Conductance phases in the quantum dots of an Aharonov-Bohm ring

A. Yahalom[1,*] and R. Englman[1,2,†]
[1]*College of Judea and Samaria, Ariel 44284, Israel*
[2]*Department of Physics and Applied Mathematics, Soreq NRC, Yavne 81800, Israel*
(Received 13 November 2005; revised manuscript received 23 June 2006; published 26 September 2006)

The regimes of growing phases (for electron numbers $N \approx 0-8$) that pass into regions of self-returning phases (for $N>8$), found recently in quantum dot conductances by Heiblum and co-workers are accounted for by an elementary Green's function formalism, appropriate to an equi-spaced ladder structure (with at least three rungs) of electronic levels in the quantum dot. The key features of the theory are physically a dissipation rate that increases linearly with the level number (and is tentatively linked to coupling to longitudinal optical phonons) and a set of Fano-like metastable levels, which disturb the unitarity, and mathematically the change-over of the position of the complex transmission amplitude zeros from the upper half in the complex gap-voltage plane to the lower half of that plane. The two regimes are identified with (respectively) the Blaschke term and the Kramers-Kronig integral term in the theory of complex variables.

DOI: 10.1103/PhysRevB.74.115328 PACS number(s): 03.65.Vf

I. TWO-PHASE REGIMES

Following a theoretical prediction in Ref. 1, a pioneering experimental determination of the phase evolution in quantum dots subject to the Aharonov-Bohm effect was made by Heiblum and co-workers in, e.g., Refs. 2–6. The same group has recently come up with an interesting development and a physical description,[7] which throw fresh light on their results (previous and recent). They showed (cf. Figs. 4–6 in Ref. 7) that as the gate voltage (V_p in their notation) increases and more electrons enter the quantum dot, the phase of the conductance evolves in the following manner. Initially, for a number of electrons N in the quantum dot up to eight ($N \leqslant 8$) the phases corresponding to each N increase in a stepwise fashion, following which, as $N>8$, the phases return continuously to their original value (make a phase lapse).

Other manifestations of the absence of "phase rigidity" (meaning the discontinuous switch of phases between 0 and $\pm\pi$, connected to unitarity) in phase-coherent transport across quantum dots were observed in Ref. 8.

On the theoretical front, the unexpected phase behavior of the experiments has resulted in numerous theoretical efforts, several of which included investigation of the Kondo effect (e.g., Ref. 9). Other works were directed at an analysis of the results in terms of the Landauer-Büttiker formalism of conductivity, which then led to consideration of the transmission amplitude $t^{QD}(U)$ as a function of the gap voltage U.[10–12] The complicated geometry of the experiments necessitated the inclusion in the theory of several channels and the couplings between these,[13] as well as a detailed analysis of the phase that was being observed.[14] A qualitative effect of changes in both the transmission probability and the phase was theoretically found when the signs were changed in some dot-lead coupling matrix elements.[15] More recently, the ingoing-outcoming coupling asymmetry was studied more comprehensively, again in a two-level system.[16] A selective choice of the experimental phase-conductance results obtained in Refs. 5 and 6 was matched with use of the Friedel sum rule in Ref. 17, without accounting for the transition between the phase growth and the phase lapse regimes.

Quantum dot ring transmission is theoretically related to transmission in quantum waveguides.[18,19] The latter was studied in Refs. 20–22, which sources noted the existence of pole-zero pairs in the transmission amplitude (as a function of the incident wave energy), and especially the changes that occurred with resonant attachments (stubs) to the waveguide. An interesting finding was made in Ref. 21, that in waveguides some geometric changes (like attaching stubs) are formally equivalent to coupling between discrete and continuum of states (the Fano effect[23]). For Aharonov-Bohm interferometers (ABIs) including quantum dots the interrelation between geometry and Fano states was formulated in Ref. 24, and recent experiments were interpreted in terms of the Fano effect.[8] We stress here this theoretical equivalence, since many previous explanations of the anomalous phase behavior in quantum dots concentrated on the geometrical aspect, whereas in the following theory the breakdown of unitarity is traced to decay of conducting levels and to the metastability of electronic states lying above the quantum dot well.

Earlier, the Fano-type behavior (including the emergence of transmission amplitude zeros) was traced to the existence of a lateral nonresonant channel in the quantum dot arm of the ABI.[25,18,19] The essential modification that the present work makes in the treatment of the Fano effect is that the higher-lying states in the quantum dot arm are assumed to possess a short lifetime, so that the imaginary part dominates the energy denominator. (Under this assumption it makes no difference whether the higher-lying states form a continuum or are discrete, as we propose for simplicity; however, the metastability of these states changes the signs of some couplings from those in the previous references. We clarify this sign change at the end of Sec. IV B.) Establishing the causes of the metastability of the high-lying states and of the level-dependent decay rate in the lower-lying states is not a primary aim of this work, but certain considerations indicate that both are due to coupling to longitudinal optical (LO) phonons and this speculative idea is described in Sec. V. Without independent support for this mechanism and for the simplified choice of some of the coupling coefficients made

in Sec. II C, the present theory has to be regarded as phenomenological.

Though apparently far removed from the physics of quantum dots, it turns out that the question of zeros and poles of the transmission amplitude or of a Green's function (the equivalence between which was demonstrated in Ref. 22) plays an important role in the explanation of the phase behavior. This was heralded in several previous works (e.g.,[26]), but the present theory does this in a more comprehensive form, namely, by use of the Hilbert transform (in Sec. III) and through a compact representation of the transmission amplitude (in Sec. IV).

