The physical significance of gauge independence and gauge covariance in quantum mechanics

1984 J. Phys. A: Math. Gen. 17 151

(http://iopscience.iop.org/0305-4470/17/1/017)

J. Phys. A: Math. Gen. **17** (1984) 151–167. Printed in Great Britain

The physical significance of gauge independence and gauge covariance in quantum mechanics†

T E Feuchtwang, E Kazes, P H Cutler and H Grotch

Physics Department, The Pennsylvania State University, University Park, Pennsylvania 16802, USA

Received 28 March 1983

Abstract. The distinction between gauge independence and gauge covariance in electrodynamics is considered for quantities of physical interest. The gauge independence of directly observable quantities, such as the spectra of operators and energy conserving transition rates, is discussed and contrasted with the gauge covariance of other quantities such as scalar products. The quasi-gauge-invariance of the classical Lagrangian and the invariance of the Euler–Lagrange equations under the addition of a total time derivative to the Lagrangian are discussed. The consequent inherent non-uniqueness of the Lagrangian and Hamiltonian formulations of classical mechanics is pointed out. A physical interpretation of the explicit gauge dependence of classical canonical momenta and of the expectation values of the corresponding quantum mechanical operators is presented. The gauge independence of energy conserving transition rates calculated in the conventional finite-order time dependent perturbation theory is discussed and illustrated. The gauge dependence of the conventional time dependent transition amplitudes in the presence of electromagnetic fields is then discussed, and gauge independent transition amplitudes are constructed. An alternative formulation of the quantum mechanics of charged particles is obtained in terms of a new gauge independent Hamiltonian, which is, as might be expected, unique only within an arbitrary canonical transformation.

1. Introduction

Several recent publications (Aharonov and Au 1981, Eckstein 1979, Kazes *et al* 1982, Kobe and Smirl 1978, Kobe and Wen 1980, Yang 1982) reflect the on-going interest in clarifying the physical significance of gauge covariance and gauge independence both in non-relativistic and relativistic quantum mechanics. The enlightening comments by Aharonov and Au (1981) did not seem to have settled the lingering confusion relating to these important issues.

The objectives of this paper are to clarify the distinction between the gauge independence and the gauge covariance‡ of quantities of physical interest, when the electrodynamic potentials are subjected to a gauge transformation, and to report a recently discovered technique for the construction of gauge independent transition amplitudes and Hamiltonian operators (Feuchtwang *et al* 1983, Kazes *et al* 1983).

We believe that the misunderstanding of the significance of gauge dependence in quantum mechanics is due to the common misconception that the wavefunction provides

† This work has been supported in part by the Office of Naval Research, Arlington, Virginia, contract no. 619-007.

‡ Covariance literally means form invariance. To be consistent with current usage we hereafter use covariance.

0305-4470/84/010151+17$02.25

a complete description of the physical state of a system. In fact, such a description requires not only the wavefunction but also the explicit representation of all observables or operators.

For the sake of simplicity we shall restrict ourselves to the case of a single non-relativistic particle. Both these restrictions are non-essential and can be relaxed†. In our study we consider also classical mechanics which, by virtue of the correspondence principle, imposes constraints on the allowed and/or required gauge dependence of quantum mechanical operators. In particular, we interpret and clarify the gauge dependence of classical canonical momenta and of the expectation values of the corresponding quantum mechanical operators. We also discuss the implication of the gauge dependence of conventional time dependent transition amplitudes in the presence of electromagnetic fields. Two alternative resolutions of the difficulties posed by this gauge dependence are offered, namely, the construction of a gauge independent transition amplitude or of a gauge independent Hamiltonian. Both constructions are presented, and the relation between them is explained.

Finally, we establish the relationship between the gauge dependence of the conventional minimally coupled quantum mechanical Hamiltonian (as well as of the corresponding wavefunctions) and a fundamental non-uniqueness of the Lagrangian formulation of classical mechanics; namely, the equations of motion are unaffected by the addition of a total time derivative to the Lagrangian.

The outline of the paper is as follows.

In § 2 we discuss the physical significance of gauge covariance of operators in general and of the Schrödinger equation in particular. The non-uniqueness of the Lagrangian and canonical momentum both in classical and quantum mechanics is demonstrated and interpreted. The connection between the gauge covariance of the Schrödinger equation and the gauge dependence of the gauge covariant Hamiltonian is pointed out.

In § 3 we discuss the relation between the gauge independence of physical quantities and their measurability. The general conclusions are then tested by an examination of the eigenvalue spectra of directly observable Hermitian operators, and of gauge independent transition amplitudes. Specifically it is demonstrated that in general the conventional definition of transition amplitudes assures the gauge independence only if they are calculated either after the electromagnetic fields have been turned off, or else if they are calculated for energy conserving transitions. The calculational consequence of these results is discussed and illustrated by a treatment of the electric dipole approximation in first-order perturbation theory.

In § 4, the gauge dependence of the conventional time dependent transition amplitudes is considered and interpreted. Two new constructions are described, either of which assures the gauge independence of transition amplitudes in the presence of electromagnetic fields. The freedom of choice of the Lagrangian, and associated canonical momentum and Hamiltonian operators, is utilised in this new theory, to define gauge independent transition amplitudes, by introducing a gauge field either into the field free Hamiltonian or into the full Schrödinger equation.

In the first procedure the solutions of the modified field-free Hamiltonian undergo the same unitary transformation as the solutions of the full Schrödinger equation when the potentials are gauge transformed. In the second procedure the 'gauge function' f makes the full Schrödinger equation gauge independent.

† See Kazes *et al* (1982) for details.

Both procedures eliminate the gauge dependence of the transition amplitudes at the cost of introducing a dependence on the path of integration specifying the gauge function f. This non-uniqueness is traced to the corresponding non-uniqueness of the Hamiltonian, which is defined only within a canonical transformation.

In § 5 we present a discussion of our results, our conclusions, and proposed extensions of our theory to quantum field theory as well as other gauge theories.

2. The physical significance of gauge covariance

The physical requirement that a change in the gauge of the electrodynamic potentials should have no dynamical consequences implies that all operators, as well as the equation of motion, should be covariant under such a transformation†. It is well known that as a result of the gauge transformation,

$$A_0(\boldsymbol{r}, t) \to A_0'(\boldsymbol{r}, t) = A_0(\boldsymbol{r}, t) - \frac{1}{c}\frac{\partial}{\partial t}\Lambda(\boldsymbol{r}, t),$$

$$\boldsymbol{A}(\boldsymbol{r}, t) \to \boldsymbol{A}'(\boldsymbol{r}, t) = \boldsymbol{A}(\boldsymbol{r}, t) + \nabla\Lambda(\boldsymbol{r}, t), \tag{2.1}$$

all state vectors are subject to the same unitary transformation‡. Specifically, if we denote gauge transforms of operators and wavefunctions by a prime, then

$$\psi(\boldsymbol{r}, t) \to \psi'(\boldsymbol{r}, t) = U_\Lambda \psi(\boldsymbol{r}, t). \tag{2.2}$$

The gauge covariance of the Schrödinger equation means that

$$[\mathscr{H}(\boldsymbol{r}, \boldsymbol{p}; A, A_0) - \mathrm{i}\hbar\partial/\partial t]\psi(\boldsymbol{r}, t) = 0$$

and

$$[\mathscr{H}(\boldsymbol{r}, \boldsymbol{p}; \boldsymbol{A}', A_0') - \mathrm{i}\hbar\partial/\partial t]\psi'(\boldsymbol{r}, t) = 0. \tag{2.3a}$$

