Broken Unitarity and Phase Measurements in Aharonov-Bohm Interferometers

O. Entin-Wohlman,[1] A. Aharony,[1] Y. Imry,[2] Y. Levinson,[2] and A. Schiller[3]

[1]*School of Physics and Astronomy, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University,
Tel Aviv 69978, Israel*
[2]*Department of Condensed Matter Physics, The Weizmann Institute of Science, Rehovot, 76100 Israel*
[3]*Racah Institute of Physics, The Hebrew University, Jerusalem 91904, Israel*

Aharonov-Bohm mesoscopic solid-state interferometers yield a conductance which contains a term $\cos(\phi + \beta)$, where ϕ relates to the magnetic flux. Experiments with a quantum dot on one of the interfering paths aim to relate β to the dot's intrinsic Friedel transmission phase α_1. For closed systems, which conserve the electron current (unitarity), the Onsager relation requires that $\beta = 0$ or π. For open systems, we show that in general β depends on the details of the broken unitarity. Although it gives information on the resonances of the dot, β is generally not equal to α_1. A direct relation between β and α_1 requires specific ways of opening the system, which are discussed.

The wave nature of an electron is reflected, e.g., by the complex amplitude of the wave transmitted through a quantum scatterer. Under appropriate conditions (discussed below), Aharonov-Bohm (AB) [1] interferometers may be regarded as analogs of the double-slit experiment [2,3], in which the transmission through two paths is $T = |t_{12}|^2$, with

$$t_{12} = t_1 + t_2 e^{i\phi}, \tag{1}$$

where $\phi = e\Phi/\hbar c$, Φ being the magnetic flux enclosed by the two paths. The path amplitudes $t_i = |t_i|e^{i\alpha_i}$ may contain the effects of obstacles [4], e.g., a quantum dot (QD) on path 1, whose nontrivial (gate voltage dependent) transmission phase α_1 can be influenced by electronic correlations [5,6]. Assuming the two-slit formula, Eq. (1), the Landauer conductance [7] through the interferometer $G = (e^2/h)T$ then includes the term [8] $\cos(\alpha_2 - \alpha_1 + \phi)$, which is sensitive to the phase difference. However, in "closed" or *"unitary"* interferometers (inside which the electron number is conserved), time-reversal symmetry implies the Onsager relation [9], $G(\Phi) = G(-\Phi)$. This relation holds for both finite and infinite systems. Hence, T must depend on ϕ via $\cos\phi$, with *no phase shift*. Here we show that *broken unitarity* does yield a term $\cos(\phi + \beta)$, where β depends in general on the rate and on the details of the electron loss. The universally assumed equality $\beta = \alpha_2 - \alpha_1$ requires special ways of opening the system, which we discuss below (in the context of some of the experiments [5,6,10,11]). Specifically, we present an exact example in which this relation does not hold, and then discuss possible conditions under which it might hold.

We consider solid-state interferometers, with narrow waveguides for the electron paths, restricted to the mesoscopic scale in order to retain the coherence of the conduction electrons [12]. AB oscillations in $G(\Phi)$ (in spite of strong impurity scattering), first suggested in Ref. [13], were subsequently observed on metallic closed systems

[14] and in semiconducting samples containing QDs near Coulomb blockade (CB) resonances [4,15]. In these experiments $G(\Phi) = G(-\Phi)$, as required by the Onsager symmetry. Further experiments [5,6,10,11] used *open systems,* in which electrons are lost via additional channels which leave the interferometer, to obtain a nonzero phase shift β. Assuming that $\beta = \alpha_2 - \alpha_1$, some of the surprising experimental results were inconsistent with the theoretical expectations for the phase α_1 of the intrinsic transmission through the QD [16–19]. Examples include the phase lapse between consecutive CB resonances [10,11] and the nonuniversal phase shifts at the Kondo resonances [5,6].

