Gauge invariance and reciprocity in quantum mechanics

P.T. Leung  
*Portland State University*

K. Young

Follow this and additional works at: https://pdxscholar.library.pdx.edu/phy_fac

Part of the Physics Commons

Let us know how access to this document benefits you.

Citation Details


This Article is brought to you for free and open access. It has been accepted for inclusion in Physics Faculty Publications and Presentations by an authorized administrator of PDXScholar. Please contact us if we can make this document more accessible: pdxscholar@pdx.edu.
Gauge invariance and reciprocity in quantum mechanics

P. T. Leung (梁培德)\textsuperscript{1,*} and K. Young (楊鋼凱)\textsuperscript{2,1}

\textsuperscript{1}Institute of Optoelectronics, National Taiwan Ocean University, Keelung, Taiwan
\textsuperscript{2}Department of Physics, The Chinese University of Hong Kong, Hong Kong, People’s Republic of China

(Received 20 December 2009; published 8 March 2010)

Reciprocity in wave propagation usually refers to the symmetry of the Green’s function under the interchange of the source and the observer coordinates, but this condition is not gauge invariant in quantum mechanics, a problem that is particularly significant in the presence of a vector potential. Several possible alternative criteria are given and analyzed with reference to different examples with nonzero magnetic fields and/or vector potentials, including the case of a multiply connected spatial domain. It is shown that the appropriate reciprocity criterion allows for specific phase factors separable into functions of the source and observer coordinates and that this condition is robust with respect to the addition of any scalar potential. In the Aharonov-Bohm effect, reciprocity beyond monoenergetic experiments holds only because of subsidiary conditions satisfied in actual experiments: the test charge is in units of $e$ and the flux is produced by a condensate of particles with charge $2e$.

DOI: 10.1103/PhysRevA.81.032107 PACS number(s): 02.20.Uw, 03.65.—w, 03.75.—b

I. INTRODUCTION

Reciprocity refers to the equivalence between the signals when the source at $\vec{r}$ and the observer at $\vec{r}'$ are interchanged. Although reciprocity is obvious in free space, it becomes nontrivial when there is a material background. The notion of reciprocity plays an important role for classical waves of reciprocity plays an important role for classical waves. The stringent condition\textsuperscript{3}, which demands the two Green’s functions (with source and observer reversed) to be equal; it is obviously not gauge invariant and hence is inappropriate. A more relaxed condition\textsuperscript{5}, labeled later in this article as (C4), only requires equality of the magnitudes (in the frequency domain); we shall show that it is in general too loose. This article analyzes in addition two intermediate proposals, labeled as (C2) and (C3) later in this article, with special attention to the question of gauge invariance. Gauge invariance, especially its extension to non-Abelian groups\textsuperscript{11}, is central to the modern concept of particle interactions.

The general concepts are of particular interest when applied to spatial domains that are not simply connected, in particular the Aharonov-Bohm (AB) effect\textsuperscript{12}, which is experimentally verified in electron diffraction\textsuperscript{13}. In this case, it turns out that reciprocity beyond monoenergetic experiments holds only on account of extra experimental conditions: the test charge is in units of $e$ and the flux is due to a superconducting condensate with charges in units of $2e$—neither of which is necessary as an \textit{a priori} condition.

II. FORMULATION

Reciprocity can be discussed in terms of the Green’s function $G(\vec{r}, \vec{r}', t)$ (or the Green’s dyadic\textsuperscript{7,14} in the case of vector fields) describing propagation from $\vec{r}'$ to $\vec{r}$ in time $t$. It is convenient to consider the corresponding frequency-domain functions\textsuperscript{3} $\tilde{G}(\vec{r}, \vec{r}', \omega)$, in particular the ratio

$$T = \frac{\tilde{G}(\vec{r}, \vec{r}', \omega)}{\tilde{G}(\vec{r}', \vec{r}, \omega)}, \tag{1}$$