This requires that the gauge transformed Hamiltonian satisfy the equation

$$\mathscr{H}(\boldsymbol{r}, \boldsymbol{p}; \boldsymbol{A}', A_0') = U_\Lambda \mathscr{H}(\boldsymbol{r}, \boldsymbol{p}; \boldsymbol{A}, A_0) U_\Lambda^{-1} - \mathrm{i}\hbar U_\Lambda \dot{U}_\Lambda^{-1} = \tilde{\mathscr{H}}(\boldsymbol{r}, \boldsymbol{p}; \boldsymbol{A}, A_0) - (e/c)\partial\Lambda/\partial t. \tag{2.3b}$$

In (2.3) and henceforth, unitary transforms are denoted by a tilde. The unitary transformation U_Λ must satisfy the relation

$$U_\Lambda = \exp[\mathrm{i}e\Lambda(\boldsymbol{r}, t)/\hbar c], \tag{2.4}$$

and the Hamiltonian is

$$\mathscr{H}(\boldsymbol{r}, \boldsymbol{p}; \boldsymbol{A}, A_0) = (2m)^{-1}(\boldsymbol{p} - (e/c)\boldsymbol{A}(\boldsymbol{r}, t))^2 + V(\boldsymbol{r}) + eA_0(\boldsymbol{r}, t). \tag{2.5}$$

Equation (2.3b) is so important in the ensuing discussion that it deserves some further elaboration. We first note that the operator on the left-hand side of (2.3b) represents the Hamiltonian in the new gauge in terms of the gauge transformed

† By gauge covariance of operators we mean that all operators which do not explicitly depend on the four-potential are unaffected by a transformation of these potentials. All other operator functions $F(\boldsymbol{r}, \boldsymbol{p}; A_\mu)$ transform into $F(\boldsymbol{r}, \boldsymbol{p}, A_\mu')$.

‡ This transformation is sometimes referred to as a gauge transformation of the first kind, in contrast to the gauge transformation of the potentials which is then referred to as a transformation of the second kind.

four-potentials of (2.1),

$$\mathscr{H}(\boldsymbol{A}', A_0') = (2m)^{-1}(\boldsymbol{p} - (e/c)\boldsymbol{A}'(\boldsymbol{r}, t))^2 + V(\boldsymbol{r}) + eA_0'(\boldsymbol{r}, t),$$

where $\boldsymbol{p}$ is the same momentum operator as in (2.5).

Next note that one can also express $\mathscr{H}(\boldsymbol{A}', A_0')$ in terms of the unitarily transformed momentum,

$$\tilde{\boldsymbol{p}} = p - (e/c)\nabla\Lambda, \tag{2.3c}$$

and the original four-potentials:

$$\mathscr{H}(\boldsymbol{A}', A_0') = \mathscr{H}'(\boldsymbol{A}, A_0) = \tilde{\mathscr{H}}(\boldsymbol{A}, A_0) - (e/c)\partial\Lambda/\partial t, \tag{2.3d}$$

with

$$\tilde{\mathscr{H}}(\boldsymbol{A}, A_0) = (2m)^{-1}(\tilde{\boldsymbol{p}} - (e/c)\boldsymbol{A}(\boldsymbol{r}, t))^2 + V(\boldsymbol{r}) + eA_0(\boldsymbol{r}, t). \tag{2.3e}$$

The second form for the gauge transformed Hamiltonian, (2.3*d*) and (2.3*e*), emphasises the equivalence of gauge and canonical transformations to be discussed below.

In conclusion, it is important to note that while the Hamiltonian operator is gauge covariant, it is not gauge independent. Furthermore, for explicitly time dependent Λ, the transformed Hamiltonian $\mathscr{H}(\boldsymbol{A}', A_0')$ is not the unitary transform of $\mathscr{H}(\boldsymbol{A}, A_0)$.

A similar analysis applies to the commutation relations. The canonical commutation relations characterise the operator algebra of the system, i.e. they specify the physical system considered. This implies that they too should be covariant under gauge transformations. In fact, the commutators are canonical invariants,

$$\begin{aligned} &[p_i, x_j] = -i\hbar\delta_{ij} = [p_i', x_j'], \qquad [p_i, p_j] = 0 = [p_i', p_j'], \\ &[x_i, x_j] = 0 = [x_i', x_j'], \qquad i,j = 1,2,3. \end{aligned} \tag{2.6}$$

The most general unitary transformation which satisfies (2.6) and keeps the coordinate operators invariant is the following (canonical) transformation of the canonical momentum operators,

$$p_j \to \tilde{p}_j = \exp(+ief(\boldsymbol{r}, t)/\hbar c)p_j \exp(-ief(\boldsymbol{r}, t)/\hbar c) = p_j - (e/c)/\partial_j f, \tag{2.7}$$

where f is an arbitrary function of $(\boldsymbol{r}, t)$. The commutation relations therefore define the canonical momentum operator only up to the gradient of an arbitrary function.

A corresponding arbitrariness is also familiar in classical mechanics. The Euler–Lagrange equations of motion are unaffected by the addition of the total time derivative of an arbitrary function of $(\boldsymbol{r}, t)$ to the Lagrangian (Goldstein 1980). Therefore the Lagrangian for a charged particle in the presence of an external electromagnetic field is

$$\mathscr{L} = \tfrac{1}{2}m\dot{\boldsymbol{r}}^2 + (e/c)\dot{\boldsymbol{r}} \cdot \boldsymbol{A} - V(\boldsymbol{r}) - eA_0 - (e/c)\mathrm{d}f(\boldsymbol{r}, t)/\mathrm{d}t \tag{2.8}$$

and thus the classical canonical momentum,

$$p_j = \partial\mathscr{L}/\partial\dot{x}_j = m\dot{x}_j + (e/c)(A_j - \partial_j f), \tag{2.9a}$$

is only defined within the arbitrariness of the vector potential and of $(e/c)\partial_j f$.

Evidently the gauge transformation,

$$A_\mu \to A_\mu' = A_\mu - \partial\Lambda/\partial x^\mu, \tag{2.10}$$

modifies the Lagrangian of (2.8),

$$\mathscr{L} \to \mathscr{L}'(\boldsymbol{r}, \dot{\boldsymbol{r}}, A'_\mu) = \mathscr{L}(\boldsymbol{r}, \dot{\boldsymbol{r}}, A_\mu) + (e/c)\mathrm{d}\Lambda/\mathrm{d}t, \tag{2.11}$$

where we have assumed that $f(\boldsymbol{r}, t)$ does not depend on A_μ. Thus, this Lagrangian is only *quasi-invariant* under a gauge transformation. Furthermore, (2.9*a*) demonstrates that the classical canonical momentum is arbitrary for two completely independent reasons, namely, the freedom in the choice of both f and the gauge for the four-potential. It is, however, remarkable that the gauge transformation $\boldsymbol{A} \to \boldsymbol{A}' = \boldsymbol{A} + \boldsymbol{\nabla}\Lambda$ for a fixed f affects the canonical momentum in the same way as the transformation $f \to f - \Lambda$ for fixed four-potentials. This confirms the gauge dependence of the conventional classical canonical momentum, and suggests a possible procedure for its elimination. These points are considered next.