Though the simple theory presented below is implemented by *ad hoc* assignment of some parameter values and needs to be amplified to fill in several physical details, it seems that it contains the answer to the question: What lies behind the strange phase behavior? A view expressed in Ref. 7 is that the so far available theories are short of providing an answer.[27] Earlier experimental data of Ref. 4 for returning phases were rather precisely correlated with the observed conductivities by the present authors, using a parameter-free method which was the precursor of the present work.[28] In Sec. VI of this work we note the differences between the theory proposed here and those of other researchers.

II. ELECTRON TRANSMISSION AMPLITUDE

The properties of the quantum dot (spinless) electron transmission function can be best understood in terms of the theory of Hackenbroich and Weidenmuller.[10] For the sake of completeness we repeat here their end result.

A. System Hamiltonian

The system under consideration is composed of three subsystems: (1) the leads; (2) the Aharanov-Bohm system not containing the dot; and (3) the quantum dot. The entire Hamiltonian of the system can be described by

$$H = H_0 + H_T. \tag{1}$$

H_0 describes the totally disconnected system and is given by

$$H_0 = \sum_{akr} \epsilon^r_{ak} c^{r\dagger}_{ak} c^r_{ak} + \sum_i \epsilon_i d^\dagger_i d_i + \sum_j E_j q^\dagger_j q_j + U_{ES}. \tag{2}$$

r denotes the leads, a runs over the channels in each lead and k over the longitudinal wave numbers, and ϵ^r_{ak} is the corresponding energy. The energies of the single-particle states within the rings and within the dot are labeled by ϵ_i and E_j, respectively. E_j is assumed to depend parametrically on U. In the formal theory to follow, U is a complex quantity, whose real part ReU is identified with the experimentally manipulated plunger voltage V_p. U_{ES} is the electrostatic charging energy of the dot.

The coupling Hamiltonian H_T has the form

$$H_T = \sum_{akir} W^r_{ai}(k) c^{r\dagger}_{ak} d_i + \sum_{ijp} V^p_{ij} q^\dagger_i d_j + \text{H.c.} \tag{3}$$

W describes the coupling between ring and leads; V describes the much smaller coupling between ring and dot. $p=L,R$ labels either side of the dot. In actuality, we should allow for more exit channels than just the two (L and R) for the dot, corresponding to the experimental arrangements in, e.g., Ref. 7 We shall account for these by including them in the postulated "high-lying" energy levels [see below in Eq. (11) and recall the discussion on the equivalence between stub and Fano-state effects in our opening section].

The transmission amplitude $t_{ab}(E)$ through the ring for an electron entering the ring via channel b in lead 2, and leaving it via channel a in lead 1, is derived in Ref. 10 We separate this as

$$t_{ab} = t^0_{ab} + t^{QD}_{ab} \tag{4}$$

into the ring transmission and the transmission t^{QD}_{ab} across the quantum dot and treat first the former.

B. Aharonov-Bohm ring transmission

The transmission matrix across the ring is expressed by

$$t^0_{ab} = -2i\pi \sum_{ik} W^1_{ai} (D^0)^{-1}_{ik} W^{2*}_{bk} \tag{5}$$

with the matrix $(D^0)_{ik}$ defined by:

$$(D^0)_{ik} = (E - \epsilon_i)\delta_{ik} + i\pi \sum_{ct} W^{t*}_{ci} W^t_{ck}. \tag{6}$$

When the ring is fed by the lead's reservoir filled up to the Fermi energy E_f, one can replace E in Eq. (6) by E_f. In the presence of a magnetic field threading the circuit, the ring transmission amplitude will acquire an Aharonov-Bohm phase factor.

C. Quantum dot transmission

We now turn our attention to the second term in Eq. (4). In the case that repeated zig-zagging of carriers between the leads can be ignored, this is the term whose magnitude and phase are obtained in an Aharonov-Bohm interference measurement.[14] For simplicity, we drop the channel labels a,b.

We model the quantum dot as an electronic system having a ladder structure, i.e., N_{el} equi-spaced levels, interacting with some dissipative reservoir, say the LO phonons in the dot[31–34]. For quantum dots typified by those in the experiments discussed, the number of available levels is of the order of 100 and their spacing is 40 μeV.[11] We shall subdivide these levels into $N_{low\ lying}$ bound states, inside the quantum well and having an equispaced ladder structure, and a set of $N_{high\ lying}$ localized, metastable ("almost bound") states, above the well.[29,30] The effect of these levels on the low-lying levels is similar to the continua that feature in the Fano effect. For simplicity, we take the number of these levels ($N_{high\ lying}$) finite.

We next write the transmission amplitude across the dot within a wideband approximation, as described in Ref. 10 The limitations in applying the Hackenbroich-Weidenmüller approach to the experiments of Refs. 2–7 have been noted in Ref. 11 (Sec. 4.3.1). On the other hand, the observed regular peak structures in some of these papers indicate that the fol-

lowing sum of Breit-Wigner terms should form an approximation to the transmission amplitude (at least close to resonances):

$$t^{QD}(U) = -iG\sum_{n=1}^{N_{el}} \frac{1}{E - E_f + U - \langle n|H_0|n\rangle - R(E - E_f, n)}. \tag{7}$$