in terms of which we shall consider four possible criteria,

$$T = 1 \quad \text{(C1)}, \tag{2a}$$

$$T = \exp 2i[\theta(\vec{r}) - \Theta(\vec{r}')] \quad \text{(C2)}, \tag{2b}$$

$$T = \exp i[\theta(\vec{r}) - \Theta(\vec{r}')] \quad \text{(C3)}, \tag{2c}$$

$$T = \exp i\Theta(\vec{r}, \vec{r}'), \quad \text{i.e., } |T| = 1 \quad \text{(C4)}, \tag{2d}$$

where $\theta$ and $\Theta$ are real phases. For (C1)–(C3) [but not (C4)], the same ratio applies to the time-domain Green’s functions $G$. The four conditions are decreasingly restrictive: (C1) $\Rightarrow$ (C2) $\Rightarrow$ (C3) $\Rightarrow$ (C4). Section III gives examples illustrating each of these conditions. Section IV analyzes key properties that support the adoption of (C2) as the

\*Present address: Department of Physics, Portland State University, Portland, OR 97207, USA.
\textsuperscript{1}Corresponding author: kyoung@cuhk.edu.hk
general condition for reciprocity, with (C3) corresponding to “accidental” reciprocity in the sense that the property is destroyed by the addition of a local real potential, while (C4) does not describe reciprocity except in the very special case of monoenergetic experiments (understood, everywhere in this article, to also include experiments with incoherent superposition of monoenergetic states).

III. EXAMPLES

Several examples will be given, with particular attention given to whether there is a nonzero magnetic field \( \vec{B} \) and/or a nonzero vector potential \( \vec{A} \); a key issue is that \( \vec{B} = 0 \) does not guarantee \( \vec{A} = 0 \).

A. No magnetic field, no vector potential

The standard representation gives, for \( t > 0 \),

\[
G(\vec{r}, \vec{r}', t) = \sum_n u_n(\vec{r})u_n^\ast(\vec{r}')\exp(-iE_nt/\hbar),
\]

in terms of eigenfunctions \( u_n \), their conjugates \( u_n^\ast \), and eigenvalues \( E_n \). In the absence of both \( \vec{B} \) and \( \vec{A} \), the Hamiltonian \( H \) (assumed to be Hermitian) is real, so the eigenfunctions can be chosen to be real; therefore, \( G \) is symmetric under the interchange of the two spatial variables, and (C1) is satisfied.

B. No magnetic field, nonzero vector potential, simply connected region

Even when \( \vec{B} = 0 \), one can have a nonzero \( \vec{A} \). Provided the domain is simply connected, it is possible to remove \( \vec{A} \) by a single-valued gauge transformation, under which

\[
\vec{A}(\vec{r}) \mapsto \vec{A}(\vec{r}) + \vec{\nabla}\xi(\vec{r}) = 0, \quad u_n(\vec{r}) \mapsto u_n(\vec{r})\exp[i\theta(\vec{r})],
\]

with \( \theta = (q/\hbar)\xi \). The transformed system with \( \vec{A} = 0 \) satisfies (C1); hence, the original system satisfies (C2)—which is then seen to be the appropriate gauge-invariant generalization of the concept of reciprocity. Incidentally, this also explains the conventional factor of 2 in the phase in (2b). For the sake of completeness, a more general proof of these properties using Green’s theorem is given in the Appendix.

C. Uniform magnetic field

An interesting example is a uniform \( \vec{B} \) field, say, \( \vec{B} \parallel \hat{z} \). The Green’s function can be obtained in closed form [15],

\[
\hat{G}(\vec{r}, \vec{r}', \omega) = e^{i(x'y'-yx')/2} \left( \frac{m}{2\pi\hbar^2} \right) \frac{\pi}{\cos(\pi\varepsilon)} \frac{1}{\Gamma(\varepsilon + 1/2)} \frac{W_{\varepsilon,0}(\rho)}{\sqrt{\rho}},
\]

where \( \varepsilon = \omega/\omega_L \) and \( \omega_L \) is the Larmor frequency, \( W_{\varepsilon,0} \) the Whittaker function, \((x, y, z)\) and \((x', y', z')\) are the Cartesian coordinates of the two position vectors, and \( \rho \equiv (\vec{r} - \vec{r}')^2/2 \). All lengths are dimensionless, in units of \( \sqrt{\hbar/m\omega_L} \). It is clear from (5) that (C2) is not satisfied, whereas (C3) is satisfied.

