

Why semiclassical electrodynamics is not gauge invariant

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

2000 J. Phys. A: Math. Gen. 33 9165

(<http://iopscience.iop.org/0305-4470/33/50/303>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 171.66.16.124

The article was downloaded on 02/06/2010 at 08:45

Please note that [terms and conditions apply](#).

Why semiclassical electrodynamics is not gauge invariant

A M Stewart

Research School of Physical Sciences and Engineering, The Australian National University,
Canberra ACT 0200, Australia

E-mail: andrew.stewart@anu.edu.au

Received 18 September 2000, in final form 30 October 2000

Abstract. It is shown that in semiclassical electrodynamics, which describes how electrically charged particles move according to the laws of quantum mechanics under the influence of a prescribed classical electromagnetic field, only a restricted class of gauge transformations is allowed. This lack of full gauge invariance, in contrast to the situation in classical and quantum electrodynamics which are fully gauge-invariant theories, is due to the requirement that the scalar potential in the Hamiltonian of wave mechanics should represent a physical potential. Probability amplitudes and energy differences are independent of gauge within this restricted class of gauge transformation.

1. Introduction

This paper is about the role of gauge in the wave mechanics of a charged particle moving non-relativistically in an externally prescribed electromagnetic field. The motion of the charged particle is described by the free-particle Schrödinger equation supplemented by minimal coupling of the charge to the potentials of the electromagnetic field.

The question of the effect of gauge transformations on the Hamiltonian of the Schrödinger equation and on transition probability amplitudes has had a long and sometimes contentious history. It was suggested by Lamb (1952) that in the hydrogenic spectrum different lineshapes would be calculated depending on whether the perturbation between the electron and the electromagnetic field was taken to be $\mathbf{A} \cdot \mathbf{p}$ or $\mathbf{E} \cdot \mathbf{r}$, the different forms of the interaction being related by a gauge transformation (Stewart 2000). Although the particular case raised by Lamb appears to have been resolved subsequently by considering the intermediate states in the radiative process, see Cohen-Tannoudji *et al* (1989) for a review (but also see the discussion of Woolley (2000)), the suggestion remained that transition amplitudes and probabilities could depend on gauge. It was also pointed out by Yang (1976) that the Hamiltonian of the problem was not independent of gauge and so a question arose as to the interpretation of the energy. There was much discussion of the issue in the literature (Aharonov and Au 1981, 1983, Feuchtwang *et al* 1984a, b, 1982, Haller 1984, Kobe 1978, 1982, 1984a, b, Kobe and Yang 1987a, b, Power 1989, Power and Thirunamachandran 1978, Rau 1996, Schlicher *et al* 1984, Yang 1981, Woolley 2000) but no clear consensus appears to have emerged.

This paper proposes a new approach to the problem in which gauge freedom is explicitly maintained throughout by the presence of an arbitrary gauge function in the equations. In section 2 we recall the derivation of the Schrödinger equation for a charged particle in an

electromagnetic field and in section 3 we obtain the basis functions that are used to describe stationary states and that are needed for use in perturbation theory when an arbitrary gauge is adopted. In section 4 we obtain matrix elements of operators in an arbitrary gauge and find that the Hamiltonian is not gauge invariant. In section 5 we realize that for the scalar potential to represent the physical potential a restricted set of gauges must be used, namely those that are a sum of functions that depend, respectively, on spatial coordinates only and on time only. In section 6 we find that probability amplitudes are essentially independent of gauge within this restricted set and in section 7 we discuss why semiclassical electrodynamics, in contrast to classical and quantum electrodynamics, is not fully gauge invariant. In the appendices a gauge transformation needed for the basis functions describing time-independent fields and a result concerning the canonical commutation relations are obtained.

2. Schrödinger equation

The Lagrangian of a particle of charge e and mass m at position \mathbf{r} and time t in the electromagnetic field described by vector potential $\mathbf{A}(\mathbf{r}, t)$ and scalar potential $\phi(\mathbf{r}, t)$ is taken to be (Doughty 1980, Schwinger *et al* 1998)

$$L = mv^2/2 + e(\mathbf{v} \cdot \mathbf{A} - \phi) \quad (1)$$

where the velocity \mathbf{v} is given by $\mathbf{v} = d\mathbf{r}/dt$.

If the electromagnetic potentials \mathbf{A} and ϕ are transformed into \mathbf{A}_χ and ϕ_χ by

$$\begin{aligned} \mathbf{A} &\rightarrow \mathbf{A}_\chi = \mathbf{A} + \nabla\chi \\ \phi &\rightarrow \phi_\chi = \phi - \partial\chi/\partial t \end{aligned} \quad (2)$$

where ∇ is the gradient operator with respect to \mathbf{r} and the subscript attached to the potentials indicates the gauge function that is used, then the electromagnetic fields \mathbf{B} and \mathbf{E} given by

$$\begin{aligned} \mathbf{B} &= \nabla \times \mathbf{A} \\ \mathbf{E} &= -\nabla\phi - \partial\mathbf{A}/\partial t \end{aligned} \quad (3)$$

are unchanged. The arbitrary scalar field $\chi(\mathbf{r}, t)$ is known as the gauge function and the transformation as a gauge transformation. The gauge function is required to satisfy the condition $\{(\partial/\partial i)(\partial/\partial j) - (\partial/\partial j)(\partial/\partial i)\}\chi = 0$ everywhere, where i and j are any pair of the coordinates x, y, z and t . Under the gauge transformation, the Lagrangian in the new gauge becomes

$$L_\chi = mv^2/2 + e\mathbf{v} \cdot (\mathbf{A} + \nabla\chi) - e(\phi - \partial\chi/\partial t) \quad (4)$$

the additional terms amounting to the total derivative $d\chi/dt$ thereby leaving the Euler-Lagrange equation

$$m d\mathbf{v}/dt = e(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \quad (5)$$

unchanged. In the new gauge the canonical momentum of the particle $\mathbf{p}_\chi = \partial L_\chi/\partial \mathbf{v}$ is given by

$$\mathbf{p}_\chi = m\mathbf{v} + e(\mathbf{A} + \nabla\chi) \quad (6)$$

and the Hamiltonian $H_\chi(\mathbf{p}_\chi, \mathbf{r}, t) = \mathbf{v} \cdot \mathbf{p}_\chi - L_\chi$ by

$$H_\chi(\mathbf{p}_\chi, \mathbf{r}, t) = \{\mathbf{p}_\chi - e(\mathbf{A} + \nabla\chi)\}^2/2m + e(\phi - \partial\chi/\partial t). \quad (7)$$

