International Journal of Theoretical Physics, Vol. 3, No. 6 (1970), pp. 467–482

Gauge Invariance, Lorentz Covariance and the Observer

EGON MARX

Drexel University,
Philadelphia, Pennsylvania

Received: 14 *May* 1970

Abstract

We discuss a number of questions related to the role of the observer in classical and quantum theories of fields, in particular electrodynamics. We find the gauge-independent parts of the electromagnetic potential, which are classical observables, both in a non-covariant manner and in a Lorentz covariant, observer-dependent way. We present an analysis of the probabilistic interpretation of relativistic quantum mechanics, similar to that of the nonrelativistic theory, and discuss the gauge invariance of the corresponding probability amplitudes.

1. *Introduction*

The requirements of gauge invariance and Lorentz covariance have shaped much of the formulations of the interactions between charged particles and the electromagnetic field, be it in the context of relativistic classical mechanics, classical electrodynamics, relativistic quantum mechanics or quantum theory of fields. The elegant four-dimensional notation of Minkowski space is very useful when we want to establish the Lorentz covariance of an equation, but it does hide the special role that the observer plays even in classical theories.

It is natural that some timelike direction has to be singled out when we want to describe the dynamical development of a system, and the performance of measurements on this system is often carried out on a hyperplane normal to this direction. We have been compelled to an explicit use of this notion of an observer, whose world line defines this timelike direction, in our studies of quantum field theory and relativistic quantum mechanics. The consideration of position and orbital angular momentum operators in the quantum theory of relativistic free fields (Marx, 1968) leads to particular choices of probability amplitudes in terms of the field variables, and these same amplitudes have a basic role in the formulation of a canonical quantization of free boson fields in terms of wave functionals (Marx, 1969a). In particular, when potentials are used as the basic field variables for the electromagnetic field, both a canonical quantization procedure (Katz, 1965; Marx, 1969a) and the special formulation developed for theories with constraints (Goldberg, 1958, 1965; Goldberg & Marx, 1968) result in the separation of a gauge-independent part, which turns out to be observer dependent. We present a detailed discussion of this point in Sections 2 and 4.

Even in classical field theories, which are usually presented in a manifestly Lorentz covariant form, we implicitly single out an observer when we specify boundary conditions on an initial hyperplane. Similarly, we find it necessary to refer to an observer when we use these classical fields as wave functions in a relativistic quantum mechanics, to which we have extended the usual probabilistic interpretation of the nonrelativistic theory. These roles of the observer are further elucidated in Section 3. Other theories, such as those described by the hyperplane formalism (Fleming, 1965, 1966), introduce the notion of a physically significant observer (or the hyperplanes perpendicular to its world line) at an earlier stage of the development of the theory. Some further aspects related to the gauge invariance of relativistic probability densities are presented in Section 5.

We use natural units and a time-favoring metric in space-time. The modified summation convention applies to repeated lower Greek indices that range from 0 to 3. Other notation is either standard or used in previous papers.

2. *The Gauge-Independent Part of the Electromagnetic Potentials*

Maxwell's equations for the electric field **E** and the magnetic field **B** in free space can be written covariantly in the form

$$F_{\mu\nu,\nu} = j_\mu \tag{2.1}$$

$$\epsilon_{\mu\nu\lambda\rho} F_{\mu\nu,\lambda} = 0 \tag{2.2}$$

where $\epsilon_{\mu\nu\lambda\rho}$ is the completely antisymmetric Levi–Civita tensor. We conclude from equations (2.2) that the antisymmetric field tensor $F_{\mu\nu}$ can be derived from a vector potential A_μ through

$$F_{\mu\nu} = A_{\mu,\nu} - A_{\nu,\mu} \tag{2.3}$$

Conversely, when $F_{\mu\nu}$ is given by equation (2.3), equations (2.2) are identically satisfied. It is usually assumed that only the fields have physical significance, and not the potentials, which are not uniquely defined by equation (2.3). The $F_{\mu\nu}$ are invariant under the gauge transformation

$$A_\mu \to \bar{A}_\mu = A_\mu + \Lambda_{,\mu} \tag{2.4}$$

where Λ is an arbitrary function of x. It is then demanded that all physically significant results be invariant under an arbitrary gauge transformation. This is obviously the case when theories are formulated directly in terms of the fields $F_{\mu\nu}$ but it is often convenient to use the potentials A_μ, especially in the classical and quantum theories of fields in their Lagrangian or Hamiltonian formulations. Nevertheless, it is possible to use the fields as basic variables even in cases where potentials are normally used (Marx, 1967, 1969a).

When we use potentials as the basic coordinates in field theory, the separation of the physically meaningful, gauge-independent part is of great help in understanding many of the results. In order to do this, we decompose the vector potential into its solenoidal and irrotational parts.