D. Aharonov-Bohm effect

The intriguing case is that of zero magnetic field in a multiply connected domain, for example, the AB effect [12]. Let there be a magnetic flux \( \Phi \) confined to a tube (“solenoid”) along the \( \hat{z} \) axis, from which a charge \( q \) is excluded; thus, the charge experiences zero field \( \vec{B} \), but the QM has to be described by a nonzero vector potential \( \vec{A} \), which moreover cannot be written as \( \vec{A} = -\vec{\nabla}\xi \) (unless one introduces multivalued \( \xi \) functions; see the Appendix). Reciprocity can be studied using the explicit form of the Green’s function and it suffices to do so for a simple case: the charge \( q \) is confined to a ring with coordinates \((r = R, \phi)\) in the \( x-y \) plane, which already captures the essential feature of the nontrivial topology; inclusion of the variables \( r \) and \( z \) will not affect the argument.

The Green’s function is given in the literature [16,17], but the key ideas are better exhibited through a simple argument. Axial symmetry is preserved if the gauge is chosen to be

\[
\vec{A} = A(r)\hat{\phi} = (\Phi/2\pi r)\hat{\phi}.
\]
The angular momentum must be quantized as \( L_z = R p = n \hbar \), where \( p \) is the azimuthal component of the conjugate momentum and \( n \) is an integer. Since \( L_z = -(i\hbar)d/d\varphi \), it follows that the eigenfunctions are \( u_n(\varphi) = (2\pi)^{-1/2} \exp(i n \varphi) \).

The energy \( E = (\vec{p} - q \vec{A})^2/(2m) \) is then quantized as \( E_n = \frac{\hbar^2}{2m R^2} (n - \sigma)^2 \), in which \( \sigma = \frac{q \Phi}{\hbar} \).

A formal solution of Schrödinger’s equation will yield the same results. Putting these into (3) then gives

\[
G(\varphi, \varphi', t) = (2\pi)^{-1} \sum_n \exp \left[ i(n \varphi - n \varphi') \right] \times \exp \left[ -i(\hbar^2 t/2m R^2)(n - \sigma)^2 \right] \tag{11}
\]

for \( t > 0 \) and

\[
G(\varphi, \varphi', \omega) = (2\pi)^{-1} \sum_n \frac{i}{\omega - (n - \sigma)^2(\hbar/2m R^2)} \tag{12}
\]

with the usual \( i \epsilon \) prescription understood to render the function causal. It is possible to adopt another approach in which \( \Phi \) is removed at the cost of introducing multivalued wave functions (Appendix).

First consider the case \( 2\sigma = \text{integer} \). Upon interchange of the two spatial coordinates \( \varphi \leftrightarrow \varphi' \), if we also reverse the dummy variable \( n \rightarrow -n \) [thus keeping the phase factor in (12) unchanged, but turning \((n - \sigma)^2 \rightarrow (n + \sigma)^2\)] and then shift \( n \rightarrow n - 2\sigma \) [which restores \((n + \sigma)^2 \rightarrow (n - \sigma)^2\)], then it is readily seen that (C2) is satisfied.

However, for \( 2\sigma \neq \text{integer} \), the shift \( n \rightarrow n - 2\sigma \) is not allowed, and the above proof of (C2) fails. Nevertheless, upon interchange of the two spatial coordinates \( \varphi \leftrightarrow \varphi' \), if the Green’s function is conjugated, the phase factor would be unchanged, and it is easy to show that

\[
G(\varphi, \varphi', \omega) = -G(\varphi', \varphi, \omega)^* \quad \text{(C4*)}, \tag{13}
\]

from which it follows that (C4) [but not (C3)] is satisfied. In fact, (13) is an interesting special case of (C4), to be denoted as (C4*).