The passage to the wave mechanics of Schrödinger is made by changing classical observables into operators and taking the Hamiltonian to be equivalent to the operator $i\hbar(\partial/\partial t)$ acting upon a wavefunction $\Psi(\mathbf{r}, t)$ which is the probability density amplitude for the particle to be at position \mathbf{r} at time t and \hbar is Planck's constant. Furthermore, from the canonical commutation relations $[r^i, p_\chi^j] = i\hbar\delta_{i,j}$ of Dirac (1947) the canonical momentum \mathbf{p}_χ becomes the differential operator $-i\hbar\nabla$ (but see appendix A). This, true in all gauges, leads to the Schrödinger wave equation

$$H_\chi(-i\hbar\nabla, \mathbf{r}, t)\Psi_\chi(\mathbf{r}, t) = i\hbar(\partial/\partial t)\Psi_\chi(\mathbf{r}, t), \quad (8)$$

the Hamiltonian H_χ in the gauge χ being given by (7).

From the readily verified identity

$$\{(\mathbf{p} - e(\mathbf{A} + \nabla\chi))^s\}^s \Psi \exp\{ie\chi/\hbar\} = \exp\{ie\chi/\hbar\} \{\mathbf{p} - e\mathbf{A}\}^s \Psi \quad (9)$$

where $\mathbf{p} = -i\hbar\nabla$, s is a positive integer and Ψ is any function of \mathbf{r} and t , it follows that if, under a gauge transformation, the wavefunction is assumed to transform according to

$$\Psi_\chi(\mathbf{r}, t) \rightarrow \Psi_\chi(\mathbf{r}, t) = \Psi_0(\mathbf{r}, t) \exp\{ie\chi(\mathbf{r}, t)/\hbar\} \quad (10)$$

then $H_\chi\Psi_\chi = i\hbar(\partial/\partial t)\Psi_\chi$ implies $H_0\Psi_0 = i\hbar(\partial/\partial t)\Psi_0$ or, in other words, the Schrödinger equation is invariant in form under a gauge transformation.

3. Construction of basis functions

The object of wave mechanics is to obtain Ψ from equation (8) given a particular set of potentials and obtain physical quantities from the expectation values of the operators that correspond to them. When equation (8) admits no simple solution the usual method of solving it is to expand $\Psi(\mathbf{r}, t)$ in a set of basis functions $\Psi_n(\mathbf{r}, t)$ that are solutions of some simpler equation as

$$\Psi_\chi(\mathbf{r}, t) = \sum_n a_{\chi,n}(t) \Psi_{\chi,n}(\mathbf{r}, t) \quad (11)$$

where the $a_{\chi,n}(t)$ form a set of expansion coefficients. The quantum numbers n are associated with a particular quantum state of the simpler system and the subscript χ is included to indicate that all quantities may, in general, depend on the gauge that is used. The probability amplitude $p'_{\chi,n}(t)$ for the system to be in state n at time t is defined to be the projection of $\Psi_{\chi,n}$ onto Ψ_χ ,

$$p'_{\chi,n}(t) = \int \Psi_{\chi,n}^*(\mathbf{r}, t) \Psi_\chi(\mathbf{r}, t) d\mathbf{r}, \quad (12)$$

where $d\mathbf{r}$ is the 3-volume element at \mathbf{r} . If the $\Psi_{\chi,n}(\mathbf{r}, t)$ are orthonormal with respect to integration over space it follows that the probability amplitude is equal to $a_{\chi,n}(t)$, depending, in general, on gauge. The probabilities themselves $|a_{\chi,n}(t)|^2$, being observables, should not depend on gauge but the amplitudes, on the basis of this argument, might depend on gauge.

Equation (8) cannot be solved by separation of variables to obtain the simpler set of functions unless the operator on the left, H_χ , is independent of time. If the fields $\mathbf{E}(\mathbf{r})$ and $\mathbf{B}(\mathbf{r})$ have no dependence on time it is always possible to find a gauge in which the potentials have no time dependence either; we call them $\mathbf{A}^0(\mathbf{r})$ and $\phi^0(\mathbf{r})$. This claim is justified by the following argument. The vector potential \mathbf{A} can, as can any 3-vector, be expressed as the sum of a solenoidal part \mathbf{A}_T , with $\nabla \cdot \mathbf{A}_T = 0$, and a longitudinal part $\nabla\alpha(\mathbf{r}, t)$ (Panofsky and Philips 1955). By making a gauge transformation with the gauge function equal to $-\alpha(\mathbf{r}, t) + \beta(\mathbf{r}) - \gamma t$, where γ is a number, the longitudinal part becomes $\nabla\beta(\mathbf{r})$, independent of time. From

