The Aharonov–Bohm Effect and Scattering Theory

S. N. M. RUJSENAARS*†

Max-Planck-Institut für Physik und Astrophysik, Munich, Federal Republic of Germany

Received November 17, 1981; revised August 11, 1982

The Aharonov–Bohm effect is studied from the viewpoint of scattering theory. A scattering experiment is proposed that could distinguish between two different dynamics that have been considered in the literature. The behavior of the S-matrices under gauge transformations is established and an anomalous cluster property is found.

1. INTRODUCTION

More than 20 years after the work of Aharonov and Bohm [1] and the subsequent experiments, the effect named after these authors (but first predicted by Ehrenberg and Siday [2]) is still a hotly debated issue. (An extensive bibliography covering the period 1959–1971 can be found in [3].) The main experimental findings concern the shift in the interference pattern due to two coherent electron beams under the influence of a tiny solenoid or magnetized whisker placed between the two beams [4, 5]. The results themselves are uncontested, but the theoretical description and interpretation of [1] have been challenged by several authors [3, 6–11]. A recent and particularly loud criticism is due to Bocchieri, Loinger and Siragusa [12–17, 53]. The main objections raised by the latter authors concern the continuity of the wave function and the influence of the magnetic field of the solenoid. More in detail, they emphasize that quantum-mechanical wave functions are not necessarily continuous, so that the explanation of the shift in terms of the phase difference around a closed loop is not cogent. Instead, they point out that the observed shift can just as well be explained by the fact that the wave function has tails extending into the region where the magnetic field is non-zero (a point that had been emphasized previously by Strocchi and Wightman [9]). Their criticism and that of Roy [18] has triggered a flurry of papers presenting counterarguments [19–31, 54–56].

It is perhaps not surprising that the physical situation present in these experiments can be analyzed theoretically in such different ways, yet lead to the same conclusions as to the observed outcome. The problem is that a really convincing mathematical model of the situation would be of such complexity as to resist an attempt at a prediction; even for the much simpler situation where the solenoid is absent (i.e., the two-slit experiment) one must take recourse to a description in terms of an optics...
analogy, instead of considering a realistic quantum-mechanical Hamiltonian. (However, Kobe [32] has performed approximate but detailed calculations based on Feynman path integrals, which shed considerable light on the problem.)

In this paper we consider a physical situation whose quantum-mechanical description is much more straightforward, viz., the scattering of electrons off a cylinder inside which there is a current-carrying solenoid. For this situation two different viewpoints on the description of excluded magnetic fields that can be found in the literature lead to two different Hamiltonians. We obtain the corresponding S-matrices and point out that an appropriate experiment could serve to verify one description and falsify the other one. (A related interference experiment, which concerns the theoretically less accessible situation of two coherent electron beams and a shielded solenoid, happens to be under way in Tubingen, as we learned from Mollenstedt.) The probative value of the scattering experiment we propose increases with decreasing electron energy and cylinder radius and therefore hinges on the availability of very slow electrons and very tiny cylinders and solenoids. Quantitatively, we find the scattering cross sections for the two dynamics and show they differ considerably for $R^2/\lambda \to 0$ (where $\lambda$ = the de Broglie wave length of the electron, $R$ = the cylinder radius).

The scattering in the idealized situation where $R = 0$ (scattering off a "thread of magnetic flux") was first considered by Aharonov and Bohm in their original paper [11], but later criticized by Feinberg [33] and Henneberger [34]. The former authors consider an incoming solution and read off the scattering amplitude from its spatial asymptotics in the traditional fashion of time-dependent scattering theory. Their choice has been challenged since its leading asymptotics is not a plane wave, but rather a plane wave multiplied by a discontinuous angle-dependent phase factor.

Here, we first reconsider this limiting situation in detail. For one dynamics we reobtain the Aharonov–Bohm amplitude both in the time-independent and in the time-dependent framework, but with an additional $\delta$-function contribution in the forward direction (not present in [1]), without which the S-matrix would not be unitary. We then study the other perturbed dynamics that has been considered in the literature [12–17, 19–24, 35–36] in connection with situations of this type (in some cases rather implicitly). Here, too, we explicitly obtain the S-matrix, which turns out to be diagonal in momentum space.

Subsequently, we study the hard-core case, which corresponds to the experimental situation described above. Again, one can explicitly find the S-matrices for the two dynamics one can envisage. In the low-energy limit, one of them leads to the anisotropic Aharonov–Bohm differential cross section, the second one to the usual isotropic hard-core cross section. We prove moreover that these results are invariant under a large class of gauge transformations.

Throughout the paper, the time-dependent (Møller operator) point of view is emphasized since this framework admits a precise description of the freedom one has in choosing a perturbed dynamics, comparison map and comparison dynamics, and is closer to physical reality in that it deals with wave packets instead of non-normalizable functions. In spite of this, it is still less widely known and used in the
physics literature than the time-independent approach. Therefore we begin in Section 2 with a concise review of the former approach and its relation to time-independent scattering theory, in a form that is most appropriate for our later requirements.

In Section 3 we define the four dynamics involved in the description of the two situations mentioned above. The two choices for each situation correspond to different interpretations of a formally given differential operator. In Section 4 we obtain the $S$-matrices associated with the various dynamics. The two $S$-matrices for the experimentally realizable hard-core case are further studied in Section 5. They form a striking illustration of the fact that the different self-adjoint extensions of a symmetric operator may lead to very different physics. In this section we also discuss our choice of boundary conditions and the "hydrodynamical" approach. Section 6 concerns the gauge invariance of the results, and in Section 7 we compare our derivation of the Aharonov–Bohm scattering amplitude with that of \cite{1} and study the long-range behavior of the various $S$-matrices and wave operators.

We have relegated some technical results to an Appendix. Here, we make use of notions that are well known in mathematical scattering theory. We should like to point out that the results have an independent interest in this context, in particular the behavior under gauge transformations and the anomalous long-range behavior of the wave and scattering operators. Also, it is rare one cannot only prove existence and completeness of the wave operators, but obtain them explicitly as well. In this regard, the mathematics involved here has some intriguing similarities to certain integrable one-dimensional systems (in particular the Calogero system) and relativistic field theories (in particular the Federbush, massless Thirring and continuum Ising models). In fact, we have occasion to use some of our results on the scattering for these field theories \cite{37} in proving an anomalous cluster property for the $S$-matrices arising in Aharonov–Bohm-like situations.

\section{Review of Scattering Theory}

Consider a quantum-mechanical system whose states $\psi$ are unit vectors in a Hilbert space $H$ and whose time evolution is described by a self-adjoint Hamiltonian $H$ acting on $H$. Time-dependent scattering theory is concerned with the question whether the states $e^{-iHt}\psi$ are scattering states, i.e., whether for $t \to \pm \infty$ they "behave like" states that are "free" in an appropriate sense. This entails the comparison of the dynamics $e^{-iHt}$ with a free dynamics $e^{-iH_0t}$ that acts in general on another Hilbert space $H_0$. Thus, one also needs a comparison map $\mathcal{F}$ from $H_0$ to $H$. A state $\psi \in H$ is a scattering state if one can find free states $g_\pm \in H_0$ such that the difference

$$e^{-iHt}\psi - \mathcal{F}e^{-iH_0t}g_\pm$$

(2.1)

goes to zero for $t \to \pm \infty$. If so, the relation between $g_\pm$ and $\psi$ can be written

$$\psi = W_\pm g_\pm,$$

(2.2)
where $W_{\pm}$ are the Møller wave operators

$$W_{\pm} = \lim_{t \to \pm \infty} e^{iH_{0}t} S e^{-iH_{0}t}. \quad (2.3)$$

The vectors $g_{+}$ and $g_{-}$ are then related by

$$g_{+} = S g_{-}, \quad (2.4)$$

where $S$ is the $S$-operator or $S$-matrix

$$S = W_{+}^{*} W_{-}. \quad (2.5)$$

The framework just described is quite general. In the sequel, we shall consider the special situation where $\mathcal{H}$ is a Hilbert space of square-integrable functions depending on points in a plane. This corresponds to the physical systems we intend to model, viz., particles scattering off cylindrical "obstacles." In this case the coordinate along the cylinder axis may be neglected, and so we need only consider scattering in a plane. In the rest of this section we therefore restrict ourselves to a situation of this kind, viz., a (non-relativistic) particle scattering off a short-range continuous electric potential $V(x, y)$. This has the advantage that the scattering theory for its three-dimensional analog is well known, so that we need only transcribe results familiar in three dimensions to the two-dimensional case.

We begin by detailing the operators and spaces introduced above for our example system. Its Hamiltonian reads

$$H = - (\partial_{x}^{2} + \partial_{y}^{2}) + V(x, y), \quad (2.6)$$

and acts on the position space $\mathcal{H} = L^{2}(\mathbb{R}^{2}, dx dy)$. (Throughout the paper we put the particle mass equal to $1/2$ and $\hbar$ equal to $1$ to unburden formulas.) As comparison Hamiltonian we take

$$H_{0} = k_{x}^{2} + k_{y}^{2}, \quad (2.7)$$

acting on the momentum space $\mathcal{H}_{0} = L^{2}(\mathbb{R}^{2}, dk_{x} dk_{y})$. The comparison operator $\mathcal{G}$ is just Fourier transformation. Adapting results for the three-dimensional case [38] one can then show that the wave operators (2.3) exist and are complete in the sense that their ranges equal the orthocomplement of the bound states. In this case, one could just as well have both $H$ and $H_{0}$ act on $\mathcal{H}$ and dispense with $\mathcal{G}$, but the advantage of the present choice is the simple relation of $W_{\pm}$ with the incoming and outgoing functions of time-independent scattering theory, and the fact that $S$ as defined by (2.5) acts on $\mathcal{H}$, where it is most easily interpreted physically.

In time-dependent scattering theory, one solves instead the time-independent Schrödinger equation (or, rather, the equivalent Lippmann–Schwinger equation) for the incoming and outgoing functions $\varphi_{-}(r, k)$ and $\varphi_{+}(r, k)$, respectively, which reduce to plane waves $e^{ik \cdot r}$ at infinity. The connection with the wave operators is very simple: $W_{\pm}$ as defined above are integral operators whose kernels are just $\varphi_{\pm}(r, k)$. 


That is, the image \((W \pm g) \in \mathcal{H}\) of \(g \in \mathcal{H}_0\) is given by

\[
(W \pm g)(r) = (2\pi)^{-1} \int dk \, \varphi \pm (r, k) \, g(k).
\]  

(2.8)

From now on we specialize to the case where \(V(x, y)\) is a spherically symmetric function \(V(r)\), and correspondingly employ polar coordinates \((r, \theta)\) and \((k, \theta)\). In time-independent scattering theory, one considers the spatial asymptotics of the incoming wave function, which for the case at hand reads

\[
\varphi_-(r, \theta; k, \theta') \sim e^{ikr \cos(\theta - \theta')} + f(k, \theta - \theta') e^{ikr/r^{1/2}}.
\]  

(2.9)

The differential scattering cross section is then given by

\[
\left(\frac{d\sigma}{d\theta}\right)(k, \theta) = |f(k, \theta)|^2.
\]  

(2.10)

To find the asymptotics of \(\varphi_-\) and hence the scattering amplitude \(f(k, \theta)\), one proceeds as follows.

On one hand (2.9) implies

\[
\varphi_-(r, \theta; k, 0) \sim \sum_{m = -\infty}^{\infty} \left[ \frac{(-)^m}{(2\pi kr)^{1/2}} e^{-im\theta} + \left( \frac{e^{-im\theta}}{(2\pi kr)^{1/2}} + \frac{f_m}{r^{1/2}} \right) e^{im\theta} \right].
\]  

(2.11)

To see this, first make a Fourier expansion of the plane wave factor,

\[
e^{ikr \cos \theta} = \sum_{l} e^{il\theta} \tilde{f}_{lm}(kr) e^{il\theta},
\]

(2.12)

and of the scattering amplitude

\[f(k, \theta) = \sum f_m(k) e^{im\theta}.
\]

(2.13)

Here as in the sequel we have denoted the Bessel function of the first kind of order \(n\) by \(J_n(x)\). Using its known behavior for \(x \to \infty\),

\[
J_n(x) \sim (2/\pi)^{1/2} x^{1/2} \cos \left( x - \frac{1}{2} n\pi - \frac{x}{4} \right),
\]

(2.14)

one obtains (2.11).