An arbitrary vector field $\mathbf{v}(\mathbf{x})$ whose magnitude vanishes at infinity sufficiently rapidly can be written as†‡

$$\mathbf{v} = \mathbf{v}_I + \mathbf{v}_S \tag{2.5}$$

where

$$\mathbf{v}_I(\mathbf{x}) = -\boldsymbol{\nabla} \int \frac{1}{4\pi} |\mathbf{x} - \mathbf{x}'|^{-1} \boldsymbol{\nabla}' . \mathbf{v}(\mathbf{x}')\, d^3x' \tag{2.6}$$

$$\mathbf{v}_S(\mathbf{x}) = \boldsymbol{\nabla} \wedge \int \frac{1}{4\pi} |\mathbf{x} - \mathbf{x}'|^{-1} \boldsymbol{\nabla}' \wedge \mathbf{v}(\mathbf{x}')\, d^3x' \tag{2.7}$$

which satisfy

$$\boldsymbol{\nabla} . \mathbf{v}_I = \boldsymbol{\nabla} . \mathbf{v}, \qquad \boldsymbol{\nabla} \wedge \mathbf{v}_I = 0 \tag{2.8}$$

$$\boldsymbol{\nabla} . \mathbf{v}_S = 0, \qquad \boldsymbol{\nabla} \wedge \mathbf{v}_S = \boldsymbol{\nabla} \wedge \mathbf{v} \tag{2.9}$$

We assume that $\mathbf{A}$ and $\mathbf{j}$ vanish sufficiently rapidly at infinity to use equations (2.6) and (2.7), for a fixed value of t. Although it is reasonable to assume that a physical system is localized, this restriction does not have to apply to unphysical fields such as the potentials. Nevertheless, we still impose it on the A_μ so that their Fourier transforms are well defined and surface terms at spatial infinity vanish in integrations by parts. This amounts to a limitation in our choice of gauges, and gauge transformations should be limited to those in which Λ vanishes at spatial infinity sufficiently rapidly too.

We express equations (2.1) in terms of these potentials, and obtain

$$\partial^2 \mathbf{A}_S = \mathbf{j}_S \tag{2.10}$$

$$\partial^2 \mathbf{A}_I + \boldsymbol{\nabla}(\dot{A}_0 + \boldsymbol{\nabla} . \mathbf{A}_I) = \mathbf{j}_I \tag{2.11}$$

$$-\nabla^2 A_0 - \boldsymbol{\nabla} . \dot{\mathbf{A}}_I = \rho \tag{2.12}$$

Any irrotational vector can be expressed as the gradient of a scalar; in particular

$$\mathbf{A}_I = -\boldsymbol{\nabla}\xi \tag{2.13}$$

where

$$\xi(\mathbf{x}, t) = \frac{1}{4\pi} \int |\mathbf{x} - \mathbf{x}'|^{-1} \boldsymbol{\nabla}' . \mathbf{A}(\mathbf{x}', t)\, d^3x' \tag{2.14}$$

† Rohrlich, F. (1965). *Classical Charged Particles*, p. 69. Addison-Wesley Publishing Company, Inc., Reading, Mass. We note that we write ∂^2 for $\partial_\mu \partial_\mu$, and not for $(n_\mu \partial_\mu)^2$ as in this reference.

‡ The decomposition can be written in slightly different forms, which in turn determine how fast the field has to go to zero at infinity. See also Sommerfeld, A. (1950). *Mechanics of Deformable Bodies*, Section 20. Academic Press, New York.

We now define

$$\chi(x) = A_0(x) - \dot{\xi}(x) \tag{2.15}$$

and substitution of equation (2.13) into equation (2.12) shows that χ satisfies

$$\nabla^2 \chi = -\rho \tag{2.16}$$

The solution of this equation is

$$\chi(\mathbf{x}, t) = \frac{1}{4\pi} \int |\mathbf{x} - \mathbf{x}'|^{-1} \rho(\mathbf{x}', t)\, d^3x' \tag{2.17}$$

Only the irrotational part of the current density is related to the charge density by conservation of charge, which implies that

$$\nabla . \mathbf{j} = \nabla . \mathbf{j}_I = -\dot{\rho} \tag{2.18}$$

We can use this equation together with equation (2.13) to show that equation (2.11) is also satisfied by the solution (2.17) for χ, while equation (2.10) for $\mathbf{A}_S$ is unrelated to the other two.

We obtain the gauge transformation for ξ from equation (2.13) or equation (2.14), which is

$$\xi \to \bar{\xi} = \xi + \Lambda \tag{2.19}$$

Hence, equation (2.15) shows that χ is gauge invariant, which can also be concluded from equation (2.17). We have thus found that χ and $\mathbf{A}_S$ are the gauge-independent parts of the potentials, while $\mathbf{A}_I$ and A_0 are changed by the transformations. In particular, equation (2.19) shows that we can always find a gauge in which ξ and, consequently, $\mathbf{A}_I$ vanish, so that $\mathbf{A}$ and A_0 are equal to $\mathbf{A}_S$ and χ, respectively. This makes the Coulomb or radiation gauge a convenient choice in many cases, and explains why calculations carried out in this gauge have essentially the same form as those in gauge-independent formulations.

When the source of the electromagnetic field is a charged field ψ, it transforms by

$$\psi \to \bar{\psi} = \psi \exp(ie\Lambda) \tag{2.20}$$

and equation (2.19) shows that the fields

$$\psi' = \psi \exp(-ie\xi) \tag{2.21}$$

are gauge independent. This is the change of variables used by Dirac (1950) and Goldberg (1965). The somewhat puzzling restriction (I. Goldberg, private communication) that excludes gauge transformations with Λ of the form

$$\Lambda(\mathbf{x}, t) = \lambda(t) \tag{2.22}$$

in these formulations is traced back to the requirement that $\Lambda(x)$ vanish at spatial infinity. The classical observables (Goldberg, 1965) for the electromagnetic fields are the two independent components of $\mathbf{A}_S$, while χ

is expressed in terms of the charged fields through equation (2.17). The gauge-independent quantization procedure then follows in a manner similar to the one carried out in a Coulomb gauge (Bjorken & Drell, 1965).