$\partial \mathbf{B} / \partial t = \mathbf{0}$ it follows that the relation $\partial(\nabla \times \mathbf{A}_T) / \partial t = \mathbf{0}$ must hold. The useful solution to this equation is $\partial \mathbf{A}_T / \partial t = \mathbf{0}$, the other solution $\nabla \times \mathbf{A}_T = \mathbf{0}$ is uninteresting as it gives zero \mathbf{B} . Accordingly, $\mathbf{A}^0(\mathbf{r}) = \mathbf{A}_T(\mathbf{r}) + \nabla \beta(\mathbf{r})$ and is a function of \mathbf{r} alone. The electric field is given in terms of the transformed potentials by $\mathbf{E} = -\nabla \phi^0 - \partial \mathbf{A}^0 / \partial t$. Since $\partial \mathbf{E} / \partial t = \mathbf{0}$ it follows that $\partial \nabla \phi^0 / \partial t = \mathbf{0}$; again the useful solution is $\partial \phi^0 / \partial t = 0$, the other solution $\nabla \phi^0 = \mathbf{0}$ is also uninteresting as it gives zero \mathbf{E} . Accordingly, ϕ^0 as well as \mathbf{A}^0 can be chosen to be a function of \mathbf{r} only. We note that other time-independent gauges $\mathbf{A}^{0'} = \mathbf{A}^0 + \nabla f(\mathbf{r})$ and $\phi^{0'} = \phi^0 - g$ can be constructed with a gauge transformation using a gauge function $\chi^0(\mathbf{r}, t) = f(\mathbf{r}) + gt$, where g is a number. Of course, time-dependent gauges for these time-independent fields may be constructed by using a gauge function that depends more generally on time, but are not appropriate for the present purpose.

We give two examples of time-independent gauges. The first is the potentials of electro- and magnetostatics expressed in terms of the time-independent charge $\rho(\mathbf{r})$ and current $\mathbf{J}(\mathbf{r})$ densities (Panofsky and Philips 1955)

$$\phi_1^0(\mathbf{r}) = \frac{1}{4\pi\epsilon_0} \int \frac{\rho(\mathbf{r}') d\mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|} \quad \text{and} \quad \mathbf{A}_1^0(\mathbf{r}) = \frac{\mu_0}{4\pi} \int \frac{\mathbf{J}(\mathbf{r}') d\mathbf{r}'}{|\mathbf{r} - \mathbf{r}'|}. \quad (13)$$

In the second example the potentials are expressed in terms of the fields in the multipolar gauge (Kobe 1982, Stewart 1999, Valatin 1954, Woolley 1973)

$$\phi_2^0(\mathbf{r}) = -(\mathbf{r} - \mathbf{R}) \cdot \int_0^1 \mathbf{E}(\mathbf{q}) du \quad \text{and} \quad \mathbf{A}_2^0(\mathbf{r}) = -(\mathbf{r} - \mathbf{R}) \times \int_0^1 \mathbf{B}(\mathbf{q}) u du \quad (14)$$

where $\mathbf{q} = u\mathbf{r} + (1-u)\mathbf{R}$. The multipolar gauge describes generally the potentials that arise from fields that depend on time, but of course can be used equally well with fields that do not. It is shown in appendix B that the two sets of potentials above are related by a gauge transformation with gauge function $\chi^0(\mathbf{r}, t) = f(\mathbf{r}) + gt$, and $f(\mathbf{r})$ and g are obtained explicitly. Of course, full gauge arbitrariness remains in expressions (13) and (14) because another arbitrary gauge function χ' can always be added to them according to equation (2). The gauge must then be fixed by setting χ' to zero, but it is only legitimate to do this if it can be demonstrated that the result of the calculation of any physical property is independent of χ' .

However, due to the presence of χ the Hamiltonian operator on the left of equation (8) is still time dependent even when the potentials are static. In this circumstance the wavefunction cannot be separated into the product of a time-dependent part and a space-dependent part. To find a solution to equation (8) with static potentials we temporarily set the gauge function to zero. This gives

$$H_0^0 \Psi_0(\mathbf{r}, t) = \{(\mathbf{p} - e\mathbf{A}^0)^2 / 2m + e\phi^0\} \Psi_0(\mathbf{r}, t) = i\hbar(\partial/\partial t) \Psi_0(\mathbf{r}, t) \quad (15)$$

with the zero superscript indicating that the potentials are time independent, i.e. that the gauge function is set to zero. The operator on the left-hand side is now independent of time so the wavefunction may be factored into a space-dependent part $\psi(\mathbf{r})$ and a time-dependent part. These separate to give

$$\Psi_{0,n}(\mathbf{r}, t) = \psi_n(\mathbf{r}) \exp(-iE_n t / \hbar) \quad (16)$$

with E_n and $\psi_n(\mathbf{r})$ being given by the eigenvalue equation

$$\{(\mathbf{p} - e\mathbf{A}^0)^2 / 2m + e\phi^0\} \psi_n(\mathbf{r}) = E_n \psi_n(\mathbf{r}) \quad (17)$$

whose solutions E_n and $\psi_n(\mathbf{r})$ are assumed to be obtainable. The $\psi_n(\mathbf{r})$ are complete and orthonormal because the operator on the left of (17) is Hermitian. The solutions of equation (8) for static potentials and arbitrary χ can then be obtained from equation (10) to be

$$\Psi_{\chi,n}(\mathbf{r}, t) = \psi_n(\mathbf{r}) \exp[i\{e\chi(\mathbf{r}, t) - E_n t/\hbar\}]. \quad (18)$$

It can be verified by substitution or from inspection of (10) that equation (18) is a solution of equation (8) with the time-independent potentials A^0 and ϕ^0 . The $\Psi_{\chi,n}(\mathbf{r}, t)$ are solutions of the time-dependent Schrödinger equation with arbitrary gauge for time-independent fields. Because of the orthonormality of the $\psi_n(\mathbf{r})$, the $\Psi_{\chi,n}$ remain orthonormal with

$$\int \Psi_{\chi,m}^*(\mathbf{r}, t) \Psi_{\chi,n}(\mathbf{r}, t) d\mathbf{r} = \delta_{m,n}. \quad (19)$$

If the Hamiltonian of equation (7) is the exact Hamiltonian operator, the solutions $\Psi_{\chi,n}$ are exact; if it is only part of the total Hamiltonian then the solutions (18) form a basis set with which perturbation theory may be carried out.