On the other hand, \(\varphi_-\) solves the Schrödinger equation and is regular at the origin. Using separation of variables one then infers

\[
\varphi_- (r, \theta; k, 0) = \sum a_m(k) \varphi_m(r, k) e^{im\theta}.
\]

(2.15)

where \(\varphi_m\) is a solution to the radial Schrödinger equation of angular momentum \(m\),

\[
\left[ -\frac{d^2}{dr^2} - \frac{1}{r} \frac{d}{dr} + \frac{m^2}{r^2} + V(r) \right] \varphi = k^2 \varphi.
\]

(2.16)
which is regular at the origin. Since \( V \) has short range, the behavior of \( \varphi_m \) for \( r \to \infty \) must read (up to a multiplicative constant),

\[
\varphi_m(r, k) \sim (2\pi)^{1/2} \cos \left( kr - \frac{1}{2} |m| \pi - \frac{\pi}{4} + \delta_m(k) \right) (kr)^{1/2},
\]

where the phase shift \( \delta_m \) measures the difference in argument as compared to the asymptotic behavior of \( J_{|m|}(kr) \), which is the solution to the radial equation with \( V = 0 \) that is regular at the origin. One normalizes \( \varphi_m \) by requiring (2.17) hold exactly, i.e., by putting the multiplicative constant equal to 1.

Now we are in a position to compare the asymptotics of \( \varphi_- \) as given by (2.11) on one hand, and by (2.15) and (2.17) on the other hand. This leads to

\[
a_m(k) = i^{|m|} e^{i\delta_m(k)}
\]

and

\[
f_m(k) = (e^{2i\delta_m(k)} - 1)/(2\pi k)^{1/2}.
\]

Thus one obtains finally

\[
f(k, \theta) = (2\pi k)^{-1/2} \sum (e^{2i\delta_m(k)} - 1) e^{im\theta},
\]

which expresses the scattering amplitude in terms of the phase shifts \( \delta_m \).

In the time-dependent approach the spherical symmetry can be taken into account similarly. One decomposes \( \mathcal{H} \) and \( \mathcal{H}_0 \) into subspaces \( \mathcal{H}_m, \mathcal{H}_{0,m} \) with corresponding projections \( P_m, P_{0,m} \). One has e.g.,

\[
(P_m \psi)(r, \theta) = (2\pi)^{-1/2} e^{im\theta} \psi_m(r),
\]

where

\[
\psi_m(r) = (2\pi)^{-1/2} \int_{-\pi}^{\pi} d\theta e^{-im\theta} \psi(r, \theta)
\]

are the angular momentum components of \( \psi \), belonging to the Hilbert space \( \mathcal{H}_m = L^2([0, \infty), r \, dr) \). The Hamiltonian \( H \) leaves \( \mathcal{H}_m \) invariant and gives rise to a sequence of channel Hamiltonians

\[
H_m = -\frac{d^2}{dr^2} - \frac{1}{r} \frac{d}{dr} + \frac{m^2}{r^2} + V(r)
\]

on \( \mathcal{H}_m \). The Fourier transformation \( \mathcal{F} \) can be decomposed as

\[
(\mathcal{F} g)(r, \theta) = (2\pi)^{-1/2} \sum e^{im\theta} (\mathcal{F}_m g_m)(r),
\]

where \( \mathcal{F}_m \) is the unitary operator

\[
(\mathcal{F}_m F)(r) = e^{im\theta} \int_0^\infty dk J_{|m|}(kr) F(k)
\]
from $\mathcal{H}_k = L^2([0, \infty), k \, dk)$ onto $\mathcal{H}_r$. Correspondingly, one considers a sequence of channel wave operators,

$$W_{\pm m} = \lim_{t \to +\infty} e^{iH_{tm} t} \mathcal{S}_m e^{-ik^2 t}.$$  \hfill (2.26)

For the potentials considered in this section one can show that $W_{\pm m}$ exist and are complete in the sense that their range equals the orthocomplement of the bound states in channel $m$. Their relation with the time-independent approach is again simple: They satisfy

$$(W_{\pm m} F)(r) = i |m| \int_0^{\infty} dk \, k \varphi_m(k, r) e^{\pi \delta_m(k)} F(k).$$  \hfill (2.27)

The corresponding $S$-matrix,

$$S_m = W_{+ m} W_{- m}^*,$$  \hfill (2.28)

is therefore multiplication by $e^{2 \pi \delta_m(k)}$ on $\mathcal{H}_k$.

The full wave operators (2.3) can now be written

$$(W_{\pm} g)(r, \theta) = (2\pi)^{-1/2} \sum e^{im\theta} (W_{\pm m} g_m)(r).$$  \hfill (2.29)

Similarly, from (2.20) one infers that the full $S$-operator satisfies

$$(g, Sh) = \sum \int_0^{\infty} dk \, k \tilde{g}_m(k) S_m(k) h_m(k)$$

$$= (g, h) + (2\pi)^{-1} \sum \int_0^{\infty} dk \, k \int_{-\pi}^{\pi} d\theta \, d\theta' \, e^{im(\theta - \theta')} \tilde{g}(k, \theta) h(k, \theta')(e^{2\pi \delta_m(k)} - 1)$$

$$= (g, h) + \int_0^{\infty} dk \, k \int_{-\pi}^{\pi} d\theta \, d\theta' \tilde{g}(k, \theta) h(k, \theta')(ik/2\pi)^{1/2} f(k, \theta - \theta'),$$  \hfill (2.30)

so that, in obvious shorthand,

$$(S - 1)(k, \theta) = (ik/2\pi)^{1/2} f(k, \theta).$$  \hfill (2.31)

An important physical feature of the $S$-matrix associated with the above potentials is its cluster property: If we introduce the translation operator $U_\alpha$ by

$$(U_\alpha g)(k, \theta) = e^{i\alpha \cdot \theta} g(k, \theta),$$  \hfill (2.32)

then one has for any $g, h \in \mathcal{H}_0$

$$\lim_{|\alpha| \to \infty} (U_\alpha g, SU_\alpha h) = (g, h).$$  \hfill (2.33)
Physically, this says that particles passing the scattering center at a great distance will not be influenced by it.

For the situations studied below, the issues discussed in this section must be reconsidered since the two Hamiltonians that we associate to them are not of the form (2.6). For one choice, there exists no natural time-independent scattering theory (as far as we can see), but it can easily be studied with the time-dependent approach sketched above. For the other choice, both approaches are possible, but since the interaction is not of the form (2.6), it is not immediately clear that the $S$-matrix is the same in both cases, or even that the wave operators exist and are complete. We shall therefore first obtain the phase shifts in the way sketched above and then prove (in the Appendix) that the time-dependent theory indeed leads to the same results.

3. DEFINITION OF DYNAMICS

Let us now pass to the situation with which this paper is concerned. We consider a magnetic field of the form

$$H(x, y, z) = (0, 0, H(r)), \quad (3.1)$$

where $r^2 = x^2 + y^2$ and where $H(r)$ vanishes for $r > R$. Denoting by $\varphi$ the magnetic flux through the $x-y$ plane,

$$\varphi = 2\pi \int_0^R dr r H(r), \quad (3.2)$$

it is clear that the magnetic potential for $r \geq R$ can be taken as

$$A(x, y, z) = \frac{\varphi}{2\pi r^2} (-y, x, 0). \quad (3.3)$$

Thus, for $r \geq R$ the corresponding formal quantum-mechanical Hamiltonian for a particle of mass $1/2$ and charge $e$ can be written in polar coordinates as (recall $\hbar \equiv 1$)

$$H(\alpha) = -\partial_r^2 - \frac{1}{r} \partial_r + \frac{1}{r^2} (i \partial_\theta - \alpha)^2. \quad (3.4)$$

Here, we have set

$$\alpha = -e\varphi/2\pi c, \quad (3.5)$$

and we have suppressed the irrelevant $z$-coordinate.

Let us first consider the limiting case where $R = 0$ (corresponding to a "thread of flux"). The space of wave functions $\mathcal{H}$ is then the usual position space in polar coordinates, $\mathcal{H}_r \otimes \mathcal{H}_\theta$, where $\mathcal{H}_r = L^2([0, \infty), \, r \, dr)$ and $\mathcal{H}_\theta = L^2((-\pi, \pi], \, d\theta)$. For $\alpha = 0$, there is no reason to interpret $H(\alpha)$ as anything else than the usual Laplacean in
polar coordinates representing the kinetic energy of a free particle, which becomes multiplication by $k^2$ on the momentum space $\mathcal{H}_k \otimes \mathcal{H}_\theta$ after a Fourier transformation. More in detail, $i\partial_\theta$ can be replaced by $-m$ on each angular momentum subspace $\mathcal{H}_m$, and the resulting differential operator $H_m$ on $\mathcal{H}$ transforms into multiplication by $k^2$ on $\mathcal{H}_k$ if one uses the unitary map $\mathcal{F}_m$ (cf. (2.25)).

For $\alpha \neq 0$, however, there are two essentially different interpretations of the formal expression (3.4). The first one is to argue that $i\partial_\theta$ should have the same meaning as in the case $\alpha = 0$, so that it can be replaced by $-m$ on $\mathcal{H}_m$. This leads to a sequence of channel Hamiltonians

$$H_{m+\alpha} = -\frac{d^2}{dr^2} - \frac{1}{r} \frac{d}{dr} + \frac{(m + \alpha)^2}{r^2}$$

(3.6)
on $\mathcal{H}$. We shall denote $H(\alpha)$ defined in this way by $H^0_\alpha$. (We are glossing over a point of rigor here that will be dealt with in the Appendix.) The second definition makes use of the following observation: if a potential $A$ satisfies $eA/c = \nabla A$, then it follows that

$$\left(\frac{\nabla i}{c} - \frac{e}{c} A\right)^2 = e^{iA}(\nabla) e^{-iA},$$

(3.7)
as can be easily verified. In the case at hand, one can take

$$A(x, y) = -\alpha \theta(x, y),$$

(3.8)
if one is willing to neglect the discontinuity of this function on the ray $\theta = \pi$. The corresponding Hamiltonian $\tilde{H}^0_\alpha$ obtained through (3.7) reads in polar coordinates

$$\tilde{H}^0_\alpha = e^{-i\alpha \theta} \left(-\partial_r^2 - \frac{1}{r} \partial_r + \frac{1}{r^2} (i\partial_\theta)^2\right) e^{i\alpha \theta},$$

(3.9)
where $i\partial_\theta$ has the same meaning as in the free case. Thus, we can also write

$$\tilde{H}^0_\alpha = e^{-i\alpha \theta} H^0_\alpha e^{i\alpha \theta}.$$ 

(3.10)
This operator has been proposed or considered in Aharonov–Bohm-like situations by many authors. It is important to notice that the two definitions differ only in their interpretation of the expression $i\partial_\theta - \alpha$ in (3.4), and that they coincide for $\alpha$ integer and for wave functions that vanish near the ray $\theta = \pi$. The difference consists in a different boundary condition at $\theta = \pi$: In the first case $i\partial_\theta$ acts on wave functions satisfying

$$\lim_{\theta \downarrow \pi} \varphi(\theta) = \lim_{\theta \uparrow -\pi} \varphi(\theta).$$

(3.11)
In the second case the wave functions in the definition domain satisfy

$$\lim_{\theta \downarrow -\pi} \varphi(\theta) = e^{2i\alpha \pi} \lim_{\theta \uparrow \pi} \varphi(\theta).$$

(3.12)
Of course, for the full wave functions $\psi(r, \theta)$ this implies a phase discontinuity on the ray $\theta = \pi$ in the $(r, \theta)$ plane if $\alpha \notin \mathbb{Z}$. It should be emphasized (as has been done by other authors) that there is no physical or mathematical principle forbidding this. (Also, this discontinuity should not be confused with “multi-valuedness” of wave functions. This phenomenon only arises if one departs from the space $\mathbb{H}$ considered here, and considers instead wave functions defined on Riemann surfaces and their projections on the plane; for this viewpoint cf. e.g., [39–41].) As usual, one is dealing here with unbounded operators that cannot be defined on all square-integrable wave functions without loosing their self-adjointness, which is the mathematical property that makes them physically interpretable. The operator $id/d\theta$, as naturally defined on smooth wave functions $\varphi(\theta)$ vanishing for $|\theta| \to \pi$, is not essentially self-adjoint, i.e., it is not yet defined on a sufficiently large set of wave functions to admit a unique extension to a self-adjoint operator. In this case there is a one-parameter family of self-adjoint extensions, characterized by the boundary condition (3.12) at $\theta = \pi$. (We mention in passing that a quite analogous situation obtains for the Thirring and Federbush quantum field theories on the unphysical sector, cf. [42, p. 374].) Let us finally mention the following argument, motivating the change in domain accompanying the second definition (which might convince those readers who consider the well-known phase shift consideration a convincing quantum-mechanical argument): If $\psi_0(t, r, \theta)$ is a continuous solution to the free Schrödinger equation, then the discontinuous wave function $\exp(-i\alpha \theta) \psi_0(t, r, \theta)$ solves the Schrödinger equation corresponding to (3.4) (if one interprets the $\theta$-derivative on the ray $\theta = \pi$ as a derivative from the left).