3. *Observer Dependence and Lorentz Covariance*

Soon after Einstein's initial formulation of the principles of special relativity, it was realized that relativistic theories of particles or fields could be advantageously formulated in terms of scalars, vectors and tensors in Minkowski's four-dimensional space-time. This way, covariance under Lorentz transformations† is manifest. Nevertheless, the dynamical development of a system has to be described in terms of an observer, who performs the experiment and records the results. We consider only inertial observers, who are in rectilinear uniform motion with respect to each other and are represented by a straight timelike world line. We specify the state of motion of the observer by means of a four-vector n, which satisfies

$$n^2 = 1, \qquad n_0 > 0 \tag{3.1}$$

The solution of a dynamical problem for a system of N point particles is completed when we determine their world lines. They are usually described by their parametric equations, and the choice of the parameters is quite arbitrary. It can be a set of Lorentz scalars, such as the N proper times λ_α defined by

$$d\lambda_\alpha{}^2 = dx_\mu^{(\alpha)}\, dx_\mu^{(\alpha)}, \qquad \alpha = 1, 2, \ldots, N \tag{3.2}$$

or a common parameter for all particles, such as the coordinate time t, or the observer time τ given by

$$x^{(\alpha)} . n = \tau \tag{3.3}$$

The equations of motion‡ can be conveniently derived from a covariant Lagrangian (Rzewuski, 1964), from a noncovariant Lagrangian (which requires special care to insure Lorentz covariance), or they can be given directly, preferably in covariant form. It is generally less clear how initial conditions should be specified; the most convenient way is probably to give the position and velocity for each particle on a spacelike surface, at least when the interactions are retarded. When they are advanced as well as retarded, as in the Feynman–Wheeler electrodynamics (Wheeler & Feynman, 1945, 1949), it is more natural to give asymptotic conditions in the infinitely remote past and/or future. In any case, we have to specify only six independent data per particle, whether we are using a relativistic or a nonrelativistic theory.

† In this discussion, we restrict ourselves to the consideration of proper orthochronous Lorentz transformations.

‡ Strictly speaking, the term 'equations of motion' should be applied only to the $3N$ equations for the $x_i^{(\alpha)}(t)$, while the $4N$ equations for the $x_\mu^{(\alpha)}(\lambda_\alpha)$ are not independent and do not give the 'motion' in space-time, but the world lines. All these formulations are, of course, equivalent.

The above-mentioned Feynman–Wheeler electrodynamics is an attempt to formulate a relativistic action-at-a-distance theory, in which no fields are needed, or where they only play an auxiliary role. This is how electric and magnetic fields were introduced in the first place, to facilitate the computation of forces between charges at rest or in motion (currents). But when Maxwell's equations were finally discovered, it was found that these fields had degrees of freedom of their own, which manifest themselves in form of electromagnetic waves or radiation. We feel (at this time) that it is convenient, and possibly necessary, to consider the electromagnetic and other fields as independent dynamical systems.

We can obtain equations of motion for a field from an action principle, using covariant Lagrangian densities, or we can start from the equations themselves.† When the field is considered as a dynamical system, we seek to determine its values at all points in space at a given time (or on a given spacelike hypersurface) from the sources and boundary values at an initial time (or on a hypersurface) when the interactions are all retarded. If we have other types of interactions, we usually need initial and/or final conditions. On the other hand, if we are only interested in the values of the field in a bounded region of space at a given time, the speed of light limits‡ the amount of information we need about the sources and initial values to those inside the backward light cones from the set of observation points, for retarded interactions.

The role of the observer is more significant when we consider the field as a dynamical system, since he defines what is meant by the *simultaneous* observations carried out over all of space to determine the state of the system. This is clearly a drastic idealization of a physical situation, both as far as the instantaneous measurement of the field at a point and the coordination needed to do this throughout all space are concerned. In a classical field theory it is possible, of course, to assume that the field vanishes outside a bounded region at the initial time.

It is not necessary to use a reference frame in Minkowski space that has its time axis parallel to the world line of the observer. When we represent the observer by the unit vector n, we assume that the hyperplanes on which the field is specified are perpendicular to n. These entities then have physical (geometrical) significance,¶ independent of the coordinates we might choose. We distinguish between Lorentz covariant quantities that are independent of the observer, such as x_μ or x^2, and others that are observer dependent, such as the observer time $x.n$.

We easily see that the specification of initial conditions on a hyperplane perpendicular to n restricts the choice of an observer (not of a reference frame), since a hyperplane perpendicular to $n' \neq n$ always cuts the one

† In certain cases, such as spinors (Case, 1957; Marx, 1970c), this turns out to be a nontrivial alternative.

‡ We do not allow for the existence of tachyons in this paper.