4. Matrix elements

It can be seen from equation (10) that for an operator V to have matrix elements that are independent of gauge and so be physically observable it must satisfy the relation

$$V_\chi = \exp(ie\chi/\hbar) V_0 \exp(-ie\chi/\hbar). \quad (20)$$

Any operator that represents a physically observable quantity must be gauge invariant in this way. Neither the operators \mathbf{A} nor \mathbf{p} are gauge invariant, but the operator $\mathbf{p} - e\mathbf{A}$ is. From equation (9) any operator consisting of positive powers of the components of $(\mathbf{p} - e\mathbf{A})$, such as the spin-orbit interaction (Frohlich and Studer 1993), will satisfy equation (20). The quantum mechanical operator identified with the particle velocity is $\mathbf{v} = d\mathbf{r}/dt = [\mathbf{r}, H]/i\hbar$, where H is the Hamiltonian. By commuting \mathbf{r} with the non-relativistic Hamiltonian $H = (\mathbf{p} - e\mathbf{A})^2/2m + e\phi$, the velocity operator \mathbf{v} is given by $(\mathbf{p} - e\mathbf{A})/m$, so for this Hamiltonian $(\mathbf{p} - e\mathbf{A})$ represents $m\mathbf{v}$, the mass times the velocity.

The semiclassical Hamiltonian of equation (7) with \mathbf{p}_χ replaced by $-i\hbar\nabla$ obeys the relation

$$H_\chi = \exp(ie\chi/\hbar) H_0 \exp(-ie\chi/\hbar) - e\partial\chi/\partial t \quad (21)$$

and so is not gauge invariant in the above sense. The first term on the right-hand side of equation (7), the kinetic energy term, does on its own, satisfy equation (20), but because of the presence of the $\partial\chi/\partial t$ term the full Hamiltonian H_χ is not gauge invariant. Accordingly, unless $\partial\chi/\partial t = 0$, the Hamiltonian cannot represent the energy which is a physical observable. This is the issue addressed by this paper whose purpose it is to explain how gauge-independent physical predictions can arise out of a Hamiltonian that is not gauge invariant in this sense and, in particular, how probability amplitudes that are essentially independent of gauge can come from a perturbation theory that uses a non-gauge-invariant Hamiltonian.

It is, of course, possible to find operators involving the Hamiltonian that are gauge invariant. $H - e\phi = (\mathbf{p} - e\mathbf{A})^2/2m$, the kinetic energy is one; another is $\{H - i\hbar(\partial/\partial t)\}$, but this is not very useful as it is identically zero.

5. Scalar potential

Since the difficulty we have encountered is associated with the scalar potential we need to ask the question of what is the appropriate gauge to use to describe the scalar potential so that it accurately represents the physical potential. The physical potential at position \mathbf{r} is the work $W(\mathbf{r}, t)$ done to bring a charge e from a point \mathbf{R} , usually at infinity, to \mathbf{r} or

$$W(\mathbf{r}, t) = -e \int_{\mathbf{R}}^{\mathbf{r}} \mathbf{E}(\mathbf{q}, t) \cdot d\mathbf{q} \quad (22)$$

where the force \mathbf{F} on the particle is given by $\mathbf{F} = m d\mathbf{v}/dt$ from equation (5) and noting that the magnetic field term vanishes identically. We see that the physical potential involves only the gauge-invariant field \mathbf{E} . There is no gauge arbitrariness in the physical potential apart from the choice of origin \mathbf{R} which just changes $W(\mathbf{r}, t)$ by the same amount everywhere, potential differences remaining unchanged.

It is now seen that an appropriate gauge for describing the physical potential is the multipolar gauge, because, from equation (14), and noting that $d\mathbf{q} = (\mathbf{r} - \mathbf{R}) du$ where the integral is along the straight line from \mathbf{R} to \mathbf{r} ,

$$\phi(\mathbf{r}, t) = -e \int_{\mathbf{R}}^{\mathbf{r}} \mathbf{E}(\mathbf{q}, t) \cdot d\mathbf{q}. \quad (23)$$

When stationary states are considered it is necessary for the physical potential to be unique which means that ϕ must be independent of path so the line integral of \mathbf{E} around any closed loop must vanish. From equation (3) the loop integral of $\nabla\phi$ is, of necessity, zero, in whatever way ϕ is defined, because ϕ is a scalar field. The loop integral of $\partial\mathbf{A}/\partial t$ converts to a surface integral of $\partial\mathbf{B}/\partial t$ and if this is to be everywhere zero it follows that $\partial\mathbf{B}/\partial t$ itself must be zero everywhere. If we consider stationary states then $\partial\mathbf{E}/\partial t$ must be zero as well.

By using equations (7), (10), (16) or (21) it follows that

$$H_{\chi}^0 \Psi_{\chi,n}(\mathbf{r}, t) = (E_n - e\partial\chi/\partial t) \Psi_{\chi,n}(\mathbf{r}, t) \quad (24)$$

and it is seen that as long as $\partial\chi/\partial t = 0$ the eigenvalues of the unperturbed Hamiltonian are unchanged by a gauge transformation. Accordingly, we conclude that in semiclassical wave mechanics the only gauge transformations that leave the eigenvalues of the physical energies unchanged are those that do not depend explicitly on time. However, if the gauge function is taken to have the form $\chi(\mathbf{r}, t) = f(\mathbf{r}) + g(t)$ then energy *differences* will remain unchanged. Since only energy *differences* are of significance in wave mechanics, in contrast to quantum field theory in which individual energies are significant, particularly when the coupling to gravity is considered, it appears that gauge functions of this more general form are acceptable in wave mechanics. Although up to now we have used gauge functions that depend on both space and time, we now realize that those that depend explicitly on time are not allowable except as mentioned above. The use of the Coulomb gauge is acceptable because the gauge function of the transformation to it does not depend on time *explicitly*, although of course, it does depend on time *implicitly* because it involves the implicitly time-dependent vector potential in the original gauge (Panofsky and Philips 1955). It is in regard to the above restriction on the gauge functions that may be used that semiclassical electrodynamics may be said to be not fully gauge invariant.

