Localization and symmetries

2007 J. Phys. A: Math. Theor. 40 3173

(http://iopscience.iop.org/1751-8121/40/12/S17)

IOP PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND THEORETICAL

J. Phys. A: Math. Theor. **40** (2007) 3173–3187 doi:10.1088/1751-8113/40/12/S17

Localization and symmetries

G Morchio[1,2] **and F Strocchi**[2,3]

[1] Dipartimento di Fisica, Universita' di Pisa, Italy
[2] INFN, Sezione di Pisa, Italy
[3] Scuola Normale Superiore, Italy

Received 5 July 2006, in final form 13 December 2006
Published 7 March 2007
Online at stacks.iop.org/JPhysA/40/3173

Abstract
The violation of the Noether relation between symmetries and charges is reduced to the time dependence of the charge associated with a conserved current. For the $U(1)$ gauge symmetry a non-perturbative control of the charge commutators is obtained by an analysis of the Coulomb charged fields. From this, in the unbroken case we obtain a correct expression for the electric charge on the Coulomb states, its superselection and the presence of massless vector bosons; in the broken case, we obtain a general non-perturbative version of the Higgs phenomenon, i.e. the absence of massless Goldstone bosons and of massless vector bosons. The conservation of the (gauge-dependent) current associated with the $U(1)$ axial symmetry in QCD is shown to be compatible with the time dependence of the corresponding charge commutators and a non-vanishing η' mass, as a consequence of the non-locality of the (conserved) current.

PACS numbers: 11.30.−j, 12.38.−t

1. Introduction

The role of continuous symmetries and their breaking in the recent developments of theoretical physics needs not to be further discussed here. However, in our opinion, the interplay between symmetries and localization properties of the fields for infinitely extended systems is not sufficiently emphasized in the textbook presentations, in particular in connection with the phenomenon of symmetry breaking in gauge theories and in Coulomb systems.

We shall focus our attention on the case of continuous symmetries which commute with space translations and with time evolution. For infinitely extended systems described by field variables, $\varphi(\mathbf{x}, t)$, and by a Lagrangian function, an *internal symmetry* is a transformation of the fields $g : \varphi(\mathbf{x}, t) \to (g\varphi)(\mathbf{x}, t)$, g independent of $\mathbf{x}$ and t, which leaves the Lagrangian density invariant. At the level of local variables and measurements the implications of such an invariance property are not as direct as they appear. In fact, the invariance of the Lagrangian under a (continuous) one parameter group of symmetries implies the existence of a *conserved current* $j_\mu(\mathbf{x}, t)$,

$$\partial_t j_0 + \mathrm{div}\,\mathbf{j} \equiv \partial^\mu j_\mu(x) = 0, \tag{1.1}$$

i.e. a *local conservation law*. However, the implications of such a local conservation, in particular the existence of a conserved charge or the existence of Goldstone bosons in the case of symmetry breaking, critically depend on the localization properties of the relevant variables and configurations.

This problem has been extensively discussed in the literature [1, 2] under the assumption that, according to the general wisdom of the Noether theorem, the time-independent symmetry transformations of the fields are generated by the space integral of the charge density j_0 of the corresponding Noether conserved current j_μ:

$$\begin{aligned} \delta A &= \mathrm{i} \lim_{R\to\infty} [Q_R, A], \\ Q_R &\equiv j_0(f_R, \alpha) \equiv \int \mathrm{d}^{s+1}x\, j_0(\mathbf{x}, t) f_R(\mathbf{x}) \alpha(x_0), \end{aligned} \tag{1.2}$$

where the smearing test functions $f_R(\mathbf{x}) = f(|\mathbf{x}|/R)$, $f, \alpha \in \mathcal{D}(\mathbf{R})$, $f(x) = 1$, for $|x| \leqslant R$, take care of the necessary ultraviolet regularization. It is enough that the limit exists for the field correlation functions. The independence of the rhs of equation (1.2) from the choice of the test function α, with the normalization condition $\int \mathrm{d}x_0\, \alpha(x_0) = 1$, is formally equivalent to the time independence of the (space) integral of the charge density, and therefore it is necessary for the validity of equation (1.2). It is also assumed that equation (1.2) holds independently of whether the symmetry is broken or not.

The validity of such assumptions follows if j_μ and A are *relatively local*, e.g. if the canonical structure is local and the time evolution of both j_μ and A is relativistically causal; in this case, the limit is reached for finite values of R, and equation (1.1) implies the independence of α and equation (1.2). The same conclusion holds if the delocalization induced by the time evolution is not worse than $r^{-2-\varepsilon}$, $\varepsilon > 0$. However, important physical phenomena are governed by time-independent symmetries for which the current conservation, equation (1.1), does not imply the generation of the symmetry by the integral of the charge density, equation (1.2), and one cannot rely on the above assumptions.

The crucial issue is the time dependence of the integral of the charge density, namely the α dependence of the rhs of equation (1.2), in spite of the conservation of the current, i.e. the failure of sufficient relative locality between j_i and the operator A. The aim of this note is to critically examine the mechanisms at the basis of such a failure and their physical consequences both in the case of an exact and of a broken symmetry.

The inevitable non-locality of the charged fields has been proved to follow if the current obeys a local Gauss law: $j_\mu = \partial^\nu F_{\nu\mu}$, with $F_{\mu\nu} = -F_{\nu\mu}$ a local field [4]; however, this does not directly imply that the charged fields are not local with respect to j_i (e.g. in the classical Maxwell–Dirac and Maxwell–Klein–Gordon the charged fields are local with respect to j_i [5]) and, even more importantly, that there is enough relative non-locality to force the α dependence of the rhs of equation (1.2) and, therefore, the *violation the Noether relation between the symmetry and the integral of the charge density of the corresponding current*, equation (1.2). In section 3, by applying the analysis of [5, 6] we shall show that such phenomenon *arises in the (physical) Coulomb gauge of an Abelian* $U(1)$ *gauge symmetry*, both in the case of unbroken and broken $U(1)$ symmetries.

In section 4, we show that in the case of unbroken symmetry, the (time-independent) generation of the $U(1)$ symmetry can be obtained by a suitable time average of the integral of the charge density, through a modified Requardt prescription [7]. This allows for a *direct* *proof of the charge superselection rule in the (physical) Coulomb gauge* (a previous proof relied on the general assumptions of the Feynman–Gupta–Bleuler gauge [8]).

In section 5, we discuss the implications of the breaking of the $U(1)$ gauge symmetry on the energy–momentum spectrum. By exploiting the Dirac–Symanzik–Steinmann (DSS) construction of Coulomb charged fields [5, 9] we shall obtain a general non-perturbative version of the Higgs phenomenon [10]: The (time-independent) $U(1)$ *gauge symmetry is generated by the integral of the charge density* discussed in section 4, and in this case unbroken, *if and only if* the Fourier transform of *the two-point function of* $F_{\mu\nu}$ *has a contribution* $\delta(k^2)$, *i.e. there are massless vector bosons*.

If the $U(1)$ gauge symmetry is broken, then it cannot be generated (in the above sense) by the current $j_\mu = \partial^\nu F_{\mu\nu}$; in this case the vacuum expectation $\lim_{R\to\infty}\langle[j_0(f_R, t), A]\rangle$, where A is a charged field with $\langle A\rangle \neq 0$, cannot vanish nor be time independent and *its Fourier spectrum coincides with the energy spectrum at* $\mathbf{k} \to 0$ *of the two-point function of* $F_{\mu\nu}$; this cannot have a $\delta(k^2)$ contribution, so that *the absence of massless Goldstone bosons coincides with the absence of massless vector bosons*.