The equations of motion of a charged particle depend only on the electric and magnetic fields. They are independent of both the gauge Λ and the function f that appears in (2.8). This, however, does not apply to the canonical momentum. The gauge independence of the classical particle velocity implies that the simultaneous gauge transformation of $A_\mu \to A'_\mu$ and the transformation $f \to f'$ necessarily induce a transformation of the classical canonical momentum such that

$$p'_j - p_j = (e/c)[A'_j - A_j - \partial_j(f' - f)]. \tag{2.12}$$

If in particular $f = f' = 0$ then a gauge transformation must also change the canonical momentum. Alternatively, if we choose p_j to be gauge independent, then we must introduce the function f into the Lagrangian, to assure the gauge independence of the right-hand side of (2.12). Proceeding from classical to quantum mechanics we note that the last term in (2.8) is often omitted, and hence (2.9*a*) becomes

$$\dot{x}_j = m^{-1}(p_j - (e/c)A_j). \tag{2.9b}$$

Classically, the velocity is directly observable and hence necessarily gauge independent. Thus the correspondence principle requires the expectation value of the velocity operator to be also gauge independent. This is demonstrated below. Proceeding from (2.2) and (2.3c) we have

$$\langle\psi|\boldsymbol{p} - (e/c)A|\psi\rangle = \langle U_\Lambda\psi|U_\Lambda(\boldsymbol{p} - (e/c)\boldsymbol{A})U_\Lambda^{-1}|U_\Lambda\psi\rangle = \langle\psi'|(\tilde{\boldsymbol{p}} - (e/c)\boldsymbol{A})|\psi'\rangle. \tag{2.13}$$

However, using (2.3*c*) and (2.1) we obtain

$$\langle\psi'|\tilde{\boldsymbol{p}} - (e/c)\boldsymbol{A}|\psi'\rangle = \langle\psi'|\boldsymbol{p} - (e/c)\boldsymbol{A}'|\psi'\rangle. \tag{2.14}$$

Combining (2.13) and (2.14), we obtain (2.15),

$$\langle\psi|\boldsymbol{p} - (e/c)\boldsymbol{A}|\psi\rangle = \langle\psi'|\boldsymbol{p} - (e/c)\boldsymbol{A}'|\psi'\rangle. \tag{2.15}$$

To conclude, (2.14) and (2.15) demonstrate that the conventional *gauge dependent* velocity operator is *gauge covariant* under a gauge transformation,

$$\dot{\boldsymbol{r}}' = \dot{\boldsymbol{r}} - (e/mc)\boldsymbol{\nabla}\Lambda, \tag{2.9c}$$

and has a *gauge independent* expectation value.

This conclusion is generally valid. The gauge independence of classical observables, such as the velocity, does not necessarily imply the gauge independence of the corresponding quantum mechanical operators. It does, however, imply the gauge independence of their expectation values. This condition implies also that for such operators

the gauge transform is equal to the unitary transform, defined by (2.4). Therefore these operators, $\hat{O}$, must satisfy the relation

$$\tilde{\hat{O}} = U_\Lambda \hat{O} U_\Lambda^{-1} = \hat{O}'. \tag{2.16}$$

In contrast, canonical variables can be defined with a considerable amount of arbitrariness, reflecting their dependence on the gauge of the four-potential and the additive total time derivative in the Lagrangian. The latter factor manifests itself also in the transition from classical to quantum mechanics and in the different realisations of the canonical commutation relations, equations (2.6). According to (2.7), the gradient of the arbitrary function f can be added to any given expression for the canonical momentum operator without any physical consequence. This illustrates the known fact that canonical variables are auxiliary constructs introduced at the convenience of the observer, rather than being unique observables specified by the phenomena studied. It is instructive to note that gauge covariance requires the canonical momentum to be gauge independent. The canonical momentum operator, for which $\boldsymbol{p}' = \boldsymbol{p}$, necessarily has a gauge dependent expectation value which, in agreement with the correspondence principle, satisfies the classical gauge transformation, (2.12), with $f = 0$. Proceeding from the correspondence principle and (2.2),

$$\langle \psi' | \boldsymbol{p}' | \psi' \rangle = \langle \psi | U_\Lambda^{-1} \boldsymbol{p} U_\Lambda | \psi \rangle = \langle \psi | \boldsymbol{p} | \psi \rangle + \langle \psi | (e/c) \nabla \Lambda | \psi \rangle, \tag{2.17}$$

where the last equation follows from (2.4) and (2.7).

The preceding emphasises that a given wavefunction does not represent the same physical state of a system in two different gauges. A complete description of a given state requires the specification of a wavefunction and of all the operators representing measurable quantities. This may require the explicit specification of the gauge, as in the case of the velocity operator, equation (2.9c). This point also is illustrated by the following example.

The wavefunctions

$$\psi(\boldsymbol{r}) = \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot\boldsymbol{r}} \tag{2.18a}$$

and

$$\psi'(\boldsymbol{r}) = \exp\{\mathrm{i}[\boldsymbol{k}\cdot\boldsymbol{r} + e\Lambda(\boldsymbol{r}, t)/\hbar c]\} \tag{2.18b}$$

refer to two different physical states if the momentum operator is defined as

$$\boldsymbol{p} = -\mathrm{i}\hbar\nabla. \tag{2.19}$$

On the other hand, these wavefunctions represent the same physical state in two different gauges, if in the second gauge the momentum operator is given by

$$\boldsymbol{p}' = U_\Lambda \boldsymbol{p} U_\Lambda^{-1} = -\mathrm{i}\hbar\nabla - (e/c)\nabla\Lambda. \tag{2.20}$$

We conclude this discussion of gauge covariance with the observation that the covariance of the Schrödinger equation, as defined by (2.2)–(2.4), inherently causes the gauge covariant Hamiltonian to be gauge dependent. We shall see below that in order to obtain a gauge independent Hamiltonian it is necessary to redefine the canonical momentum and the Hamiltonian, within the Lagrangian formalism, in terms of the arbitrary function f so that their dependence on the potentials will only involve the gauge independent combination $A_\mu + \partial f/\partial x^\mu$.

3. Gauge independence and measurability

3.1. The significance of gauge independent eigenvalue spectra of direct observables

The correspondence principle suggests that all directly observable† operators must have gauge independent expectation values and conversely, a gauge dependent expectation value cannot represent the outcome of a measurement. The prediction of the outcome of measurements, by two different observers, on identical systems must therefore be identical, whether or not they communicate to each other their respective choice of gauge. Furthermore, if calculated predictions concerning the outcome of measurements are based on different gauges, then it is not necessary to exhibit the gauge transformation relating them, in order to assure their identity. For instance, the canonical momentum, $-i\hbar\nabla$, which has gauge dependent expectation values is not a directly observable operator (Cohen-Tannoudji 1977, Aharonov and Au 1981).