6. Time development of probability amplitudes

We now take the potentials to have the more complicated form $\mathbf{A} = \mathbf{A}^0(\mathbf{r}) + \mathbf{A}^1(\mathbf{r}, t)$ and $\phi = \phi^0(\mathbf{r}) + \phi^1(\mathbf{r}, t)$, where \mathbf{A}^0 and ϕ^0 are the time-independent potentials describing static

fields discussed in the previous sections of the paper and \mathbf{A}^1 and ϕ^1 , which are not necessarily small, describe perturbing fields that generally depend on time. In this approach the gauge is first *fixed*, namely specific mathematical forms are prescribed for \mathbf{A}^0 , ϕ^0 , \mathbf{A}^1 and ϕ^1 (for example, $\mathbf{A}^0 = -\mathbf{r} \times \mathbf{B}^0/2$, $\phi^0 = -e/(4\pi\epsilon_0 r)$, etc), and then *unfixed* by adding in the arbitrary gauge function χ which, as discussed, essentially depends explicitly only on \mathbf{r} . It is necessary to show that the physical results obtained do not depend on $\chi(\mathbf{r})$.

By expanding the square in equation (7), namely $H_\chi = \{\mathbf{p} - e(\mathbf{A}^0 + \mathbf{A}^1 + \nabla\chi)\}^2/2m + e(\phi^0 + \phi^1 - \partial\chi/\partial t) = H_\chi^0 + V_\chi$ the wave equation (8) may be written formally as

$$\{H_\chi^0 + V_\chi - i\hbar(\partial/\partial t)\}\Psi_\chi(\mathbf{r}, t) = 0 \quad (25)$$

where

$$H_\chi^0 = \{\mathbf{p} - e(\mathbf{A}^0 + \nabla\chi)\}^2/2m + e(\phi^0 - \partial\chi/\partial t) \quad (26)$$

and

$$V_\chi = -\mathbf{A}^1 \cdot \{\mathbf{p} - e(\mathbf{A}^0 + \nabla\chi)\}e/m + e^2(\mathbf{A}^1)^2/2m + ie\hbar(\nabla \cdot \mathbf{A}^1)/2m + e\phi^1 \quad (27)$$

and where the relation $(\mathbf{p} \cdot \mathbf{A}^1 - \mathbf{A}^1 \cdot \mathbf{p}) = -i\hbar\nabla \cdot \mathbf{A}^1$ has been used. The operator H_χ^0 has the form of an unperturbed Hamiltonian in the gauge χ and apart from χ has no dependence on time. The basis functions arising from it are obtained as described in a previous section. The perturbation V_χ of equation (27) depends on time through \mathbf{A}^1 and ϕ^1 as well as χ . However, the operator V_χ is gauge invariant and satisfies equation (20) because the operator $\mathbf{p} = -i\hbar\nabla$ generates the $e\nabla\chi$ term from the phase factor $(-ie\chi/\hbar)$. Are the two sets of potentials (\mathbf{A}^0, ϕ^0) and (\mathbf{A}^1, ϕ^1) considered to each have their own gauge function χ^0 and χ^1 ? If they are, then since the sum of two arbitrary functions is another arbitrary function, the two may just be added together to give χ , which may be incorporated in equations (26) and (27) as shown above. The whole of the time-independent perturbation series arising from equations (25)–(27) is independent of gauge; this is because the matrix elements are because V_χ is gauge invariant and the energy denominators are because they are energy differences.

To obtain an explicit expression for the probability amplitudes the wavefunction is expressed as a linear combination of the $\Psi_{\chi,n}$ according to equation (11) and as a result of the presence of the perturbation V_χ transitions take place between the basis states causing the $a_{\chi,n}$ to vary with time. By substituting equation (11) into (25) we obtain

$$(H_\chi^0 + V_\chi) \sum_n a_{\chi,n} \Psi_{\chi,n} = i\hbar \sum_n a_{\chi,n}(t) (d/dt) \Psi_{\chi,n} + i\hbar \sum_n \Psi_{\chi,n} (d/dt) a_{\chi,n}. \quad (28)$$

Because the $\Psi_{\chi,n}(\mathbf{r}, t)$ satisfy equation (8) with the time-independent Hamiltonian from which the basis states are obtained, the first terms on the left- and right-hand sides of equation (28) cancel. After multiplying the residual from the left by $\Psi_{\chi,m}^*(\mathbf{r}, t)$, integrating over \mathbf{r} and making use of orthonormality, equation (28) becomes

$$i\hbar \frac{da_{\chi,m}(t)}{dt} = \sum_n a_{\chi,n}(t) \int d\mathbf{r} \Psi_{\chi,m}^*(\mathbf{r}, t) V_\chi(\mathbf{r}, t) \Psi_{\chi,n}(\mathbf{r}, t). \quad (29)$$

By making use of equations (18) and (20) the integral over \mathbf{r} may be carried out to give

$$i\hbar \frac{da_{\chi,m}(t)}{dt} = \sum_n a_{\chi,n}(t) V_{m,n}(t) \exp\{i(E_m - E_n)t/\hbar\} \quad (30)$$

where

$$V_{m,n}(t) = \int d\mathbf{r} \psi_m^*(\mathbf{r}) V_0(\mathbf{r}, t) \psi_n(\mathbf{r}). \quad (31)$$

The important point is that $V_{m,n}(t)$ is independent of gauge. Consequently, the equations of motion (30) for the $a_{\chi,m}(t)$ are precisely the same for all gauges. Since the $a_{\chi,m}(t)$ have to satisfy the requirement $\sum_n |a_{\chi,n}(t)|^2 = 1$ for the normalization of $\Psi_\chi(\mathbf{r}, t)$ to be preserved the only way that the $a_{\chi,m}(t)$ may differ with gauge is by a phase that is independent of the gauge function, by \mathbf{r} or by t , in other words a number or a global phase factor. It follows that there is no loss of generality in taking the probability amplitudes to be independent of gauge, namely $a_{\chi,m}(t) = a_{0,m}(t)$. Consequently, the procedure that has always been followed to carry out perturbation calculations, namely setting the gauge function to zero, is justified. This is the main result of the paper.