The strict analogy of the above mechanism with the evasion of the Goldstone theorem in non-relativistic Coulomb systems is discussed in section 6. In section 7 we discuss the $U(1)$ problem in QCD; we show (on the basis of local gauges) that *the axial* $U(1)$ *transformations define a symmetry of the observable field algebra* and argue that its spontaneous breaking is *not accompanied by massless Goldstone bosons as a consequence of the time dependence of the corresponding charge commutators in the (physical) Coulomb gauge*.

2. Locality and symmetries in quantum field theory

In the Wightman formulation of quantum field theory (QFT) [11] one of the basic assumptions is that the field algebra $\mathcal{F}$ satisfies *microscopic causality*, also called *locality*; this means that fields commute or anticommute at space-like separations (depending on their spin). While microscopic causality is a must for the subalgebra $\mathcal{F}_{\rm obs}$ generated by observable fields, there is no cogent physical reason for the locality of the whole field algebra, which typically involves non-observable fields (e.g. fermion fields or charged fields).

The locality condition for $\mathcal{F}$ may be read as a statement about the *localization of the states* obtained by applying $\mathcal{F}$ to the vacuum. Following Doplicher, Haag and Roberts (DHR) [12], a state ω, defined by its expectations $\omega(A)$ on the observables, $A \in \mathcal{A}_{\rm obs}$, is *localized* in the (bounded) spacetime region $\mathcal{O}$ (typically a double cone), if for all observable A localized in the spacetime complement $\mathcal{O}'$ of $\mathcal{O}$, i.e. in the set of points which are space-like to every point of $\mathcal{O}$, briefly $\forall A \in \mathcal{A}_{\rm obs}(\mathcal{O}')$, ω coincides with the vacuum state ω_0:

$$\omega(A) = \omega_0(A), \qquad \forall\, A \in \mathcal{A}_{\rm obs}(\mathcal{O}').$$

Now, for any bounded region $\mathcal{O}$ the unitary operators $U_{\mathcal{O}}$ constructed in term of fields localized in $\mathcal{O}$, typically $\varphi(f)$, with supp $f \subseteq \mathcal{O}$, give states $\omega(A) \equiv (U_{\mathcal{O}}\Psi_0, AU_{\mathcal{O}}\Psi_0)$ which are localized in $\mathcal{O}$ in the DHR sense. In fact, thanks to the locality of the field algebra, $U_{\mathcal{O}}$ commutes with $\mathcal{A}_{\rm obs}(\mathcal{O}')$ and therefore $\omega(A) = \omega_0(A), \forall A \in \mathcal{A}_{\rm obs}(\mathcal{O}')$.

It is important to stress that in general the vacuum sector $\mathcal{H}_0$, obtained by applying the observable field algebra $\mathcal{F}_{\rm obs}$ to the vacuum, does not exhaust the physically interesting states and the non-observable fields of $\mathcal{F}$ play the important role of producing from the vacuum the physical states which do not belong to $\mathcal{H}_0$. The properties of the non-observable fields are therefore physically interesting and worthwhile to study in view of the states they produce; technically the unitary operators constructed in terms of non-observable fields intertwine between the vacuum sector and the other physically relevant representations of the observable algebra. The locality property of $\mathcal{F}$ guarantees that such intertwiners are localizable and so are the corresponding states.

In general, a one parameter group of *internal symmetries* β^λ, $\lambda \in \mathbf{R}$, is a group of field transformations, technically a one parameter group of *-automorphisms of the field algebra, which commutes with the spacetime translations $\alpha_{\mathbf{x},t}$. The relation between symmetries and localization is formalized by the following property: β^λ is *locally generated* on the field algebra $\mathcal{F}$ if

(i) there exists a conserved current field j_μ, $\partial^\mu j_\mu(x) = 0$,
(ii) the infinitesimal transformation of the field algebra is given by

$$\delta F = \mathrm{i} \lim_{R\to\infty} [Q_R, F], \qquad \forall\, F \in \mathcal{F}, \tag{2.1}$$

where Q_R is morally the integral of the charge density j_0 in the sphere of radius R, suitably regularized to cope with the (possible) distributional UV singularities of j_μ, see equation (1.2).

The existence of a conserved current may be taken as equivalent to the invariance of the Lagrangian (or the action); however, condition (ii) is in general not obvious, even if it is often taken for granted. Here, locality plays a crucial role. In fact, if the field algebra is local both the limit $R \to \infty$ exist *and* it is independent of the time smearing, equivalently $\lim_{R\to\infty}[j_0(f_R, t), F]$ is independent of time:

$$\lim_{R\to\infty} \partial_t [j_0(f_R, t), F] = 0, \tag{2.2}$$

so that equation (2.1) may be checked by (canonical) equal-time commutators.

If the vacuum expectations of the fields, which, by the cluster property, describe their mean behaviour at space infinity, are invariant under β^λ, the symmetry is globally realized in the universe described by the given vacuum, i.e. one has a global conservation law, whereas if some expectation is not invariant, $\langle \delta F\rangle \neq 0$, the symmetry is spontaneously broken and there is no global charge associated with the current continuity equation.

For locally generated symmetries, the spontaneous symmetry breaking implies a strong (non-perturbative) constraint on the energy–momentum spectrum, namely the existence of massless particles, called *Goldstone bosons*, with the same (conserved) quantum numbers of the current and of the field F with non-invariant vacuum expectation (*symmetry breaking order parameter*).

The original proof of the theorem [13] applies to the case in which the non-symmetric order parameter is given by a scalar (elementary) field φ and exploits the Lorentz covariance of the two-point function $\langle j_\mu(x)\varphi(y)\rangle$, but it was later realized that the crucial property is the *relative locality* between j_μ and the symmetry breaking order parameter F (which needs not to be one of the basic or elementary fields, but may be a polynomial of them) [1]. The lack of appreciation of this point has been at the basis of discussions and attempts for evading the Goldstone theorem, which eventually led to the Higgs mechanism and to the standard model of elementary particles.

3. Locality and symmetries in gauge theories

Gauge field theories exhibit very distinctive features, with fundamental experimental consequences, such as spontaneous symmetry breaking with energy gap (Higgs mechanism) in apparent contradiction with the Goldstone theorem, quark confinement and linearly rising potential in contrast with the cluster property, axial current anomaly, asymptotic freedom etc[4].

[4] For the general structure and properties of gauge field theories see [14]. For the lack of locality and the violation of cluster property see [15]; for a non-perturbative discussion of the evasion of the Goldstone theorem in gauge theories see [16, 17].

It is natural to try to understand such departures from standard quantum field theory in terms of general ideas independently of the specific model. The original motivation by Yang and Mills, namely that quantum numbers or charges associated with gauge transformations have only a local meaning does not have a direct experimental interpretation since, as a consequence of confinement and symmetry breaking, the observed physical states do not carry non-Abelian gauge charges. More generally, by the definition gauge transformations reduce to the identity on the observables, so that they can be defined only by introducing non-observable fields. The role of gauge symmetries has therefore been regarded [12, 18] as that of providing a classification of the (inequivalent) representations of the observable algebra, through the action of the charged fields. It is still unexplained why only states corresponding to one-dimensional representations of the gauge groups (which include the non-Abelian gauge group of permutations of identical particles) occur in nature.

For the meaning of local gauge invariance, we recall that the standard characterization of gauge field theories is that they are formulated in terms of (non-observable) fields which transform non-trivially under the group $\mathcal{G}$ of *local gauge transformations* leaving the Lagrangian invariant.