¶ This implies that n transforms like any other four-vector. A different point of view is taken by Rohrlich (1965, p. 72 ff.).

perpendicular to n and necessarily includes earlier events, where the field is not determined by the observer n. In this sense, it is not possible to describe the same *system* by two different observers. This observer dependence of the initial values is also reflected in the so-called invariant functions, such as

$$\Delta^{(i)}(x) = -(2\pi)^{-4} \int d^4k \exp(-ik.x)(k^2 - m^2)^{-1} \tag{3.4}$$

where we have to specify the integration path in the complex k_0-plane (this integration is assumed to be performed first) in order to obtain the different types of functions $\Delta^{(i)}$. This implies a choice of time axis, and $\Delta(x)$, for instance, satisfies

$$(\partial^2 + m^2)\Delta(x) = 0 \tag{3.5}$$

$$\Delta(\mathbf{x}, 0) = 0 \tag{3.6}$$

$$\partial_0 \Delta(\mathbf{x}, 0) = -\delta(\mathbf{x}) \tag{3.7}$$

and from equation (3.6) we obtain†

$$\partial_i \Delta(\mathbf{x}, 0) = 0 \tag{3.8}$$

Equation (3.5) is covariant, but equations (3.3)–(3.5) are not. Their covariant analogues are

$$\Delta(\tilde{x}, 0) = 0 \tag{3.9}$$

$$\partial_\mu \Delta(\tilde{x}, 0) = -n_\mu \delta^{(3)}(\tilde{x}) \tag{3.10}$$

where $\tilde{x}$ is the covariant space part of the vector x, defined by

$$\tilde{x} = x - x.nn \tag{3.11}$$

this indicates that $\Delta(\tilde{x}, \tau)$ depends on n to some extent.

We thus conclude that the observer plays a somewhat marginal, but by no means negligible, role in classical relativistic theories. On the other hand, the situation is quite different in quantum mechanics and quantum theory of fields, and consideration of the observer is basic for their physical interpretation.

In nonrelativistic quantum mechanics, the space coordinates x_i are operators, while the time t is a parameter. If we consider a simple system, such as a single particle, its state is represented by a unit vector (or a ray) in a Hilbert space, that of square-integrable functions of $\mathbf{x}$, for instance, with the usual scalar product

$$(f, g) = \int d^3x f^*(\mathbf{x}) g(\mathbf{x}) \tag{3.12}$$

† We do not derive our results for these singular functions with mathematical rigor, but more in the formal sense customary in papers of this nature. We know that such manipulations can lead to contradictions such as those associated with Schwinger terms (Schwinger, 1959; Goldberg & Marx, 1967), which arise from ambiguities in the definition of products of such functions.

If we formulate a dynamical problem in the Schrödinger picture, in which the physical interpretation is more easily understood, we have to specify the state vector at the initial time and the potential at all later times, then the Schrödinger equation of motion,

$$i\dot{\Psi} = H\Psi \tag{3.13}$$

allows us to find the state vector at all later times.† We can imagine the time development of the system represented by a curve on the unit sphere in this Hilbert space, parametrized by t.

Although the subject of relativistic quantum mechanics is not well developed, we have proposed a formulation that interprets the Klein–Gordon equation for bosons (Marx, 1969b) and a *modified* Dirac equation for fermions (Marx, 1970b) in terms of probability amplitudes for particles in an external electromagnetic field. This interpretation is patterned closely after the generally accepted nonrelativistic one, and reduces to it in a natural way for kinetic energies small compared to the mass.

We still work in the Hilbert space of functions of a *three*-vector variable,‡ but in a relativistic theory we have to consider two vectors, both of which vary with time, and which represent particle and antiparticle states. We can still write the equations of motion in the form of a Schrödinger equation (3.13), where Ψ now has two or four components. Our probabilistic interpretation is based on the conservation of charge, and requires a somewhat unusual specification of boundary conditions. We consider a finite process, from time t_i to t_f, and we *either* give the particle amplitude, normalized to 1, at the initial time t_i and set the antiparticle amplitude equal to 0 at the final time t_f, *or* we specify the normalized antiparticle amplitude at t_f and make the particle amplitude vanish at t_i. The equations of motion then allow us to find both amplitudes at times t such that

$$t_i \leqslant t \leqslant t_f \tag{3.14}$$

These can be called causal boundary conditions, and they arise from the use of the causal Green function or Feynman propagator. In particular, we

† It is important for our subsequent discussion to distinguish clearly between the roles of the spatial coordinates x_i and the time t. Given a vector f in this Hilbert space, we can determine both $x_i f$ and $\partial_i f$, which are different vectors, while tf is essentially (up to the norm) the same vector and $\partial_0 f$ cannot be determined at all, unless we are given a family of vectors as a function of the parameter t. We also have to distinguish between $i\partial_0$ and the Hamiltonian operator H, which is usually expressed in terms of $\mathbf{x}$ and $\mathbf{\nabla}$ and can change in time (be explicitly time dependent). In particular, we often find reference in the literature on nonrelativistic quantum mechanics to an uncertainty relation between time and energy, which is at best a misnomer. It usually refers to an inequality satisfied by the product of the lifetime of an unstable state and the natural line width of the radiation emitted in its decay (Davydov, 1965). It might be possible to formulate quantum mechanics in the vector space of functions of four variables x_μ, where t and $\partial/\partial t$ are operators, but we see no reason to do this. On the contrary, since an inner product in this space would be unrelated to the physical interpretation of the theory.