7. Why is semiclassical electrodynamics not gauge invariant?

It is a truth universally acknowledged that the measurable predictions of any physical theory must be independent of gauge. This being the case, it is necessary to provide a justification at a fundamental level for the assertion made in this paper that in semiclassical wave mechanics only a restricted set of gauge transformations is allowable.

The issue of the gauge dependence of the quantum Hamiltonian, equation (7) with $\mathbf{p}_\chi = -i\hbar\nabla$, has been discussed by Yang (1976) who defined what he called an energy operator which consists of the kinetic energy term and a potential energy term which has no gauge properties. However, it is difficult to understand the validity of this approach as the potential energy term in the cases that he considers, which involve the binding of an electron to an atom, is a gauge potential and must be treated as a gauge potential as is done in the present paper. On the other hand, Park (1990) considers that the inconsistency is due to the non-relativistic nature of the theory in which time appears not as a quantum variable but as a numerical parameter.

In this paper we take a different view. Semiclassical wave mechanics describes the quantum motion of a particle in a prescribed external electromagnetic field. The field affects the motion of the particle but the particle is not supposed to affect the field. The feature of wave mechanics discussed in this paper, that the Hamiltonian is not fully gauge invariant, is a consequence of this asymmetry in the theory. The same is found in the wave mechanics of the single-particle Dirac equation (Stewart 1997a, b).

In other theories of electrodynamics, the symmetry of interaction between particles and fields is preserved. In non-relativistic classical electrodynamics (Schwinger *et al* 1998) the Hamiltonian is

$$H = (\mathbf{p} - e\mathbf{A})^2/2m + e\phi + \int d\mathbf{r} \{ \varepsilon_0 \mathbf{E}^2/2 + \mathbf{B}^2/2\mu_0 - \rho\phi \} \quad (32)$$

where ρ is the charge density. The first two terms give the energy of the particle moving in the field; they correspond to the wave mechanical Hamiltonian equation (7). The remaining terms give the energy of the field in the presence of the particle. It is clear that when the two identical terms involving the scalar potential are cancelled out this Hamiltonian (a) is positive definite, (b) does not contain the scalar potential explicitly and (c) is gauge invariant. Therefore, the lack of gauge invariance of the wave mechanical Hamiltonian is not necessarily due to its non-relativistic nature.

The situation is similar in quantum electrodynamics. Its Lagrangian 4-density in standard notation (Doughty 1980) is

$$\mathcal{L} = i\bar{\psi} \not{\partial} \psi / 2 - m\bar{\psi} \psi + e\bar{\psi} A_\nu \gamma^\nu \psi - F^{\mu\nu} F_{\mu\nu} / 4 \quad (33)$$

where the parameters and operators now refer to the bare particles. The coupled Euler–Lagrange equations arising from equation (33) are

$$\partial_\nu \partial^\nu A^\mu - \partial^\mu (\partial \cdot A) = -e \bar{\psi} \gamma^\mu \psi \quad \text{and} \quad (i \not{\partial} - m) \psi = -e A_\nu \gamma^\nu \psi \quad (34)$$

where the equation on the left incorporates the two Maxwell equations involving the sources and the one on the right is the Dirac equation for a particle in an electromagnetic field. The energy–momentum tensor for the coupled Maxwell–Dirac fields is

$$T^{\mu\nu} = \bar{\psi} \gamma^\mu (i \partial^\nu - e A^\nu) \psi + F^{\mu\sigma} F_\sigma^\nu + g^{\mu\nu} F^{\lambda\sigma} F_{\lambda\sigma} / 4 \quad (35)$$

and the Hamiltonian 3-density $\mathcal{H} = T^{00}$ is

$$\mathcal{H}(x^\nu) = \psi^\dagger \{ \alpha \cdot (-i \nabla - e \mathbf{A}) + m \beta \} \psi + (\mathbf{B}^2 + \mathbf{E}^2) / 2 \quad (36)$$

where ψ^\dagger is the adjoint of the Dirac field operator ψ which transforms as in equation (10) under a gauge transformation. The operators act on occupation number states that have no gauge properties. The Hamiltonian is the integral of equation (36) over coordinate space and like the Hamiltonian of classical electrodynamics it does not contain the scalar potential explicitly as does the Hamiltonian equation (7) of wave mechanics. It is readily checked that all equations (33)–(36) are gauge invariant, they are unchanged by a gauge transformation. The independence of gauge of the S -matrix, which corresponds to the probability amplitudes of wave mechanics, is also well established. Consequently, although quantum electrodynamics suffers from serious unresolved difficulties such as the renormalization needed to deal with divergent self-energies, its fundamental equations are all gauge invariant. Wave mechanics consists of using only the second of equations (34) (in this paper the non-relativistic version of it) and as a result the symmetry and consistency of quantum field theory is lost. That the Hamiltonian of wave mechanics for a particle in an electromagnetic field, equation (7), depends on gauge is hardly surprising since it is derived from a Lagrangian, equation (4), that also depends on gauge. However, since a gauge transformation causes the Lagrangian to change by a term that is a total time derivative, namely $e(d\chi/dt)$, the resulting equation of motion, the Schrödinger equation, retains the same form under a gauge transformation. The underlying reason why semiclassical wave mechanics is not a fully gauge-invariant theory is that it is not a properly formulated theory of the interaction between particles and fields. It tells only half the story. It has, of course, other shortcomings such as the inability to describe particle creation and destruction at high energies and these shortcomings led to the development of quantum field theory. However, it has been shown in this paper that despite the shortcomings of wave mechanics the probability amplitudes and energies that it predicts are independent of gauge providing it is realized that to maintain a correspondence with the physical energy only certain gauge transformations that preserve energy differences are permissible.