In classical field theory, the invariance of the Lagrangian or of the Hamiltonian under a (n-dimensional) Lie group G of spacetime-independent field transformations can be checked by considering the infinitesimal variation of the fields (for simplicity we take G compact and include the coupling constants in the generators):

$$\delta\varphi_i(x) = \mathrm{i}\varepsilon_a t^a_{ij}\varphi_j(x) \equiv \mathrm{i}(\varepsilon t\varphi)_i(x), \tag{3.1}$$

$$\delta A^a_\nu(x) = \mathrm{i}\varepsilon_c T^c_{ab} A^b_\nu(x) \equiv \mathrm{i}(\varepsilon T A_\nu)^a(x), \qquad T^a_{bc} = \mathrm{i} f^b_{ac}, \tag{3.2}$$

where $a = 1, \ldots, n$, $i = 1, \ldots, d$, summation over repeated indices is understood, ε are the infinitesimal group parameters, t is the (d-dimensional) matrix representation of the generators of the group G, provided by the fields φ_i and f are the Lie algebra structure constants.

The *local gauge group* $\mathcal{G}$ *associated with* G (called the *global group*), is the infinite-dimensional group obtained by letting the group parameters to be regular *localized* functions $\varepsilon(x)$ of the spacetime points, typically $\varepsilon \in \mathcal{D}(\mathbf{R}^4)$ or $\in \mathcal{S}(\mathbf{R}^4)$, with the result of an additional term $\partial_\nu\, \varepsilon^a(x)$ in equation (3.2).

It is very important to keep separate the Lie algebra $L(G)$ of G and the infinite-dimensional algebra corresponding to $\mathcal{G}$, briefly denoted by $L(\mathcal{G})$. It would be improper to consider the first as a finite-dimensional subalgebra of the second, both from a mathematical and for a physical points of view. In particular, trivial representations of the $L(\mathcal{G})$ need not to be trivial representations of $L(G)$, and in fact the construction of gauge invariant charged fields is one of the strategies for the analysis of gauge theories; an example of such a construction is the DSS construction in the Abelian case of the DSS fields [9].

A physically very important consequence of the invariance under a local gauge group is that one gets a stronger form of the local conservation laws $\partial^\mu J^a_\mu(x) = 0$, implied by the invariance under the *global group* G. In fact, by the second Noether theorem, the conserved currents,

$$J^{a\mu} \equiv -\mathrm{i}\frac{\delta\mathcal{L}}{\delta\partial_\mu\varphi_i}(t^a\varphi)_i - \mathrm{i}\frac{\delta\mathcal{L}}{\delta\partial_\mu A^b_\nu}(T^a A_\nu)^b \equiv j^{a\mu}(\varphi) + j^{a\mu}(A), \tag{3.3}$$

satisfy the additional equation

$$J^b_\mu = \frac{\delta\mathcal{L}}{\delta A^\mu_b} = \partial^\nu G^b_{\mu\nu} + E[A]^b, \qquad G^{\mu\nu}_b \equiv -\frac{\delta\mathcal{L}}{\delta\partial_\mu A^b_\nu} = -G^{\nu\mu}_b, \tag{3.4}$$

where $E[A]_b = 0$ are the Euler–Lagrange (EL) equations of motion of A.

Equations (3.4) encode the invariance under the local gauge group $\mathcal{G}$, briefly *gauge invariance*, and can be taken as a characterization of such an invariance property; they shall be called *local Gauss' laws*, since Gauss' theorem represents their integrated form. The current continuity equation is trivially implied without using the EL equations for the matter fields.

The validity of local Gauss' laws appears to have a more direct physical meaning than the gauge symmetry, which is non-trivial only on non-observable fields. It is therefore tempting to regard the validity of local Gauss' laws as the basic characteristic feature of gauge field theories, and to consider gauge invariance merely as a useful recipe for writing down Lagrangian functions which automatically lead to the validity of local Gauss' laws.

Actually, for the canonical formulation of gauge field theories one has to exploit the freedom of fixing a gauge, typically by adding a gauge fixing term in the Lagrangian (irrelevant for the physical implications), and this can be done even at the expense of totally breaking the gauge invariance of the Lagrangian (as e.g. in the so-called unitary gauge). Thus, the gauge invariance of the Lagrangian is not so crucial from a physical point of view, whereas so is the validity of local Gauss' laws, which is preserved under the addition of a gauge fixing and the corresponding subsidiary condition [14].

As we shall see below, the local Gauss' law is at the basis of most of the peculiar features of gauge quantum field theories, with respect to standard quantum field theories[5].

From a structural point of view, a first consequence of the local Gauss' law is that, if the local charges Q^a_R, equation (1.2), generate the global group G, the *charged fields cannot be local* [4]. In fact, the field F is charged with respect to the ath one parameter subgroup of G if $\delta^a F \neq 0$, whereas if F is *local* with respect to the conserved current J^a_μ,

$$\lim_{R\to\infty} \left[Q^a_R, F\right] = \int \mathrm{d}^3x\, \mathrm{d}t \nabla_i f_R(\mathbf{x})\alpha(t)[G_{0i}(\mathbf{x}, t), F] = 0, \tag{3.5}$$

since supp $\nabla_i f_R \alpha \subset \{R \leqslant |\mathbf{x}| \leqslant R(1+\varepsilon)\}$ becomes space-like with respect to any bounded region for R large enough. The rhs also vanishes for more general time smearing, $\alpha_R(t)$ with support in $[-R(1-\varepsilon), R(1-\varepsilon)]$ (see below).

Since the local generation of G follows from the locality of the time evolution and of the equal time canonical commutators, equation (3.5) implies that the charged fields are not local with respect to G_{i0}.

The physical reason is that Gauss' law establishes a tight link between the local properties of the solutions and their behaviour at infinity; e.g. the charge of a solution of the electrodynamics equation can be computed either by integrating the charge density, i.e. a local function of the charge carrying fields, or by computing the flux of the electric field at space infinity.

This result has very strong implications at the level of structural properties of gauge quantum field theories: *the field algebra* generated by $G^a_{\mu\nu}$ and the charged fields *cannot be local*.

This may appear as a mere gauge artefact with no physical relevance, [21] since charged fields are not observable fields. However, charged fields play the important role of generating from the vacuum charged states and describing (even neutral) states in terms of charged particles. The non-locality of the charged fields, as implied by the local Gauss' law, has therefore the important physical consequence that the charged states cannot be local in the DHR sense.

[5] The recognition of local Gauss' laws as the basic characteristic features of gauge field theories has been argued and stressed (also in view of the quantum theories) in [19] and later re-proposed [20], ignoring the above references.

4. Gauss' law and local generation of symmetries

Since the fields which transform non-trivially under the global group G cannot be local, the local generation of G becomes problematic, namely both the existence of the limit $R \to \infty$ in equation (1.2), as well as its time independence are in question and, as far as we know, no general conclusion follows directly from Gauss' law. As we shall show, both questions can be answered for Coulomb charged fields, by (crucially) exploiting their construction in terms of the local charged fields of the Feynman–Gupta–Bleuler (FGB) gauge.

Proposition 4.1. *In the Coulomb gauge of QED, $\forall\Psi, \Phi \in \mathcal{F}_C\Psi_0$, with $\mathcal{F}_C$ the field algebra of the Coulomb gauge and Ψ_0 the vacuum vector, the limits*

$$\lim_{R\to\infty} (\Psi, [j_0(f_R\alpha), F]\Phi), \qquad F \in \mathcal{F}_c \tag{4.1}$$

exist; however, they are (generically) α dependent if F is a charged field and therefore equation (1.2) fails.