G is a single parameter characterizing the scattering across the dot and is equivalent to $2\pi W^2$ introduced above, E_f (as before) is the Fermi energy in the leads, U is the gap voltage parameter in suitable units whose real part V_p is the experimentally manipulated depth of the dot well (however, we shall occasionally use U also when we mean its real part), $\langle n|H_0|n\rangle = n$ is the electronic level energy in suitable units, and R is the complex self-energy of the nth dot level, including also the coupling of the electrons to the environment (erstwhile, the phonons and the stubs). Note that U is not the Hubbard repulsion parameter, which will not be explicitly taken into account, except for its presence in the self-energy R, which will also incorporate off-diagonal terms.[15]

For the self-energy $R = R' + iR''$ we now introduce our main assumption that its imaginary part scales linearly for low-lying levels with the electronic level height

$$R'' = -\gamma n \quad (\gamma > 0). \tag{8}$$

For higher lying levels we assume that the phonon-electron coupling mechanism is so efficient that $|R''| \gg |U - \langle n|H_0|n\rangle - R|$. The width of these levels is extremely large, so that the dependence of U on the contribution by those levels to $t^{QD}(U)$ is negligible. (This is different from the usual treatment of the Fano effect in which the contribution of the continuum is energy dependent.[23])

The $t^{QD}(U)$ terms can thus be dissected into two terms as follows:

$$t^{QD}(U) = t_h^{QD} + t_l^{QD}(U), \tag{9}$$

in which we artificially ignore the intermediate cases. In the above equation,

$$t_l^{QD}(U) = -iG \sum_{n=1}^{N_{low\ lying}} \frac{1}{E - E_f + U - \langle n|H_0|n\rangle - R(E - E_f, n)} \tag{10}$$

and

$$t_h^{QD} = G \sum_{n=N_{low\ lying}+1}^{N_{high\ lying}} \frac{1}{R''(E - E_f, n)}. \tag{11}$$

We next use Eq. (9) to calculate $t^{QD}(U)$, the quantum dot transmission coefficient as function of the the gap voltage U. Figures 1 and 2 show the results, with the following choice of parameters (having put $E = E_f$):

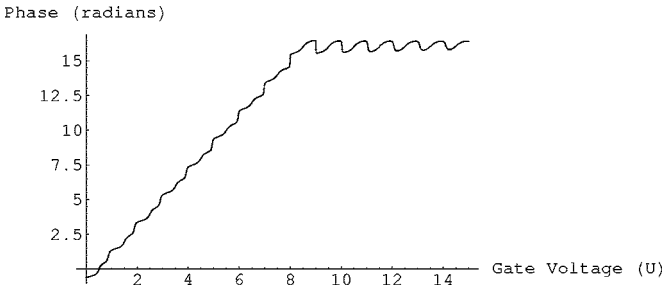

FIG. 1. The phase of the transmission amplitude for the parameters given by Eq. (12).

$$N_{low\ lying} = 34, \quad \frac{t_h^{QD}}{G} = 1.35, \quad \gamma = 0.0086, \quad R' = -8.5. \tag{12}$$

The figures show clearly the peaked structure of the absolute value of the transmission amplitude (the visibility or scaled |conductance|) at subsequent electron fillings and the radical change of character in the phase behavior. Due to our chosen fitting of the energy shift parameter (−8.5) and of $\gamma = 0.0086 \approx 0.01$ in Eq. (12), this change occurs just at the experimental value of Ref. 7

D. Numerical properties of the transmission amplitude

Inserting the numerical parameters from Eq. (12) into Eq. (9) we obtain the quantum dot transmission amplitude as

$$F(U) = t^{QD}(U)/G \approx 1.35 - i\sum_{n=1}^{34} \frac{1}{U + 8.5 - (1 - 0.0086i)n}$$

$$= 1.35 - i\sum_{n=1}^{34} \frac{1}{U - U_n + 0.0086ni}. \tag{13}$$

The poles (resonances) occur at such half-integral values of U:

$$U_n = n - 8.5. \tag{14}$$

The widths increase linearly with n. The height of each resonance is given for small γ approximately as

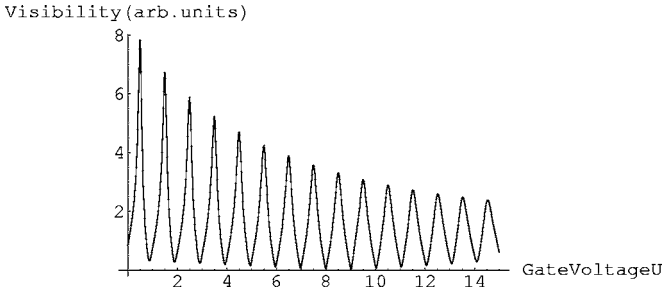

FIG. 2. The absolute value of the transmission amplitude for the parameters given by Eq. (12).

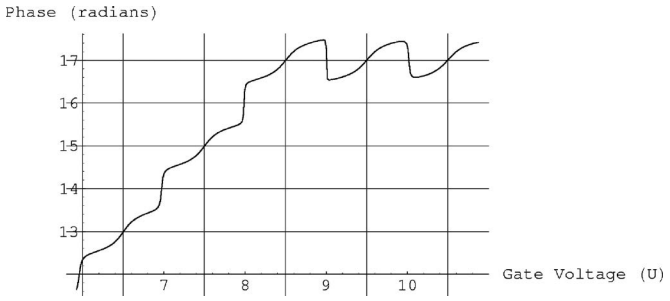

FIG. 3. The phase of the transmission amplitude for the parameters given by Eq. (12). A small section of the previous phase figure is enlarged in order to allow one to follow the details more closely. The perpendicular grid lines are located at resonances.