Acknowledgments

Dr R Ball and Professor A R P Rau are thanked for helpful discussions.

Appendix A

The most general solution of the canonical commutation relations $[r^i, p_\chi^j] = i\hbar \delta_{i,j}$ and $[p_\chi^i, p_\chi^j] = 0$ is

$$\mathbf{p}'_\chi = -i\hbar \nabla + \hbar \nabla m(\mathbf{r}, t) \quad (A1)$$

where $m(\mathbf{r}, t)$ is an arbitrary dimensionless scalar function of \mathbf{r} and t (Dirac 1947). Substituting this into (7) gives

$$H'_\chi = \{-i\hbar\nabla + \hbar\nabla m(\mathbf{r}, t) - e(\mathbf{A} + \nabla\chi)\}^2/2m + e(\phi - \partial\chi/\partial t). \tag{A2}$$

If the wavefunction is taken to be of the form

$$\begin{aligned} \Psi'_\chi(\mathbf{r}, t) &= \Psi_0(\mathbf{r}, t) \exp[i\{e\chi(\mathbf{r}, t)/\hbar - m(\mathbf{r}t)\}] \\ &= \Psi_\chi(\mathbf{r}, t) \exp\{-im(\mathbf{r}, t)\} \end{aligned} \tag{A3}$$

then it follows that

$$\begin{aligned} H'_\chi \Psi'_\chi(\mathbf{r}, t) &= \exp\{-im(\mathbf{r}, t)\} H_\chi \Psi_\chi \\ &= \exp\{-im(\mathbf{r}, t)\} \{-i\hbar\nabla - e(\mathbf{A} + \nabla\chi)\}^2/2m + e(\phi - \partial\chi/\partial t) \Psi_\chi. \end{aligned} \tag{A4}$$

Accordingly, the quantity $m(\mathbf{r}, t)$ in (A1) can be absorbed by the phase factor $\exp\{-im(\mathbf{r}, t)\}$ in the wavefunction (A3) and so can be ignored because, unlike in the case of the electromagnetic gauge, it does not give rise to an extra factor analogous to $\partial\chi/\partial t$ in the Hamiltonian.

Appendix B

We derive the relations $\phi_2^0(\mathbf{r}) = \phi_1^0(\mathbf{r}) - g$ and $\mathbf{A}_2^0(\mathbf{r}) = \mathbf{A}_1^0(\mathbf{r}) + \nabla f(\mathbf{r})$ between the potentials of equations (13) and (14) and obtain explicit expressions for $f(\mathbf{r})$ and g . For simplicity we take $\mathbf{R} = \mathbf{0}$. First, we deal with the scalar potential. Using $\mathbf{E}(\mathbf{y}) = -\nabla_y\phi(\mathbf{y})$, where ∇_y is the 3-gradient with respect to \mathbf{y} , we obtain

$$\mathbf{E}_1(\mathbf{y}) = -\frac{1}{4\pi\epsilon_0} \int d\mathbf{r}' \rho(\mathbf{r}') \nabla_y \frac{1}{|\mathbf{y} - \mathbf{r}'|} \tag{B1}$$

and so from equation (14)

$$\phi_2^0(\mathbf{y}) = \frac{1}{4\pi\epsilon_0} \int d\mathbf{r}' \rho(\mathbf{r}') \int_0^r d\mathbf{y} \cdot \nabla_y \frac{1}{|\mathbf{y} - \mathbf{r}'|}. \tag{B2}$$

Performing the integration over \mathbf{y} and putting in the limits $\mathbf{y} = \mathbf{0}$ and \mathbf{r} we obtain $\phi_2^0(\mathbf{r}) = \phi_1^0(\mathbf{r}) - \phi_1^0(\mathbf{0})$, so $g = \phi_1^0(\mathbf{0})$, where

$$\phi_1^0(\mathbf{0}) = \frac{1}{4\pi\epsilon_0} \int \frac{\rho(\mathbf{r}') d\mathbf{r}'}{|\mathbf{r}'|}. \tag{B3}$$

In a similar way

$$\mathbf{A}_2^0(\mathbf{r}) = \frac{\mu_0}{4\pi} \int d\mathbf{r}' \int_0^1 u du \mathbf{r} \times \left\{ \mathbf{J}(\mathbf{r}') \times \mathbf{E}_{ur} \frac{1}{|u\mathbf{r} - \mathbf{r}'|} \right\}. \tag{B4}$$

The triple vector product may be expanded as

$$\left[\mathbf{J}(\mathbf{r}') \left\{ \mathbf{r} \cdot \nabla_{ur} \frac{1}{|u\mathbf{r} - \mathbf{r}'|} \right\} - \{ \mathbf{J}(\mathbf{r}') \cdot \mathbf{r} \} \nabla_{ur} \frac{1}{|u\mathbf{r} - \mathbf{r}'|} \right] \tag{B5}$$

to give two terms in (B4). When the relations between derivatives $u\nabla_{ur}h(u\mathbf{r}) = \nabla h(u\mathbf{r})$ and $(\mathbf{r} \cdot \nabla)h(u\mathbf{r}) = u(\partial h(u\mathbf{r})/\partial u)$ are used (Stewart 1999, 2000) where h is any function of $u\mathbf{r}$, in this case $h(u\mathbf{r}) = |u\mathbf{r} - \mathbf{r}'|^{-1}$, the first term in the integrand becomes $u(\partial h/\partial u) = (\partial/\partial u)(uh) - h$ and (B4) is

$$\mathbf{A}_2^0(\mathbf{r}) = \frac{\mu_0}{4\pi} \int d\mathbf{r}' \left[\frac{\mathbf{J}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} - \int_0^1 du \left\{ \frac{\mathbf{J}(\mathbf{r}')}{|u\mathbf{r} - \mathbf{r}'|} + (\mathbf{J}(\mathbf{r}') \cdot \mathbf{r}) \nabla_r \frac{1}{|u\mathbf{r} - \mathbf{r}'|} \right\} \right]. \tag{B6}$$