Proof. The field algebra $\mathcal{F}_C$ is generated by the vector potential A_C^i and the elementary charged fields φ_C, so that it is enough to discuss the case $F = \varphi_C$ and a basic ingredient is the DSS construction [9] of the Coulomb charged fields φ_C in terms of the local fields φ, A_μ of the FGB gauge:

$$\varphi_C(y) = \exp(\mathrm{i}e(-\Delta^{-1}\partial^j A_j)(y))\varphi(y). \tag{4.2}$$

The necessary ultraviolet regularization of equation (4.2) has been discussed by Steinmann [9] within the perturbative expansion. A regularized version, which only uses the existence of the FGB correlations and is constructed by using in the exponential fields smeared in space and time, has been given by Buchholz *et al* [5]. In this framework, the space asymptotic of the correlation functions of the commutator $[F_{\mu\nu}(x), \varphi_C(y)]$ is given, at all orders in the expansion of the exponential entering in equation (4.2), with corrections $O(|\mathbf{x}|^{-4})$, by

$$[F_{\mu\nu}(x), \varphi_C(y)] \sim \frac{-\mathrm{i}e}{4\pi}\int \mathrm{d}^3z\, \partial_z^j \frac{1}{|\mathbf{z}-\mathbf{y}|}\langle[F_{\mu\nu}(x), A_j(\mathbf{z}, y_0)]\rangle\varphi_C(y). \tag{4.3}$$

Since $\langle[F_{\mu\nu}(x), A_j(z)]\rangle = \mathrm{i}(\partial_\nu g_{\mu j} - \partial_\mu g_{\nu j})K(x-z)$, with K being the commutator function of the electromagnetic field, one has, for $R \to \infty$,

$$[j_0(f_R, x_0), \varphi_C(y)] = [\partial^i F_{0i}(f_R, x_0), \varphi_C(y)] \sim -e\partial_0 \int \mathrm{d}^3x\, f_R(\mathbf{x})K(x-y)\varphi_C(y). \tag{4.4}$$

By the support properties of $K(x) = -\mathrm{i}\int \mathrm{d}\rho(m^2)\varepsilon(k_0)\delta(k^2-m^2)\,\mathrm{e}^{-\mathrm{i}kx}$, the charge density is integrable and, in all correlation functions,

$$\lim_{R\to\infty} [j_0(f_R, x_0), \varphi_C(y)] = e\int \mathrm{d}\rho(m^2)\cos(m(x_0-y_0))\varphi_C(y). \tag{4.5}$$

The rhs is independent of time if and only if $\mathrm{d}\rho(m^2) = \lambda\delta(m^2)$, i.e. if $F_{\mu\nu}$ is a free field.

The same conclusions are obtained if instead of equation (4.2) one uses the regularized version of [5], since in this case equations (4.3)–(4.5) get changed only by a convolution with a test function $h(y_0) \in \mathcal{D}(\mathbf{R})$. □

The time dependence of $\lim_{R\to\infty}[j_0(f_R, t), F]$, $F \in \mathcal{F}_C$ is compatible with the conservation of the current because the above analysis gives

$$\begin{aligned} &[j_i(x), \varphi_C(y)] = (e/4\pi)\int \mathrm{d}^3z\, \partial_z^i|\mathbf{z}-\mathbf{y}|^{-1}\partial_0^2 K(x-z), \qquad z_0 = y_0, \\ &\lim_{R\to\infty} [\dot{Q}_R(x_0), \varphi_C(y)] = [\operatorname{div}\mathbf{j}(f_R, x_0), \varphi_C(y)] \neq 0. \end{aligned} \tag{4.6}$$

The time dependence of the commutator of equation (4.5) is at the basis of the appearance of an infinite renormalization constant in the equal time commutator of the charge density $j_0 = \partial^i F_{0i}$ and the Coulomb charged field φ_C:

$$[j_0(x), \varphi_C(y)]_{x_0=y_0} = e(Z_3)^{-1}\delta(\mathbf{x}-\mathbf{y})\varphi_C(y),$$

(all fields being renormalized fields and e the renormalized charge), as it appears by comparing the integrated form of the above equal time commutator and equation (4.4). For such a phenomenon the vacuum polarization due to fermionic loops plays a crucial role, so that the semi-classical approximation does not provide relevant information, and in fact the phenomenon does not appear in the classical theory.

Proposition 4.1 shows that, in contrast with the local case, the equal time commutators are misleading for the charge commutators and, contrary to statements in the literature, the $U(1)$ charge group of QED is not locally generated by the integral of the charge density, in the sense of equation (4.1). Thus, the heuristic argument that if the symmetry commutes with the time translations, equivalently if the current continuity equation holds, then the generating charge commutes with the Hamiltonian and is therefore independent of time is not correct. Time independence of the charge commutator holds provided one has (relative) locality at all times between j_μ and the charged fields; now, even if the equal time commutators have a sufficient localization, the time evolution may induce a delocalization leading to a failure of equation (2.2).

5. Electric charge and its superselection

The results of the previous section leave open the question of whether a modification of equation (1.2) may yield a relation between a gauge symmetry and the charge density of the corresponding Noether current. As we shall discuss below, if the gauge symmetry is unbroken a time average of equation (1.2), similar to that proposed by Requardt [7], provides the required relation.

Actually, equation (4.4) gives the renormalized charge for any time smearing $\alpha_{T(R)}(x_0) \equiv \alpha(x_0/T(R))/T(R)$, with $T/R \to 0$ as $R \to \infty$ [6], if $\mathrm{d}\rho(k^2)$ has a $\delta(k^2)$ contribution. Moreover, for a certain class of functions $T(R)$, which depends on the infrared behaviour of $k^2\,\mathrm{d}\rho(k^2)$, $j_0(f_R\alpha_{T(R)})\Psi_0$ converges strongly to zero [6]. A smearing, which gives both results independently of any information on the above infrared behaviour, is given by taking $T(R) = \delta R$, with $\delta \to 0$ after the limit $R \to \infty$.

Proposition 5.1. *In the Coulomb gauge the $U(1)$ gauge symmetry is generated by the integral of the charge density*

$$\delta F = \mathrm{i} \lim_{\delta\to 0} \lim_{R\to\infty} [Q_{R\delta}, F], \qquad F \in \mathcal{F}_c \tag{5.1}$$

$$Q_{R\delta} \equiv j_0(f_R\alpha_{\delta R}), \qquad \alpha_{\delta R}(x_0) \equiv \alpha(x_0/(\delta R))/(\delta R) \tag{5.2}$$

if and only if $\mathrm{d}\rho(k^2)$ has a $\delta(k^2)$ contribution, i.e. there are massless photons.

