unbroken forms in such a way that there is a substantial physical difference between the two cases (that is, a phase transition that separates them), these symmetries are the bearers of non-trivial physical properties and, therefore, must be real. Although there may be disagreement about the sense in which gauge symmetries are supposedly established as "real" according to this line of thought, it seems clear from the considerations presented in the previous sections that Weinberg's argument fails, whatever exactly it is supposed to show, for several reasons. *Local* gauge symmetry, as we know from Elitzur's theorem, is never broken in quantized gauge theories, so phase transitions such as the electroweak phase transition cannot be described in terms of its breaking and the existence of phase transitions cannot have any implications whatsoever for the reality of local gauge symmetries. Remnant global subgroups of local gauge groups, on the other hand, may break spontaneously, but their breaking is ambiguous in that it depends on the gauge and is not necessarily accompanied by a qualitative change in physical properties. It seems therefore problematic to regard these global symmetries as the true bearers of physical properties and thus as "real" in a more substantial sense than the original, local, symmetries. The standard view of gauge symmetries as purely formal symmetries which do not have physical instantiations, in particular, is not in the least called into question by the result that there are phase transitions in gauge theories at high temperatures which for certain choices of gauge fixing are accompanied by a restoration of remnant gauge symmetry.

8 Conclusion

The aim of this paper has been to clarify the status and significance of gauge symmetry breaking in gauge theories. While local gauge symmetry itself cannot break spontaneously in quantized gauge theories according to Elitzur's theorem, this does not hold for remnant global gauge symmetries under which the action of a gauge theory typically remains invariant after gauge fixing. The physical significance of these instances of symmetry break-

[37]See ([Weinberg, 1974] p. 3359).

ing was considered by investigating their relation to transitions between distinct physical phases. Based on the results of [Caudy and Greensite, 2008] it was argued that there seems to be no general fixed connection between remnant gauge symmetry breaking and phase transitions in that, first, a transition between broken and unbroken remnant gauge symmetry can exist without any accompanying discontinuous change in the expectation values of observables and, second, the breaking of remnant gauge symmetry may depend on the choice of gauge fixing made.[38]

With respect to the Higgs mechanism the following two conclusions can be drawn from the considerations presented: The first is that the standard textbook characterization of the Higgs mechanism as a spontaneously broken *local* gauge symmetry is misleading (even though useful from a heuristic point of view) in that it is valid only for the classical, not for the quantum, case. The second is that while remnant *global* gauge symmetries may indeed be broken in regimes that exhibit the typical features of a "Higgs-phase", it does not suffice to detect the breaking of a remnant global symmetry to establish that these features actually hold. A more direct inspection of objective, that is, gauge-invariant, quantities remains necessary.

Acknowledgements

I would like to thank Kerry McKenzie, Robert Harlander, Dennis Lehmkuhl, Holger Lyre, Michael Kobel, Michael Krämer, Michael Stöltzner and Ward Struyve for very helpful comments on earlier drafts of this paper. Furthermore, I am grateful to Gernot Münster and Franco Strocchi for useful answers to questions I raised.

References

Buchmüller, W., Fodor, Z., and Hebecker, A., Gauge invariant treatment of the electroweak phase transition, *Physics Letters B* 331:131-6 (1994),

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

Caudy, W. and Greensite, J., Ambiguity of spontaneously broken gauge symmetry, *Physical Review D* 78:025018 (2008),

[38] Or, equivalently, on the choice of gauge transformation $g(x;A)$ as in Eq. (16), used to define a remnant subgroup of the original local gauge group.