‡ We also have to allow for the two spin states in the case of spin-$\frac{1}{2}$ fermions, but this is a trivial complication.

determine this way the particle amplitude at t_f and the antiparticle amplitude at t_i, and charge conservation implies that the norms of these vectors add up to 1, which makes the probabilistic interpretation possible.†

It is quite obvious that this formulation of relativistic quantum mechanics relies heavily on the separation of space and time variables, which with little effort can be performed covariantly in terms of the observer vector n. Thus, we conclude that, although an equation such as the Klein–Gordon equation might be relativistically covariant and independent of the observer, the specification of initial and final conditions and the physical interpretation have to be carried out in terms of an arbitrary but given observer. There also are a number of operators, such as position and orbital angular momentum, that are explicitly observer dependent.‡

We also have extended the above theory to several identical particles (Marx, 1970a). It is basically a theory with a fixed number of particles, although it allows for pair creation and annihilation. We use the many-time formalism, with one time variable for each particle, but they still are used as parameters, and the general nature of the interpretation remains unaltered. The space of state vectors is the Hilbert space of functions of several three-vector variables, or, to be more precise, there are several components in different subspaces with the right symmetry properties.

We now consider the theory of a quantized field, where we again encounter this marked difference between space and time; this is most clearly seen in the Schrödinger picture, although it is equally true in the Heisenberg picture. The classical system to be quantized is a field (or several fields) specified throughout space, that is, the basic generalized coordinate is itself a vector in the Hilbert space of functions of a three-vector variable. In a quantum theory, we no longer look for the coordinate as a function of time, but we deal with the time development of a probability amplitude for all possible coordinates. In a field theory of integer spin (bosons), we can represent this amplitude explicitly by wave functionals (Marx, 1969a), which are vectors in another Hilbert space, where the inner product is defined with the help of functional integration (Berezin, 1966; Rzewuski, 1969),

$$(F, G) = \int F^*[f]\, G[f]\, \delta f \tag{3.15}$$

† Charge conservation does not require that the norms of the vectors remain less than 1 at intermediate times. If they do, they can be interpreted as probability amplitudes, but if they do not, this would not be too difficult to 'explain'. It might correspond to the possibility of finding more than one particle at intermediate times, due to pair creation, although this would again lead us away from the one-particle theory which we have been seeking. A more reasonable assumption is to exclude observations at intermediate times, since it is well known that observations in quantum mechanics disturb the system, which is particularly undesirable in a problem when boundary conditions at the final time have to be specified. The whole problem of observation of antiparticles by observers formed by matter requires further elucidation.

‡ Although the discussion in an earlier work (Marx, 1968) was carried out for quantized fields, most of the results can be reinterpreted in terms of the classical fields of relativistic quantum mechanics.

Operators are defined in this space in terms of multiplication by the coordinate field and the corresponding functional derivative, and the time development of the system makes the wave functional time dependent. When it is given at the initial time, a functional differential Schrödinger equation determines it at later times. Space and time coordinates appear in completely different ways, and it is difficult to see how to relate the descriptions of two different observers of a single system. It is frequently stated that it is simpler to discuss the Lorentz invariance of a quantum field theory in the Heisenberg picture, where the field operators depend both on space and time. Actually, the transformation from one picture to the other does not change the nature of the Hilbert space and operators, which still operate on functionals of functions of $\mathbf{x}$; in the Heisenberg picture, the operators that represent physical quantities are different at different times. The role of the x_i is that of continuous indices, and that of t, of a parameter; in particular, we often find it advantageous to go over to momentum space, and use functions of $\mathbf{k}$ instead of those of $\mathbf{x}$. The proofs of Lorentz invariance found in the literature usually correspond closely to those for the classical field, especially when this is done in the language of canonical transformations and Poisson brackets.

The above discussion in terms of wave functionals is appropriate only for bosons, which require commutators of canonical coordinates and momenta. For fermion fields, we usually refer to no specific representation of state vectors and operators, and use anticommutation relations for the latter; this way we do not obtain a simple physical interpretation in terms of a probability amplitude for the classical states of the field. We have suggested (Marx, 1970b) that it should be possible to treat fermions within the context of relativistic quantum mechanics, without going to a 'second' quantization. These would then interact via classical or quantized boson fields, which can also have dynamical degrees of freedom of their own, as discussed in Section 2 in connection with the electromagnetic field.

Thus, the observer plays an important role in the description of relativistic processes, and space and time have to be carefully distinguished in order to obtain physical interpretations. What special relativity tells us is that this separation can be carried out in many ways, by choosing an arbitrary timelike direction in Minkowski space to represent the state of motion of the observer.

We should still mention that theories in which only asymptotic states are considered, be it of particles or fields, are those that most easily can dispense with the notion of an observer. On the other hand, processes that take an infinite time and monochromatic plane waves are clearly very drastic idealizations, which might be responsible for some of the divergences that beset quantum field theories. This is the approach used in the so-called S-matrix theory, but it has not been shown that this is a consistent and complete dynamical theory. Furthermore, it explicitly excludes the possibility of making observations during the time an interaction is present, which is probably an unnecessary limitation. Other difficulties are related to unstable particles, excited states and bound states.