We now return to the case $R > 0$. We assume that the magnetic field (3.1) is shielded by a hard core at $r = R$. Correspondingly, the dynamics $H(\alpha)$ only acts on wave functions vanishing at $r = R$ (Dirichlet boundary conditions). For the position space $\mathbb{H}$, we therefore take $\mathbb{H}_R \otimes \mathbb{H}_\theta$, where $\mathbb{H}_R = L^2([R, \infty), r \, dr)$. Again, we can turn $H(\alpha)$ into a self-adjoint operator in two ways. The first definition, denoted $H_R^\alpha$, is the analog of $H_\alpha^0$; it corresponds to a sequence of radial Hamiltonians $H_{m+\alpha}^R$ on $\mathbb{H}_R$, given by the right-hand side of (3.6), supplemented by the Dirichlet condition at $r = R$ on the wave functions in their domains. The second definition, denoted $\tilde{H}_R^\alpha$, is the transform under $e^{i\alpha \theta}$ of the hard-core Laplacean. That is, it is given by the right-hand side of (3.9) viewed as an operator on $\mathbb{H}_R \otimes \mathbb{H}_\theta$ by taking $\psi(R, \theta) = 0$ for the wave functions in its domain. Thus we have

$$\tilde{H}_R^\alpha = e^{-i\alpha \theta} H_0^R e^{i\alpha \theta}. \quad (3.13)$$

Before closing this section it may not be out of place to emphasize that we have tried to paraphrase the arguments of other authors concerning the second version of the dynamics, as given by $\tilde{H}_R^\alpha$, $R \geq 0$. In our opinion the second definition does not have the cogency of the first one. Apart from the arguments needing some act of faith on our part, they lead to a dynamics that is not rotationally invariant. There appears to be no reason to prefer having a discontinuity on the ray $\theta = \pi$ above having it on any other ray. In any case, the four dynamics involved are well-defined mathematical
THE AHARONOV-BOHM EFFECT

objects, and in the next sections we shall study the scattering theory for all of them. It will turn out that from a physical point of view the scattering associated with $\hat{H}_R^0$ is rotationally invariant, though the $S$-matrix is not. Also, it will be seen that there is a remarkable connection between the $S$-matrices for $H_R^0$ and $\hat{H}_R^0$ (cf. Section 7).

4. The $S$-Matrices for $H_R^0$ and $\hat{H}_R^0$, $R \geq 0$

The dynamics $H_R^0$, $R \geq 0$, lead to channel Hamiltonians $H_{m+a}^0$ and $\hat{H}_{m+a}^0$ that are of the form (2.23), except that in these cases $V(r)$ is an $m$-dependent and for $R = 0$ quite singular perturbation of the kinetic energy corresponding to angular momentum $m$. However, it is easy to solve the radial Schrödinger equations and obtain the asymptotics of the solutions regular at the origin and vanishing at $r = R$, respectively, since the equations are of the well-known Bessel form. Thus, one can formally find the phase shifts $\delta_m$ and hence the $S$-matrix $e^{2i\theta_m}$ as sketched in Section 2. What is not clear is whether in the case at hand this procedure can be motivated by and leads to the same results as a treatment in terms of the temporal asymptotics of normalizable wave packets (as is the case for the potentials considered in Section 2). We shall therefore also find the $S$-matrix in the time-dependent framework. As we have explained in Section 2, this entails that we have to select an appropriate free comparison dynamics $H_0$ on a free wave packet space $\mathcal{H}_0$ and a comparison map $\mathcal{S}$. A minimal requirement for "appropriateness" is that the wave operators (2.3) exist, but this does not uniquely determine $H_0$ and $\mathcal{S}$. However, their choice can be further restricted by physical considerations about how perturbed wave packets are expected to behave far from the scattering center. In our case, the potential $A$ can be gauged to zero by a smooth gauge transformation, except on some ray, where singularities must occur. Thus, one expects that wave packets moving according to one of the four dynamics defined in Section 3 will behave like free wave packets "far away" from the origin. This common sense expectation leads us to choose the same $H_0$, $\mathcal{H}_0$ and $\mathcal{S}$ as in Section 2, except that in the case of $H_R^0$ and $\hat{H}_R^0$ we take the hard core into account by taking instead of $\mathcal{S}$ the comparison map

$$\mathcal{S}_R = P_R \mathcal{S},$$

where $P_R$ is the natural projection from $\mathcal{H}_R$ onto $\mathcal{H}_R$, i.e.,

$$(P_R F)(r) = F(r), \quad r \geq R.$$ (4.2)

It turns out that with these choices the wave operators indeed exist and are unitary in all four cases. Moreover, in the case of $H_R^0$, $R \geq 0$, the wave operators are related to the solution of the radial Schrödinger equation in the same way as for the familiar situation considered in Section 2, so that the $S$-matrix found with the time-independent "phase shift" approach is confirmed in the time-dependent "wave packet" framework. We shall now detail these results, considering the four dynamics successively.
Case $A_0$. The Dynamics $H^0_\alpha$. The radial Schrödinger equation for angular momentum $m$ reads

$$
\left( -\frac{d^2}{dr^2} - \frac{1}{r} \frac{d}{dr} + \frac{(m + \alpha)^2}{r^2} \right) \phi = k^2 \phi. \tag{4.3}
$$

The solution regular at the origin is the Bessel function $J_{|m+\alpha|}(kr)$, whose asymptotic behavior is given by

$$
J_{|m+\alpha|}(x) \sim (2/\pi)^{1/2} \cos \left( x - \frac{1}{2} \left| m + \alpha \right| \pi - \frac{\pi}{4} \right) / x^{1/2}. \tag{4.4}
$$

Comparing with (2.17), we conclude that the corresponding phase shift depends only on $\alpha$ and is given by

$$
\delta_m(\alpha) = \frac{1}{2} \pi (|m| - |m + \alpha|). \tag{4.5}
$$

(This result was obtained first by Henneberger [43].) Thus, the $S$-matrix reads

$$
e^{2i\delta_m(\alpha)} = e^{-i\pi \alpha}, \quad m \geq -\alpha
= e^{i\pi \alpha}, \quad m \leq -\alpha. \tag{4.6}
$$

Notice that it equals 1 or $-1$ for $\alpha$ an even or odd integer, respectively. Relation (4.6) can be intuitively understood as follows: Positive (negative) angular momentum wave functions "go around" the origin in an anticlockwise (clockwise) direction. The potential $A$ creates "vorticity" $-\alpha$, which induces wave functions with $m$ greater (smaller) than $-\alpha$ to go around the origin in anticlockwise (clockwise) fashion; therefore such wave functions pick up a phase $e^{-i\pi \alpha} e^{i\pi \alpha}$. (Recall that in a uniform magnetic field a positively charged particle describes a circle that may be taken as a field line for the corresponding vector potential $A$; the particle velocity and $A$ then point in opposite directions along the orbit.)

To obtain the scattering amplitude associated with the phase shifts (4.5), we first note that the corresponding $S$-operator on $\mathcal{H}_0$ satisfies

$$
(S^0_\alpha g)(k, \theta) = \int_{-\pi}^{\pi} d\theta' s_\alpha(\theta - \theta') g(k, \theta'), \tag{4.7}
$$

where

$$
s_\alpha(\theta) = \delta(\theta) \cos \pi \alpha + i \frac{\sin \pi \alpha}{\pi} e^{-i\pi \alpha}/ e^{i\theta} - 1. \tag{4.8}
$$

(Here, $[\alpha]$ denotes the greatest integer less than or equal to $\alpha$, and $P$ denotes the principal value.) To see this, remember that we have

$$
(S^0_\alpha g)(k, \theta) = (2\pi)^{-1} \sum_m \int_{-\pi}^{\pi} d\theta' e^{im(\theta - \theta')} e^{2i\delta_m(\alpha)} g(k, \theta'), \tag{4.9}
$$
so that we need only show the Fourier coefficients of \( s_\alpha(\theta) \) are just \( e^{2i\delta_m(\alpha)} \). But this follows from the relation

\[
\int_{-\pi}^{\pi} d\theta \frac{e^{-ik\theta}}{1 - e^{i\theta}} = \frac{1}{2i} \lim_{\epsilon \to 0} \left[ \oint_C \frac{dz}{1 + \epsilon - z} + \oint_C \frac{dz}{1 - \epsilon - z} \right] = \pi, \quad k \geq 0
\]
\[
= -\pi, \quad k \leq -1. \tag{4.10}
\]

(Here, \( C \) is the unit circle in the complex plane, and a simple transformation and the residue theorem have been used.) The scattering amplitude now follows from (2.31), (4.7) and (4.8): It is given by

\[
f_\alpha(k, \theta) = \left( \frac{2\pi}{ik} \right)^{1/2} \left[ \delta(\theta)(\cos \pi\alpha - 1) + i \frac{\sin \pi\alpha}{\pi} e^{-i|\alpha|\theta} \frac{P}{e^{i\theta} - 1} \right]. \tag{4.11}
\]

Hence, for \( \theta \neq 0 \) the differential scattering cross section reads

\[
\left( \frac{d\sigma}{d\theta} \right)(k, \theta) = \frac{1}{2\pi k} \frac{\sin^2 \pi\alpha}{\sin^2 \theta/2}. \tag{4.12}
\]

Note that the total cross section is infinite for \( \alpha \in \mathbb{Z} \) and vanishes for \( \alpha \notin \mathbb{Z} \).

As mentioned above, these results are confirmed in the time-dependent approach. In the Appendix we shall prove that the wave operators

\[
W_{\pm \alpha, m} = \lim_{t \to \pm \infty} e^{iH_{m+\alpha}t} \mathcal{F}_m e^{-iklt} \tag{4.13}
\]

exist and are unitary maps from \( \mathcal{H}_k \) onto \( \mathcal{H}_r \), satisfying

\[
(W_{\pm \alpha, m} F)(r) = i^{|m|} \int_{0}^{\infty} dk J_{|m+\alpha|}(kr) e^{\mp i\delta_m(\alpha)} F(k) \tag{4.14}
\]

(cf. Theorem A.1). Thus, the corresponding "time-dependent" \( S \)-operator for angular momentum \( m \) is given by (4.6), too.