De Angelis, G. F., De Falco, D., and Guerra, F., Note on the Abelian Higgs-Kibble model on a lattice: Absence of spontaneous magnetization, *Physical Review D* 17:1624-28 (1978),

Earman, J., Laws, symmetry, and symmetry breaking: Invariance, conservation principles, and objectivity, *Philosophy of Science* 71:1227-41 (2004),

Elitzur, S., Impossibility of spontaneously breaking local symmetries, *Physical Review D* 12:3978-82 (1975),

Fradkin, E. and Shenker, S. H., Phase diagrams of lattice gauge theories with Higgs fields, *Physical Review D* 19:3628-97 (1979),

Friederich, S. (2011), How to spell out the epistemic conception of quantum states, *Studies in History and Philosophy of Modern Physics*, 42(3):149-157,

Fröhlich, J., Morchio, G., and Strocchi, F. Higgs phenomenon without symmetry breaking order parameter, *Nuclear Physics B* 190:553-82 (1981),

Greensite, J., *An Introduction to the Confinement Problem*, Springer: Berlin (2011),

Healey, R., *Gauging what's real: The conceptual foundations of contemporary gauge theories*, Oxford University Press: New York (2007),

Itzykson, C. and Drouffe, J.-M., *Statistical Field Theory. Vol.1: From Brownian motion to renormalization and lattice gauge theory*, Cambridge University Press: Cambridge (1989),

Liu, C. and Emch, G. G., Explaining quantum spontaneous symmetry breaking, *Studies in History and Philosophy of Modern Physics* 36:137-63 (2005),

Lyre, H., Holism and structuralism in $U(1)$ gauge theory, *Studies in History and Philosophy of Modern Physics* 35:643-70 (2004),

Lyre, H., Does the Higgs mechanism exist?, *International Studies in the Philosophy of Science* 22:119-33 (2008),

Kajantie, K., Laine, M., Rummukainen, K., and Shaposhnikova, M., Is There a Hot Electroweak Phase Transition at mH $\gtrsim$ mW?, *Physical Review Letters* 77:2887-90 (1996), Kosso, P., The epistemology of spontaneously broken symmetries, *Synthese* 122:359-76 (2000),

Leggett, A. J., *Quantum Liquids*, Oxford University Press: Oxford (2006),

Morrison, M., Spontaneous symmetry breaking: theoretical arguments and philosophical problems, in: [Brading and Castellani, 2003], pp. 347-63, (2003),

Münster, G. and Walzl, M, Lattice gauge theory—a short primer, Lectures given at the PSI Zuoz Summer School 2000, available at: `http://arxiv.org/abs/hep-lat/0012005`, (2000),

Noether, E., Invariante Variationsprobleme, *Nachrichten der königlichen Gesellschaft der Wissenschaften zu Göttingen, Mathematisch-physikalische Klasse* Book 2:235-57 (1918),

Redhead, M., The interpretation of gauge symmetry, in: Kuhlmann, M., Lyre, H., and Wayne, A. (eds.) *Ontological Aspects of Quantum Field Theory*, World Scientific: Singapore (2002),

Ruetsche, L., *Interpreting Quantum Theories*, Oxford University Press: Oxford (2011),

Sewell, G. L., *Quantum Theory of Collective Phenomena*, Clarendon Press: Oxford (1986),

Smeenk, C., The elusive Higgs mechanism, *Philosophy of Science* 73:487-99 (2006),

Strocchi, F., *Elements of Quantum Mechanics of Infinite Systems* , World Scientific: Singapore (1985),

Strocchi, F., *Symmetry Breaking*, 2nd edition, Springer: Berlin, Heidelberg (2008),

Struyve, W., Gauge invariant accounts of the Higgs mechanism, *arXiv.org:1102.0468* (2011),

't Hooft, G., The conceptual basis of quantum field theory, in: Butterfield, J. and Earman, J. (eds.) *Philosophy of Physics*, Elsevier: Amsterdam (2007),

Weinberg, S., Gauge and global symmetry at high temperature, *Physical Review D* 9:3357-78 (1974), Wegner, F., Duality in generalized Ising models and phase transitions without local order parameter, J. Math. Phys. 12:2259-72 (1971),

Wilson, K., Confinement of quarks, *Physical Review D* 10:2445-59 (1974).