$$|F(U_n)| = \left| 1.35 - \frac{1}{\gamma n} \right| = \left| 1.35 - \frac{1}{\gamma(U_n + 8.5)} \right|, \quad (15)$$

hence (for $n < 1/1.35\gamma \approx 86$) the height of the resonances is decreasing, as is apparent from Fig. 2. The location of the valley between the resonances U_n and U_{n+1} can be approximated for small γ by taking into account only the two close resonances to the valley as

$$\hat{U}_n = \frac{1}{2}(U_n + U_{n+1}) = U_n + \frac{1}{2}. \quad (16)$$

Adding symmetrical resonances on both sides of the valley does not change its location. However, valleys that do not have the same number of resonances on both sides are shifted slightly. Consider for example the valleys $\hat{U}_{n-1}, \hat{U}_n, \hat{U}_{n+1}$ located between four resonances $U_{n-1}, U_n, U_{n+1}, U_{n+2}$; their locations will be

$$\hat{U}_{n-1} \simeq U_n - 0.62, \quad \hat{U}_n \simeq U_n + 0.5, \quad \hat{U}_{n+1} \simeq U_{n+1} + 0.62. \quad (17)$$

To study more clearly the phase evolution we enlarge part of Fig. 1. This is given in Fig. 3. From the figure it is clear that the phases change continuously across the resonances and change abruptly across the valleys. Across each resonance the phase increases by π, as usually across Breit-Wigner resonances. Across each valley an abrupt change of π also occurs. This can be positive for $\hat{U}_n < 8.5$ or negative for $\hat{U}_n > 8.5$. The above behavior can be associated with the existence of (complex) zeros, passing which either adds or subtracts π to (from) the phase, depending on whether the zero lies in the upper or the lower half U plane. The leading term [1.35 in Eq. (13)], whose source is the higher-lying states, is essential for the existence of the zeros. It turns out that to obtain the zeros around any U_n (or $n-8.5$), it is necessary to consider two more terms in the sum, one on each side of the resonance. (One neighboring term is insufficient; three or more terms are qualitatively unnecessary. This numerical aspect distinguishes the present approach from several previous ones, e.g., Ref. 15, which considered only two resonances. Some exceptions are, Refs. 14 and 35 which, however, do not include the fast-decaying levels.) It thus emerges that, with only three resonances, for $U < 8.5$ one finds three zeros in the upper half U plane which make up a total 6π increase over three resonances; whereas, for $U > 8.5$ one finds just one zero in the lower half U plane, which leads to a total of 2π phase change. Restoring all the resonances yields the curves shown in Figs. 1 and 2.

Changes in some parameters can alter, e.g., the relative magnitudes of the peaks. The slope of the phase change across a valley is proportional to the height of the valley above zero. (This property was first predicted in Ref. 28 and rediscovered in several subsequent papers.) The width is therefore π/ [slope].

III. GENERAL SIGNIFICANCE OF COMPLEX ZEROS

We now describe the formal basis of the above result, showing that the change of behavior is not accidental, but rather required by simple mathematical properties of the transmission amplitude $t^{QD}(U)$ regarded as a function of the variable U. The underlying reason is that just such behavior of phases is expected for a quantity $t^{QD}(U)$ that has the following properties [in addition to $t^{QD}(U)$ satisfying certain formal, analytical properties[36,37]]: $t^{QD}(U)$ has zeros in the upper half of the complex U plane for $\mathrm{Re}U \leqslant 8$ and has zeros in the lower half of the U plane for $\mathrm{Re}U > 8$. (As before, we have identified the real part of U with a scaled gate voltage V_p. The gate voltage V_p increases the number n of bound electrons in the quantum dot.)

Why is this so straightforward?

As shown immediately below, the phase evolution can be expressed as a sum of (essentially) two terms: an integral term and the (so called) Blaschke terms. The former shows structure (wiggles) or phase return, but no net gain (i.e., it returns to the starting value) and the latter shows net gains, phase growth (and no structure). Precisely, the Blaschke terms arise from singularities of $\ln t^{QD}(U)$ in the upper half plane and the structure in the integral comes from singularities of $\ln t^{QD}(U)$ in the lower half plane (due to continuity). Furthermore, both the wiggles and the gain (in the phase) are tied to maxima in the visibility [$|t^{QD}(U)|$], as in the experiments.

Thus the minimal property required of $t^{QD}(U)$ is that its complex zeros lie in the upper half plane for $\mathrm{Re}(U)$ less than 8 and in the lower half plane for $\mathrm{Re}(U)$ larger than 8. In the sequence we shall build up at least one simple function $t^{QD}(U)$ that has these properties, but there are obviously others, too.

A. The Blaschke terms

Let us explain the Blaschke terms. These arise if the well-known Kramers-Kronig (KK) relations are applied to the logarithm of a regular function $t^{QD}(U)$ of its argument U, rather than to $t^{QD}(U)$ itself, as is usual. Then the zeros of $\ln t^{QD}(U)$ add singularities to the KK integrand and these have to be subtracted in a manner that does not affect adversely the conditions that are the basis of the KK relations. As a consequence (for real values of U) one can express the argument (phase) of this function as

$$\arg t^{QD}(U) = -\frac{1}{\pi}\mathrm{P}\int_{-\infty}^{\infty}\frac{dU'}{U'-U}\ln|t(U')| + \Phi_B(U) \quad (18)$$

where P represents the principal part of the singular integral and $\Phi_B(U)$ is the Blaschke phase given as the sum of terms[38,39]

$$\Phi_B(U) = -i\sum_j \ln\frac{U-U^j}{U-U^{j*}}. \quad (19)$$

Here U^j are those zeros of $t^{QD}(U)$ that lie in the upper half of the complex U plane and U^{j*} are their complex conjugates. [Actually, Eq. (19) can be generalized for poles and branch points in the upper half plane by suitably attaching (negative and fractional) weights to each term in the sum, but since we shall find that there are no poles or branch points in the transmission amplitude for the range of interest, we can disregard these possibilities.]