While this paper solves specific theoretical models, the results can be cast in terms of the various energy scales
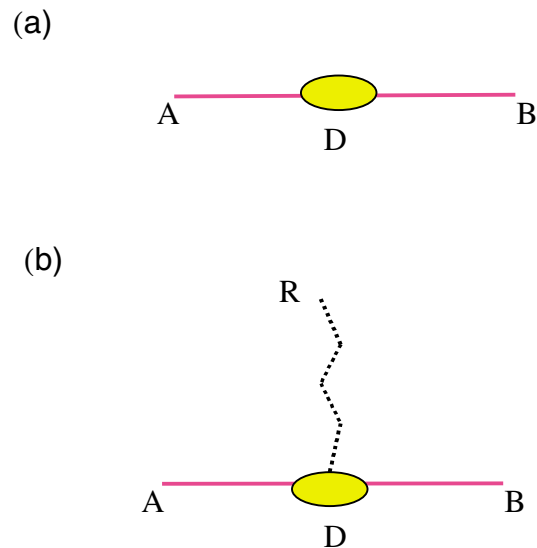
(a)

(b)

FIG. 1 (color). Simple model for a QD (denoted by D) connected to one-dimensional leads. (a) A closed system, with no electron losses. (b) An open system, with a third lead, which connects the QD to a fully absorbing reservoir R.

(e.g., decay widths) characterizing the system. Thus they are much more general than the models employed. Below we expound the underlying model-independent physical principles behind these results.

We first consider a single path, and then connect two paths into an AB interferometer. The QD transmission is typically [16,17] defined by the geometry in Fig. 1a: a dot D is placed on a one-dimensional conductor (described below by a tight-binding model), which models the narrow electronic waveguides ("leads"). An electron wave with amplitude 1 coming from A (or B) generates a transmitted wave with amplitude t_1 (or t_1'), and a reflected wave with amplitude r_1 (or r_1'). This is described by the 2×2 scattering matrix, $S_2 = \begin{pmatrix} r_1 & t_1' \\ t_1 & t_1' \end{pmatrix}$, mapping the two-component vector of incoming amplitudes onto those of outgoing ones. Unitarity implies that the determinant of S_2 is $r_1 r_1' - t_1 t_1' = e^{2i\alpha_1}$, and α_1 is defined (for the specific geometry of Fig. 1a) as the *intrinsic Friedel phase* [16,17] of the QD. At zero temperature, and for electrons on the Fermi surface, α_1 is equal to the phase of the Green function on the QD [17],

$$G_D^{(a)} = 1/[\epsilon_q - \epsilon_0 + e^{iqa} X_{LR}/J], \qquad (2)$$

where $X_{LR} = J_L^2 + J_R^2$, J_L, J_R represent the quantum hopping into D from left or right, J is the (tight binding) hopping between neighboring sites on the leads (with lattice constant a), $\epsilon_q = -2J \cos(qa)$ is the energy (taken equal to the Fermi energy) of an electron with wave vector q, and ϵ_0 denotes the potential energy on the dot, determined by the gate voltage. The real parameters J_L, J_R, and $(\epsilon_q - \epsilon_0)$ may be renormalized by the Coulomb interactions on the dot, so that α_1 contains the effects of the interactions [20]. Whenever the gate voltage yields a resonance, i.e., when the real part of the denominator changes sign, α_1 increases by π. The width of this jump, given by the imaginary part of $[G_D^{(a)}]^{-1}$, is determined by X_{LR}, $\Gamma_R = \sin(qa) X_{LR}/J$.

A particular way to break unitarity between A and B is described in Fig. 1b: a third lead connects the QD to an absorbing electron reservoir R [21] (i.e., with a chemical potential which is slightly lower than that on the emitting source, similar to that of the absorbing sink). This QD is described by a unitary 3×3 scattering matrix S_3, related to the leads from D to A, B and R. However, the 2×2 matrix S_2, which is now a submatrix of S_3, need not be unitary. An explicit calculation with such a hopping Hamiltonian yields that the transmission phase now becomes $\overline{\alpha_1}$, equal to the phase of the renormalized Green function, $G_D^{(b)} = 1/[(G_D^{(a)})^{-1} - \Sigma]$, where the complex self-energy Σ depends on details of the absorbing lead. In the simplest case where D is connected to R by the hopping amplitude V_1, we have $\Sigma = -(V_1^2/J)e^{iqa}$. In particular, its imaginary part, which is proportional to the rate of electron losses through that lead, contributes to the total width of the resonance. Thus, the phase $\overline{\alpha_1}$ measured in this case is in general *not* the intrinsic transmission phase of