Before passing to the next case, we should like to comment on the Born approximation to the scattering amplitude, as this might clear up some confusion on this point in the literature. One readily verifies that to first order in \( \alpha \) the Born amplitude is given by the formal expression

\[
f_{\alpha, b}(k, \theta) = -i\alpha(2\pi k)^{-1/2} \int_{0}^{\infty} dr \int_{-\pi}^{\pi} d\theta' \sin \theta' \exp(ir[\cos(\theta' - \theta) - \cos \theta']) \tag{4.15}
\]
To render the integral convergent we insert a factor $e^{-e^r}$, and then do the $r$-integration and consider the limit $e \to 0$ of the resulting $\theta'$-integral. It is easily seen this limit exists provided $\theta \neq 0$, and the limit can be readily evaluated if one uses a bit of trigonometry. Explicitly, one gets

$$f_{\alpha, \theta}(k, \theta) = \alpha \left( \frac{\pi}{2ik} \right)^{1/2} \frac{\cos \theta/2}{\sin \theta/2}, \quad \theta \neq 0,$$

(4.16)

a result that was also obtained by Feinberg [33] and Corinaldesi and Rafeli [44]. In comparison, the exact amplitude (4.11) satisfies

$$f_\alpha(k, \theta) = \alpha \left( \frac{\pi}{2ik} \right)^{1/2} e^{i\theta/2} + O(\alpha^2), \quad 0 < \alpha < 1, \theta 
eq 0.$$

(4.17)

Thus, the Born amplitude is simply an average of the two different amplitudes (4.17). As such, it leads of course to the wrong lowest-order cross section. In our opinion this is hardly a matter of concern however. Rather, it may be considered as surprising that the Born approximation exactly catches the ambiguous behavior evinced by (4.17) since the perturbation at hand violates the conditions for the Born series to make sense and converge both at the origin and at infinity.

Clearly, it is this singular character of the perturbation that can be considered to be the source of the non-analyticity of the exact scattering amplitude near $\alpha = 0$. In the angular momentum channel picture it shows up in the non-analyticity of the $m = 0$ phase shift (cf. (4.5)). Here, too, the Born approximation does more than what one can count on: Using the known integrals

$$\int_0^\infty dr \frac{J_m(r)^2}{r} = (2m)^{-1}, \quad m = 1, 2, 3,\ldots$$

(4.18)

one infers it yields the correct first order phase shift for $m \neq 0$. For $m = 0$ it leads to a divergent integral, viz., the left hand side of (4.18) with $m = 0$ (as noted first by Purcell and Henneberger [45]), but in this case the contribution linear in $\alpha$ vanishes since the perturbation is quadratic in $\alpha$ for $m = 0$. This was pointed out first by Corinaldesi and Rafeli [44], who also noted that the absence of the $m = 0$ contribution accounts for the difference between (4.17) with $0 < \alpha < 1$ and (4.16) (cf. (4.5)). Like Feinberg [33], they overlooked the non-analyticity near $\alpha = 0$ and the fact that the Born term averages the two analytic functions involved. In view of this fact and the formal character of the Born series for the present case we consider the result (4.16) a success rather than a failure.

**Case $A_R$. The Dynamics $H^k_\alpha$, $R > 0$.** The solution of (4.3) vanishing at $r = R$ is given by

$$\phi^\alpha_m(k, r) = F^\alpha_m(k, R) \{J_{|m+\alpha|}(kr) N_{|m+\alpha|}(kR) - N_{|m+\alpha|}(kr) J_{|m+\alpha|}(kR)\},$$

(4.19)
where \( N_v \) denotes the Neumann function of order \( v \) and where \( F_m^\alpha(k, R) \) is to be determined by the requirement (2.17). Using the asymptotics of \( J_v \) and \( N_v \), we obtain

\[
\varphi_m^\alpha(k, r) \sim (2/\pi)^{1/2} F_m^\alpha(k, R) \left[ \cos \left( kr - \frac{1}{2} |m + \alpha| \pi - \frac{\pi}{4} \right) \right. \\
\left. \cdot \frac{N_{|m + \alpha|}(kR)}{(kr)^{1/2}} - \sin \left( kr - \frac{1}{2} |m + \alpha| \pi - \frac{\pi}{4} \right) \frac{J_{|m + \alpha|}(kR)}{(kr)^{1/2}} \right].
\]  

(4.20)

Some trigonometry now shows that (2.17) is satisfied provided

\[
F_m^\alpha(k, R) = [N_{|m + \alpha|}(kR)^2 + J_{|m + \alpha|}(kR)^2]^{-1/2},
\]  

(4.21)

and that the \( S \)-matrix is given by

\[
e^{2i\delta_m^\alpha(x; a)} = -e^{2i\delta_m^\alpha(a)} H_{|m + \alpha|}^{(2)}(kR)/H_{|m + \alpha|}^{(1)}(kR),
\]  

(4.22)

where

\[
H_{\nu}^{(1)}(x) = J_\nu(x) \pm i N_\nu(x)
\]  

(4.23)

are the Hankel functions and where \( \delta_m^\alpha(a) \) is the phase shift of case \( A_0 \), given by (4.5). The corresponding scattering amplitude will be studied in more detail in the next section.

As in case \( A_0 \), we shall prove in the Appendix that the wave operators

\[
W^{\pm, a, m}_k = \lim_{t \to \pm \infty} e^{iH_{|m + \alpha|}^{\pm} t} P_R \mathcal{F}_m e^{-iH_0 t}
\]  

(4.24)

exist and are unitary maps from \( \mathcal{H}_k \) onto \( \mathcal{H}_R^a \) (cf. Theorem A2). Explicitly, they satisfy

\[
(W^{\pm, a, m}_k F)(r) = i^{m_l} \int_0^{2\pi} dk \varphi_m^\alpha(k, r) e^{\pm i\delta_m^\alpha(k, \alpha)} F(k),
\]  

(4.25)

so that we also obtain (4.22) for the \( S \)-matrix of time-dependent scattering theory.

**Case \( B_0 \). The Dynamics \( \bar{W}^{\pm, a}_k \).** In this case we see no convincing time-independent way to obtain the scattering. However, Theorem A3 below implies that the wave operators

\[
\bar{W}^{\pm, a}_k = \lim_{t \to \pm \infty} e^{iH_0^+ t} \mathcal{F} e^{-iH_0^+ t}
\]  

(4.26)

exist and are unitary. Since

\[
e^{\pm iH_0^+ t} \mathcal{F} = \mathcal{F} e^{\pm iH_0^+ t}
\]  

(4.27)
by definition, we obtain explicitly from (3.10) and (A31),
\[ \tilde{W}^{0}_{+\alpha} = e^{-i\alpha \theta} \hat{S} e^{i\alpha \theta} \]  
and from (A32),
\[ \tilde{W}^{0}_{-\alpha} = \tilde{W}^{0}_{+\alpha} e^{-i\alpha \varepsilon(\theta)}, \]
where \( \varepsilon(\theta) \) is the function
\[ \varepsilon(\theta) = 1, \quad \theta > 0 \]
\[ = -1, \quad \theta < 0. \]

It follows that the corresponding \( S \)-operator on \( \mathcal{H} \) is given by
\[ \tilde{S}^{0}_{\alpha} = e^{-i\alpha \varepsilon(\theta)}. \]

Thus, \( \tilde{S}^{0}_{\alpha} \) is diagonal in momentum space and therefore leads to a vanishing cross section in non-forward directions. According to the standard interpretation of scattering theory (based on the fact that scattering states as prepared and observed in a laboratory are sharply peaked around a definite momentum) such a scattering operator cannot be observed. (Note that it does not give rise to time delay either, since it is energy-independent.)

**Case B. The Dynamics \( \hat{H}_{\alpha}^{R}, R > 0 \).** As in case B, we only consider the wave operator approach. They are now defined by
\[ \tilde{W}^{R}_{\pm\alpha} = \lim_{t \to \pm \infty} e^{i\hat{R}_{\alpha}^{R} t} \mathcal{F}_{R} e^{-iH_{0} t} \]
(recall \( \mathcal{F}_{R} = P_{R} \mathcal{F} \)). Theorems A3 and A2 below imply they exist and are unitary operators from \( \mathcal{H} \) onto \( \mathcal{H} = \mathcal{H}_{r}^{R} \otimes \mathcal{H}_{\theta} \). Explicitly, they are given by
\[ \tilde{W}^{R}_{+\alpha} = e^{-i\alpha \theta} W^{R}_{+0} e^{i\alpha \theta} \]
and
\[ \tilde{W}^{R}_{-\alpha} = e^{-i\alpha \theta} W^{R}_{-0} e^{i\alpha \theta} e^{-i\alpha \varepsilon(\theta)}, \]
where \( W^{R}_{\pm0} \) are the wave operators of case \( A_{R} \) for \( \alpha = 0 \), given by (4.25) with \( \alpha = 0 \). As a consequence, the \( S \)-matrix satisfies
\[ \tilde{S}^{R}_{\alpha} = e^{-i\alpha \theta} S^{R}_{0} e^{i\alpha \theta} e^{-i\alpha \varepsilon(\theta)}, \]
where \( S^{R}_{0} \) is the \( S \)-matrix of case \( A_{R} \), whose angular momentum \( m \) projections are given by (4.22) with \( \alpha = 0 \). The scattering amplitude for this case (denoted \( \tilde{f}_{\alpha}^{R} \) is
therefore related to the hard-core scattering amplitude (denoted $f_{hc}$ and explicitly given by (2.20) and (4.22) with $\alpha = 0$) by

$$
\tilde{f}_a^R(k, \theta, \theta') = e^{-ia\theta} f_{hc}(k, \theta - \theta') e^{ia\theta'} e^{-iax_i(\theta')}.
$$

(4.36)

The scattering is therefore indistinguishable from hard-core scattering: The differential cross sections are the same. Hence, as in case B, we find that in spite of the lack of rotational symmetry of the dynamics the observable consequences do have rotational symmetry and do not depend on the location of the ray of discontinuity.

5. Experimental Predictions

As we have seen in the Introduction, the last two cases considered in the preceding section correspond to an experimentally realizable physical situation. Let us therefore begin this section by studying the corresponding phase shifts and cross sections in more detail. As we have seen, $H^R_a$ for integer $a$ and $H^R_{a}$ for any $a$ lead to the hard-core cross section

$$
\left(\frac{d\sigma}{d\theta}\right)_{hc}(k, \theta) = \frac{2}{\pi k} \left| \sum_m \frac{J_{|m|}(kR)}{H_{|m|}^{(1)}(kR)} e^{im\theta} \right|^2.
$$

This cross section also describes ($r$ times) the intensity of electromagnetic waves scattered off a cylinder, and has been studied in [46, (pp. 1376–1382)]. For $kR \ll 1$ it reduces to the classical cross section for scattering of point particles off a cylinder,

$$
\left(\frac{d\sigma}{d\theta}\right)_{hc}(k, \theta) \simeq \frac{R}{2} \sin \frac{\theta}{2}, \quad kR \gg 1,
$$

(5.2)

plus an additional contribution concentrated around the forward direction, which describes the diffraction shadow. (Note this is analogous to the well-known hard-sphere scattering in three dimensions.) For $kR \ll 1$, one can use the small-$x$ behavior of $J_{|m|}(x)$ and $N_{|m|}(x)$ to conclude that only the $m = 0$ term gives rise to a noticeable contribution, leading to the isotropic cross section

$$
\left(\frac{d\sigma}{d\theta}\right)_{hc}(k, \theta) \simeq \frac{\pi}{2k} (\ln kR)^{-2}, \quad kR \ll 1.
$$

(5.3)

(Note that the "effective width" of the cylinder diverges for $k \rightarrow 0$, in contrast to the hard-sphere total cross section, which becomes equal to $\pi R^2$ for $kR \ll 1$.) As regards the intermediate regime $kR \sim 1$, the coefficients $J_{|m|}(kR)/H_{|m|}^{(1)}(kR)$ converge rapidly to 0 for $|m| > kR$, which renders a numerical evaluation of the cross section feasible in this regime.
Let us now consider the scattering associated with $H_R^\alpha$ for $\alpha \in \mathbb{Z}$. As we have seen above, the scattering amplitude reads

$$f_R^\alpha(k, \theta) = (2\pi ik)^{-1/2} \sum \left[ -e^{2i\delta_{m\alpha}} H_{im+\alpha}^{(2)}(kR)/H_{im+\alpha}^{(1)}(kR) - 1 \right] e^{im\theta}. \quad (5.4)$$