Now if we look at the integral term, we see that it tends to 0 for both $U\to-\infty$ and $U\to\infty$ (provided the log function has no singularities on the real U axis). Therefore, as claimed, this term cannot cause a net gain of the phase, only some structure. Such structure will indeed occur when $|t^{QD}(U)|$ becomes small at some value of U. It will have the form of a very sharp peak whenever a zero of $t^{QD}(U)$ is very close to the real axis. This will occur when γ is very small compared to the level spacing, as in Eq. (12), for which the level spacing was unity.

A different story are the terms in the Blaschke phase. Each term will cause a step of 2π in the phase.

In a subsequent drawing (Fig. 5), we show $|t^{QD}(U)|$ and $\frac{1}{\pi}\arg t^{QD}(U)$ both vs U on the same graph. In the visibility $|t^{QD}(U)|$ one sees the peak structure and in the phase the initial steps (up to $U=8$), followed (for U above this value) by the rise and lapse of the phases.

IV. COMPACT FORM OF THE TRANSMISSION AMPLITUDE

We now rewrite the preceding expression for the total transmission matrix $t^{QD}(U)$ (making only an approximation that will turn out to have almost no effect on the results) and obtain a compact, closed expression. From this we can deduce the relevant analytic properties of $t^{QD}(U)$ almost by inspection.

Because we expect that for a given value of U only a few (nearly resonant) terms in Eq. (10) will contribute, we extend the sum in Eq. (9) to $-\infty$ and ∞. The resulting series can be summed to take the simple form

$$\frac{t^{\infty}(U)}{t^{\infty}(U\to-\infty)} = \frac{1+Ae^{-2\pi i(U-U_c)/(1-i\gamma)}}{1-Be^{-2\pi i(U-U_c)/(1-i\gamma)}}. \quad (20)$$

The algebra is based on the result[40]

$$\frac{1}{e^z-1} = -\frac{1}{2} + \sum_{n=-\infty}^{\infty}\frac{1}{z-2n\pi i}, \quad (21)$$

from which follows the expansion of $t^{\infty}(U)$ as the series

$$\frac{t^{\infty}(U)}{t^{\infty}(U\to-\infty)} = \frac{1}{2}\left(1-\frac{A}{B}\right) - i\left(1+\frac{A}{B}\right)\frac{(1-i\gamma)}{2\pi}\sum_{n=-\infty}^{\infty}\frac{1}{U-n-[U_c+(\gamma/2\pi)\ln B]+i[\gamma n+(1/2\pi)\ln B]}. \quad (22)$$

Recalling now from Eq. (9) $t^{QD}(U)=t_h^{QD}+t_l^{QD}(U)$, and noting the expression for t_l^{QD} in Eq. (10), we can make the following replacements:

$$t_h^{QD} = \frac{t^{\infty}(U\to-\infty)}{2}\left(1-\frac{A}{B}\right), \quad 2\pi G \simeq t^{\infty}(U\to-\infty)\left(1+\frac{A}{B}\right)$$

$$R' = -U_c - \frac{\gamma}{\pi}\ln B, \quad R'' = -\gamma n - \frac{1}{2\pi}\ln B \quad (23)$$

(In the second equation we have neglected the small and unimportant quantity $-i\gamma$ before the sum.) Equation (23) will lead to the following proportion between A and B:

$$\frac{A}{B} = \frac{\pi G - t_h^{QD}}{\pi G + t_h^{QD}}. \quad (24)$$

The following values of the four parameters (A, B, U_c, γ) in the function $t^{QD}(U)$ are compatible with the choices of the parameters in Eq. (12):

$$A=1, \quad B=2.5, \quad U_c=8.5, \quad \gamma=0.0086. \quad (25)$$

The plotted $t^{\infty}(U)$ with these parameters is shown in Fig. 5. The result is virtually identical with that obtained for $t^{QD}(U)$ from the restricted sum in Eq. (7), in the gap voltage range of Figs. 1 and 2. As already noted, the reason is that the contributions to the infinite sum outside the restricted range are negligible. The signal advantage of the compact form in Eq. (20) over the partial sum in Eq. (7) is that the zeros and poles of the transmission amplitude can be derived from the former considerably simpler. We now obtain these with the parameters chosen in Eq. (25).