Among the manifestly gauge independent quantities we have the eigenvalue spectra of all directly observable Hermitian operators. This is the case because corresponding wavefunctions in different gauges are related by unitary transformations and directly observable operators satisfy (2.16). In reference to the energy spectrum, one has to recall that the unitary transform of the time independent Hamiltonian $\mathscr{H}$ is not the conventional gauge transform of this operator, $\mathscr{H}' = \mathscr{H}(\Lambda)$. The conventional definition of the gauge transformed Hamiltonian follows from the transformed Schrödinger equation,

$$[\mathscr{H}(\Lambda) - i\hbar\partial/\partial t]\psi'_\varepsilon(t) = 0, \tag{3.1}$$

where, by virtue of (2.3b),

$$\mathscr{H}(\Lambda) = U_\Lambda \mathscr{H} U_\Lambda^{-1} - (e/c)\,\partial\Lambda/\partial t, \tag{3.2}$$

and thus $\mathscr{H}(\Lambda)$ is not the unitary transform of $\mathscr{H}$. However, writing

$$\psi'_\varepsilon(t) = e^{-i\varepsilon t/\hbar}\phi'_\varepsilon(t) = e^{-i\varepsilon t/\hbar}U_\Lambda\phi_\varepsilon, \tag{3.3}$$

where ϕ_ε is time independent, we obtain the 'time independent' form of (3.1),

$$(U_\Lambda \mathscr{H} U_\Lambda^{-1} - \varepsilon)\phi'_\varepsilon(t) = 0. \tag{3.4}$$

Equation (3.4) is evidently the gauge transform of

$$(\mathscr{H} - \varepsilon)\phi_\varepsilon = 0. \tag{3.5}$$

Thus the gauge transformation corresponds, as far as the 'time independent' Schrödinger equation is concerned, to a unitary transform of (3.5) into (3.4). Therefore, in this limited sense the energy spectrum is explicitly gauge independent.

3.2. Gauge independent transition amplitudes

Another important class of gauge independent quantities are the transition amplitudes discussed below.

† By directly observable operators we mean operators corresponding to variables which can be measured directly without the need for calculations involving theoretical constructs, e.g. the position and velocity, in contrast to the canonical momentum, which requires for its determination the evaluation of $\boldsymbol{A}(\boldsymbol{r})$, as well as a definite choice of the Lagrangian.

3.2.1. The general time dependent overlap matrix elements. Consider the field-free Hamiltonian,

$$\mathcal{H}_0 = p^2/2m + V(\boldsymbol{r}), \tag{3.6}$$

and the corresponding eigenfunctions $\{\phi_\varepsilon\}$,

$$(\mathcal{H}_0 - \varepsilon)\phi_\varepsilon = 0. \tag{3.7}$$

The effect of switching on an electromagnetic field at $t = t_0$ is to transform a given eigenfunction $\phi_\varepsilon(\boldsymbol{r})$ of the field-free Hamiltonian into the wavefunction $\psi_A(\boldsymbol{r}, t; \varepsilon)$ which satisfies the wave equation

$$[\mathcal{H}(\boldsymbol{A}, A_0) - \mathrm{i}\hbar\partial/\partial t]\psi_A(\boldsymbol{r}, t; \varepsilon) = 0, \tag{3.8}$$

subject to the initial condition

$$\psi_A(\boldsymbol{r}, t_0; \varepsilon) = \phi_\varepsilon(\boldsymbol{r}, t_0) = \exp(-\mathrm{i}\varepsilon t_0/\hbar)\phi_\varepsilon(\boldsymbol{r}). \tag{3.9}$$

Neither the Hamiltonian $\mathcal{H}(\boldsymbol{A}, A_0)$ nor the wavefunction $\psi_A(\boldsymbol{r}, t; \varepsilon)$ is uniquely determined by the physical state of the system, because the potentials are only specified within a gauge transformation. One might expect this non-uniqueness to affect the calculation of transition amplitudes to be considered next.

Let ψ_{A_1} and ψ_{A_2} be two wavefunctions related by a gauge transformation,

$$\psi_{A_2}(\boldsymbol{r}, t; \varepsilon) = \exp[\mathrm{i}e\Lambda(\boldsymbol{r}, t)/\hbar c]\psi_{A_1}(\boldsymbol{r}, t; \varepsilon), \tag{3.10}$$

where

$$\boldsymbol{A}_2 = \boldsymbol{A}_1 + \boldsymbol{\nabla}\Lambda, \qquad A_{0,2} = A_{0,1} - c^{-1}\partial\Lambda/\partial t. \tag{3.11a,b}$$

Evidently the overlap matrix elements of the wavefunctions ψ_{A_i} and the field-free eigenstates $\{\phi_\varepsilon\}$ are, in general, gauge dependent. That is,

$$\langle\phi_{\varepsilon'}|\psi_{A_1}(t; \varepsilon)\rangle \neq \langle\phi_{\varepsilon'}|\psi_{A_2}(t; \varepsilon)\rangle. \tag{3.12}$$

This observation is unrelated to the known gauge invariance of scalar products, which guarantees that the evaluation of an overlap matrix, specified by a given choice of potentials, can subsequently be carried out in any gauge, i.e.,

$$\langle\phi_{\varepsilon'}|\psi_{A_1}(t; \varepsilon)\rangle = \langle U_\Lambda\phi_{\varepsilon'}|U_\Lambda\psi_{A_1}(t; \varepsilon)\rangle. \tag{3.13}$$

The significance of (3.12) was considered by Aharonov and Au (1981) (AA). They noted that the time dependent matrix element $\langle\phi_{\varepsilon'}|\psi_{A_i}(t; \varepsilon)\rangle$ is the probability amplitude for the measurement of the quantity $\mathcal{H}_0 = p^2/2m + V(\boldsymbol{r})$ to yield, in the state ψ_{A_i}, the value ε'. This implies that the fraction of times for which the repeated evaluation of the quantity

$$\mathcal{H}_0 = (2m)^{-1}(m\boldsymbol{v} + (e/c)\boldsymbol{A}_i(\boldsymbol{r}))^2 + V(\boldsymbol{r}),$$

in terms of basic gauge invariant measurable quantities, $\boldsymbol{r}$ and $\boldsymbol{v}$, yields ε' is $|\langle\phi_{\varepsilon'}|\psi_{A_i}(t; \varepsilon)\rangle|^2$. Aharanov and Au stress that, because of their dependence on $\boldsymbol{A}_i$, the field-free Hamiltonians above represent distinct and gauge dependent functions of $\boldsymbol{r}$ and $\boldsymbol{v}$. Hence the dependence of the corresponding probability amplitudes on $\boldsymbol{A}_i$, and the inequality of (3.12), have a clear physical meaning.

We shall extend and complement the analysis of AA, by demonstrating the precise conditions under which it applies. We shall distinguish three cases:

(a) transition amplitudes calculated for times after the fields are switched off;

(b) energy conserving transition amplitudes;
(c) transition amplitudes in the presence of fields.

In the first two cases we shall show that the transition amplitudes are explicitly gauge independent. In the third case, to be considered in § 4, we shall present an explicit construction of gauge independent transition amplitudes†.