The first term of this is just $A_1^0(\mathbf{r})$. Now consider the gradient with respect to \mathbf{r} of the scalar function $f(\mathbf{r})$ with

$$f(\mathbf{r}) = -\frac{\mu_0}{4\pi} \int d\mathbf{r}' \{ \mathbf{J}(\mathbf{r}') \cdot \mathbf{r} \} \int_0^1 \frac{du}{|u\mathbf{r} - \mathbf{r}'|}. \quad (\text{B7})$$

Noting that $\nabla\{\mathbf{J}(\mathbf{r}') \cdot \mathbf{r}\} = \mathbf{J}(\mathbf{r}')$ we obtain

$$\nabla f(\mathbf{r}) = -\frac{\mu_0}{4\pi} \int d\mathbf{r}' \int_0^1 du \left[\frac{\mathbf{J}(\mathbf{r}')}{|u\mathbf{r} - \mathbf{r}'|} + \{\mathbf{J}(\mathbf{r}') \cdot \mathbf{r}\} \nabla \frac{1}{|u\mathbf{r} - \mathbf{r}'|} \right]. \quad (\text{B8})$$

The two terms in (B8) correspond to the second and third terms in (B6) to give $A_2^0(\mathbf{r}) = A_1^0(\mathbf{r}) + \nabla f(\mathbf{r})$ as required. Accordingly, the gauge function for the transformation is $\chi(\mathbf{r}, t) = f(\mathbf{r}) + gt$, where g and f are given by equations (B3) and (B7).

The integral over u in equation (B7) may be done by standard methods to give

$$f(\mathbf{r}) = -\frac{\mu_0}{4\pi} \int d\mathbf{r}' \frac{\{\mathbf{J}(\mathbf{r}') \cdot \mathbf{r}\}}{r} \ln\{1 + x\} \quad (\text{B9})$$

where

$$x = \frac{(1 - r/r')}{(1 - \cos\theta)} \left[\left\{ 1 + \frac{2r(1 - \cos\theta)}{r'(1 - r^2/r'^2)} \right\}^{1/2} - 1 \right] \quad (\text{B10})$$

and r and r' are the magnitudes of the vectors and θ is the angle between them.

References

- Aharonov Y and Au K 1981 *Phys. Lett. A* **86** 269–72
 —1983 *Phys. Lett. A* **95** 412–4
 Cohen-Tannoudji C, Dupont-Roc J and Gilbert G 1989 *Photons and Atoms* (New York: Wiley)
 Dirac P A M 1947 *Quantum Mechanics* (Oxford: Clarendon)
 Doughty N A 1980 *Lagrangian Interaction* (Sydney: Addison-Wesley)
 Feuchtwang T E, Kazes E and Cutler P H 1984a *J. Phys. A: Math. Gen.* **17** 1157–61
 Feuchtwang T E, Kazes E, Grotch H and Cutler P H 1982 *Phys. Lett. A* **93** 4–6
 —1984b *J. Phys. A: Math. Gen.* **17** 151–67
 Frohlich J and Studer U M 1993 *Rev. Mod. Phys.* **65** 733–802
 Haller K 1984 *Quantum Electrodynamics and Quantum Optics* ed A O Barut (New York: Plenum) pp 373–92
 Kobe D H 1978 *Am. J. Phys.* **46** 342–8
 —1982 *Am. J. Phys.* **50** 128–33
 —1984a *J. Phys. A: Math. Gen.* **17** 1141–50
 —1984b *Quantum Electrodynamics and Quantum Optics* ed A O Barut (New York: Plenum) pp 393–8
 Kobe D H and Yang K H 1987a *Eur. J. Phys.* **8** 236
 —1987b *Am. J. Phys.* **56** 549–50
 Lamb W E 1952 *Phys. Rev.* **85** 259–76
 Panofsky W K H and Phillips M 1955 *Classical Electricity and Magnetism* (Cambridge, MA: Addison-Wesley)
 Park D 1990 *Classical Dynamics and Its Quantum Analogues* (Berlin: Springer)
 Power E A 1989 *New Frontiers in Quantum Electrodynamics and Quantum Optics* ed A O Barut (New York: Plenum) pp 555–62
 Power E A and Thirunamachandran T 1978 *Am. J. Phys.* **46** 370–8
 Rau A R P 1996 *Phys. Rev. A* **54** 717–20
 Schlicher R R, Becker W, Bergou J and Scully M O 1984 *Quantum Electrodynamics and Quantum Optics* ed A O Barut (New York: Plenum) pp 405–41
 Schwinger J, DeRaad L L Jr, Milton K A and Wu-yang T 1998 *Classical Electrodynamics* (Reading, MA: Perseus)
 Stewart A M 1997a *Aust. J. Phys.* **50** 869–78
 —1997b *Aust. J. Phys.* **50** 1061–81
 —1999 *J. Phys. A: Math. Gen.* **32** 6091–6

- 2000 *Aust. J. Phys.* **53** 613–29
Valatin J G 1954 *Proc. R. Soc. A* **222** 93–108
Woolley R G 1973 *J. Phys. B: At. Mol. Phys.* **6** L97–9
—2000 *Proc. R. Soc. A* **456** 1803–19
Yang K H 1976 *Ann. Phys.* **101** 62–118
—1981 *Phys. Lett. A* **84** 165–8