A. Analysis of zeros and poles

(1) Zeros of Eq. (20): These occur when the second term in the numerator is −1, so that

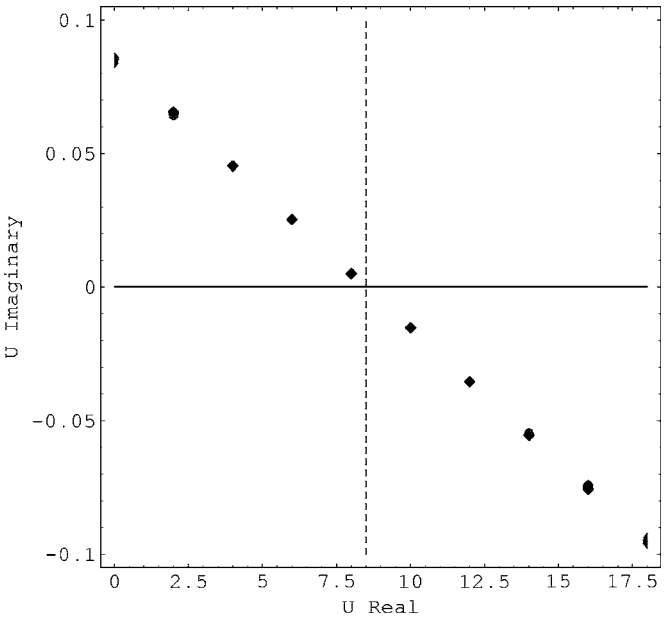

FIG. 4. Zeros of $t^{\infty}(U)$ [see Eq. (22) below] with parameter values $A=1$, $B=2.5$, $U_c=8.5$, $\gamma=0.01$.

$$U = U_c + n + \frac{1}{2} + \frac{\gamma \ln A}{2\pi} - i\gamma\left[\left(n+\frac{1}{2}\right) - \frac{\ln A}{2\gamma\pi}\right]$$
$$\times (n = 0 \text{ or a signed integer})$$
$$= 8.5 + \left(n+\frac{1}{2}\right) - 0.0086i\left(n+\frac{1}{2}\right)$$
$$= V_p(n) + 0.0086i[8.5 - V_p(n)], \quad (26)$$

where in the second equality we have inserted the parameter values from Eq. (25) and in the third equality we have written $V_p(n)$ for the real part of the nth zero. For a small value of the decay rate γ this will be the value of the observed gap voltage at the position of the minimum. It is now apparent that for minima at gap voltages below 8.5 the zeros will be at positive imaginary parts of U, while for gap voltages above 8.5 the imaginary part will be negative. (This is shown in Fig. 4.)

(2) Poles of Eq. (20): For these the second term in the denominator must be 1, giving

$$U = U_c + m + \frac{\gamma \ln B}{2\pi} - i\gamma\left(m - \frac{\ln B}{2\gamma\pi}\right)$$
$$\times (m = 0 \text{ or a signed integer})$$
$$= 8.5 + m - 0.0086i(m - 17)$$
$$= V_p(m) + 0.0086i[25.5 - V_p(m)] \quad (27)$$

where again (in the second equality) we have substituted the parameters and then have rewritten the equation in terms of the observational gap voltages $V_p(n)$ at the maxima. [The small quantity $(\gamma \ln B)/2\pi \approx 0.001$ has been neglected.] It is clear now that the poles lie in the upper half of the U plane for all gap voltages below 25.5. Larger gap voltages than this are outside the range of interest for the discussion of the experiments.

B. Deductions from the compact form

The essential features of this form are that for values of the gate voltage U that are experimentally measured (1) there are no singularities (i.e., denominator zeros) in the upper complex U half plane, and (2) for Re $U > U_c$ the zeros of $t^{QD}(U)$ are only in the lower half of the complex U plane (this is the phase-lapse regime, identified with the integral part regime), whereas for Re $U < U_c$ there are zeros in the upper half of the complex U plane (this is the increasing-phase regime, identified with the Blaschke phase regime). Important in Eq. (20) are the parameter U_c (=8.5) and that B (=2.5) $\gg A$ (=1). The latter requirement removes the poles from the wrong half plane and, by Eq. (24), translates immediately to the physical one that the high-lying states' transmission amplitude t_h^{QD} is real and positive.

Contrasting with our zeros, the zeros that were found in both Refs. 22 and 12 were real. In the last work it was indeed pointed out (on p. 106602-4) that the reality was due to the time-reversal invariance of the Hamiltonian, tied to an infinitely sharp π phase jump, whereas a finite-width phase jump could be achieved by interlevel thermal excitation. Alternatively, it could be obtained with a nonzero γ due to inelastic electron-phonon interactions, which is the possibility envisaged here, and $B/A > 1$, and which, by Eq. (24), is contingent on the virtual excitation to metastable states.

(At this stage one may want to compare the functions of Table II and the figures in the earlier paper[28] by the present authors, in which the decay parameter γ was 0. A more significant difference is that in the functions of Ref. 28 the assumed region of analyticity was the opposite to that in the present paper. The former choice is the natural one if U is identified with a "timelike" variable, whereas the present choice is the proper one if U is energy- or frequencylike.)

Comparing the present formalism to previous one-electron, many-level theories,[18,19,25] noted in the introductory section, one can write the transmission amplitude, Eq. 2 in Ref. 18, in the form of the expression in Eq. (20) (apart from constant proportionality factors), with the following substitutions:

$$A = -1, \quad B = \frac{1}{\sqrt{1-2\varepsilon}}, \quad \pi(U - U_c)/(1 - i\gamma) = kL$$

The complex and sign-changing "kL" is an essential ingredient of the present theory. There is also a phenomenological requirement on the ratio A/B to be positive, in order to obtain Fig. 5.