Moreover, one has

$$strong - \lim_{R\to\infty} j_0(f_R\alpha_{\delta R})\Psi_0 = 0, \tag{5.3}$$

so that, if there are massless photons one can express the electric charge Q, i.e. the generator of the $U(1)$ symmetry, as an integral of the charge density j_0 not only in the commutators with charged fields, but also in the matrix elements of the Coulomb charged states $\Phi, \Psi \in \mathcal{F}_C\Psi_0$:

$$(\Phi, Q\Psi) = \lim_{\delta\to 0} \lim_{R\to\infty} (\Phi, j_0(f_R\alpha_{\delta R})\Psi). \tag{5.4}$$

Proof. The time smearing of equation (4.4) with $\alpha_{\delta R}(x_0)$ gives

$$[j_0(f_R\alpha_{\delta R}), \varphi_C(y)] = e\int \mathrm{d}\rho(m^2)\,\mathrm{d}^3q\,\tilde{f}(\mathbf{q})\mathrm{Re}[\mathrm{e}^{-\mathrm{i}\omega_R(q,m)y_0}\tilde{\alpha}(\delta\sqrt{\mathbf{q}^2 + R^2m^2})]\varphi_C(y),$$

where $\omega_R(\mathbf{q}, m) \equiv \sqrt{\mathbf{q}^2R^{-2} + m^2}$. Then, since α is of fast decrease, by the dominated convergence theorem the rhs vanishes if the $\mathrm{d}\rho(m^2)$ measure of the point $m^2 = 0$ is zero, i.e. if there is no $\delta(m^2)$ contribution to $\mathrm{d}\rho$. In general, if the point $m^2 = 0$ has measure λ, one gets $\lambda e\varphi_C(y)$; finally the renormalization condition of the asymptotic electromagnetic field gives $\lambda = 1$.

For the proof of equation (5.3) one has ($\mathrm{d}\Omega_m(\mathbf{k}) \equiv \mathrm{d}^3k(2\sqrt{\mathbf{k}^2 + m^2})^{-1}$)

$$\begin{aligned}\|Q_{R\delta}\Psi_0\|^2 &= \int \mathrm{d}\rho(m^2)m^2\,\mathrm{d}\Omega_m(\mathbf{k})|\mathbf{k}\tilde{f}_R(\mathbf{k})\tilde{\alpha}(\delta R\sqrt{\mathbf{k}^2 + m^2})|^2\\ &= \int \mathrm{d}\rho(m^2)\,\mathrm{d}\Omega_m(\mathbf{q}/R)m^2R|\tilde{\alpha}(\delta\sqrt{\mathbf{q}^2 + m^2R^2})\mathbf{q}\tilde{f}(\mathbf{q})|^2.\end{aligned}$$

Now, $m^2R|\tilde{\alpha}(\delta\sqrt{\mathbf{q}^2 + m^2R^2})|^2$ converges pointwise to zero for $R \to \infty$ and since $\mathrm{d}\rho(m^2)$ is tempered and α is of fast decrease the rhs of the above equation converges to zero by the dominated convergence theorem. □

One of the basic Dirac–Von Neumann axioms of quantum mechanics is that the states of a quantum mechanical system are described by vectors of a Hilbert space $\mathcal{H}$ and that every vector describes a state, equivalently all projections and therefore all (bounded) self-adjoint operators represent observables (briefly $\mathcal{A}_{\mathrm{obs}} = \mathcal{B}(\mathcal{H})$). It was later realized [22] that, typically for systems with infinite degrees of freedom, the physical states may belong to a direct sum of irreducible representations of the observable algebra, and therefore one cannot measure coherent superpositions of vectors belonging to inequivalent representations of the observable algebra. This means that if $\mathcal{H} = \oplus\mathcal{H}_j$, each $\mathcal{H}_j$ carrying an irreducible representation of $\mathcal{A}_{\mathrm{obs}}$, a linear combination $\alpha\Psi_1 + \beta\Psi_2$ of vectors Ψ_1, Ψ_2 belonging to different $\mathcal{H}_j$ is not a physically realizable (pure) state and it rather describes a mixture with the density matrix $|\alpha|^2\Psi_1 \otimes \Psi_1 + |\beta|^2\Psi_2 \otimes \Psi_2$.

The impossibility of measuring such relative phases is equivalent to the existence of operators Q, called *superselected charges*, which commute with all the observables (and have a denumerable spectrum if the Hilbert space is separable)[6].

Wick, Wightman and Wigner (WWW) proved that rotation and time reversal invariance imply that the operator $Q_F = (-1)^{2J} = (-1)^F$, where J is the angular momentum and F is the fermion number modulo 2 is a superselected charge (*univalence superselection rule*, also called *fermion-boson superselection rule*). It was later shown that only rotational invariance was needed for the proof [23]. WWW also suggested that the electric charge and possibly the baryon number define superselected charges.

The superselection rule for the electric charge was later questioned and debated [25]. The proof may be dismissed as trivial by arguing that observables must be gauge invariant and that gauge invariance implies zero charge, but as stressed before such an argument is not correct, since the latter implication is contradicted by the Dirac–Symanzik–Steinmann field operator [9] showing that gauge invariant operators need not to commute with the electric charge.

The superselection of the electric charge Q may be shown to be a consequence of the locality of the observables and the Gauss law, provided one can express Q as an integral

[6] The superselected charges are often called *gauge charges*, but we prefer the name of superselected charges. The gauge group which classify the representations of the observable algebra defined by DHR localized states has been proved to be compact [18].

of $j_0 = \partial^i F_{0i}$. A proof of the charge superselection rule has been given by using a local gauge quantization of QED, e.g. the Feynman–Gupta–Bleuler gauge, and by identifying Q with the generator of the global gauge transformations of the local fields [8]. In this gauge, the construction of the DSS operators [6] makes clear that *invariance under the local gauge transformations does not imply invariance under the global gauge transformations for non-local operators*.

By exploiting proposition 5.1 one can get a direct proof of the charge superselection rule in the *physical* Coulomb gauge.

Proposition 5.2. *The electric charge Q, defined in the Coulomb gauge by*

$$Q\Psi_0 = 0, \qquad [Q, \varphi_C(y)] = e\varphi_C(y),$$

commutes with the observables (on the Coulomb states)

$$(\Phi, [Q, A]\Psi) = \lim_{\delta\to 0}\lim_{R\to\infty}(\Phi, [j_0(f_R\alpha_{\delta R}), A]\Psi) = 0, \qquad \forall\Phi, \Psi \in \mathcal{F}_C\Psi_0, \tag{5.5}$$

and it is therefore superselected.

Proof. The proof follows from equation (5.4), which relates the electric charge Q and the electric flux at infinity, by the argument which exploits the relative locality of the observables with respect to the (observable) electromagnetic field, as required by Einstein causality, [8, 24] so that the rhs vanishes by the same argument of equation (3.5). Actually the rhs of equation (5.5) vanishes independently of the adopted time smearing by locality. □

The superselection of electromagnetic fluxes at space-like infinity has been discussed by Buchholz [24], under the assumption of weak convergence. The special choice of spacetime smearing $j_0(f_R\alpha_{\delta R})$ adopted above guarantees the strong convergence on the vacuum, convergence in expectations on Coulomb charged states and the relation between the corresponding electric flux and the electric charge.

6. Gauge symmetry breaking and the energy–momentum spectrum. The Higgs mechanism

The Higgs mechanism, relative to the breaking of the global group G in a gauge quantum field theory, plays a crucial role in the standard model of elementary particle physics. The standard discussion of this mechanism is based on the perturbative expansion and, in particular, the evasion of the Goldstone theorem is checked at the tree level with the disappearance of the massless Goldstone bosons and the vector bosons becoming massive [26]. This is displayed by the Higgs–Kibble (Abelian) model of a (complex) scalar field φ interacting with a real gauge field A_μ, defined by the following Lagrangian ($\rho(x) \equiv |\varphi(x)|$):

$$\mathcal{L} = -\tfrac{1}{4}F_{\mu\nu}{}^2 + \tfrac{1}{2}|D_\mu\varphi|^2 - U(\rho), \qquad D_\mu = \partial_\mu - \mathrm{i}eA_\mu \tag{6.1}$$

invariant under the $U(1)$ gauge group: $\beta^\lambda(\varphi) = \mathrm{e}^{\mathrm{i}\lambda}\varphi$, $\beta^\lambda(A_\mu) = A_\mu$ and under local gauge transformations.