The examples are summarized in Table I.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>(C1)</td>
<td>No vector potential</td>
</tr>
<tr>
<td>(C2) but not (C1)</td>
<td>No magnetic field in a simply-connected region, but nonzero vector potential; or AB effect with ( 2\sigma = \text{integer} )</td>
</tr>
<tr>
<td>(C3) but not (C2)</td>
<td>Uniform magnetic field</td>
</tr>
<tr>
<td>(C4) but not (C3)</td>
<td>AB effect with ( 2\sigma \neq \text{integer} )</td>
</tr>
</tbody>
</table>

TABLE I. Summary of examples.

IV. EXPERIMENTAL MANIFESTATIONS

It is obvious that (C1) is not gauge invariant and cannot be adopted as the condition of reciprocity when there is a vector potential and/or if one allows gauge transformations. Noting that gauge transformations alter phases and experiments only measure the absolute square of the amplitudes, it was proposed [5] that (C4) be adopted as the condition for reciprocity. Indeed, it is true that for monoenergetic experiments, (C4) will guarantee that no difference is observed when the source and observation points are reversed. However, one can contemplate experiments involving coherent superposition of different energies with relative amplitudes \( \rho(\omega) \), in which case the amplitude is schematically

\[
\int \rho(\omega) \tilde{G}(\vec{r}, \vec{r}', \omega) d\omega. \tag{14}
\]

[The case of the time-domain Green’s function is included by taking \( \rho(\omega) = \exp(-i\omega t) \). The absolute square of this quantity is not invariant under the interchange of the two spatial coordinates, if only (C4) is satisfied. Therefore, it is possible to design an experiment that would in principle observe a difference between the two situations; the AB effect with \( 2\sigma \neq \text{integer} \) falls into this category.

For (C3) [and \textit{a fortiori} also for (C2)], an amplitude such as (14) would have its absolute square preserved under the interchange of the two spatial coordinates, so no experiment can detect a difference between the two situations. A uniform magnetic field and the AB effect with \( 2\sigma = \text{integer} \) fall into this category; these examples would be physically reciprocal.

Yet there is an important difference between (C3) and (C2). Start with a Hamiltonian \( H \) and embed it into a broader class \( H' = H + V \), where \( V \) is any local real potential energy (e.g., an electrostatic potential energy). The Green’s function becomes

\[
\tilde{G}' = \tilde{G} + (i\hbar) \tilde{G} V \tilde{G} + (i\hbar)^2 \tilde{G} V \tilde{G} V \tilde{G} + \cdots \tag{15}
\]

in which we adopt the obvious shorthand, for example,

\[
(\tilde{G} V \tilde{G})(\vec{r}, \vec{r}', \omega) = \int \tilde{G}(\vec{r}, \vec{r}_1, \omega)V(\vec{r}_1)\tilde{G}(\vec{r}_1, \vec{r}', \omega) d^3r_1. \tag{16}
\]

It is easy to see that under \( H \rightarrow H' \), (C2) would be preserved but not (C3). In this sense, (C2) is said to be \textit{robust} while (C3) describes a kind of reciprocity that is “accidental” in that the property is not shared by the broader class of systems \( H' \) in which \( H \) is embedded, and a generic perturbation on \( H \) would destroy the reciprocity. This is exactly the situation for a uniform magnetic field, as was already discussed with reference to the classical limit. Incidentally, (C4*) [but not (C4) in general] is also robust, that is, preserved under (15). We therefore expect that the classification of the AB effect as (C2) for \( 2\sigma = \text{integer} \) or as (C4) for \( 2\sigma \neq \text{integer} \) should survive the addition of an electrostatic potential; this will be shown explicitly in the Appendix.