4. *Gauge Independence and Observer Dependence*

The separation of the vector potential **A** into irrotational and solenoidal parts is clearly not a Lorentz covariant procedure, but it can be made covariant through the introduction of the observer n. For instance, we find

$$A_{I\mu}(\tilde{x},\tau) = -\tilde{\partial}_\mu \int \frac{1}{4\pi}[-(\tilde{x}-\tilde{x}')^2]^{-1/2}\,\tilde{\partial}'.\tilde{A}(\tilde{x}',\tau)\,d^3\tilde{x}' \tag{4.1}$$

$$A_{S\mu}(\tilde{x},\tau) = \epsilon_{\mu\nu\lambda\rho}\,\epsilon_{\rho\alpha\beta\gamma}\,n_\nu n_\alpha\,\partial_\lambda \int \frac{1}{4\pi}[-(\tilde{x}-\tilde{x}')^2]^{-1/2}\,\partial'_\beta A_\gamma(\tilde{x}',\tau)\,d^3\tilde{x}' \tag{4.2}$$

which, together with the component of A along n, satisfy

$$A = A_I + A_S + n.An \tag{4.3}$$

$$n.A_I = n.A_S = 0 \tag{4.4}$$

$$\epsilon_{\mu\nu\lambda\rho}\,n_\nu\,\partial_\lambda A_{I\rho} = 0, \qquad \tilde{\partial}.A_S = 0 \tag{4.5}$$

The solenoidal part A_S was determined (Goldberg & Marx, 1968) to be the classical observable that is then quantized according to the procedure due to Bergmann (Bergmann & Goldberg, 1955). The gauge-independent part related to the Coulomb interaction,

$$\chi = n.A - \frac{\partial\xi}{\partial\tau} \tag{4.6}$$

where ξ is the integral in equation (4.1), can be expressed in terms of the charged fields, or we can incorporate this interaction to the gauge-independent fields defined by equation (2.21).

It is obvious that, once we take into account the observer dependence of our equations explicitly through the vector n, the choice of a reference frame in Minkowski space is immaterial. On the other hand, there is no simple way to relate the gauge-independent parts of the potentials for different observers, since their determination is nonlocal; integrations such as those in equations (4.1) and (4.2) are carried out over different hyperplanes.

Even Maxwell's equations (2.1) and (2.2), although written in tensor form independent of the observer, have to be separated into true equations of motion and constraints, and this separation is clearly dependent on the observer. If the initial values of the fields are given on a hyperplane perpendicular to n, they have to satisfy the equations of constraint

$$n_\mu\,\tilde{\partial}_\nu F_{\mu\nu}(\tilde{x},\tau_0) = n_\mu j_\mu(\tilde{x},\tau_0) \tag{4.7}$$

$$n_\rho\,\epsilon_{\mu\nu\lambda\rho}\,\tilde{\partial}_\lambda F_{\mu\nu}(\tilde{x},\tau_0) = 0 \tag{4.8}$$

which are the covariant generalizations of

$$\nabla . \mathbf{E}(\mathbf{x}, t_0) = \rho(\mathbf{x}, t_0) \tag{4.9}$$

$$\nabla . \mathbf{B}(\mathbf{x}, t_0) = 0 \tag{4.10}$$

The true equations of motion and conservation of charge then insure that equations (4.7) and (4.8), or equations (4.9) and (4.10), are satisfied at later times.

We can also separate **E** into solenoidal and irrotational parts (**B** is always solenoidal, of course), and $\mathbf{E}_S$ is found (Goldberg, 1965) to be the momentum canonically conjugate to **A** when gauge-independent charged fields are used. We obtain similar results in the corresponding covariant, observer-dependent formulation (Goldberg & Marx, 1968).

When we use equation (2.7) for **A**, we find that it expresses $\mathbf{A}_S$ in terms of **B**,

$$\mathbf{A}_S(\mathbf{x}, t) = \frac{1}{4\pi} \nabla \wedge \int |\mathbf{x} - \mathbf{x}'|^{-1} \mathbf{B}(\mathbf{x}', t) \, d^3x' \tag{4.11}$$

which indicates that the gauge-independent formulation can be carried out in terms of the electromagnetic field alone, and is nonlocal in space.

The same difference in the roles of the different parts or components of the potential is in evidence when the free electromagnetic field is quantized in terms of wave functionals (Katz, 1965; Marx, 1969a). Again there is no relationship between the states of the quantized electromagnetic field as described by different observers.

Once we realize the observer dependence of the dynamical part of the electromagnetic potential, as exemplified by equation (4.11), we no longer feel compelled to eliminate from consideration an observer-dependent interaction, such as the one we have proposed for the two-component spinor field (Marx, 1970c). In the case of an external electromagnetic field, our starting point is the Lagrangian density

$$\mathscr{L} = (D_\mu^* \chi_{\dot{A}}^*) N_{\mu\nu}^{\dot{A}B} D_\nu \chi_B - m^2 \chi_{\dot{A}}^* n^{\dot{A}B} \chi_B \tag{4.12}$$

where

$$N_{\mu\nu}^{\dot{A}B} = g_{\mu\nu} n^{\dot{A}B} + i\epsilon_{\mu\nu\alpha\beta} n_\alpha \sigma_\beta^{\dot{A}B} \tag{4.13}$$

The charge density in this theory is indefinite, not positive definite as is the one we obtain for the unquantized Dirac field; consequently, we can extend to this spin-$\frac{1}{2}$ field our probabilistic interpretation of relativistic quantum mechanics (Marx, 1969b).