3.2.2. *Transition amplitudes after the fields have been turned off.* Consider the case where two gauge equivalent potentials have been used to describe a system. After the electromagnetic fields have been turned off, the wavefunctions $\psi_{A_i}(\boldsymbol{r}, t; \varepsilon)$ satisfy the field-free wave equations,

$$[\mathscr{H}_0^{(+)}(\Lambda_i) - \mathrm{i}\hbar\partial/\partial t]\psi_{A_i}(\boldsymbol{r}, t; \varepsilon) = 0, \tag{3.14}$$

where

$$\mathscr{H}_0^{(+)}(\Lambda_i) = (2m)^{-1}(\boldsymbol{p} - (e/c)\boldsymbol{\nabla}\Lambda_i)^2 + V(\boldsymbol{r}) - (e/c)\,\partial\Lambda_i/\partial t, \qquad i = 1, 2. \tag{3.15}$$

Equation (3.15) shows that the freedom in the choice of gauge destroys the uniqueness of the Hamiltonian not only in the presence of fields, but also in their absence. However, the gauge used to describe a given system in the absence of fields is not completely arbitrary, because one and the same gauge function $\Lambda(\boldsymbol{r}, t)$ must be used at all times. The terms $\boldsymbol{\nabla}\Lambda_i$ and $c^{-1}\partial\Lambda_i/\partial t$ in (3.15) represent the '*pure gauge*' or '*residual gauge*' fields which survive after the electromagnetic fields are turned off. A similar consideration applies to $\mathscr{H}_0^{(-)}$, the field-free Hamiltonian describing the system before the electromagnetic fields were turned on. Therefore, at a given fixed time at which the fields are off, both asymptotic forms of the Hamiltonians are given by

$$\mathscr{H}_0(\Lambda_i) = U_{\Lambda_i}\mathscr{H}_0 U_{\Lambda_i}^{-1} - (e/c)\partial\Lambda_i/\partial t. \tag{3.16}$$

Consequently the physically meaningful transition amplitudes from a state after the fields are turned off to states of the field-free Hamiltonian at the same time are

$$\langle\psi'_{\varepsilon'}(t)|\psi_A(t; \varepsilon)\rangle = \langle U_\Lambda \exp(-\mathrm{i}\varepsilon' t/\hbar)\phi_{\varepsilon'}|\psi_A(t; \varepsilon)\rangle \tag{3.17a}$$

$$= \exp(\mathrm{i}\varepsilon' t/\hbar)\langle\phi_{\varepsilon'}|U_\Lambda^{-1}\psi_A(t; \varepsilon)\rangle \tag{3.17b}$$

where $\psi'_{\varepsilon'}(t)$ is the solution to the field-free version of the Schrödinger equation defined by (3.1). Equation (3.17*a*) follows from (3.3) and (3.4), and (3.17*b*) is an identity for unitary operators U_Λ. However,

$$U_{\Lambda_1}^{-1}\psi_{A_1}(t; \varepsilon) = U_{\Lambda_2}^{-1}\psi_{A_2}(t; \varepsilon), \tag{3.18}$$

since both functions are solutions of the same wave equation and they satisfy the same initial conditions. The latter statement holds because at any time t_0 after the fields have been turned off,

$$\boldsymbol{A}_1 - \boldsymbol{A}_2 = \boldsymbol{\nabla}\Lambda = \boldsymbol{\nabla}\Lambda_1 - \boldsymbol{\nabla}\Lambda_2, \tag{3.19}$$

therefore

$$\psi_{A_1} = U_\Lambda\psi_{A_2} = U_{\Lambda_1}U_{\Lambda_2}^{-1}\psi_{A_2}. \tag{3.20}$$

But (3.20) is equivalent to (3.18). Thus we conclude that the physically meaningful transition amplitudes $\langle\phi_{\varepsilon'}|U_\Lambda^{-1}\psi_A(t; \varepsilon)\rangle$ are independent of the gauge of the potentials $(\boldsymbol{A}, A_0)$, in contrast to the conventional amplitudes of (3.12).

† A preliminary report of this result was presented by Feuchtwang *et al* (1983).

3.2.3. Energy conserving transition probabilities. The preceding discussion of the gauge independence of the transition amplitudes observed after the fields are switched off is perfectly general. However, it is more an existence theorem than a tool in actual calculations. In general, one does not have available the exact wavefunction $\psi_A(\boldsymbol{r}, t; \varepsilon)$, but rather an approximation of the exact asymptotic state ψ_∞ by an nth-order perturbation theory. Here we define ψ_∞ by the limit

$$\lim_{t\to\infty} \psi_A(\boldsymbol{r}, t; \varepsilon) = \psi_\infty(\boldsymbol{r}; \varepsilon, \Lambda). \tag{3.21}$$

The effect of an electromagnetic field switched on adiabatically at $t = -\infty$ and off at $t = \infty$ is now described by the transitions it induces between the field-free eigenfunctions $\{\phi_\varepsilon\}$ of $\mathscr{H}_0$. In calculating these transition probabilities by perturbation theory it is impractical to ensure the gauge independence explicitly in each order. As shown by Kazes *et al* (1982) the transition amplitudes are gauge independent, i.e. satisfy

$$\langle \phi_{\varepsilon'} | \psi_\infty(\varepsilon, \Lambda = 0)\rangle = \langle \phi_{\varepsilon'} | \psi_\infty(\varepsilon, \Lambda)\rangle = \langle \phi_{\varepsilon'} | U_\Lambda \psi_\infty(\varepsilon, \Lambda = 0)\rangle, \tag{3.22}$$

if and only if:

(1) the transition conserves energy, i.e. the energy difference $(\varepsilon - \varepsilon')$ is transferred to (from) the particle from (to) the radiation field†;

(2) the time dependent gauge function satisfies the condition

$$\lim_{t\to-\infty} \Lambda(\boldsymbol{r}, t) = \lim_{t\to\infty} \Lambda(\boldsymbol{r}, t). \tag{3.23}$$

Note that this condition corresponds to equation (3.16).

The gauge independence of the overlap matrix elements in (3.22) is to be contrasted with the familiar gauge invariance of scalar products, i.e.

$$\langle \phi_{\varepsilon'} | \psi_\infty(\varepsilon, \Lambda = 0)\rangle = \langle U_\Lambda \phi_{\varepsilon'} | U_\Lambda \psi_\infty(\varepsilon, \Lambda = 0)\rangle \tag{3.24a}$$

and

$$\langle \phi_{\varepsilon'} | \psi_\infty(\varepsilon, \Lambda)\rangle = \langle \phi_{\varepsilon'} | U_\Lambda \psi_\infty(\varepsilon, \Lambda = 0)\rangle = \langle U_\Lambda^{-1} \phi_{\varepsilon'} | \psi_\infty(\varepsilon, \Lambda = 0)\rangle. \tag{3.24b}$$

Equations (3.24) indicate that since $U_\Lambda \neq I$, equation (3.22) should not be expected to hold for arbitrary ε and ε'.

It should be emphasised that the gauge independence of energy conserving transition probabilities applies not only to the exact probabilities, but also order by order in perturbation theory (Kazes *et al* 1982). This is particularly important in practical calculations, which of necessity are often restricted to a finite-order perturbative treatment. In this case, one is concerned with the transition rates induced by the time dependent perturbation $\mathscr{H}_1$,

$$\mathrm{d}P^{(n)}_{\varepsilon'\varepsilon}/\mathrm{d}t = (2\pi/\hbar)|\langle \phi_{\varepsilon'} | \mathscr{H}_1^{(n)} | \phi_\varepsilon\rangle|^2, \tag{3.25}$$

where n is the order (in e) of the perturbation theory and $\mathscr{H}_1^{(n)}$ is the nth iterated or effective interaction operator. These rates are gauge independent if the transition conserves energy, i.e. if

$$\varepsilon' - \varepsilon + \sum_{i=1}^{n} \hbar\omega_i = 0, \tag{3.26}$$

† In fact Kazes *et al* (1982) dealt only with closed systems including the radiation field, for which the gauge independent transition amplitudes were shown to involve states of equal energy. However, that result is equivalent to the one above in which the radiation field is treated as an external field. The importance of energy conservation in assuring the gauge independence of matrix elements has been noted by Fried (1973), Haller and Landowitz (1970), Aharonov and Au (1979) and Kazes *et al* (1982).

where the photon frequency ω_i is taken positive (negative) if it labels an absorbed (emitted) photon.