Returning to the AB effect, we see that reciprocity beyond monoenergetic experiments depends critically on the value of \( 2\sigma \). If either \( q \) or \( \Phi \) is allowed to take on continuous values, then reciprocity does not hold; the early experiment by Chambers [18] did not make use of a quantized flux, though its
interpretation in terms of the indispensable role of the vector potential was less convincing because the magnetic field was not completely confined. It is more interesting to consider the case where (i) $q$ is quantized in units of $q_0 = \alpha e$ (where $e$ is the charge of the electron) and (ii) the confined flux is produced by a supercurrent due to a Bose condensate of particles with charge $\beta e$, so that $\Phi$ is quantized in units of $\Phi_0 = h/(\beta e)$ [19]. Then the condition for (C2) becomes
\[
\frac{2\alpha}{\beta} = \text{integer.} \quad (17)
\]
In actual experiments [13] with electrons as test particles and the flux produced by Cooper pairs, $\alpha = 1$ and $\beta = 2$, so reciprocity does hold. However, this result depends on subsidiary conditions which are not required by the fundamental principles of electromagnetism, since one can imagine free quarks ($\alpha = 1/3$) and/or a condensate of Cooper quadruplets ($\beta = 4$).

V. CONCLUSION

Four possible proposals for reciprocity have been analyzed: (C1) is not gauge invariant, and (C4) does not ensure reciprocity except in monoenergetic experiments. Under both (C2) and (C3), reversing the source and detection points would lead to no detectable consequences, but only (C2) is robust, that is, would preserve this property for the broader class of Hamiltonians $H' = H + V$—so that reciprocity for $H$ is not “accidental.” In the case of the AB effect, the result other than for monoenergetic experiments is different depending on whether $2\sigma$ is an integer; in the most convincing actual experiments [13], $2\sigma$ is indeed an integer, but this is not guaranteed by the laws of electromagnetism and QM. In a hypothetical AB experiment with free quarks, say, one would observe nonreciprocity in the probability. This would be another manifestation of the effect of the vector potential even when the charge never encounters a magnetic field.

A more realistic proposal is an experiment such as that of Chambers [18], provided the complications due to the leaked magnetic flux can be interpreted convincingly.

ACKNOWLEDGMENTS

P.T.L. acknowledges the partial support received from the Fulbright Scholarship Program when this work was first begun at National Taiwan University. We thank T. P. Cheng, G.-J. Ni, A. Tonomura, C. N. Yang, and M. H. Yung for correspondence and discussions.

APPENDIX: DERIVATION USING GREEN’S THEOREM

In this appendix, adopt the shorthand $G \equiv \hat{G}(\vec{R}, \vec{r}, \omega), G' \equiv \hat{G}(\vec{r}', \omega)$. First consider the case $\vec{B} = 0$, $\vec{A} = 0$, and use the relation
\[
(H - E)G = \frac{\hbar}{i} \delta(\vec{R} - \vec{r}), \quad (H - E)G' = \frac{\hbar}{i} \delta(\vec{R} - \vec{r}').
\]
where $H$ is understood to act on the first variable in $G$. Multiply the first equation by $G'$ and the second by $G$, take the difference, and integrate $\vec{R}$ over the domain $\Omega$ of the problem, say the annulus $R_1 < r < R_2$ in the $x$-$y$ plane (inclusion of the $z$ coordinate is trivial); we obtain
\[
\int_\Omega \{G[(H - E)G'] - G'[(H - E)G]\} dV = \frac{i\hbar}{2} [\hat{G}(\vec{r}', \vec{r}, \omega) - \hat{G}(\vec{r}, \vec{r}', \omega)]. \quad (A2)
\]
All terms without derivatives cancel on the left-hand side, and the only relevant term in $H$ is $H = -\hbar^2/(2m)\nabla^2$, which upon integration by parts gives [3]
\[
\frac{\hbar^2}{2m} \int_S \{G(\partial G' / \partial n) - G'(\partial G / \partial n)\} dS = \frac{i\hbar}{2} [\hat{G}(\vec{r}', \vec{r}, \omega) - \hat{G}(\vec{r}, \vec{r}', \omega)], \quad (A3)
\]
where $S$ is the boundary of $\Omega$ (in this case, the circles at $R_1$ and $R_2$) and $n$ is the coordinate along the outward normal. If the boundary condition for the first variable in $G$ is [3]
\[
\alpha G + \beta(\partial G / \partial n) = 0, \quad (A4)
\]
on these circles, then the left-hand side of (A3) vanishes and we prove (C1) in this case. This derivation [in contrast to the one based on (3) and eigenfunctions being real] remains valid even for a non-Hermitian Hamiltonian with a complex scalar potential. Next, consider the case $\vec{B} = 0$, $\vec{A} \neq 0$ in a simply connected domain. We transform to a new gauge with $\vec{A} = 0$, in which (C1) is established as shown previously; upon transforming back, we prove (C2) in the original gauge.