the QD, α_1. In fact, for $V_1^2 \gg X_{LR}$ this contribution of the imaginary self-energy will be larger than the intrinsic one. It is only when $V_1^2 \ll X_{LR}$ that $\overline{\alpha_1} \approx \alpha_1$. This distinction is similar to the one obtained in the usual two-slit diffraction experiment [2] in the following circumstance: inserting an isotropic resonance scatterer in the upper slit causes the upper beam to acquire an additional phase shift. Connecting the source, the scatterer and the screen via a narrow waveguide produces qualitatively similar results, except that the width Γ_t is now replaced by the typically much smaller width Γ of the resonance against decay into the waveguide. Γ is modified whenever one changes the channels through which the scatterer can decay.

We next place either Fig. 1a or 1b as path 1 in the AB interferometer, as in Figs. 2a or 2b, and calculate the transmission amplitude t for an electron going from X to Y. For simplicity, we include only one (real, except for the AB phase ϕ) hopping matrix element between the sites AD (J_L), DB (J_R), and AB (V). In the unitary case (Fig. 2a), we find

$$t = C G_D [V(\epsilon_q - \epsilon_0) - J_L J_R e^{i\phi}], \qquad (3)$$

where G_D is the *fully renormalized* Green function of the dot (containing the effects of all the leads),

$$[G_D]^{-1} = \epsilon_q - \epsilon_0 + \frac{J X_{LR} + 2 V J_L J_R \cos\phi \, e^{iqa}}{J^2 e^{-iqa} - V^2 e^{iqa}}, \qquad (4)$$

and $C = 2iJ \sin(qa)/[V^2 e^{iqa} - J^2 e^{-iqa}]$ is a smooth function of the parameters.

Note that α_1 dropped out from the square brackets in Eq. (3), which represent the interference: the coefficients inside the brackets are real, and $T = |t|^2$ depends on ϕ only through $\cos\phi$, as expected from Onsager. G_D does depend on α_1 and on ϕ, but its dependence on ϕ is also only via $\cos\phi$. The coefficient of $\cos\phi$ in T, which has contributions from both $J_L J_R V(\epsilon_q - \epsilon_0)$ and the expansion of G_D in a Fourier series in ϕ, changes sign as ϵ_0 increases, yielding a sharp jump of the phase shift by π. The vanishing width of this jump is independent of the dot's intrinsic Friedel phase α_1.

We now break unitarity, as in Fig. 2b. Our calculation yields a similar expression, except that ϵ_0 is now replaced by the complex $\epsilon_0 + \Sigma$. Thus, the absolute value of the square brackets in Eq. (3) now contains a term proportional to $\cos(\beta + \phi)$, with

$$\tan\beta = -\mathrm{Im}\Sigma/(\epsilon_q - \epsilon_0 - \mathrm{Re}\Sigma). \qquad (5)$$