At the classical level, one may argue that by a local gauge transformation

$$\varphi(x) = \mathrm{e}^{\mathrm{i}\theta(x)}\rho(x) \to \rho(x), \qquad A_\mu(x) \to A_\mu(x) + e^{-1}\partial_\mu\theta(x) \equiv W_\mu(x)$$

one may eliminate the field θ from the Lagrangian, which becomes

$$\mathcal{L} = -\tfrac{1}{4}F_{\mu\nu}{}^2 + \tfrac{1}{2}e^2\rho^2W_\mu^2 + \tfrac{1}{2}(\partial^\mu\rho)^2 - U(\rho). \tag{6.2}$$

If the (classical) potential U has a non-trivial (absolute) minimum $\rho = \overline{\rho}$ one can consider a semiclassical approximation based on the expansion $\rho = \overline{\rho} + \sigma$, treating $\overline{\rho}$ as a classical constant field and σ as small. At the lowest order, keeping only the quadratic terms in σ and W_μ one has

$$\mathcal{L}^{(2)} = -\tfrac{1}{4}{F_{\mu\nu}}^2 + \tfrac{1}{2}e^2\overline{\rho}^2 W_\mu^2 + \tfrac{1}{2}(\partial^\mu\sigma)^2 - \tfrac{1}{2}U''(\overline{\rho})\sigma^2. \tag{6.3}$$

This Lagrangian describes a massive vector boson and a massive scalar with (square) masses $M_W^2 = e^2\overline{\rho}^2$, $m_\sigma^2 = U''(\overline{\rho})$, respectively. This argument is commonly taken as an evidence that there are no massless particles in the theory described by the Lagrangian $\mathcal{L}$.

This argument, widely used in the literature, is not without problems, because already at the classical level, for the equivalence between the two forms of the Lagrangian, equations (6.1) and (6.2), one must add the constraint that ρ is positive, a property which is in general spoiled by the time evolution given by the Lagrange equations for the variables ρ and W_μ. For the variables of the quadratic Lagrangian (6.3), one should require that the time evolution of σ keeps it bounded by $\overline{\rho}$, a condition which is difficult to satisfy. Thus, the constrained system is rather singular and its mathematical control is doubtful. The situation becomes obviously more critical for the quantum version, since the definition of $|\varphi(x)|$ is very problematic also for distributional reasons. In conclusion, ρ is a very singular field and one cannot consider it as a genuine Lagrangian (field) variable.

A better alternative is to decompose the field $\varphi = \varphi_1 + \mathrm{i}\varphi_2$ in terms of Hermitian fields, and to consider the semiclassical expansion $\varphi_1 = \overline{\varphi} + \chi_1, \varphi_2 = \chi_2$, treating $\chi_i, i = 1, 2$, as small. By introducing the field $W_\mu \equiv A_\mu + e^{-1}\partial_\mu\chi_2$, one eliminates χ_2 from the quadratic part of the so-expanded Lagrangian, which gets exactly the same form of equation (6.2), with $\overline{\rho}$ replaced by $\overline{\varphi}$ and σ by χ_1.

If indeed the fields χ_i can be treated as small, by appealing to the perturbative (loop) expansion one has that $\langle\varphi\rangle \sim \overline{\rho} \neq 0$, i.e. the vacuum expectation of φ is not invariant under the $U(1)$ charge group (*symmetry breaking*). Thus, the expansion can be seen as an expansion around a (symmetry breaking) mean field ansatz, and it is very important that a renormalized perturbation theory based on it exists and yields a non-vanishing symmetry breaking order parameter $\langle\varphi\rangle \neq 0$ at all orders. This is the standard (perturbative) analysis of the Higgs mechanism.

The extraordinary success of the standard model motivates an examination of the Higgs mechanism from a general non-perturbative point of view. In this perspective, one of the problems is that mean field expansions may yield misleading results about the occurrence of symmetry breaking and the energy spectrum[7]. Actually, a non-perturbative analysis of the Euclidean functional integral defined by the Lagrangian of equation (6.1) gives symmetric correlation functions and in particular $\langle\varphi\rangle = 0$ (*Elithur–De Angelis–De Falco–Guerra (EDDG) theorem* [27]). This means that the mean field ansatz is incompatible with quantum effects and the approximation leading to equation (6.3) is not correct[8].

The same negative conclusion would be reached if (as an alternative to the transformation which leads to equation (6.2)) by means of a gauge transformation one reduces $\varphi(x)$ to a real, not necessarily positive, field $\varphi_r(x)$. This means that the local gauge invariance has not been completely eliminated and the corresponding Lagrangian, of the same form (6.2) with ρ replaced by φ_r, is invariant under a residual Z_2 local gauge group. An easy adaptation of the proof of the EDDG theorem gives $\langle\varphi_r\rangle = 0$.

[7] For example, the mean field ansatz on the Heisenberg spin model of ferromagnetism gives a wrong critical temperature and an energy gap. For a discussion of the problems of the mean field expansion see e.g. [28].

[8] The crux of the argument is that gauge invariance decouples the transformations of the fields inside a volume V (in a Euclidean functional integral approach) from the transformation of the boundary, so that the boundary conditions are ineffective and cannot trigger non-symmetric correlation functions. For a simple account of the argument see e.g. [29].

In order to avoid the vanishing of a symmetry breaking order parameter one must reconsider the problem by adding to the Lagrangian (6.1) a gauge fixing $\mathcal{L}_{\mathrm{GF}}$ which breaks local gauge invariance. Thus, the discussion of the Higgs mechanism is necessarily gauge fixing dependent; this should not appear strange, since the vacuum expectation of φ is a gauge-dependent quantity.

The important physical properties at the basis of the Higgs mechanism are particularly clear in the so-called physical gauges, like the Coulomb gauge. Since the charged Coulomb fields cannot be local, the local generation of the symmetry, required for the applicability of the Goldstone theorem, is in question. Actually, in the Abelian case by using the results of section 5 one has a non-perturbative proof of the characterization of the Higgs phenomenon given by Weinberg on the basis of the perturbative expansion [14]. By proposition 5.1, the (time-independent) $U(1)$ gauge symmetry is generated by the integral of the charge density, equation (5.1), and in this case unbroken, if and only if the Fourier transform of the two-point function of $F_{\mu\nu}$ has a contribution $\delta(k^2)$, i.e. there are massless vector bosons.