The much more interesting case is $\vec{B} = 0$, $\vec{A} \neq 0$ in a multiply connected domain, say the AB effect as discussed in Sec. III D. If the vector potential (7) is to be removed by the gauge transformation (4), we need $\xi = -\Phi/2\pi\varphi, \theta = -\sigma\varphi$. To avoid the wave functions becoming multivalued, we need to impose a cut $C$, say at $\varphi = 2\pi$, across which the wave functions are discontinuous,
\[
\psi(\rho, 2\pi) = \exp(-2i\pi\sigma)\psi(\rho, 0), \quad (A5)
\]
and likewise for the first argument in the Green’s function’s, say,
\[
\hat{G}(R, 2\pi; r, \varphi; \omega) = \exp(-2i\pi\sigma)\hat{G}(R, 0; r, \varphi; \omega). \quad (A6)
\]
Now the boundary $S$ consists of not just the circles at $R_1$ and $R_2$, but also the two sides of the cut. The contribution from the former to (A3) vanishes as before, while the latter gives
\[
\frac{\hbar^2}{2m} \int_C \Delta[G(\partial G' / \partial y) - G'(\partial G / \partial y)] d^2 R = \frac{i\hbar}{2} [\hat{G}(\vec{r}', \vec{r}, \varphi) - \hat{G}(\vec{r}, \vec{r}', \varphi)], \quad (A7)
\]
where $n = \mp y$ on the two sides of the cut (i.e., $\varphi = 0$ and $\varphi = 2\pi$) and $\Delta$ denotes the discontinuity across it. Because there are two factors of $G$ in each term of the surface integral, this discontinuity is proportional to
\[
\exp(-4i\pi\sigma) - 1, \quad (A8)
\]
which then establishes (C2) if $2\sigma$ is integer. This derivation lumps all the complications into the cut in a way that does not require knowledge of the wave function or Green’s function [except in the relative phase as in (A5) or (A6)] and consequently has the advantage that it remains valid for (i) a generalized (i.e., noncircular) annulus, $R_1(\varphi) < r < R_2(\varphi)$, provided that the magnetic field remains axially...
symmetric; (ii) inclusion of the third dimension; (iii) addition of an electrostatic potential [since potential energy terms obviously cancel in (A2)]; and (iv) likewise the addition of an imaginary part to the scalar potential. Incidentally, this provides another proof that in this case (C2) is robust; that is, it survives the addition of a potential $V$. These generalizations are important for application to the Chambers experiment [18], in which an electric field is involved.

If $2\sigma \neq$ integer, (A8) is nonzero and (C2) cannot be proved. However, we can do something different: in the derivation, replace $G$ with its conjugate, in which case (A7) becomes, for a Hermitian Hamiltonian $H$,

$$\frac{\hbar^2}{2m} \int \Delta[G^*(\partial G/\partial n) - G'(\partial G^*/\partial n)] d^2R = i\hbar[\tilde{G}(\vec{r}', \vec{r}, \omega)^* + \tilde{G}(\vec{r}, \vec{r}', \omega)].$$

(A9)

Note the relative plus sign on the right-hand side. Since $G$ and $G^*$ have opposite phase discontinuities, the left-hand side of (A9) does not contain a discontinuity, so we prove (C4*) as in (14). Again, this proof survives the generalizations (i), (ii), and (iii) [but not (iv)] cited at the end of the preceding paragraph.