5. *Gauge Invariance of Probability Amplitudes*

The usual approach to a relativistic quantum mechanics of free spin-0 and spin-$\frac{1}{2}$ particles starts from Lagrangian densities that lead to the Klein–

Gordon and Dirac equations respectively. Electromagnetic interactions are then introduced through the gauge-invariant substitution†

$$\partial_\mu \to D_\mu = \partial_\mu \pm ieA_\mu \tag{5.1}$$

and the addition of the free-field Lagrangian density for the electromagnetic field,

$$\mathscr{L}_{\text{e.m.}} = -\tfrac{1}{4}F_{\mu\nu}F_{\mu\nu} \tag{5.2}$$

when the electromagnetic field is dynamical, rather than external. The gauge invariance of the theory is then a consequence of

$$\begin{aligned}(\partial_\mu - ie\bar{A}_\mu)\bar{\psi}_j &= (\partial_\mu - ieA_\mu - ie\Lambda_{,\mu})[\exp(ie\Lambda)\psi_j] \\ &= \exp(ie\Lambda)(\partial_\mu - ieA_\mu)\psi_j\end{aligned} \tag{5.3}$$

where ψ_j are the components of the charged field.

For the charged scalar field, we define the probability amplitudes for particles and antiparticles by (Marx, 1969b)

$$g^{(\pm)}(x) = \left(\frac{\tilde{E}}{2}\right)^{1/2} \phi(x) \pm i(2\tilde{E})^{-1/2} D_0(x)\phi(x) \tag{5.4}$$

where $\tilde{E}$ is the integral operator

$$\tilde{E} = (-\nabla^2 + m^2)^{1/2} \tag{5.5}$$

We have not replaced ∇^2 by $\mathbf{D}^2$, which would have introduced many new complications in this operator, and its counterpart in momentum space,

$$k_0 = (\mathbf{k}^2 + m^2)^{1/2} \tag{5.6}$$

Consequently, the probability amplitudes (5.4) do not transform in a simple manner under gauge transformations, and the probability densities are not gauge invariant.

There are at least two reasons why this is not a particularly serious problem. As we have pointed out, the quantities $N^{(+)}(t)$ and $N^{(-)}(t)$ do not necessarily remain less than 1 at intermediate times, so that they might not correspond to probabilities, and probability densities do not have to be measurable at intermediate times. Hence, it is sufficient to assume that the fields are free at t_i and t_f to avoid problems with gauge invariance. Alternatively, we can use the gauge-independent fields (2.21) throughout, together with the operator (Marx, 1970c)

$$D'_\mu = \partial_\mu \pm ieA'_\mu \tag{5.7}$$

where

$$A'_\mu(\tilde{x}, \tau) = \frac{1}{4\pi}\int d^3\tilde{x}' \, [-(\tilde{x} - \tilde{x}')^2]^{-1/2} \, \tilde{\partial}'_\alpha F_{\mu\alpha}(\tilde{x}', \tau) \tag{5.8}$$

† The sign of the second term depends on one's choice of the charge of the particle, represented by the positive frequency part of the field.

is a gauge-independent, observer-dependent part of the potential. Then

$$g'^{(\pm)}(x) = \left(\frac{\tilde{E}}{2}\right)^{1/2} \phi'(x) \pm i(2\tilde{E})^{-1/2} D_0'(x)\, \phi'(x) \tag{5.9}$$

are gauge-invariant probability amplitudes.

There is another important reason (Marx, 1970c) why we should use gauge-independent fields in conjunction with a dynamical electromagnetic field. This has to do with the equation of constraint (4.7) or (4.9), which has to be satisfied by the given values of the electromagnetic field† either at t_i or at t_f. But the charged field is not specified completely at either time, so that we do not know the charge density and we cannot specify the electromagnetic field consistently *a priori*. When we use gauge-independent fields, the static field is incorporated into the particle field; the equation of constraint for the solenoidal part of the electric field or the gauge-independent part of the potential is homogeneous, and can be trivially satisfied. This whole procedure, of course, is observer dependent.

In the case of a charged spin-$\frac{1}{2}$ field, straightforward application of the same substitution does not lead to a consistent theory, since the charge density is positive definite. We then have to use a somewhat different theory, such as the ones proposed by Marx (1970b, 1970c); in either case, the definition of the probability amplitudes in terms of the original field is not gauge invariant unless we use gauge-independent fields at the outset.

The above formulations correspond closely to Feynman's propagator approach to quantum electrodynamics (Feynman, 1949; Bjorken & Drell, 1964), although the details and the interpretation may differ.‡

The relativistic quantum mechanics of charged particles in an external electromagnetic field has the general form of the nonrelativistic theory; in particular, the Schrödinger equation is linear in the field. When we consider a dynamical electromagnetic field, the equations of motion are no longer linear, which makes the use of Green functions somewhat questionable. Nevertheless, the probability amplitudes are still in a Hilbert space, and the operators that correspond to the observables are linear operators.

6. *Concluding Remarks*

In the preceding sections, we have discussed a number of problems related to Lorentz covariance, gauge invariance and observer dependence. Rather than seeking new answers, we have concentrated on exploring the questions and pointing out the shortcomings of the usual treatments.