Proposition 6.1. *If the (time-independent) $U(1)$ gauge symmetry is broken, then it cannot be generated by the integral of the charge density, equation (5.1), of the associated Noether current $j_\mu = \partial^\nu F_{\nu\mu}$; in this case the vacuum expectation $\lim_{R\to\infty}\langle[j_0(f_R, t), A]\rangle$, where $A \in \mathcal{F}_c$ is a charged field with $\langle A\rangle \neq 0$, cannot vanish nor be time independent and its Fourier spectrum coincides with the energy spectrum at $\mathbf{k} \to 0$ of the two-point function of the vector boson field $F_{\mu\nu}$, which cannot have a $\delta(k^2)$ contribution, so that the absence of massless Goldstone bosons coincides with the absence of massless vector bosons.*

Proof. The first part follows trivially from proposition 5.1, which also states that the symmetry generated according to equation (5.1) cannot be broken. Thus, in the broken case the spectral measure of $F_{\mu\nu}$ cannot have a $\delta(k^2)$ contribution and therefore $\lim_{\delta\to 0}\lim_{R\to\infty}[j_0(f_R\alpha_{\delta R}), F] = 0, \forall F \in \mathcal{F}_C$. The vacuum expectation $\lim_{R\to\infty}\langle[j_0(f_R, t), A]\rangle$, where A is a charged field with $\langle A\rangle \neq 0$, is obtained from equations (4.4), (4.5) and the relation with the Fourier spectrum of $F_{\mu\nu}$ follows. □

In conclusion, the above discussion shows that the evasion of the Goldstone theorem crucially depends on the non-locality of the charged fields. The local structure of the canonical commutation relations, in particular of the commutator $[\mathbf{j}, \varphi_C]$, is not stable under the time evolution induced by the electromagnetic interactions, as displayed by the Coulomb gauge. This is possible in a relativistic theory because in this case the field algebra does not satisfy manifest covariance. For these reasons, no reliable information can be inferred from the equal time commutators and the check of the basic assumptions of the Goldstone theorem becomes interlaced with the dynamical problem, as it happens for non-relativistic systems. The failure of locality leading to equation (4.6), rather than the lack of manifest covariance, is the crucial structural property which explains the evasion of the Goldstone theorem in the Higgs mechanism as well as in Coulomb systems and in the $U(1)$ problem, as discussed below.

7. Coulomb delocalization and symmetries in many-body theory

A natural question, following by the above discussion of symmetry breaking, is the general characterization of the dynamics which induces a delocalization leading to the failure of equation (2.2), so that the symmetry is not locally generated and one may have symmetry breaking with energy gap. In this perspective, whenever the field algebra is not manifestly covariant, *instantaneous* interactions are possible and there is no longer a deep distinction between relativistic and non-relativistic systems. Actually, both the Coulomb gauge in QED and the non-relativistic Coulomb systems are characterized by the instantaneous Coulomb interaction:

$$H_{\mathrm{int}} = \frac{1}{2}e^2 \int \mathrm{d}^3x\, \mathrm{d}^3y\, j_0(\mathbf{x}) V(\mathbf{x}-\mathbf{y}) j_0(\mathbf{y}). \tag{7.1}$$

As argued by Swieca [3], for two-body instantaneous interactions the range of the potential characterizes the delocalization induced by the dynamics: if $V(\mathbf{x})$ falls off like $|\mathbf{x}|^{-d}$, for $|\mathbf{x}| \to \infty$, then the unequal time commutators generically decay with the same power and

$$\lim_{\mathbf{x}\to\infty} |\mathbf{x}|^{d+\varepsilon}[A_{\mathbf{x}}, B_t] = 0, \tag{7.2}$$

at least order by order in a perturbative expansion in time.

The above equation suggests that the delocalization needed for equation (2.2) in three space dimensions is that given by a potential fall-off like $|\mathbf{x}|^{-2}$. Actually, the current $\mathbf{j}$ involves space derivatives and the critical decay turns out to be $|\mathbf{x}|^{-1}$, i.e. that of the Coulomb potential.

This unifies the mechanism of symmetry breaking with the energy gap of the Higgs phenomenon and that of non-relativistic Coulomb systems (typically the breaking of the Galilei symmetry in the jellium model or the breaking of the electron $U(1)$ symmetry in the BCS model of superconductivity) and provides a clarification of the analogies proposed by Anderson [30]. For a general discussion of the energy gap associated with symmetry breaking for long range dynamics see [17].

In gauge theories relative locality may fail because either the order parameter (as in the Higgs phenomenon) or the conserved current (as discussed below) associated with the symmetry of the Lagrangian are *non-local* fields. This is the case of the $U(1)$ problem in QCD.

8. Axial symmetry breaking and the $U(1)$ problem

The debated problem of $U(1)$ axial symmetry breaking in quantum chromodynamics without massless Goldstone bosons can be clarified by the realization of the non-locality of the associated axial current. As clearly shown by Bardeen [31], the $U(1)$ axial symmetry gives rise to a conserved, gauge-dependent, current:

$$J^5_\mu = j^5_\mu - (2\pi)^{-2}\varepsilon_{\mu\nu\rho\sigma}\mathrm{Tr}[A^\nu\partial^\rho A^\sigma - (2/3)\mathrm{i}A^\nu A^\rho A^\sigma] \equiv j^5_\mu + K^5_\mu,$$

where j^5_μ is the gauge invariant point splitting regularized fermion current $\overline{\psi}\gamma_\mu\gamma_5\psi$. The current j^5_μ is not conserved because of the anomaly, which is equivalent to the conservation of J^5_μ.

In the usual discussion of the $U(1)$ problem (see e.g. [32]), the current J^5_μ has been discarded on the blame of its gauge dependence, and the lack of conservation of j^5_μ has been taken as the evidence that the axial $U(1)$ is not a symmetry of the field algebra and therefore the problem of its spontaneous breaking does no longer exist. Such a conclusion would imply that time-independent $U(1)$ axial transformations cannot be defined on the field algebra $\mathcal{F}$ and not even on its observable subalgebra $\mathcal{F}_{\mathrm{obs}}$, which contains the relevant order parameter. However, as argued by Bardeen on the basis of perturbative renormalization (in local gauges), the axial $U(1)$ transformations define a time-independent symmetry of the field algebra and of its observable subalgebra. This also follows from the conservation of J^5_μ (equivalent to the anomaly of j^5_μ), since in local renormalizable gauges J^5_μ is a local operator, so that the standard argument (see section 2) applies, i.e. equations (2.1), (2.2) hold. This implies that (at least at the infinitesimal level) the rhs of equation (2.1) defines in this case a symmetry of the field algebra and in particular of the gauge invariant observable subalgebra $\mathcal{F}_{\mathrm{obs}}$. Therefore, there is no logical reason for *a priori* rejecting the use of the gauge-dependent current J^5_μ and of its associated Ward identities; one should only keep in mind that in physical gauges J^5_μ is a non-local function of the observable (gauge-independent) fields.

The existence of axial $U(1)$ transformations of the observable subalgebra $\mathcal{F}_{\mathrm{obs}}$ implies that the absence of parity doublets is a problem of spontaneous symmetry breaking, and the absence of massless Goldstone bosons is reduced to the discussion of local generation of the symmetry, equations (2.1), (2.2), as in the case of the Higgs phenomenon.

In the local (renormalizable) gauges the time-independent $U(1)$ axial symmetry is generated by J^5_μ (and not by j^5_μ) and the problem of massless Goldstone modes does not arise because, as indicated by the perturbative expansion and also by the Schwinger model [33], *the correlation functions of the (local) field algebra* $\mathcal{F}$ *are axial* $U(1)$ *invariant*. However, the invariance of the vacuum functional Ψ_0, which defines the local gauge quantization, does not mean that the symmetry is unbroken in the irreducible representations of the observable subalgebra $\mathcal{F}_{\mathrm{obs}}$. In fact, Ψ_0 gives a reducible representation of $\mathcal{F}_{\mathrm{obs}}$ (as signalled by the failure of the cluster property by the corresponding vacuum expectations), with a non-trivial centre which is generated by the large gauge transformations T_n and is not pointwise invariant under $U(1)$ axial transformations [33]. Thus, *the symmetry is broken in each pure physical phase* (θ*-vacuum sectors*) obtained by the diagonalization of T_n (in the technical terminology by a *central decomposition* of the observables) in the subspace $\mathcal{F}_{\mathrm{obs}}\Psi_0$. It should be stressed that the so-obtained (gauge invariant) θ-vacua do *not* provide well-defined representations of the field algebra $\mathcal{F}$, since the latter transforms non-trivially under T_n. This is at the origin of the difficulties (and paradoxes) arising in the discussion of the chiral Ward identities (corresponding to the conservation of J^5_μ) in θ-vacua expectations [34]. In the θ sectors a conserved axial current may be constructed as a non-local operator, typically by using for J^5_μ its (non-local) expression in terms of the observable fields in a physical gauge. The above discussion, in particular the lack of time independence in equation (2.1) as a consequence of the failure of relative locality between the current and the order parameter, applies to such non-local currents.