By relabeling one sees that

$$f_R^\alpha(k, \theta) = (-)^n e^{-in\theta} f_{\alpha-a}^R(k, \theta) + (2\pi ik)^{1/2} \left[ (-)^n - 1 \right] \delta(\theta), \quad n \in \mathbb{Z}, \quad (5.5)$$

so that the (non-forward) cross section,

$$\left( \frac{d\sigma}{d\theta} \right)_\alpha^R(k, \theta) = (2\pi)^{-1} \sum_m \left[ e^{2i\delta_{m\alpha}} H_{im+\alpha}^{(2)}(kR)/H_{im+\alpha}^{(1)}(kR) + 1 \right] e^{im\theta}, \quad (5.6)$$

is periodic in $\alpha$ with period 1. Thus, we may (and will henceforth, to ease the notation) assume $0 < \alpha < 1$. To further study $f_R^\alpha$, it is convenient to introduce a function $f_r$ by writing

$$f_R^\alpha(k, \theta) = f_\alpha(k, \theta) + f_r(k, \theta), \quad (5.7)$$

where $f_\alpha$ is the scattering amplitude for $R = 0$, given by

$$f_\alpha(k, \theta) = (2\pi ik)^{-1/2} \sum (e^{2i\delta_{m\alpha}} - 1) e^{im\theta}. \quad (5.8)$$

Thus the remainder function $f_r$ can be written

$$f_r(k, \theta) = -\left( \frac{2}{\pi ik} \right)^{1/2} \sum e^{2i\delta_{m\alpha}} \frac{J_{im+\alpha}(kR)}{H_{im+\alpha}^{(1)}(kR)} e^{im\theta}. \quad (5.9)$$

From the large-$\nu$ behavior of the Bessel functions it follows that the Fourier coefficients of $f_r$ go to zero for $|m| \to \infty$ faster than any inverse power of $m$. Thus, $f_r$ is infinitely differentiable for $k \neq 0$, in contrast to $f_\alpha$, which is singular in the forward direction and satisfies

$$f_\alpha(k, \theta) = \frac{\sin \pi \alpha}{(2\pi ik)^{1/2}} \frac{e^{-i\theta/2}}{\sin \theta/2}, \quad \theta \neq 0 \quad (5.10)$$

(cf. (4.11)). Note also that the series vanishes uniformly in $\theta$ for $kR \to 0$ in view of the small-$x$ behavior of the Bessel functions.

These considerations have several useful consequences:

(i) The total cross section is infinite in contrast to the hard-core total cross section;

(ii) For small values of $kR$ one gets the differential cross section (4.12),

$$\left( \frac{d\sigma}{d\theta} \right)_\alpha^R(k, \theta) \approx \frac{1}{2\pi k} \frac{\sin^2 \pi \alpha}{\sin^2 \theta/2}, \quad kR \ll 1; \quad (5.11)$$
(iii) In the regime $kR \sim 1$ $f_r$ is well approximated by taking only a small number of Fourier coefficients into account. Thus, $f_r$ and hence $(do/d\theta)^R_\alpha$ can be evaluated numerically in this range.

Clearly, (5.11) is drastically different from the isotropic low-energy hard-core cross section (5.3). Thus, if one could perform the scattering experiment in the regime $kR \ll 1$, one would have a precise quantitative test of the physical reality of the potential $A$, as embodied in $H^R_\alpha$. Unfortunately, this region is presently inaccessible, as we learned from Möllenstedt. However, for $kR \sim 1$ (a regime which seems to be experimentally very difficult, but possibly feasible) it seems plausible that the cross section (5.6) still deviates considerably from the hard-core cross section (5.1).

In contrast, we expect that the regime $kR \gg 1$ leads again to the classical cross section (5.2) with a deviation in a small cone around the forward direction. This expectation is based on two heuristic considerations. First, the limit $k \to \infty$ can be regarded as a classical limit $\hbar \to 0$ in view of the relation $k = p/\hbar$. Second, from (4.22) one concludes that

$$e^{2i\delta_m(k,\alpha)} \sim -i(-)^m e^{-2ikR}, \quad kR \to \infty,$$

just like in the hard-core case $\alpha = 0$, where the expected behavior indeed occurs [46]. For large values of $kR$ it is therefore probably very difficult to distinguish between $H^R_\alpha$ and $\tilde{H}^R_\alpha$.

In both cases we have taken the presence of the cylinder into account by requiring that the wave functions in the domain of the Hamiltonian vanish at the cylinder surface. Physically, this Dirichlet boundary condition seems the obvious one, but from a mathematical standpoint it is not the only possibility. In fact, there is an infinite-parameter ambiguity involved here since each of the channel Hamiltonians acting on smooth wave functions $F(r)$ vanishing for $r \leq R$ admits a one-parameter family of self-adjoint extensions (cf. the Appendix). This state of affairs clearly renders a further analysis cumbersome, so that we will forego this here. Instead, we just point out that if observed cross sections would turn out to disagree both with the hard-core cross section (5.1) and with the cross section (5.6), a Hamiltonian with unorthodox boundary conditions might still lead to the observed results.

We do want to consider briefly the converse question, viz., whether verification of the cross section (5.1)/(5.6) excludes the possibility that the same cross section can be obtained by using $H^R_\alpha/\tilde{H}^R_\alpha$ as defined with anomalous boundary conditions. In view of the infinite-parameter freedom involved it is difficult to give a definitive answer to this "inverse scattering" question. For a fixed value of $k$ it certainly seems likely one can choose the boundary conditions in each channel so that the same phase shifts and hence cross section results. It is even possible that different sets of phase shifts lead to the same cross section. (In three-dimensional central potential scattering such phase-shift ambiguities already occur if one only takes a few partial waves into account [47–48].) However, in both cases it is very unlikely that equality of cross sections persists for other $k$-values. Verification of the cross section (5.6) for several energies would therefore in our opinion constitute a convincing proof that a
correct quantum Hamiltonian must include the vector potential in Aharonov–Bohm-like situations along the lines of cases \( A_0 \) and \( A_R \) (assuming, of course, \( (5.6) \) differs significantly from the hard-core cross section \( (5.1) \) for the energies involved).

We would like to close this section by considering the question what the "hydrodynamical" framework predicts concerning the situation at hand. In this approach the Schrödinger equation is written in terms of the current density \( j \) and the charge density \( \rho \) and the electric and magnetic fields only. Since no potentials occur in this setup, it would seem natural to adopt as the Hamiltonian in the external region the usual free one \( \mathcal{H}_0^R \). However, as first emphasized by Strocchi and Wightman [9], one must supplement the hydrodynamical equations with appropriate boundary conditions if excluded regions are present. For the case of a cylinder these authors state that the requirement that the tangential component of \( \frac{j}{\rho} \) integrated along the cylinder mantle equal the magnetic flux enclosed would produce the Aharonov-Bohm effect. This was further clarified in a recent paper by Casati and Guarneri [10]. They point out that one of the hydrodynamical equations, \( \nabla \times \left( \frac{j}{\rho} \right) = 0 \), admits a solution that is the sum of a gradient and a term \( CA \). Here, \( A \) is the potential \( (3.3) \) and the constant \( C \) is fixed by the boundary condition mentioned above. (Such a solution is admissible since \( A \) is smooth in the external region.) The point is now that this solution corresponds to a wave function that satisfies the Schrödinger equation \textit{with the potential} \( A \text{ included} \), and not the free Schrödinger equation. Thus, in our case the boundary condition translates into adoption of the Hamiltonian \( \mathcal{H}_e^R \), not the free one (up to the ambiguity at \( r = R \) discussed in the preceding paragraphs). Since this boundary condition is moreover obtained as the limit of an identity satisfied by \textit{any} solution in the case of a finite potential wall surrounding the same flux, the hydrodynamical viewpoint most naturally leads to the cross section \( (5.6) \).

6. Gauge Invariance

In cases \( A_0 \) and \( A_R \) above we have defined the dynamics by first making the obvious choice for \( A \), then making the gauge-invariant substitution \( p \rightarrow p - (e/c)A \), and, finally, by defining the operator \( i\partial_\theta \) as in the free case. There is a natural question that now arises: Do the physical consequences depend on the choice of \( A \)? In our context this can be rephrased as: Are the cross sections \( (5.1) \) and \( (5.6) \) gauge-invariant? We shall now answer this question for a large class of gauge transformations.

First, let us consider gauge transformations \( eA/c \rightarrow eA/c - \nabla A \), where \( A \) is such that

\[
\int dx \, dy \left| \exp(iA(x, y)) - 1 \right| < \infty. \tag{6.1}
\]

Thus, crudely speaking, \( e^{iA} \) approaches 1 at infinity. We take the gauge transformation into account by defining a new dynamics,

\[
\mathcal{H}_{\alpha}^R(A) = e^{-iA} \mathcal{H}_{\alpha}^R e^{iA}, \quad R \geq 0. \tag{6.2}
\]
and new wave operators,

\[ W^R_{\pm \alpha}(A) = \lim_{t \to -\infty} e^{iH^R_{\pm}(A)t} P_{\pm} e^{-iH_0t}, \quad R \geq 0. \] (6.3)

It then follows from (A35) below that

\[ W^R_{\pm \alpha}(A) = e^{iA} W^R_{\pm \alpha} \] (6.4)

if (6.1) is satisfied. As a consequence one has

\[ S^R_\alpha(A) \equiv W^R_{+\alpha}(A)* W^R_{-\alpha}(A) = S^R_\alpha. \] (6.5)

Thus, for such \( A \) the \( S \)-matrix is exactly gauge-invariant.

Condition (6.1) excludes gauge transformations that only depend on the polar angle \( \theta \). However, this is actually the most interesting case. Indeed, by picking an appropriate smooth function \( \Lambda(\theta) \) one can obtain a new \( \Lambda \) that is smooth and vanishes everywhere except in a cone with arbitrarily small opening angle. If one allows discontinuous \( \Lambda(\theta) \), one can even gauge \( \Lambda \) away except on a ray (take, e.g., \( \Lambda(\theta) = e^{-i\alpha \theta} \)). Fortunately, this class of gauge transformations can also be handled: Let \( \Lambda(\theta) \) be any (measurable) function defined on \((-\pi, \pi]\), and define new dynamics and wave operators by (6.2) and (6.3). It then follows from (A31) below that

\[ W^R_{+\alpha}(A) = e^{-i\Lambda} W^R_{-\alpha} e^{i\Lambda}. \] (6.6)

and from (A32) that

\[ W^R_{-\alpha}(A) = e^{i\Lambda} W^R_{-\alpha} e^{i\Lambda}. \] (6.7)

Here, \( \Lambda_\pi \) denotes the function \( \Lambda \) "rotated" by \( \pi \) in the "circle picture" of the polar angle. That is,

\[ \Lambda_\pi(\theta) = \Lambda(\theta + \pi), \quad -\pi < \theta < 0 \]
\[ = \Lambda(\theta - \pi), \quad 0 < \theta < \pi. \] (6.8)

Therefore the corresponding \( S \)-matrix satisfies

\[ S^R_\alpha(A) = e^{-i\Lambda} S^R_\alpha e^{i\Lambda}. \] (6.9)

Thus, the \( S \)-matrix is not strictly gauge invariant for such \( A \). However, (6.9) implies the corresponding scattering amplitudes are related by

\[ f_{\Lambda}(\theta, \theta') = e^{-i\Lambda(\theta)} f(\theta - \theta') e^{i\Lambda(\theta')}, \] (6.10)

and consequently the \textit{cross section} is gauge invariant for this class of gauge transformations as well. Thus, once one has opted for taking \( \Lambda \) into account as in cases \( A_0 \) and \( A_R \), the physical consequences are stable under gauge transformations, even transformations that put \( \Lambda \) equal to zero except on a ray.
The considerations above also shed more light on the dynamics of cases $B_0$ and $B_R$. Indeed, comparing (6.2) with (3.10) and (3.13) one concludes that

$$H_0^R(e^{i\alpha \theta}) = \tilde{H}_R^\alpha, \quad R \geq 0. \quad (6.11)$$

The explicit expressions for the wave and scattering operators of cases $B_0$ and $B_R$ are therefore just special cases of (6.6), (6.7) and (6.9). It is important to realize that this implies the dynamics of cases $B_0$ and $B_R$ is not essentially different from the free one. The choice $e^{-i\alpha \theta}$ for the gauge function seems to capture the special features of the potential $A$, and this is presumably another reason this sort of dynamics has been often considered in the literature. However, the above clearly shows that this choice has no other physical consequences than the free dynamics as far as scattering is concerned.