C. Phase-step magnitude

The phase shown in the drawing for the initial (step-up) regime is not the same as in, e.g., Ref. 4 or 7, in that we predict a net phase gain of 2π per peak, whereas the experimental phase steps seem (in most cases) to be less than this. If the discrepancy really exists, the present interpretation may have to be withdrawn or be changed in a way not clear

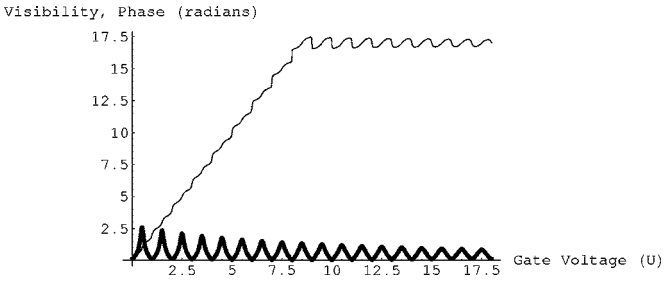

FIG. 5. Visibility amplitude (in arbitrary units) and phase (in radians) as functions of gap voltage ReU. $|t^{\infty}(U)|$ and arg $t^{\infty}(U)$ are plotted from Eq. (20) with parameter values $A=1$, $B=2.5$, $U_c=8.5$, $\gamma=0.0086$.

to us just now. However, it seems that the experimental phases are not traced quite precisely throughout the step. Thus, when the visibility is near zero, the phase changes may not be properly recorded, but rather connected together in a continuous fashion so that part of the rise is lost.

V. SPECULATING ON THE DECAY MECHANISM: ELECTRON–LO-PHONON COUPLING?

Assuming a ladderlike structure for the low-lying electronic levels in the quantum dot, with level separation of unity (when expressed in units of U), our expression for $t^{\infty}(U)$ with $A=0$ can be simply understood as the Green's function of broadened, regularly spaced electronic states. The preexponential factor B and the $i\gamma$ part in the exponent then represent the broadening of low-lying levels. Were the former 1 and the latter 0, we would have the Green's function for a series of equidistant, infinitely sharp electronic levels. However, our main interest is in the zeros of the numerator. These arise because $A \neq 0$.

The phonon bottleneck or its absence has long been under consideration for the mechanism of decay of discrete electronic levels in quantum dots.[31–34] It is generally supposed that LO phonons in the dot of energy $\hbar\omega_{LO}$ couple to the levels. It has also been noted that when the electron level structure at some rung in the ladder gets into near coincidence with the phonon energy, then a Rabi splitting takes place. The physical meaning of this is that the near-coincidental excited electron level gets strongly admixed with the ground electronic level in which one LO phonon is excited. As a result, two admixture levels are formed, which are separated by roughly the coupling energy between the electron and the LO phonon. The condition for coincidence to occur at the n_R (R for Rabi) electronic level is that

$$n_R\Delta \approx \hbar\omega_{LO} \tag{28}$$

where Δ is the electronic energy separation.

We speculate that the decay in $t^{QD}(U)$ reflects this resonance condition, i.e., $\hbar\omega_{\mathrm{LO}}=n_R\Delta \approx N_{low\ lying}\Delta$, since above the low-lying levels commences the LO-phonon decay mechanism. We have not calculated the transmission matrix of the coupled electron–LO-phonon excitation (constituting a polaron), along the lines of Refs. 41–44. With the estimates of Ref. 11 that $\Delta=40$ μeV (which may be a minimal estimate) and that there are $N_{low\ lying} \approx 200$ electronic states up to the brim of the quantum dot well, one obtains 8 meV for the height of electronic levels, at which the phonon coupling causes an effective admixture. This is about a quarter of $\hbar\omega_{LO}=36$ meV where we would expect the electron-phonon coupling to be felt in GaAs.[34] Our computations have stopped at $N_{low\ lying}=34$, since the experimental range of scanned levels is considerably below this. No observable difference would be felt by extending the sum to $N_{low\ lying}$ placed in the hundreds.

VI. CONCLUSION

The two distinct regimes in the electron-transmission phase of an Aharonov-Bohm arrangement containing a quantum dot, already present in earlier experiments in Refs. 2–6 but recently definitively established in Ref. 7, have been explained by a model based on a ladder of electronic levels with increasingly faster decay from higher levels up to a metastable continuum (or bunch of levels) with very short lifetimes. The decay mechanism is tentatively surmised as due to LO phonons in the dot. Though a Hamiltonian is postulated, its implementation in the transmission amplitude is phenomenological. The main additional features not present in several previous theories are that at least three particle states are needed to reproduce the observed phase behavior; geometrical effects (side arms in the ring) are treated on the same footing as the admixture with metastable states in disrupting the unitarity. In formal terms, the two regimes of phases, those increasing across the resonance and those returning to former values are identified with zeros (but not the poles) of the complex transmission lying (respectively) in the lower and upper half planes of the complex energy (or gap voltage) variable. The absence of poles is connected to the metastable state, but the zeros do not arise from the usual Fano form or from cancellation between adjacent resonances or between resonances and conducting states, but from interference with the metastable levels.

ACKNOWLEDGMENTS

We thank Amnon Aharony, Yuval Gefen, and Yuval Oreg for valuable remarks on our work and the first for suggestions.