This behavior is portrayed in Fig. 3 (plotted with parameters for which the dependence of G_D on $\cos\phi$ is weak). Note that β is fully determined by the electron loss into the reservoir R, and it has *no dependence* on the intrinsic QD transmission phase α_1, which follows from Eq. (2). Nevertheless, β will change by π across any resonance, where $(\epsilon_q - \epsilon_0 - \mathrm{Re}\Sigma)$ changes sign (up to a shift due to the harmonics of G_D). The width of this change is
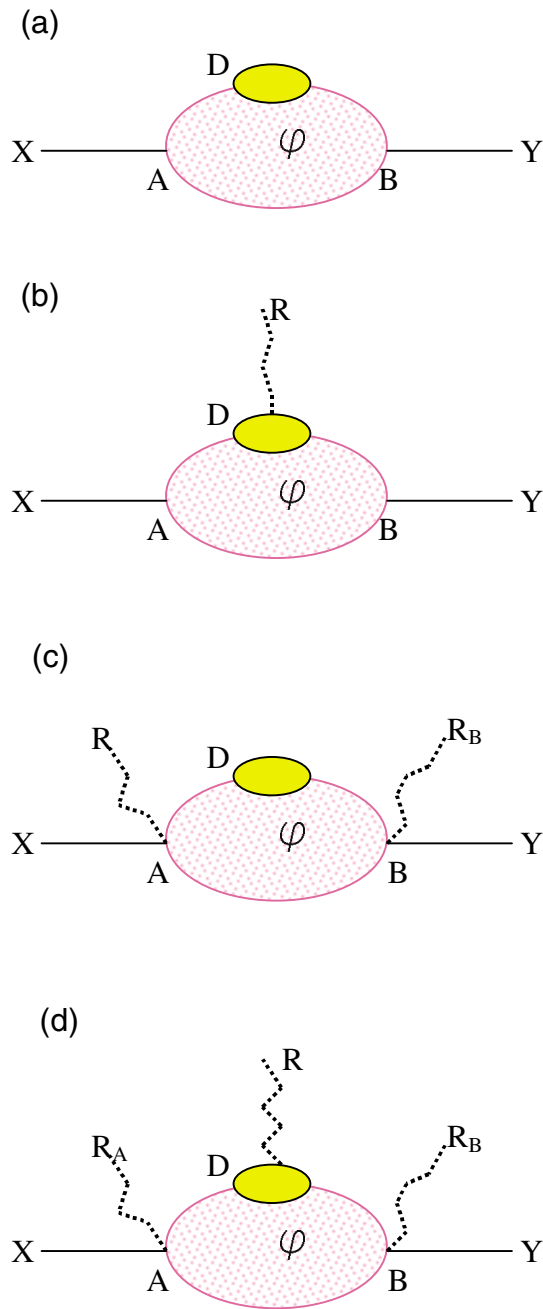
(a)

(b)

(c)

(d)

FIG. 2 (color). AB interferometers, with a magnetic flux ϕ inside the ring. The text describes calculations of the transmission amplitude of a wave from terminal X to terminal Y, for a tight-binding model with single real hopping matrix elements between A and D (J_L), D and B (J_R), and on the lower path, from A to B (V). (a) A closed system. (b) Electrons are lost from the QD via a link to the absorbing reservoir R. (c) Electrons are lost via links to the absorbing reservoirs R_A and R_B. In the case discussed in the text, these losses "conspire" to have full reflection of waves approaching A and B from either path within the ring. (d) Same as (c), with the additional loss from D into R.

FIG. 3. The transmission from X to Y in Fig. 2b, as function of the AB phase ϕ; $J = 1$, $J_L = J_R = 0.1$, $V = 0.01$, $qa = \pi/2$, $\epsilon_0 = 0.1$. The extrema of the curve shift by the phase β (indicated by a point on each curve), increasing as V_1, which measures the rate of electron losses to the reservoir R, grows from 0 to 0.35 (in steps of 0.05). Note that the total magnitude of the transmission decreases with V_1, reflecting the same losses.

determined by ImΣ, i.e., by the rate of electron loss from the QD, and *not* by the intrinsic properties of the dot. In a similar fashion, the phase shift β will exhibit a plateau near $\pi/2$ whenever $|\text{Im}\Sigma| \gg |\epsilon_q - \epsilon_0 - \text{Re}\Sigma|$. Such a
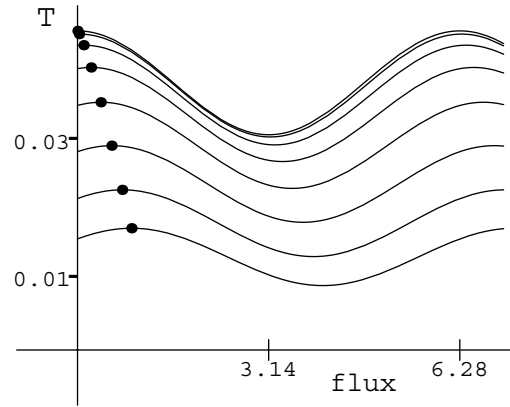
plateau is a hallmark [5] of the Kondo effect. However, establishing its connection to Kondo physics requires more evidence (such as the enhanced conductance in the CB valley, found in [5]).