† A real field, such as the electromagnetic field, has to satisfy either retarded or advanced boundary conditions. Causal boundary conditions would overspecify the field, since either the positive or the negative frequency part determine both the real field and its time derivative, which cannot be given at two different times. This is related to the fact that the causal Green function is complex, while the retarded and advanced ones are real.

‡ For instance, we emphasize a description of processes that take a finite time, and an extension to infinite times, to obtain a scattering matrix, might lead to mathematical problems.

We have identified χ and $\mathbf{A}_S$ as the parts of the electromagnetic potential A_μ that are independent of the gauge. We have found that χ can be expressed in terms of the charge density ρ, and $\mathbf{A}_S$ in terms of the magnetic field $\mathbf{B}$, which makes the usefulness of the potential questionable in certain areas of physics, especially quantum field theory.† We have also pointed out the reason why the Coulomb gauge has definite advantages over the covariant Lorentz gauge in similar situations.

We have emphasized the difference between a change of coordinates in space-time, or Lorentz transformation, and a change of observer. The usual assumption has the observer at rest in the Lorentz frame, which makes this distinction disappear; it is to a certain extent a matter of taste whether one finds this distinction useful. On the other hand, the acceptance of the dependence of the preparation and outcome of experiments on the observer is of physical significance. It is quite possible that the consideration of fields as physical systems with an infinite number of degrees of freedom, which extend over all space,‡ is an idealization of experimental situations that leads to unavoidable mathematical difficulties; in that case we would have to resort to a different description of a field, or to one of the action-at-a-distance theories.

We conclude that the observer plays an important role, especially in quantum theory of fields and relativistic quantum mechanics. Manifest covariance independent of the observer has only a limited usefulness, and it quite possibly tends to obscure the physical differences between spacelike and timelike directions.

The fact that observers and instruments also obey the laws of physics, and that they interfere with the system during a measurement, was a significant contribution to the understanding of quantum mechanics. We feel it is now necessary to clarify the role of the observer in relativistic theories of particles and fields.

References

Berezin, F. A. *The Method of Second Quantization*. Academic Press, New York.

Bergmann, P. G. and Goldberg, I. (1955). *Physical Review*, **98**, 531.

Bjorken, J. D. and Drell, S. D. (1964). *Relativistic Quantum Mechanics*. McGraw-Hill Book Company, New York.

Bjorken, J. D. and Drell, S. D. (1965). *Relativistic Quantum Fields*. McGraw-Hill Book Company, New York.

† We are, of course, aware of the Gupta–Bleuler formalism in quantum electrodynamics, in which all four components of A are quantized independently, at the cost of introducing subsidiary conditions on physical state vectors and an indefinite metric in Hilbert space. No physical significance has been attached to this procedure, and it is quite possible that here too an elegant formalism hides physical difficulties.

‡ The alternative of putting the system into a large rectangular box, and imposing periodic boundary conditions seems even more unphysical to us, in addition of being *too* asymmetric in space and time. The other possibility that is often mentioned, of having to consider the global properties of spacetime, that is, general relativity, is too discouraging to be accepted before other alternatives of a particle and field theory in flat spacetime are fully explored and found wanting.

Case, K. M. (1957). *Physical Review*, **107**, 307.
Davydov, A. S. (1965). *Quantum Mechanics*, p. 313. Addison-Wesley Publishing Company, Inc., Reading, Mass.
Dirac, P. A. M. (1950) *Proceedings of the Royal Society*, **235**, 138.
Feynman, R. P. (1949). *Physical Review*, **76**, 749, 769.
Fleming, G. N. (1965). *Physical Review*, **137**, B188.
Fleming, G. N. (1966). *Journal of Mathematical Physics*, **7**, 1959.
Goldberg, I. (1958). *Physical Review*, **112**, 1361.
Goldberg, I. (1965). *Physical Review*, **139**, B1665.
Goldberg, I. and Marx, E. (1967). *Nuclear Physics*, **B3**, 25.
Goldberg, I. and Marx, E. (1968). *Nuovo cimento*, **57B**, 485.
Katz, A. (1965). *Nuovo cimento*, **37**, 342.
Marx, E. (1967). *American Journal of Physics*, **35**, 14.
Marx, E. (1968). *Nuovo cimento*, **57B**, 43.
Marx, E. (1969a). *Nuovo cimento*, **60A**, 683.
Marx, E. (1969b). *Nuovo cimento*, **60A**, 669.
Marx, E. (1970a). *Nuovo cimento*, **67A**, 129.
Marx, E. (1970b). *Relativistic Quantum Mechanics of Fermions. Int. J. Theor. Phys.* (in press).
Marx, E. (1970c). *The Two-Component Spinor Field. Physica* (in press).
Rzewuski, J. (1964). *Field Theory.* Part I: *Classical Theory*, p. 123 ff, p. 208 ff. Polish Scientific Publishers, Warsaw.
Rzewuski, J. (1969). *Field Theory.* Part II: *Functional Formulation of S-Matrix Theory.* Polish Scientific Publishers, Warsaw.
Schwinger, J. (1959). *Physical Review Letters*, **3**, 276.
Wheeler, J. A. and Feynman, R. P. (1945). *Reviews of Modern Physics*, **17**, 157.
Wheeler, J. A. and Feynman, R. P. (1949). *Reviews of Modern Physics*, **21**, 425.