The physical significance of our result is illustrated by the following example of the perturbative treatment of the interaction of electromagnetic radiation with matter in the electric dipole approximation. In this approximation, the spatial variation of the fields is neglected over the region in which the overlap of the initial and final (unperturbed) states is significant. That is, the long wavelength limit is taken, so that

$$\boldsymbol{A}(\boldsymbol{r}, t) = \boldsymbol{A}(0)\, \mathrm{e}^{-\mathrm{i}\omega t} + \mathrm{cc}$$

and

$$\boldsymbol{E}(\boldsymbol{r}, t) = \boldsymbol{E}(0)\, \mathrm{e}^{-\mathrm{i}\omega t} + \mathrm{cc}, \tag{3.27}$$

where $\boldsymbol{A}(0)$ and $\boldsymbol{E}(0)$ are constant vectors.

The system can be described by two equivalent Hamiltonians. The first includes only a vector potential,

$$\mathscr{H}(\boldsymbol{A}, 0) = (2m)^{-1}[\boldsymbol{p} - (e/c)\boldsymbol{A}]^2 + V(\boldsymbol{r}) = \mathscr{H}_0 + \mathscr{H}_1(\boldsymbol{A}, 0), \tag{3.28}$$

where

$$\mathscr{H}_1(\boldsymbol{A}, 0) = -(e/2mc)(\boldsymbol{p} \cdot \boldsymbol{A} + \boldsymbol{A} \cdot \boldsymbol{p}) + (e^2/2mc^2)A^2 \tag{3.29}$$

and

$$\boldsymbol{E} = -c^{-1}\partial \boldsymbol{A}/\partial t. \tag{3.30}$$

The second Hamiltonian includes only a scalar potential,

$$\mathscr{H}(0, A_0') = \boldsymbol{p}^2/2m + V(\boldsymbol{r}) + eA_0' = \mathscr{H}_0 + \mathscr{H}_1(0, A_0'), \tag{3.31}$$

where

$$\mathscr{H}_1(0, A_0') = eA_0' \tag{3.32}$$

and

$$\boldsymbol{E} = -\boldsymbol{\nabla} A_0'. \tag{3.33}$$

Evidently the two sets of potentials,

$$\boldsymbol{A} = \boldsymbol{A}(0) \cos \omega t = (c/\omega)\boldsymbol{E}(0) \cos \omega t, \qquad A_0 = 0, \tag{3.34a, b}$$

and

$$\boldsymbol{A}' = 0, \qquad A_0' = -e\boldsymbol{E}(0) \cdot \boldsymbol{r} \sin \omega t, \tag{3.35a, b}$$

are related by the gauge transformation defined by

$$\Lambda(\boldsymbol{r}, t) = -(c/\omega)\boldsymbol{E}(0) \cdot \boldsymbol{r} \cos \omega t. \tag{3.36}$$

The wave equations are also gauge transforms of each other,

$$U_\Lambda(\mathscr{H}(\boldsymbol{A}, 0) - \mathrm{i}\hbar\partial/\partial t)\, U_\Lambda^{-1} = \mathscr{H}(0, A_0') - \mathrm{i}\hbar\partial/\partial t. \tag{3.37}$$

The interaction operators $\mathscr{H}_1(\boldsymbol{A}, 0)$ and $\mathscr{H}_1(0, A_0')$ are clearly not unitarily equivalent, and hence exhibit generally different matrix elements. Thus, in agreement with the general argument of Kobe (1978) and Yang (1981), the first-order transition amplitudes are generally gauge dependent, e.g.

$$\langle \phi_{\varepsilon'} | \mathscr{H}_1(0, A_0) | \phi_\varepsilon \rangle = [\hbar\omega/(\varepsilon' - \varepsilon)] \langle \phi_{\varepsilon'} | \mathscr{H}_1(\boldsymbol{A}, 0) | \phi_\varepsilon \rangle. \tag{3.38}$$

However, the energy conserving first-order transitions can only occur between states for which

$$\varepsilon' - \varepsilon = \hbar\omega, \tag{3.39}$$

and in this case the corresponding observable transition amplitudes are strictly gauge independent.

4. Gauge dependence of transition amplitudes and the construction of gauge independent amplitudes in the presence of fields

4.1. Interpretation of gauge dependence of some transition amplitudes

In our discussion of gauge independence we emphasised that no directly observable quantity can be gauge dependent. We have noted that the time dependent transition amplitudes in the presence of electromagnetic fields may be gauge dependent, and we have outlined an argument due to Aharonov and Au (1981) showing that this is to be expected. It is therefore important to realise that while gauge dependent transition amplitudes may occur in an intermediate stage of a calculation, the gauge dependence always drops out of the final result, representing the prediction of the outcome of an experiment, since such results are inherently gauge independent. Stated differently, gauge dependent transition amplitudes are never necessary for the interpretation of an experiment. This point is illustrated by the following discussion concerning the interpretation of the Lamb shift experiment.

A possible approach would be to interpret this experiment as an observation of the radiative transition of a hydrogen atom from its $2S$ to its $2P$ manifold. The corresponding transition amplitude in the presence of the microwave field is gauge dependent. However, this calculated result is strictly irrelevant to the interpretation of the experiment. The actual observation consists of the detection of photons. Since the photon number operator has a gauge independent expectation value, neither the actual nor the calculated prediction of the outcome of the experiment is gauge dependent.

We now proceed to a more detailed analysis of the gauge dependence of time dependent transition amplitudes in the presence of fields. This will enable us to derive a gauge independent transition amplitude.

One might interpret the overlap matrix $\langle\phi_{\varepsilon'}|\psi_A(t;\varepsilon)\rangle$ as the conditional probability amplitude at time t, for finding the system in the field-free state $\phi_{\varepsilon'}$ if it is known to be in the state $\psi_A(t;\varepsilon)$. However, we note that:

(i) there is not given a precise operational specification of how we are to determine that the system is in the state $\phi_{\varepsilon'}$ at time t (Aharonov and Au 1981);

(ii) neither the field-free Hamiltonian $\mathcal{H}_0$ nor its eigenfunctions $\{\phi_\varepsilon\}$ is uniquely defined by the physical field-free state of energy ε.

In fact, both (2.7) and the non-uniqueness of the Lagrangian within a total time derivative of an arbitrary function (see (2.8)) imply that $\mathcal{H}_0$ is a functional of f,

$$\mathcal{H}_0[f] = (2m)^{-1}(\boldsymbol{p} - (e/c)\nabla f)^2 + V(\boldsymbol{r}) - (e/c)\partial f/\partial t. \tag{4.1}$$

In order to obtain (4.1) we have, in contrast to (2.8), added to (rather than subtracted from) the conventional Lagrangian the term $(e/c)\mathrm{d}f/\mathrm{d}t$, where as noted before f is a completely arbitrary function of $(\boldsymbol{r}, t)$.