7. Comparison with Ref. [1] and Long-Range Behavior

It is instructive to compare our results for case $A_0$ with those of [1]. In [1] Aharonov and Bohm studied the scattering associated with $H_0^A$, taking apparently $0 < \alpha < 1$. The scattering amplitude they obtained looks different from (4.11), but this is due to another angle convention: The only real difference is that they did not mention the $\delta$-function contribution in the forward direction (which is essential for the unitarity of the $S$-matrix). Their approach appears quite different from ours, too, but again this is not borne out by more careful examination. They consider the asymptotics of an incoming solution, which is of the form (2.9) with $\theta' = 0$, except that the plane wave is multiplied by a factor $e^{-i\alpha \theta}e^{i\pi \delta(\theta)}$ with our angle convention (which is more natural since $\theta$ is really a difference angle if one puts $\theta' = 0$). With their convention this factor looks different: It is given by $e^{-i\alpha \delta}$. (Their $\delta$ is related to our $\theta$ by $\theta = \tilde{\theta} + \pi$ for $0 < \theta < \pi$ and by $\theta = \tilde{\theta} - \pi$ for $-\pi < \theta < 0$.) They motivate the extra factor by taking into account that the current density $j$ has an extra term due to the vector potential $A$. The factor is needed to ensure $j$ is constant and in the $x$-direction. In our time-independent approach this problem did not arise since we only considered the radial equations, which led to the phase shift and scattering amplitude almost without effort (in contrast to the considerable labor of [1]).

At first sight, this seems to imply our "total" incoming wave function $\varphi_-(r, \theta; k, \theta')$, as given by (2.15), must be different from theirs. Indeed, as we sketched in Section 2, the phase shift method takes (2.15) as its starting point, and it would seem one can reverse all steps to conclude the asymptotic behavior of our $\varphi_-$ is given by (2.9), so that its dominant behavior is that of a plane wave. But this is false. Our $\varphi_-$ only differs from that of [1] by the different angle convention mentioned above, so that their calculation implies the dominant asymptotics of our $\varphi_-$ is given by

$$\varphi_-(r, \theta; k, 0) \sim e^{ikr \cos \theta} e^{-i\alpha \theta}e^{i\pi \delta(\theta)}, \quad \theta \neq 0. \quad (7.1)$$
The fallacy in the above argument is that one cannot reverse the steps: The asymptotics of a series cannot be read off from the asymptotics of its terms. For example, the dominant spatial asymptotics of the Fourier coefficients of the function $e^{ikr\cos\theta}e^{-2\sqrt{n}}$ with $n \in \mathbb{Z}$ does not depend on $n$, as can be seen from (2.12) and (2.14). Note that $\phi_-(r, \theta; k, 0)$ for $\alpha = 2n$ is actually equal to the function in this example. This readily follows from the radial Schrödinger equations: Changing $a$ by one unit is equivalent to shifting $m$ by one unit. In this regard the perturbation is quite different from the electric potentials of Section 2, and this dependence on $m$ may be regarded as the mathematical cause of the long-range distortion of the plane wave.

At this point the reader may still feel the behavior (7.1) is nevertheless somewhat unsettling. For one thing, he might object that the representation of the scattering amplitude in terms of phase shifts was derived in Section 2 by assuming (2.9), which turns out to be false in our case. For another, to require a priori that the incoming function satisfy (7.1) seems a little strange; if one throws out discontinuous functions from the domain of the dynamics, one would prefer not to have them return through the back door.

As concerns the first point, we calculated the phase shifts and scattering amplitude without worrying about this, since the procedure is well known and quickly leads to the correct result. (By “correct” we mean here that the result could be confirmed in the less well-known time-dependent framework, which we regard as more trustworthy, since it only deals with the temporal asymptotics of normalizable wave packets.) As regards the second point, in our approach the behavior (7.1) only appears a posteriori as an interesting fact about the kernel of $W_-$, which can be at least partly understood from the considerations in the preceding paragraph, and which we shall soon explain in yet another way. Furthermore, the discontinuous asymptotics of the incoming function does not imply the function itself is discontinuous on the ray $\theta = \pi$. As a matter of fact, the incoming function is infinitely differentiable (except possibly at $kr = 0$). This easily follows from the rapid decrease of $J_n(x)$ as $n \to \infty$ for fixed $x > 0$ and the well-known recursion relations for Bessel functions.

Anomalous long-range behavior is also exhibited by the $S$-matrices of Section 4. As we have indicated in Section 2, the $S$-matrix usually “looks like” the identity operator at large distances from the scattering center, in the sense of (2.33). However, this does not hold for the $S$-matrices $S^0_\alpha$ and $S^R_\alpha$. For $S^0_\alpha$ this is obvious, and in the case of $S^R_\alpha$ it follows from the fact $S^R_\alpha$ does have the usual cluster property. What is more surprising is that $S^0_\alpha$ and $S^R_\alpha$ do not have the usual cluster property either. Indeed, states translated to the far right “see” an $S$-matrix $e^{-ia\pi\xi(\theta)}$, in the sense that

$$
\lim_{a \to \infty} \left( U_a g, S^R_\alpha U_a h \right) = \int_0^\infty dk \int_0^\pi d\theta \tilde{g}(k, \theta) \tilde{h}(k, \theta) e^{-ia\pi\xi(\theta)}, \quad R \geq 0, \quad (7.2)
$$

where $U_a$ is multiplication by $e^{-iak\cos\theta}$. The corresponding result for other directions can be obtained by using rotations. These assertions follow from a more general cluster property proven in the Appendix (Theorem A4).
It is important to note that the asymptotic $S$-matrix occurring in (7.2) is equal to the $S$-matrix $S_0^0$ of case $B_0$ (cf. (4.31)). Intuitively, this can be understood by realizing that at the far right the dynamics $H^{\delta}_0$ looks like $H^{\delta}_B$ since the wave packets no longer feel the cylinder and the difference in the boundary conditions on the ray $\theta = \pi$. This heuristic argument can also be used to tie in the abnormal asymptotics of the incoming and outgoing functions corresponding to $H^{\delta}_B$ with out results on the wave operators $\tilde{W}_{{\pm}0}$ for case $B_0$ (cf. Section 4).

This is most easily seen for the outgoing function (i.e., $2\pi$ times the kernel of $W^{0}_{+0}$), which satisfies

$$\phi_+(r, \theta; k, \theta') \sim e^{-ia\theta} e^{ikr \cos(\theta - \theta')} e^{i\alpha \theta'}, \quad |\theta - \theta'| < \pi,$$

according to the calculation of Aharonov and Bohm [1]. The right-hand side (without the restriction on $|\theta - \theta'|$) is just the kernel of $2\pi \tilde{W}^0_+ \alpha$ (cf. (4.28)). This can be understood as follows: If $g(k, \theta)$ is a wave packet sharply peaked around $(k_0, 0)$, then the wave operators $W^0_{+0}$ and $\tilde{W}^0_+ \alpha$ acting on $e^{-ik2\theta} g(k, \theta)$ lead to wave packets in position space that are at the far right in the distant future. The approximate coincidence of the kernels in this region and the resulting nearly identical propagation just reflect the “loss of knowledge” about the boundary conditions on the ray $\theta = \pi$.

For the incoming function one has instead

$$\phi_-(r, \theta; k, \theta') \sim e^{-ia\theta} e^{ikr \cos(\theta - \theta')} e^{i\alpha \theta'} e^{i\alpha \pi \epsilon(\theta - \theta')}, \quad \theta \neq \theta',$$

as compared to the kernel of $2\pi \tilde{W}^0_{-0}$, which reads

$$e^{-ia\theta} e^{ikr \cos(\theta - \theta')} e^{i\alpha \theta'} e^{-i\alpha \pi \epsilon(\theta')}$$

according to (4.29) (note no restriction on the range of $|\theta - \theta'|$ is needed in (7.4)). Now the kernels coincide approximately when $(k, \theta')$ is in the left half-plane, $(r, \theta)$ in the right one, and $kr \gg 1$. Picking a wave packet sharply peaked around $(k_0, \pi)$ leads through $W^0_{-0}$ and $\tilde{W}^0_{-0}$ to wave packets in position space that were at the far right in the distant past. Their equal evolution for asymptotic times can now be explained by realizing they do not yet see the different boundary conditions on the ray $\theta = \pi$.

In view of the above we consider it plausible that the asymptotic behavior given by (7.3) and (7.4) is generic for situations of this type. We conjecture that it is in particular present in case $A_k$ for $R > 0$.

**APPENDIX**

In this appendix we present a number of results whose proofs involve some functional analysis. For more details on the concepts and language we refer to the books by Reed and Simon [38, 49, 50]. Theorems A1 and A2 establish the connection between the time-independent and time-dependent viewpoints for cases $A_0$ and $A_k$, respectively (cf. Section 4). Theorem A3 has been used to obtain the wave
operators for cases $B_0$ and $B_k$ and to prove invariance under gauge transformations (cf. Sections 4 and 6), and Theorem A4 concerns the cluster properties of a class of $S$-matrices containing in particular $S_0^0$ and $S_k^0$ (cf. Section 7). To facilitate the notation we assume $0 < \alpha < 1$ in this appendix.

First of all, we need to define $H_\alpha$ more precisely. For this, it suffices to define the channel Hamiltonians

$$H_\alpha = -\frac{d^2}{dr^2} - \frac{1}{r} \frac{d}{dr} + \frac{\nu^2}{r^2}$$

as self-adjoint operators on $\mathcal{H}_f \equiv L^2(R^+, r \, dr)$. Well-known results (cf. [50, p. 161]) imply $H_\nu$ is essentially self-adjoint on $C_0^\infty((0, \infty))$ for $|\nu| \geq 1$, while for $|\nu| < 1$ it admits a one-parameter family of self-adjoint extensions. In the case at hand this is easily seen explicitly: Vectors $\varphi_{\nu \pm}$ in the kernel of $(H_{\nu} \upharpoonright C_0^\infty)^* \pm i$ are weak solutions on $(0, \infty)$ to the ODE

$$\left[-\frac{d^2}{dr^2} - \frac{1}{r} \frac{d}{dr} + \frac{\nu^2}{r^2}\right] \varphi_{\nu \pm} = \mp i \varphi_{\nu \pm},$$

and thus by hypo-ellipticity classical solutions. The requirement that $\varphi_{\nu \pm}(r)$ be in $\mathcal{H}_f$ at infinity them implies (up to a constant)

$$\varphi_{\nu \pm}(r) = H_\nu^{(\nu)}(e^{\pm i\nu/4} r),$$

where $H_\nu^{(\nu)}$ are the Hankel functions (4.23). But for $|\nu| \geq 1$ the right-hand side is not in $\mathcal{H}_f$ at 0, while it is in $\mathcal{H}_f$ for $|\nu| < 1$. From this the above statement follows.

Fortunately, there is a natural choice for a self-adjoint extension of $H_\nu$ for $|\nu| < 1$. Indeed, the extensions are parameterized by the number $\chi \in [0, 2\pi)$ in the relation

$$H_\nu(\varphi_{\nu +} + e^{i\chi} \varphi_{\nu -}) \equiv -i \varphi_{\nu +} + ie^{i\chi} \varphi_{\nu -},$$

where $\varphi_{\nu \pm}$ are given by (A3). The requirement that $\varphi_{\nu +} + e^{i\chi} \varphi_{\nu -}$ vanish at the origin is easily seen to fix $\chi$ uniquely. The self-adjoint operators thus defined will be denoted again by $H_\nu$. We are now in a position to state Theorem A1, which establishes the analog of the relations (2.27) for our “potentials.” Note that for $|m + \alpha| < 1$ this hinges on our choice for $H_{m + \alpha}$, which corresponds to the “obvious” decision made in Section 4 (below (4.3)) to discard solutions singular at the origin.