The resulting mechanism for the solution of the $U(1)$ problem can be made explicit in the Coulomb gauge. In the Schwinger model, in the Coulomb gauge one has $K_0 = (e/\pi)A_1 = 0$, $K_1 = (e/\pi)A_0$, so that $J^5_0 = j^5_0$ and the (θ-)vacuum expectations of the commutators $\left[J^5_0(f_R, t), A\right]$, $\left[j^5_0(f_R, t), A\right]$, $A \in \mathcal{F}_{\mathrm{obs}}$, coincide and describe the same mass spectrum; however, the time dependence in the limit $R \to \infty$, in the first case can be ascribed to the non-locality of the conserved axial current, whereas in the second case it reflects the non-conservation of j^5_μ.

References

[1] Kastler D, Robinson D W and Swieca J A 1966 *Commun. Math. Phys.* **2** 108

[2] Swieca J A 1970 Goldstone theorem and related topics (*Cargése Lectures in Physics* vol 4) ed D Kastler (London: Gordon and Breach)

[3] Swieca J A 1967 *Commun. Math. Phys.* **4** 1

[4] Ferrari R, Picasso L E and Strocchi F 1974 *Commun. Math. Phys.* **35** 25

[5] Buchholz D, Doplicher S, Morchio G, Roberts J E and Strocchi F 2001 *Ann. Phys.* **290** 53

[6] Morchio G and Strocchi F 2003 *J. Math. Phys.* **44** 5569

[7] Requardt M 1976 *Commun. Math. Phys.* **50** 259
Buchholz D 1986 *Phys. Lett.* B **174** 331

[8] Strocchi F and Wightman A S 1974 *J. Math. Phys.* **15** 2198

[9] Dirac P A M 1955 *Can. J. Phys.* **33** 650
Symanzik K Lectures on Lagrangian quantum field theory Desy Report T-71/1

D' Emilio E and Mintchev M 1984 *Fortsch. Phys.* **32** 473
Steinmann O 2000 *Perturbative Quantum Electrodynamics and Axiomatic Field Theory* (Berlin: Springer)
[10] Weinberg S 1971 *Phys. Rev. Lett.* **27** 1688
[11] Streater R F and Wightman A S 1980 *PCT, Spin and Statistics and All That* (New York: Benjamin-Cummings)
[12] Haag R 1996 *Local Quantum Physics* (Berlin: Springer)
[13] Goldstone J, Salam A and Weinberg S 1962 *Phys. Rev.* **127** 965
[14] Weinberg S 1996 *The Quantum Theory of Fields* vol 2 (Cambridge: Cambridge University Press)
[15] Strocchi F 1978 *Phys. Rev.* D **17** 2010
[16] Strocchi F 1977 *Commun. Math. Phys.* **56** 57
[17] Morchio G and Strocchi F 1989 Removal of the infrared cutoff, seizing of the vacuum and symmetry breaking in many body and in gauge theories, invited talk at the *9th Int. Congr. Mathematical Physics* ed B Simon *et al* (Bristol: Hilger) p 490 (and references therein)
Strocchi F 2005 *Symmetry Breaking* (Berlin: Springer)
[18] Doplicher S and Roberts J 1990 *Commun. Math. Phys.* **131** 51
Doplicher S 1991 *Proc. Int. Congr. Mathematicians (Kyoto)* (Berlin: Springer) p 1319
[19] Strocchi F and Wightman A S 1974 *J. Math. Phys.* **15** 2198
Strocchi F 1981 Gauss' law in local quantum field theory *Field Theory, Quantization and Statistical Physics* (Dordrecht: Reidel) p 227
Strocchi F 1985 *Elements of Quantum Mechanics of Infinite Systems* (Singapore: World Scientific) part C, chapter II
[20] Karatas D L and Kowalski K L 1990 *Am. J. Phys.* **58** 123
Al-Kuwari H A and Taha M O 1990 *Am. J. Phys.* **59** 363
Nakanishi N and Ojima I 1990 *Covariant Operator Formalism of Gauge Theories and Quantum Gravity* (Singapore: World Scientific)
Brading K and Brown H R 2000 *Preprint* hep-th/0009058
[21] Hurst C A *Math. Rev.* MR2053089
[22] Wick G C, Wightman A S and Wigner E P 1952 *Phys. Rev.* **88** 101
[23] Hegerfeldt G C, Kraus K and Wigner E P 1968 *J. Math. Phys.* **9** 2029
For a nice review on the superselection rules see Wightman A S 1995 *Nuovo Cimento* B **110** 751
[24] Buchholz D 1986 *Phys. Lett.* B **174** 331
[25] Mirman R 1969 *Phys. Rev.* **186** 1380
Aharonov Y and Susskind L 1967 *Phys. Rev.* **155** 1428
Kershaw D and Woo C H 1974 *Phys. Rev. Lett.* **33** 918
Wick G C, Wightman A S and Wigner E P 1970 *Phys. Rev.* D **1** 3267
[26] Higgs P W 1964 *Phys. Lett.* **12** 132
Higgs P W 1964 *Phys. Rev. Lett.* **13** 508
Higgs P W 1966 *Phys. Rev.* **145** 1156
Guralnik G S, Hagen C R and Kibble T W 1964 *Phys. Rev. Lett.* **13** 585
Guralnik G S, Hagen C R and Kibble T W 1968 Broken Symmetries and the Goldstone theorem *Advances in Particle Physics* vol 2, ed R L Cool and R E Marshak (New York: Interscience) p 567
Kibble T W 1966 Broken Symmetries *Proc. Oxford Int. Conf. on Elementary Particles* (Oxford: Oxford University Press) p 19
Kibble T W 1966 *Phys. Rev.* **155** 1554
Englert F and Brout R 1964 *Phys. Rev. Lett.* **13** 321
[27] Elitzur S 1975 *Phys. Rev.* D **12** 3978
De Angelis G F, De Falco D and Guerra F 1978 *Phys. Rev.* D **17** 1624
[28] Strocchi F 2005 *Symmetry Breaking* (Berlin: Springer) chapters 10, 11
[29] Strocchi F 1985 *Elements of Quantum Mechanics of Infinite Systems* (Berlin: World Scientific) part C, section 2.5
[30] Anderson P W 1964 *Phys. Rev.* **130** 439
[31] Bardeen W A 1974 *Nucl. Phys.* B **75** 246
[32] Coleman S 1985 *Aspects of Symmetry* (Cambridge: Cambridge University Press) chapter 7
[33] Morchio G, Pierotti D and Strocchi F 1988 *Ann. Phys.* **188** 217
Strocchi F 1993 *Selected Topics on the General Properties of Quantum Field theory* (Singapore: World Scientific) section 7.4
[34] Crewther R J 1980 *Field Theoretical Methods in Particle Physics* ed W Rühl (Dordrecht: Reidel) pp 529