From (3.1)–(3.3) we conclude that the eigenfunctions $\{\phi_\varepsilon[f]\}$ are defined in terms of the solutions of the corresponding Schrödinger equation,

$$(\mathscr{H}_0[f] - i\hbar\partial/\partial t)\psi_\varepsilon[f(t), t] = 0, \tag{4.2}$$

which have the form

$$\psi[f(t), t] = e^{i\varepsilon t/\hbar}\phi_\varepsilon[f] = e^{i\varepsilon t/\hbar}U_f\phi_\varepsilon[f \equiv 0], \tag{4.3}$$

where

$$(U_f\mathscr{H}_0[f \equiv 0]U_f^{-1} - \varepsilon)\phi_\varepsilon[f] = 0. \tag{4.4}$$

Thus, if we interpret 'being in the field-free state ϕ_ε' as 'having an unperturbed energy ε', then it follows that in the presence of fields the overlap matrix $\langle\phi_{\varepsilon'}[f]|\psi_A(t; \varepsilon)\rangle$ is a functional of the arbitrary function $f(\boldsymbol{r}, t)$ used to define the field-free Hamiltonian. An operational definition of the field-free state with energy ε evidently requires a definite choice of f. One possible choice is to make $f(\boldsymbol{r}, t)$ a functional of the four-potential A_μ, so that the overlap matrix becomes independent of any particular choice of A_μ (Feuchtwang *et al* 1983). Here we note that the adiabatic switching on of the fields and of $f(\boldsymbol{r}, t)$ guarantees the independence of $\psi_A(\boldsymbol{r}, t; \varepsilon)$ from $f(\boldsymbol{r}, t)$. An alternative to the procedure described above is to make the Hamiltonian itself gauge independent (Kazes *et al* 1983). That is, following the reasoning that suggested the definition of $\mathscr{H}_0[f]$, but using (2.8), we now obtain

$$\mathscr{H}[f] = (2m)^{-1}[\boldsymbol{p} - (e/c)(\boldsymbol{A} - \boldsymbol{\nabla}f)]^2 + V(\boldsymbol{r}) + e(A_0 + c^{-1}\partial f/\partial t). \tag{4.5}$$

Here again f is a functional of A_μ which now cancels the gauge dependence of the conventional Hamiltonian $\mathscr{H}[f \equiv 0]$, rather than just of the transition amplitudes. As we shall show below, the two approaches are quite similar, and, as might be expected, they involve the same functional $f[A_\mu]$.

4.2. Determination of the functional $f[A_\mu]$

As noted above, we can obtain gauge independent transition amplitudes either by introducing into the field-free Hamiltonian an appropriate gauge dependence, or by defining the full Hamiltonian in a gauge independent fashion. In either case we introduce into the Lagrangian the total time derivative $\mathrm{d}f/\mathrm{d}t$ of a function $r(\boldsymbol{r}, t)$ which we shall now construct.

We shall first determine the arbitrary function $f(\boldsymbol{r}, t)$ as a functional of A_μ so that

$$\mathscr{H}_0[A_\mu + \partial_\mu\Lambda] = U_\Lambda\mathscr{H}_0[A_\mu]U_\Lambda^{-1} - i\hbar U_\Lambda\dot{U}_\Lambda^{-1}, \tag{4.6}$$

and hence also

$$\phi_\varepsilon[A_\mu + \partial_\mu\Lambda] = U_\Lambda\phi_\varepsilon[A_\mu]. \tag{4.7}$$

Equations (4.5) and (4.6) imply that the same overlap matrix is obtained with two different choices of the four-potential, related by a gauge transformation. That is, if

$$A_{\mu,2} = A_{\mu,1} + \partial_\mu\Lambda, \qquad \psi_{A_2}(t; \varepsilon) = U_\Lambda\psi_{A_1}(t; \varepsilon), \tag{4.8}$$

then according to (4.7)

$$\langle\phi_{\varepsilon'}[A_{\mu,2}]|\psi_{A_2}(t; \varepsilon)\rangle = \langle U_\Lambda\phi_{\varepsilon'}[A_{\mu,1}]|U_\Lambda\psi_{A_1}(t; \varepsilon)\rangle = \langle\phi_{\varepsilon'}[A_{\mu,1}]|\psi_{A_1}(t; \varepsilon)\rangle. \tag{4.9}$$

It is easily verified that $f[A_\mu]$ has to satisfy the following functional relation:

$$f[A_\mu + \partial_\mu \Lambda] = f[A_\mu] + \Lambda + \text{constant}. \tag{4.10}$$

This functional equation is solved by any of the following four-dimensional path integrals.

$$f = \oint_0^{(r,t)} A_\mu(x) \mathrm{d}x^\mu. \tag{4.11}$$

We note that an alternative path $\bar{c}$ between '0' and $(\boldsymbol{r}, t)$ will define a function $\bar{f}$ such that the difference $f - \bar{f}$ is given by a closed contour integral of A_μ. Such an integral is clearly gauge independent, because the closed contour integral of a four-gradient vanishes. Thus we can write

$$f = \oint_0^{(r,t)} A_\mu(x)\, \mathrm{d}x^\mu + g(\boldsymbol{r}, t), \tag{4.12}$$

where c_1 is an arbitrary but fixed 'standard' path, and $g(\boldsymbol{r}, t)$ is an arbitrary gauge independent function (Kazes *et al* 1982). Here it should be emphasised that in general A_μ is not a gradient and hence

$$\partial_\mu f \neq A_\mu. \tag{4.13}$$

However, after the field has been switched off, then over a simply connected region A_μ reduces to a pure gauge field,

$$\partial_\mu f = A_\mu, \tag{4.14}$$

and the present case reduces to that discussed in § 3.2.2.

Here we note that (4.11) is not the only class of functionals satisfying (4.10). Another example is

$$f = \partial^\mu A_\mu / \Box, \tag{4.15}$$

where $\Box$ is the d'Alembertian, and the operator $\partial^\mu/\Box$ is to be interpreted in terms of its Fourier transform, provided A_μ has such a transform.

We now turn to the alternative procedure for the determination of gauge independent transition amplitudes. As noted before, the conventional minimally coupled Lagrangian is not gauge independent, rather it is *gauge quasi-invariant.* However, an appropriate choice of the function f in (2.8) makes the Lagrangian *gauge independent.* Consequently also the Hamiltonian and the canonical momentum become gauge independent. If the full Hamiltonian $\mathscr{H}[f]$ is to be gauge independent, then it follows from (4.5) and (4.10) that both

$$\boldsymbol{A}' - \boldsymbol{\nabla} f[\boldsymbol{A}', A_0'] = \boldsymbol{A} - \boldsymbol{\nabla} f[\boldsymbol{A}, A_0] \tag{4.16}$$

and

$$\boldsymbol{A}_0' + c^{-1}(\partial/\partial t) f[\boldsymbol{A}', \boldsymbol{A}_0'] = A_0 + c^{-1}(\partial/\partial t) f[\boldsymbol{A}, A_0]. \tag{4.17}$$

These functional equations for $f[A_\mu]$ are evidently equivalent to (4.10). Thus, the discussion following (4.10) applies equally to the present procedure. We recall that in the absence of electromagnetic fields and, for a simply connected region, A_μ is a pure gauge field. Thus in this case, the field-free Hamiltonian corresponding to (4.5) is the conventional $\mathscr{H}_0$.