**Theorem A1.** The wave operators

$$W_{\pm \alpha, m} = \text{lim}_{\nu \to \pm \alpha} e^{iH_{m + \alpha} t} \mathcal{F}^{-1} e^{-ik^2 t}$$

exist and are isometric maps from $\mathcal{H}_k$ onto $\mathcal{H}_f$. Explicitly, they are given by

$$(W_{\pm \alpha, m} F)(r) = i^{|\alpha|} \text{i.m.} \int_{N=0}^{N} dk \, k J_{|m + \alpha|}(kr) e^{\pm i\delta_m(n)} F(k), \quad \forall F \in \mathcal{H}_k,$$
where

\[ \delta_m(\alpha) \equiv \frac{\pi}{2} (|m| - |m + \alpha|). \quad (A.7) \]

**Proof.** In the proof of this theorem and the next one we shall use some ideas that appeared before in a paper by the author and Bongaarts [51], where analogous problems were considered. We shall only consider \( W_+ \), the proof for \( W_- \) being similar. Also, the case \( \alpha = 0 \) is trivial, so we take \( 0 < \alpha < 1 \) from now on.

First, let \( F \in \mathcal{H}_k \) have support in \( (\varepsilon, N) \) and define an operator \( U_{a,m} \) by setting

\[ (U_{a,m} F)(r) = \int_{\varepsilon}^{N} dk \, k J_{|m+a|}(kr) e^{i\delta_m(\alpha)} F(k). \quad (A.8) \]

Clearly, the integral at the right-hand side is absolutely convergent and defines a function that is continuous and \( O(r^{1+|m+a|}) \) near the origin. Furthermore, \( (U_{a,m} F)(r) \) is in \( \mathcal{H}_r \) and satisfies a bound

\[ \| U_{a,m} F \| \leq C_{N,\varepsilon} \| F \|, \quad (A.9) \]

where \( C_{N,\varepsilon} \) only depends on \( N \) and \( \varepsilon \). To see this, write

\[ J_{|m|}(kr) = (2/\pi)^{1/2} \cos \left( kr - \frac{1}{2} |m| \pi - \frac{\pi}{4} \right) (kr)^{1/2} + R_{|m|}(kr), \quad (A.10) \]

and note that the first term gives rise to a bounded operator from \( \mathcal{H}_k \) to \( \mathcal{H}_r \), while the remainder function is bounded and \( O((kr)^{-3/2}) \) at infinity. We now claim that \( U_{a,m} F \) is actually in the domain of \( H_{m+a} \) and satisfies

\[ H_{m+a} U_{a,m} F = U_{a,m} k^2 F. \quad (A.11) \]

To prove this, let \( G(r) \in C^{\infty}_0((0, \infty)) \). By virtue of Fubini's theorem we may then write

\[ (H_{m+a} G, U_{a,m} F) = i^{|m|} e^{i\delta_m(\alpha)} \int_{\varepsilon}^{N} dk \, k F(k) \]

\[ \cdot \left[ (m+a)^2 \int_{0}^{\infty} \frac{dr}{r} J_{|m+a|}(kr) \bar{G}(r) \right. \]

\[ - \int_{0}^{\infty} dr \, r J_{|m+a|}(kr) \left( \frac{d^2}{dr^2} + \frac{1}{r} \frac{d}{dr} \right) \bar{G}(r) \]. \quad (A.12) \]

Integrating the second term in brackets by parts we conclude that

\[ (H_{m+a} G, U_{a,m} F) = (G, U_{a,m} k^2 F). \quad (A.13) \]
But for \( m > 0 \) and \( m < -1 \), \( C^\infty_0 \) is a core for \( H_{m+1} \), so that (A11) follows for these \( m \)-values. To get a core for \( H_{m} \), \( |v| < 1 \), we can supplement \( C^\infty_0 \) with

\[
G_\nu \equiv \varphi_\nu^+ + e^{i\chi_\nu} \varphi_\nu^- ,
\]

where \( \chi_\nu \) is the \( \chi \)-value corresponding to \( H_{-\nu} \). It is readily seen that (A12) and (A13) also hold for \( m = -1 \) and 0 if \( G \) is replaced by \( G_{-1} \) and \( G_{1} \), resp., so that (A11) also holds true for \( m = -1, 0 \).

We now iterate (A11) and use an analytic vector argument based on (A9) to conclude that

\[
e^{-iH_{m+1}t} U_{-\alpha, m} F = U_{-\alpha, m} e^{-ik^2 t} F .
\]

Therefore we can write, assuming from now on \( F \in C^\infty_0 ((\epsilon, N)) \),

\[
\| e^{iH_{m+1}t} F - U_{-\alpha, m} F \|_2^2 = \| (F_m - U_{-\alpha, m}) e^{-ik^2 t} F \|_2^2
\]

\[
= \int_0^\infty dr r \left| \int_\epsilon^N dk \, k e^{-ik^2 t} F(k) \left[ J_{|m|}(kr) - J_{m+\alpha}(kr) e^{i\beta_m(a)} \right] \right|^2 . \tag{A16}
\]

Splitting the \( r \)-integral into \( \int_0^1 dr + \int_1^\infty dr \) it follows from the Riemann–Lebesgue lemma and dominated convergence that the first integral vanishes for \( |t| \to \infty \). The second one can be written

\[
\int_0^\infty dr r \left| \int_\epsilon^N dk \, k e^{-ik^2 t} F(k) \left[ C e^{ikr} / (kr)^{1/2} + O((kr)^{-3/2}) \right] \right|^2
\]

\[
\leq 2 |C|^2 \int_1^\infty dr \left| \int_\epsilon^N dk \, k^{1/2} F(k) e^{-ik^2 t + ikr} \right|^2
\]

\[
+ \int_1^\infty dr \, r^{-2} \left| \int_\epsilon^N dk \, k^{-1/2} e^{-ik^2 t} F(k) R(kr) \right|^2 , \tag{A17}
\]

where \( C \) is a constant and \( R \) is bounded on the integration region. Now for \( |t| \to \infty \) the second term at the right-hand side vanishes by the Riemann–Lebesgue lemma and dominated convergence. Replacing the phase factor in the first one by

\[
(-2ikt + ir)^{-1} \partial_k e^{-ik^2 t + ikr} , \tag{A18}
\]

integrating by parts and estimating in the obvious way, it follows by dominated convergence that this term vanishes as well, provided \( t \to -\infty \).

Since \( C^\infty_0 ((0, \infty)) \) is dense, it follows that the wave operator \( W_{-\alpha, m} \) exists and equals the operator \( U_{-\alpha, m} \). As \( W_{-\alpha, m} \) is isometric, it remains to prove its range is \( \mathcal{H}_r \). For this we need only show its adjoint is isometric. But this follows from the fact that \((-i)^{|m|} e^{-i\beta_m(a)} U_{-\alpha, m} \), viewed as an operator from \( \mathcal{H}_r \) to \( \mathcal{H}_r \) (by identifying \( \mathcal{H}_k \) with \( \mathcal{H}_r \) in the obvious way) is self-adjoint.
The next theorem concerns the scattering associated with the channel Hamiltonians $H^R_v$ on $\mathcal{H}^R_v \equiv L^2([R, \infty), r \, dr)$, given formally by the right-hand side of (A1). In this case these differential operators are not essentially self-adjoint on $C_0^\infty((R, \infty))$ for any $v$, since the functions $\varphi_{v, \pm}(r)$ (cf. (A3)) are in $\mathcal{H}^R_v$ for any real $v$ if $R > 0$. Another way to see this is to note that the potential term is bounded on $\mathcal{H}^R_v$, so that we need only consider the differential operator $-(d/dr)^2 - r^{-1}d/dr + 1/4r^2$. But this operator transforms into $-(d/dr)^2$ under the unitary map $U: \mathcal{H}^R_v \rightarrow L^2([R, \infty), dr)$, given by

$$ (UF)(r) \equiv r^{1/2}F(r), \quad (A19) $$

and $-(d/dr)^2$ is not essentially self-adjoint on $C_0^\infty((R, \infty))$ since the functions $\exp(-\rho^{-1}r)$ are in $L^2([R, \infty), dr)$. However, $-(d/dr)^2$ is essentially self-adjoint on the subspace

$$ D \equiv \{ F \in C^\infty([R, \infty)) \mid F(R) = 0 \text{ and supp } F \text{ compact} \}, \quad (A20) $$

as is readily seen (and well known). Since $U^{-1}D = D$, this also holds for the differential operators $H^R_v$. We designate the self-adjoint closures by the same symbol from now on. We are now prepared to state Theorem A2.

**Theorem A2.** The wave operators

$$ W^R_{\pm, a, m} = s \cdot \lim_{t \to \pm \infty} e^{iH^R_{\alpha, a}t} P_{R, a} e^{-ik^2t} \quad (A21) $$

exist and are isometric maps from $\mathcal{H}_k$ onto $\mathcal{H}^R_v$. Explicitly, they are given by

$$ (W^R_{\pm, a, m}F)(r) = i^{\left\lvert m \right\rvert} \lim_{N \to \infty} \int_{0}^{N} dk \varphi^a_m(k, r) e^{\pm i \delta_m(k, \alpha)} F(k), \quad \forall F \in \mathcal{H}_k. \quad (A22) $$

Here, $\varphi^a_m$ and the phase shift are given by (4.19), (4.21) and (4.22).

**Proof.** We consider only $W^R_{-a, m}$. Proceeding as in the proof of Theorem A1, we define an operator $U^R_{-a, m}$ through the right-hand side of (A22), with $F$ having support in $(\epsilon, N)$. Using (A10) and its analog for $N_{\left\lvert v \right\rvert}(kr)$ one then infers that the analog of (A9) holds true. Also, taking $G$ in the analog of (A12) as an element of the core $D$, one obtains the analog of (A11). (Indeed, the integration by parts again does not lead to boundary terms since both the kernel and $G$ vanish at the lower limit $r = R$.) It then follows as before that the analog of (A15) holds true. Using this and the asymptotics of $\varphi^a_m$ we can write

$$ \left\lVert e^{iH^R_{-a, a}t} P_{R, a} e^{-ik^2t} - U^R_{-a, m}F \right\rVert^2 = \int_{R}^{\infty} dr \left\lvert \int_{0}^{N} dk ke^{-i\delta_m(k, \alpha)} F(k) \cdot \left[ C_1(kR) e^{ikr/kr^{1/2}} + C_2(kR) O((kr)^{-3/2}) \right] \right\rvert^2, \quad (A23) $$
where $C_1$ and $C_2$ are $C^\infty$-functions. From this it follows as before that the wave operator $W^{R}_{-\alpha, m}$ exists and satisfies (A22).