*Electronic address: asya@yosh.ac.il
†Electronic address: englman@vms.huji.ac.il

[1] Y. Gefen, Y. Imry, and M. Ya. Azbel, Phys. Rev. Lett. **52**, 129 (1984).
[2] A. Yacoby, M. Heiblum, D. Mahalu, and H. Shtrikman, Phys. Rev. Lett. **74**, 4047 (1995); A. Yacoby, R. Schuster, and M. Heiblum, Phys. Rev. B **53**, 9583 (1996).
[3] E. Buks, R. Schuster, M. Heiblum, D. Mahalu, V. Umansky, and H. Shtrikman, Phys. Rev. Lett. **77**, 4664 (1996).
[4] R. Schuster, E. Buks, M. Heiblum, D. Mahalu, V. Umansky, and H. Shtrikman, Nature (London) **385**, 417 (1997).
[5] Y. Ji, M. Heilblum, D. Sprinzak, D. Mahalu, and H. Shtrikman, Science **290**, 779 (2000).
[6] Y. Ji, M. Heilblum, and H. Shtrikman, Phys. Rev. Lett. **88**, 076601 (2002).
[7] M. Avinun-Kalish, M. Heiblum, O. Zarchin, D. Mahalu, and V. Umansky, Nature (London) **436**, 529 (2005).
[8] K. Kobayashi, H. Aikawa, S. Katsumoto, and Y. Iye, Phys. Rev. Lett. **88**, 256806 (2002); H. Aikawa, K. Kobayashi, A. Sano, S. Katsumoto, and Y. Iye, *ibid.* **92**, 176802 (2004).
[9] Y. Oreg and Y. Gefen, Phys. Rev. B **55**, 13726 (1997).
[10] G. Hackenbroich and H. A. Weidenmüller, Phys. Rev. Lett. **76**, 110 (1996); Phys. Rev. B **53**, 16379 (1996); Europhys. Lett. **38**, 129 (1997).
[11] G. Hackenbroich, Phys. Rep. **343**, 463 (2001).
[12] P. G. Silvestrov and Y. Imry, Phys. Rev. Lett. **85**, 2565 (2000); **90**, 106602 (2003).
[13] O. Entin-Wohlman, A. Aharony, and V. Kashcheyevs, J. Phys. Soc. Jpn. **72**, 277 (2003).
[14] A. Aharony, O. Entin-Wohlman, B. I. Halperin, and Y. Imry, Phys. Rev. B **66**, 115311 (2002).
[15] A. Silva, Y. Oreg, and Y. Gefen, Phys. Rev. B **66**, 195316 (2002).
[16] D. I. Golosov and Y. Gefen, cond-mat/0601342 (unpublished).
[17] A. Jerez, P. Vitushinsky, and M. Lavagna, Phys. Rev. Lett. **95**, 127203 (2005).
[18] H. Xu and W. Sheng, Phys. Rev. B **57**, 11903 (1998).
[19] C.-M. Ryu and Y. S. Cho, Phys. Rev. B **58**, 3572 (1998).
[20] P. J. Price, Phys. Rev. B **38**, 1994 (1988); IEEE Trans. Electron Devices **39**, 520 (1992).
[21] W. Porod, Z. Shao, and C. S. Lent, Phys. Rev. B **48**, 8495 (1993).
[22] Z. Shao, W. Porod, and C. S. Lent, Phys. Rev. B **49**, 7453 (1994).
[23] U. Fano, Phys. Rev. **124**, 1866 (1961).
[24] O. Entin-Wohlman, A. Aharony, Y. Imry, amd Y. Levinson, J. Low Temp. Phys. **126**, 1251 (2002).
[25] S. Y. Cho, T. Choi, and C.-M. Ryu, Int. J. Mod. Phys. B **10**, 3569 (1996).
[26] P. S. Deo, Physica E (Amsterdam) **1**, 301 (1997); Solid State Commun. **107**, 69 (1998).
[27] "Expected 'mesoscopic' features in the phase, related to the dot's shape, spin degeneracy or to exchange effects, were never observed. Presently, there is no satisfactory explanation for the observed phase 'universality'." (Taken from the opening paragraph of Ref. 7.)
[28] R. Englman and A. Yahalom, Phys. Rev. B **61**, 2716 (2000).
[29] G. Baym, *Lectures on Quantum Mechanics* (Benjamin, New York, 1969), pp. 104–113.
[30] K. Hirose, Y. Meir, and N. S. Wingreen, Phys. Rev. Lett. **90**, 026804 (2003).
[31] U. Bockelmann and G. Bastard, Phys. Rev. B **42**, 8947 (1990).
[32] X. Q. Li, H. Nakayama, and Y. Arakawa, Phys. Rev. B **59**, 5069 (1999).
[33] T. Inoshita and H. Sakaki, Physica B **227**, 373 (1996); Phys. Rev. B **56**, R4355 (1997).
[34] T. Stauber, R. Zimmermann, and H. Castella, Phys. Rev. B **62**, 7336 (2000).
[35] R. Baltin and Y. Gefen, Phys. Rev. Lett. **83**, 5094 (1999).
[36] R. A. E. C. Paley and N. Wiener, *Fourier Transforms in the Complex Domain* (APS, New York, 1934), p. 15, Theorem XII.
[37] E. C. Titchmarsh, *Introduction to the Theory of Fourier Integrals* (Clarendon Press, Oxford, 1948), Chap. V.
[38] N. G. van Kampen, Phys. Rev. **89**, 1072 (1953).
[39] L. Mandel and E. Wolf, *Coherence and Quantum Optics* (Cambridge University Press, Cambridge, U.K., 1995), Sec. 7.3.2.
[40] T. A. MacRobert, *Functions of a Complex Variable* (MacMillan, London, 1938), p. 128.
[41] L. I. Glazman and R. I. Shekhter, Sov. Phys. JETP **67**, 163 (1988).
[42] A.-P. Jauho, N. S. Wingreen, and Y. Meir, Phys. Rev. B **50**, 5528 (1994).
[43] O. Entin-Wohlman, Y. Imry, and A. Aharony, Phys. Rev. Lett. **91**, 046802 (2003).
[44] P. Machnikowski and L. Jacak, Phys. Rev. B **71**, 115309 (2005).