The inherent similarity of the two procedures is most easily recognised by noting that $\partial_\mu f[A_\nu]$ corresponds to a pure gauge field. It can be inserted at will into $\mathscr{H}$ or

$\mathscr{H}_0$. It can thus either be used to specify the proper (asymptotic) gauge for $\mathscr{H}_0$ to be the field-free Hamiltonian corresponding to a particular $\mathscr{H}$, or else obviate the preceding selection of the proper $\mathscr{H}_0[A_\mu]$ by eliminating the gauge dependence of $\mathscr{H}$ itself.

4.3. The gauge independent transition amplitudes

The path dependence of the functional $f[A_\mu]$, which can be expressed in terms of the gauge independent function $g(\boldsymbol{r}, t)$ defined by (4.12), causes a corresponding path dependence of the gauge independent transition amplitudes. Namely, a change Δg in the arbitrarily chosen g corresponds to a redefinition of the canonical momentum, and hence also of the field free state $\phi_\varepsilon[A_\mu]$,

$$\phi_\varepsilon[A_\mu] \to U_{\Delta g}\phi_\varepsilon[A_\mu]. \tag{4.18}$$

However, the consequent non-uniqueness of the gauge independent transition amplitudes,

$$\langle \phi_{\varepsilon'}[A_\mu]\psi_A(t;\varepsilon)\rangle \to \langle U_{\Delta g}\psi_{\varepsilon'}[A_\mu]|\psi_A(t,\varepsilon)\rangle, \tag{4.19}$$

is not an artifact of our theory, rather it reflects the fundamental non-uniqueness of the classical Lagrangian as well as the Hamiltonian, which were discussed in § 2, and are familiar from classical mechanics. Thus the problem posed in quantum mechanics by the dependence of wavefunctions on the arbitrary gauge of the potentials is traced to the inherent non-uniqueness of the Lagrangian and Hamiltonian formulations of classical mechanics, even in the absence of electromagnetic fields. However, just as the latter has no physical consequence, so also the path dependence of the Hamiltonian cannot have any physical consequence. That is, any directly observable quantity must be uniquely specified. It cannot depend either on the choice of path or alternatively, on the choice of the gauge independent function $g(\boldsymbol{r}, t)$, either of which specify the functional $f[A]$. This is analogous to the independence of the calculated predictions of observations from the choice of the canonical variables.

5. Discussion and conclusions

In this paper we have drawn attention to some issues pertaining to gauge covariance and gauge independence in non-relativistic physics. We have noted that for classical charged particles, which in the conventional fashion are minimally coupled to the electromagnetic field, the velocity is gauge independent, whereas the canonical momentum is gauge dependent. The corresponding results were also derived for quantum systems. Applying the correspondence principle, we obtain a quantum mechanical description in which the wavefunction depends on the gauge. This gauge dependence of the wavefunction manifests itself in a multiplicative space–time dependent phase factor, and leads to a velocity operator which is explicitly gauge dependent. However, in agreement with the correspondence principle, its matrix elements are not gauge dependent. This is in contrast to the momentum operator, which is not gauge dependent, but which has a gauge dependent expectation value. Finally, we also note that the conventional, minimally coupled quantum mechanical Hamiltonian is explicitly gauge dependent.

We noted on purely physical grounds that no direct measurement can depend on the gauge used to calculate the prediction of its outcome. Hence, the same must also

apply to any prediction based on quantum mechanics. Purely local measurements, such as that of position, are gauge independent, whereas some non-local quantities, such as the momentum, can have gauge dependent expectation values. We have remarked that canonical variables, canonical momenta and Hamiltonians are not uniquely defined in classical mechanics even when the trajectories of the particle are known. This fact was traced to the freedom in the choice of the Lagrangian: the addition of a total time derivative does not change the Euler–Lagrange equations of motion, but because this changes the Lagrangian, the canonical variables are non-unique and can be chosen appropriately. This non-uniqueness reflects itself in a corresponding non-uniqueness of wavefunctions. In particular, the specification of the energy of a system does not determine the wavefunction. Wavefunctions that describe the same physical state of a system can differ by space–time dependent phase factors. Consequently the question 'what is the probability of finding a system, interacting with an electromagnetic field, in a given field-free state of energy ε' has no unique answer because this probability can be gauge dependent. However, we have demonstrated that a suitable choice of the free particle Hamiltonian eliminates this gauge dependence.

We have used the freedom of adding a total time derivative of a function f to the Lagrangian to obtain a gauge independent Hamiltonian formalism. We constructed gauge independent momenta, Hamiltonians and wavefunctions. We noted that this gauge dependent definition of the arbitrary function 'f' still allows the addition of the total time derivative of a gauge independent function 'g' to the Lagrangian. Thus, our procedure leads to a gauge independent, but path dependent, canonical formulation of non-relativistic quantum mechanics, which is similar to Mandelstam's (1968) reformulation of relativistic gauge field theories. We emphasise that this path dependence cannot be eliminated, because it is equivalent to the quasi-invariance of the Lagrangian, i.e. the freedom of adding a gauge independent total time derivative to the Lagrangian. That is, although we could eliminate all traces of gauge dependence, we cannot eliminate the underlying freedom of performing gauge independent canonical transformations. The dependence of the Hamiltonian on the arbitrary gauge independent function $g(\boldsymbol{r}, t)$ does not pose any practical difficulty. Just as we expect direct physical predictions to be independent of the gauge, we also expect the choice of canonical variables not to affect any direct physical predictions. This is in fact one of the tests of the reliability of any physical prediction.

Finally, we note that our gauge invariant Lagrangian and the corresponding Hamiltonian contain an unconventional feature: they are non-local. For instance, the Lagrangian at a given time depends on the vector potential at previous times. In field theory, the Lagrangian density will also be non-local in space–time. The consequent non-locality due to an externally applied electromagnetic field is irrelevant. However, when this formalism is applied to quantum electrodynamics and to the other gauge field theories the consequences of this new feature must be considered carefully. Mandelstam's work leads us to believe that such an extension of our theory is feasible and that his results can now be obtained in the canonical procedure which we have developed for non-relativistic systems.

References

Aharonov Y and Au C K 1979 *Phys. Rev.* A **20** 1553
——1981 *Phys. Lett.* **86A** 269

Cohen-Tannoudji C, Diu B and Laloe F 1977 *Quantum Mechanics* (Paris: Herman/Wiley)
Eckstein H 1979 *Nuovo Cimento* **54** 305
Feuchtwang T E, Kazes E, Grotch H and Cutler P H 1983 *Phys. Lett.* **93A** 4
Fried Z 1973 *Phys. Rev.* A **8** 2835
Goldstein H 1980 *Classical Mechanics* 2nd ed (Reading, Mass.: Addison-Wesley) p 33, prob 14
Haller K and Landowitz L F 1970 *Phys. Rev.* D **2** 1498
Kazes E, Feuchtwang T E, Grotch H and Cutler P H 1982 *Ann. Phys.* **142** 80
—— 1983 *Phys. Rev. D* **27** 1388
Kobe D H 1978 *Phys. Rev. Lett.* **40** 538
Kobe D H and Smirl A L 1978 *Am. J. Phys.* **46** 624
Kobe D H and Wen E C-T 1980 *Phys. Lett.* **80A** 121
Mandelstam S 1962 *Ann. Phys.* **19** 1
——1968 *Phys. Rev.* **175** 1580
Yang K H 1981 *Phys. Lett.* **84A** 165
——1982 *J. Phys. A: Math. Gen.* **15** 1201