It remains to show it is isometric and has range $\mathcal{R}_r$. Its isometry follows by taking $F \in \mathcal{H}_k$ with compact support and noting that

$$\|W^{R}_{-\alpha, m}F\| = \lim_{t \to -\infty} \|P_{R}\mathcal{S}_m e^{-ik^2 t}F\| = \lim_{t \to -\infty} \|\mathcal{S}_m e^{-ik^2 t}F\| = \|F\|. \quad (A24)$$

(Indeed, the integral $\int_0^\infty dr \left| (\mathcal{S}_m e^{-ik^2 t}F)(r) \right|^2$ vanishes for $|t| \to \infty$ by virtue of the Riemann–Lebesgue lemma and dominated convergence.) To show its range is $\mathcal{R}_r$, it suffices to show its adjoint is isometric. For this we need only prove that

$$F(r) = \text{l.i.m.} \int_{N=0}^{N=\infty} dk \left[ J_\nu^2(kR) + N_\nu^2(kR) \right]^{-1} D_\nu(kR, kr) \cdot \int_{r}^{\infty} dr' r' D_\nu(kR, kr') F(r'), \quad (A25)$$

where $F \in C^\infty_0((R, \infty))$ and $\nu \geq 0$, and where we have set

$$D_\nu(x, y) \equiv N_\nu(x) J_\nu(y) - J_\nu(x) N_\nu(y). \quad (A26)$$

This relation follows from Weber and Orr’s formula [52, p. 74, (66)], but for completeness we sketch a proof: Using the fact that the Wronskian of $J_\nu(z)$ and $N_\nu(z)$ equals $2/\pi z$, it readily follows that the resolvent

$$R_E \equiv (E - H_v)^{-1}, \quad |\text{Im} \ E| > 0, \quad (A27)$$

is given by

$$\left( R_E F \right)(r) = \pi \left[ 2H_\nu^{(1/2)}(E^{1/2} R) \right]^{-1} \left[ D_\nu(E^{1/2} R, E^{1/2} r) \cdot \int_{r}^{\infty} dr' r' H_\nu^{(1/2)}(E^{1/2} r') F(r') + H_\nu^{(1/2)}(E^{1/2} r) \right] \cdot \int_{r}^{\infty} dr' r' D_\nu(E^{1/2} R, E^{1/2} r') F(r'), \quad \text{Im} \ E \geq 0. \quad (A28)$$

Now recall Stone’s formula for the spectral projection of $H_v$ on $[a, b]$ (note $H_v$ has no bound states),

$$P_{[a, b]} F = s \cdot \lim_{\delta \to 0} F_\delta, \quad (A29)$$

where

$$F_\delta \equiv (2\pi)^{-1} \int_a^b dE (R_{E-i\delta} - R_{E+i\delta})F. \quad (A30)$$
Using dominated convergence one concludes from this by a straightforward
calculation that \( P_{[a,b]} = 0 \) for \([a,b] \subset (-\infty, 0)\), while \( (P_{[1/N,N]} f)(r) \) is a.e. equal to
the function defined by the integral at the right-hand side of (A25). Thus, (A25)
follows, so the theorem is proven. \( \square \)

The following theorem concerns properties of two-dimensional Fourier transform-
ation that we have used above to find wave operators and prove gauge invariance.
For notation used, cf. Section 2.

**Theorem A3.** Let \( \Lambda(\theta) \) be a measurable real-valued function on \((-\pi, \pi]\). Then
one has for any \( g \in \mathcal{H}_0 \)
\[
\lim_{t \to \infty} \|(e^{i\Lambda(\cdot)} \mathcal{F} - \mathcal{F} e^{i\Lambda(\cdot)}) e^{-ik^2 t} g\| = 0, \tag{A31}
\]
and
\[
\lim_{t \to -\infty} \|(e^{i\Lambda(\cdot)} \mathcal{F} - \mathcal{F} e^{i\Lambda(\cdot)}) e^{-ik^2 t} g\| = 0. \tag{A32}
\]
Here,
\[
\Lambda_\ast(\theta) = \Lambda(\theta + \pi) \quad (\text{mod } 2\pi). \tag{A33}
\]

Let \( \Lambda(r, \theta) \) be a measurable real-valued function on \( \mathbb{R}^+ \times (-\pi, \pi] \) with the property
\[
\int_0^\infty dr \int_{-\pi}^\pi d\theta |e^{i\Lambda(r, \theta)} - 1| < \infty, \tag{A34}
\]
Then one has for any \( g \in \mathcal{H}_0 \)
\[
\lim_{|t| \to \infty} \|(e^{i\Lambda(\cdot)} \mathcal{F} - \mathcal{F} e^{i\Lambda(\cdot)}) e^{-ik^2 t} g\| = 0. \tag{A35}
\]

**Proof.** It suffices to show (A31) for a total set. We may therefore take
\( g \in P_{m,0} \mathcal{H}_0 \) and \( g_m(k) \in C_0^\infty((0, \infty)) \). Since \( e^{i\Lambda(\theta)} \) is square-integrable over \((-\pi, \pi]\),
one has
\[
e^{i\Lambda(\theta)} = \sum a_n e^{in\theta}, \tag{A36}
\]
where \( \{a_n\} \in l_2 \). Hence,
\[
\|(e^{i\Lambda(\cdot)} \mathcal{F} - \mathcal{F} e^{i\Lambda(\cdot)}) e^{-ik^2 t} g\|^2 = \sum_n \|P_n(e^{i\Lambda(\cdot)} \mathcal{F} - \mathcal{F} e^{i\Lambda(\cdot)}) e^{-ik^2 t} P_{m,0} g\|^2
\]
\[
= \sum_n |a_{n-m}|^2 \|(\mathcal{F}_m - \mathcal{F}_n) e^{-ik^2 t} g_m\|^2, \tag{A37}
\]
since, e.g.,

\[ P_n e^{iA(\theta)} P_m = a_{n-m} e^{i(n-m)\theta} P_m. \]  

(A38)

Arguing now in the same way as after (A16), we conclude that the norm at the right-hand side of (A37) vanishes for \( t \to \infty \). But \( \{|a_n|^2\} \in l_1 \), so that (A31) then follows by dominated convergence.

To prove (A32) we observe that

\[ e^{iA(\theta)} = \sum a_n e^{in(\theta + \pi)} = \sum a_n(-)^n e^{in\theta}. \]

(A39)

Thus, replacing the second \( A \) at the left-hand side of (A37) by \( A_\pi \) entails that the term at the right-hand side containing \( \mathcal{F}_\pi \) gets an additional factor \( (-)^{n-m} \). This factor ensures that the norm vanishes for \( t \to -\infty \) by the same reasoning as above, so that (A32) follows, too.

To prove (A35), let \( g \) have compact support. Then we may write

\[ (e^{iA(\cdot)} - 1) \mathcal{F} e^{-ik^2t} g = C e^{-ik^2 t} h, \]

(A40)

where

\[ C \equiv (e^{iA(\cdot)} - 1) \mathcal{F} (k^2 + 1)^{-1} \]

and

\[ h \equiv (k^2 + 1) g. \]

(A41)

(A42)

Now \( e^{iA(\cdot)} - 1 \) is bounded and in \( L^1 \) by assumption, and hence in \( L^2 \). Since \( (k^2 + 1)^{-1} \) is in \( L^2 \), \( C \) is a Hilbert–Schmidt operator. But \( e^{-ik^2 t} h \) weakly converges to zero for \( |t| \to \infty \) by the Riemann–Lebesgue lemma, so that \( Ce^{-ik^2 t} h \) strongly converges to zero, implying (A35).

We should like to point out that (A35) and its proof have an obvious generalization to higher dimensions, leading in particular to invariance of the \( S \)-operator in three dimensions under gauge transformations vanishing at infinity in the sense that \( e^{iA(x)} - 1 \) is in \( L^1(R^3, dx) \). It would be of interest to establish whether three-dimensional analogs of (A31) and (A32) exist as well. If so, this would lead to invariance of three-dimensional differential cross sections under angular gauge transformations.

Our last theorem concerns the cluster property for a class of \( S \)-operators containing the operators \( S_\alpha^0 \) and \( S_\alpha^R \) of Section 4. To state it, we introduce the multiplication operator

\[ (S_{(\cos, \sin)} g)(k, \theta) \equiv e^{-i\pi \alpha} g(k, \theta), \theta \equiv \theta \mod \pi \]  

(A43)

Thus, \( S_\alpha \) is multiplication by \( e^{-i\pi \alpha} \) in one half-plane bounded by the line \( \lambda e, \lambda \in R \), and multiplication by \( e^{i\pi \alpha} \) in the other half-plane.
**Theorem A4.** Let $S$ be a unitary operator on $\mathcal{H}_0$ such that

$$(g, Sh) = \int_0^\infty \int_{-\pi}^\pi dk \, d\theta \, g(k, \theta) \, h(k, \theta') [s_\alpha(\theta - \theta') + s_\alpha(k, \theta, \theta')]$$

where

$$s_\alpha(\theta) = -\frac{1}{\pi} \int_0^\infty \frac{P}{e^{it} - 1} \sin \pi a \, \delta(\theta) \cos \pi a + \frac{\sin \pi a}{\pi} \frac{P}{e^{it} - 1},$$

and where $s_\alpha$ is a function satisfying

$$|s_\alpha(k, \theta, \theta')| \leq C_N < \infty, \quad \forall (k, \theta, \theta') \in [0, N] \times (-\pi, \pi]^2.$$  

Then one has

$$s \cdot \lim_{\lambda \to \infty} U_{(\lambda \cos \gamma, \lambda \sin \gamma)}^\pi S U_{(\lambda \cos \gamma, \lambda \sin \gamma)} = S_{(\cos \gamma, \sin \gamma)},$$

where $U_\gamma$ is the translation operator (2.32).

**Proof.** By using a rotation over $\gamma$ one concludes that it suffices to prove (A47) for $\gamma = 0$. Also, since the operators involved are unitary, we need only show that (A47) holds weakly on a dense set. Thus, noting that $S_{(1, 0)}$ is multiplication by $e^{-i\alpha \pi \delta(\theta)}$, it readily follows from (A45) and (A46) that we need only prove the relations

$$\lim_{\lambda \to \infty} \int_0^\infty \int_{-\pi}^\pi dk \, d\theta \, d\theta' \, B(k, \theta, \theta') \exp[i\lambda k(\cos \theta - \cos \theta')] = 0,$$

where $B$ is bounded on the integration region, and

$$\lim_{\lambda \to \infty} \int_0^\infty \int_{-\pi}^\pi dk \, d\theta \, d\theta' \, (e^{i(\theta - \theta')} - 1)^{-1} \, \bar{g}(k, \theta) \, h(k, \theta') \cdot \exp[i\lambda k(\cos \theta - \cos \theta')]$$

$$= -\pi \int_0^\infty \int_{-\pi}^\pi dk \, d\theta \, c(\theta) \, \bar{g}(k, \theta) \, h(k, \theta),$$

where $g$ and $h$ belong to a dense set of $C_0^\infty$-functions we shall presently specify.

To prove (A48), we note that by Fubini's theorem we may do the $k$-integration first. By virtue of the Riemann–Lebesgue lemma the resulting function of $\lambda$, $\theta$ and $\theta'$ converges to zero for $\lambda \to \infty$ provided $\cos \theta \neq \cos \theta'$. Since the exceptional set has measure zero, (A48) now follows by dominated convergence.

The integral at the left-hand side of (A49) is similar to a class of integrals studied in [37, Chap. 5, Lemmas 5.1 and 5.2]. To follow the reasoning there, we assume from now on that $g$ and $h$ have support in the interior of a quadrant. If the quadrants are different, the integral is of the form (A48), so that (A49) holds true. Let us therefore consider the case that the supports are in the same quadrant, e.g., the first one. We
may replace \((e^{i(\theta - \theta')} - 1)^{-1}\) by \(-i(\theta - \theta')^{-1}\) since the difference of these two terms gives rise to an integrand of the type occurring in (A48). After a change of variables the resulting integral can then be written

\[
\int \frac{dk}{k} \int \frac{dq}{q} F(k, p + \frac{1}{2}q, p - \frac{1}{2}q) \exp(ik|\omega(p + \frac{1}{2}q) - \omega(p - \frac{1}{2}q)|),
\]

(A50)

where \(F\) is a \(C^0\)-function satisfying

\[
\text{supp } F(k, \theta, \theta') \subseteq \left(\frac{1}{N}, N\right) \times \left(\frac{\pi}{2} - \epsilon\right)^2
\]

(A51)

for suitable positive numbers \(N\) and \(\epsilon\), and where we have introduced \(F\) and \(\omega(x) \equiv \cos x\) to facilitate comparison with [37]. The point is that \(\omega''(p + s)\) is bounded away from zero for \(|s| \leq \frac{1}{2}|q|\) on \(\text{supp } F\), so that the argument in the proof of Lemma 5.1 of [37] leads to the conclusion the \(q\)-integration may be restricted to the region \(|q| < \|k\|^\delta\), where \(-1 < \delta < 0\). We can then argue as in the proof of Lemma 5.2 of [37] to infer that (A49) holds true.

ACKNOWLEDGMENTS

We are indebted to E. Seiler, who contributed much in the initial stages of this work. It is a pleasure to thank him and other colleagues at the Max-Planck-Institut for useful criticism. We would also like to thank G. Möllenstedt for an enlightening conversation on experimental aspects.

Note added in proof: The Tübingen experiment referred to in the Introduction has been performed, meanwhile. The results are reported in [57].

REFERENCES