The physical reason for the Onsager symmetry is clear: the electron wave encircles the interferometer and is reflected from the junctions at A and B many times, complicating the simple two-slit formula, Eq. (1). Indeed, our derivation of Eq. (3) shows that the cancellation of the phase difference $\alpha_2 - \alpha_1$ from inside the square brackets occurs at each order in the summation over *all* of these reflections. As already hinted in Ref. [2], the two-slit formula requires *total absorption* on the junction B (for waves approaching it from the two paths in the AB ring), thus breaking unitarity at or before this point. In fact, a sufficient condition for this formula is that there be *no reflections* from B backwards to D and A, and similarly from A back towards D and B. One theoretical way to achieve this is shown in Fig. 2c: attach to each junction an additional lead to a fully absorbing reservoir. The full four-link point is now described by a unitary 4×4 scattering matrix. One possibility for such a matrix at point B is

$$S_4 = \begin{pmatrix} 0 & 0 & \cos\omega & -\sin\omega \\ 0 & 0 & \sin\omega & \cos\omega \\ \cos\omega & \sin\omega & 0 & 0 \\ -\sin\omega & \cos\omega & 0 & 0 \end{pmatrix}, \quad (6)$$

in which the rows represent R_B, Y, A, and D. Such a matrix would arise, e.g., for a semitransparent mirror placed at B, at 45° with the four orthogonal links. Clearly, the 3×3 submatrix corresponding to Y, A, and D is not unitary; however, its zeros ensure no reflections back into the ring. Introducing a similar matrix at A then yields the two-slit formula, Eq. (1). Nonetheless, note that the above matrix S_4 has not been derived from a microscopic model (however, similar elements do exist for microwaves [22]). Such a derivation for electrons on single-channel leads may

require more absorbing leads (i.e., a larger initial scattering matrix), or more complicated elements. Furthermore, this matrix has a special and restricted form, and it is not obvious how to achieve it experimentally. Finally, the two-slit formula so obtained contains the transmission amplitudes t_1 and t_2 of the two individual paths, and these depend on all the internal details of these paths, including losses (e.g., as shown in Fig. 2d). The amplitude t_1 will have the desired intrinsic phase α_1 of the QD only when, in addition to the total absorption on junctions A and B, the width of the dot's resonating state against losses to all available channels is much smaller than the intrinsic width of the resonance, Γ_R.

In real experiments [3,5,10,11], additional leads are attached to the ballistic arms of the interferometer, between the dot and the "forks" of the interferometer [23]. These leads are "lossy," as reflected by the small fraction of the current coming out of the interferometer. When the losses occur within the back-and-forth reflections of the resonance itself, then the measured phase will be mainly due to those losses, similar to our calculations for Fig. 2b. In that case, the AB phase shift β continues to grow with V_1, with no connection to α_1. Alternatively, one could have many weakly coupled absorbing leads along the conducting paths between the QD and the junctions A or B, outside of this "rattling" region. Under appropriate conditions, the reflections from A and B back into the ring (through the junctions to these leads) become negligible, the two-slit limit is reached and β saturates at the intrinsic QD transmission phase α_1 for a large number of such leads [24]. Thus, an appropriate specific design of the unitarity breaking in the experiments should recover the two-path interference. Considering some of the qualitative results found in Ref. [11] and in consecutive work, it is quite possible that these experiments did contain such a design. A quantitative measurement of the dependence of the measured phase shift β on the strength of the losses could confirm this possibility.

