Gauge Invariance and Mass

JULIAN SCHWINGER
Harvard University, Cambridge, Massachusetts, and University of California, Los Angeles, California

(Received July 20, 1961)

It is argued that the gauge invariance of a vector field does not necessarily imply zero mass for an associated particle if the current vector coupling is sufficiently strong. This situation may permit a deeper understanding of nucleonic charge conservation as a manifestation of a gauge invariance, without the obvious conflict with experience that a massless particle entails.

DOES the requirement of gauge invariance for a vector field coupled to a dynamical current imply the existence of a corresponding particle with zero mass? Although the answer to this question is invariably given in the affirmative,[1] the author has become convinced that there is no such necessary implication, once the assumption of weak coupling is removed. Thus the path to an understanding of nucleonic (baryonic) charge conservation as an aspect of a gauge invariance, in strict analogy with electric charge,[2] may be open for the first time.

One potential source of error should be recognized at the outset. A gauge-invariant system is not the continuous limit of one that fails to admit such an arbitrary function transformation group. The discontinuous change of invariance properties produces a corresponding discontinuity of the dynamical degrees of freedom and of the operator commutation relations. No reliable conclusions about the mass spectrum of a gauge-invariant system can be drawn from the properties of an apparently neighboring system, with a smaller invariance group. Indeed, if one considers a vector field coupled to a divergenceless current, where gauge invariance is destroyed by a so-called mass term with parameter m_0, it is easily shown[3] that the mass spectrum must extend below m_0. The lowest mass value will therefore become arbitrarily small as m_0 approaches zero. Nevertheless, if m_0 is exactly zero the commutation relations, or equivalent properties, upon which this conclusion is based become entirely different and the argument fails.

If invariance under arbitrary gauge transformations is asserted, one should distinguish sharply between numerical gauge functions and operator gauge functions, for the various operator gauges are not on the same quantum footing. In each coordinate frame there is a unique operator gauge, characterized by three-dimensional transversality (radiation gauge), for which one has the standard operator construction in a vector space of positive norm, with a physical probability interpretation. When the theory is formulated with the aid of vacuum expectation values of time-ordered operator products, the Green's functions, the freedom of formal gauge transformation can be restored.[4] The Green's functions of other gauges have more complicated operator realizations, however, and will generally lack the positiveness properties of the radiation gauge.

Let us consider the simplest Green's function associated with the field $A_\mu(x)$, which can be derived from the unordered product

$$\langle A_\mu(x)A_\nu(x')\rangle = \int \frac{(dp)}{(2\pi)^3} e^{ip(x-x')} dm^2\, \eta_+(p)\delta(p^2+m^2)A_{\mu\nu}(p),$$

where the factor $\eta_+(p)\delta(p^2+m^2)$ enforces the spectral restriction to states with mass $m\geq 0$ and positive energy. The requirement of non-negativeness for the matrix $A_{\mu\nu}(p)$ is satisfied by the structure associated with the radiation gauge, in virtue of the gauge-dependent asymmetry between space and time (the time axis is specified by the unit vector n_μ):

$$A_{\mu\nu}{}^R(p)=B(m^2)\left[g_{\mu\nu}-\frac{(p_\mu n_\nu+p_\nu n_\mu)(np)+p_\mu p_\nu}{p^2+(np)^2}\right].$$

Here $B(m^2)$ is a real non-negative number. It obeys the sum rule

$$1=\int_0^\infty dm^2\, B(m^2),$$

which is a full expression of all the fundamental equal-time commutation relations.

The field equations supply the analogous construction for the vacuum expectation value of current products $\langle j_\mu(x)j_\nu(x')\rangle$, in terms of the non-negative matrix

$$j_{\mu\nu}(p)=m^2B(m^2)(p_\mu p_\nu-g_{\mu\nu}p^2).$$

The factor m^2 has the decisive consequence that $m=0$ is not contained in the current vector's spectrum of vacuum fluctuations. The latter determines $B(m^2)$ for $m>0$, but leaves unspecified a possible delta function contribution at $m=0$,

$$B(m^2)=B_0\delta(m^2)+B_1(m^2).$$

The non-negative constant B_0 is then fixed by the sum rule,

$$1=B_0+\int_0^\infty dm^2\, B_1(m^2).$$

---
[1] For example, J. Schwinger, Phys. Rev. **75**, 651 (1949).
[2] T. D. Lee and C. N. Yang, Phys. Rev. **98**, 1501 (1955).
[3] K. Johnson, Nuclear Phys. **25**, 435 (1961).
[4] J. Schwinger, Phys. Rev. **115**, 721 (1959).

We have now recognized that the vacuum fluctuations of the vector A_μ are composed of two parts. One, with $m>0$, is directly related to corresponding current fluctuations, while the other part, with $m=0$, can be associated with a pure radiation field, which is transverse in both three- and four-dimensional senses and has no accompanying current. Imagine that the current vector contains a variable numerical factor. If this is set equal to zero, we have $B_1(m^2)=0$ and $B_0=1$ or, just the radiation field. For a sufficiently small nonzero value of the parameter, B_0 will be slightly less than unity, which may be the situation for the electromagnetic field. Or it may be that the electrodynamic coupling is quite considerable and gives rise to a small value of B_0, which has the appearance of a fairly weak coupling. Can we increase further the magnitude of the variable parameter until $\int dm^2\, B_1(m^2)$ attains its limiting value of unity, at which point $B_0=0$, and $m=0$ disappears from the spectrum of A_μ? The general requirement of gauge invariance no longer seems to dispose of this essentially dynamical question.

Would the absence of a massless particle imply the existence of a stable, unit spin particle of nonzero mass? Not necessarily, since the vacuum fluctuation spectrum of A_μ becomes identical with that of j_μ, which is governed by all of the dynamical properties of the fields that contribute to this current. For the particularly interesting situation of a vector field that is coupled to the current of nucleonic charge, the relevant spectrum, in the approximate strong-interaction framework, is that of the states with $N=Y=T=0$, $R_T=-1$, $J=1$, and odd parity. This is a continuum, beginning at three pion masses.[5] It is entirely possible, of course, that $B(m^2)$ shows a more or less pronounced maximum which could be characterized approximately as an unstable particle.[6] But the essential point is embodied in the view that the observed physical world is the outcome of the dynamical play among underlying primary fields, and the relationship between these fundamental fields and the phenomenological particles can be comparatively remote, in contrast to the immediate correlation that is commonly assumed.

[5] The very short range of the resulting nuclear interaction together with the qualitative inference that like nucleonic charges are thereby repelled suggests that the vector field which defines nucleonic charge is also the ultimate instrument of nuclear stability.

[6] *Note added in proof.* Experimental evidence for an unstable particle of this type has recently been announced by B. C. Maglić, L. W. Alvarez, A. H. Rosenfeld, and M. L. Stevenson, in Phys. Rev. Letters 7, 178 (1961).