Two final comments. First, note that in the unitary case, the interference part of the transmission [square brackets in Eq. (3)] is *real* at zero flux. It may therefore be tuned to vanish as a function of a single control parameter. Such vanishing may result in a sharp jump of the phase shift measured in the experiment, from 0 to π or vice versa [25]. This entails the same physical mechanism as the one appearing in the Fano line shape [26] (see, e.g., Refs. [18,19] for related suggestions). These considerations may explain some of the aforementioned experimental puzzles. Second, unitarity would also be broken with emitting, rather than absorbing, additional channels. In view of the lossy experiments, we preferred to concentrate on the latter.

[1] Y. Aharonov and D. Bohm, Phys. Rev. **115**, 485 (1959).
[2] E.g., R. P. Feynmann, R. B. Leighton, and M. Sands, *The Feynmann Lectures on Physics* (Addison-Wesley, Reading, 1970), Vol. III, Chap. 1.
[3] A. Yacoby *et al.,* Phys. Rev. Lett. **66**, 1938 (1991).
[4] A. Yacoby, M. Heiblum, V. Umansky, H. Shtrikman, and D. Mahalu, Phys. Rev. Lett. **73**, 3149 (1994).
[5] Y. Ji, M. Heiblum, D. Sprinzak, D. Mahalu, and H. Shtrikman, Science **290**, 779 (2000).
[6] Y. Ji, M. Heiblum, and H. Shtrikman, Phys. Rev. Lett. **88**, 076601 (2002).
[7] R. Landauer, Philos. Mag. **21**, 863 (1970).
[8] E.g., F. Schwabl, *Quantum Mechanics* (Springer-Verlag, Berlin, 1990), Sec. 7.5.2.
[9] L. Onsager, Phys. Rev. **38**, 2265 (1931).
[10] E. Buks *et al.,* Phys. Rev. Lett. **77**, 4664 (1996).
[11] R. Schuster *et al.,* Nature (London) **385**, 417 (1997).
[12] Y. Imry, *Introduction to Mesoscopic Physics* (Oxford University Press, Oxford, 2002), 2nd ed.
[13] Y. Gefen, Y. Imry, and M. Ya. Azbel, Phys. Rev. Lett. **52**, 129 (1984).
[14] R. A. Webb, S. Washburn, C. P. Umbach, and R. B. Laibowitz, Phys. Rev. Lett. **54**, 2696 (1985).
[15] A. Yacoby, M. Heiblum, D. Mahalu, and H. Shtrikman, Phys. Rev. Lett. **74**, 4047 (1995).
[16] D. C. Langreth, Phys. Rev. **150**, 516 (1966).
[17] T. K. Ng and P. A. Lee, Phys. Rev. Lett. **61**, 1768 (1988).
[18] B. R. Bulka and P. Stefanski, Phys. Rev. Lett. **86**, 5128 (2001).
[19] W. Hofstetter, J. Koenig, and H. Schoeller, Phys. Rev. Lett. **87**, 156 803 (2001).
[20] E.g., K. Kang, S. Y. Cho, J.-J. Kim, and S.-C. Shin, Phys. Rev. B **63**, 113 304 (2001).
[21] M. Büttiker, Phys. Rev. B **32**, 1846 (1985).
[22] R. Levi, Adv. Microwaves **1**, 155 (1966).
[23] Four additional lossy channels, connected to the arms AD, DB, and AB, were studied within a Breit-Wigner resonance approximation, by G. Hackenbroich and H. A. Weidenmüller, [Europhys. Lett. **38**, 129 (1997)]. However, the dependence of β on the opening V_1 and its possible relation to α_1 were not discussed.
[24] A. Aharony, O. Entin-Wohlman, B. I. Halperin, and Y. Imry (unpublished)
[25] O. Entin-Wohlman, A. Aharony, Y. Imry, and Y. Levinson, J. Low Temp. Phys. **126**, 1251 (2002)
[26] U. Fano, Phys. Rev. **124**, 1866 (1961).