Rate equations from the Keldysh formalism applied to the phonon peak in resonant-tunneling diodes

Roger Lake, Gerhard Klimeck, and Supriyo Datta
School of Electrical Engineering, Purdue University, West Lafayette, Indiana 47907
(Received 6 August 1992; revised manuscript received 20 October 1992)

Starting from the Keldysh formalism, general analytical expressions are derived for the current and the occupation of the well in the presence of inelastic scattering, both at the main peak and at the phonon peak. These expressions are then evaluated from a continuous coordinate representation of a double-barrier potential profile and also from a tight-binding model of a weakly coupled central site. The resulting expressions are similar, and the analytical expressions derived from the continuous coordinate representation compare well with the results obtained from numerical simulations. The analytical expressions and the numerical results show that unlike the main peak, the phonon peak is normally independent of the collector transmissivity. But with very opaque collector barriers, the resonant level fills up and the current decreases because the inelastic scattering is suppressed by the exclusion principle. An alternative but equivalent point of view is that the effective coupling Γ'_E between the incident energy in the emitter and the resonant energy in the well is $g'\Gamma_E$ where g' is the effective phonon coupling constant and Γ_E is $\hbar$ times the tunneling rate through the emitter barrier. The total low-temperature inelastic current at the phonon-peak bias is $(2e/\hbar)\Gamma'_E\Gamma_C/(\Gamma'_E+\Gamma_C)$. Since $g'<1$, the effective coupling Γ'_E determines the current until Γ_C is reduced to $\sim\Gamma'_E$. The "backflow" correction to the current due to absorption of phonons is derived, interpreted, and limiting cases discussed. The approach described here could be applied to other problems involving resonant tunneling in the presence of inelastic scattering.

I. INTRODUCTION

The effect of inelastic scattering on resonant tunneling is a question of great interest from both basic and applied points of view.[1] A very clear manifestation of this effect is the appearance of a phonon peak[2-22] in the valley current of resonant-tunneling diodes when the resonant level E_r is one phonon energy $\hbar\omega_0$ below the energy of the incoming electrons from the emitter [Fig. 1(a)]. It is well known that the magnitude of the main peak is proportional to $T_E T_C/(T_E+T_C)$, where $T_{E(C)}$ is the transmission probability of the emitter (collector) barrier. The phonon peak on the other hand can be nearly independent of T_C over a range of values. In this paper, we present a simple analytical model that explains this behavior. The model is based on the Keldysh formalism and is in good agreement with detailed numerical simulations. The model is quite general and should prove useful in other problems involving quantum transport in the presence of inelastic scattering.

The fact that the off-resonant current at the phonon peak is independent of the collector transmission probability may seem counterintuitive since one expects the current to decrease as the collector barrier is made thicker. Suppose we gradually reduce the transmissivity of the collector so that T_C becomes smaller while T_E remains unchanged. The main peak is reduced, but the phonon peak is not reduced; under reverse bias, however, both peaks are reduced since the roles of emitter and collector are interchanged. This is evident from our detailed numerical calculations based on the Keldysh formalism

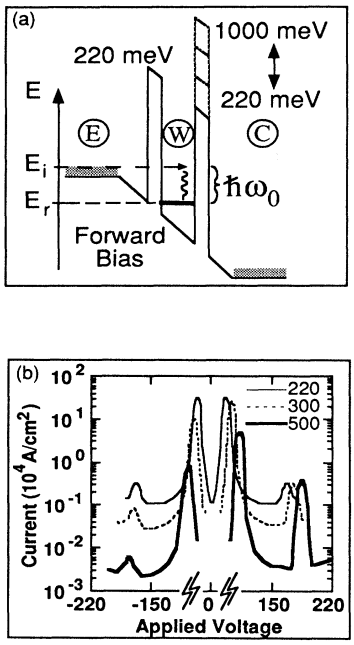

FIG. 1. The asymmetric resonant tunneling diode. $T=4.2$ K; $m^*=0.067m_0$, where m_0 is the free-electron mass. The cross-sectional area is (2 nm)^2. (a) Potential profile used in numerical calculations. The conduction-band discontinuity of the emitter barrier is 220 meV. A series of seven devices was studied with collector barrier conduction-band discontinuities of 220, 300, 400, 500, 600, 700, and 1000 meV. (b) *I-V* characteristics for three of these devices having collector barrier heights of 220, 300, and 500 meV.

[Fig. 1(b)]. Thus, phonon peaks should be more easily observed if an asymmetric structure is biased such that the collector (rather than the emitter) is weakly transmitting. This is in agreement with experimental evidence.[3–6]

However, this does not mean that we can make the collector transmissivity arbitrarily small and still have a phonon-peak current. That is clearly impossible since the current ultimately has to flow out through the collector. When the collector transmissivity gets very small, the resonant level fills up. This reduces the scattering rate for electrons coming in from the emitter since there are fewer empty states into which the electrons can scatter. Consequently the inelastic current is reduced.

An alternative and equivalent point of view is that the effective coupling T_E' between the emitter and the resonance is given by $g'T_E$, where g' is the effective phonon coupling constant. For $g' \ll 1$, the current is determined by T_E' alone, unless the collector barrier is made large enough so that T_C is reduced to T_E'. This is also the point at which the resonance begins to fill.

Experimental observations of the phonon peak occur at low temperatures so that "backflow" due to absorption of phonons by electrons at the resonant energy is negligible. Under such conditions, calculation of the current reduces to a calculation of the scattering rate of electrons injected from the emitter into the well.[12,19–22] At higher temperatures, when absorption cannot be ignored, a "backflow" term appears in the current expression.

The purpose of this paper is threefold. First, we present and evaluate a simple analytical model, derived rigorously from the Keldysh formalism,[23] whose range of applicability is broader than the specific problem of the phonon peak considered here. Second, the previous discussion concerning the effect of barrier asymmetry on the phonon peak is put on a firm quantitative footing. Finally, the "backflow" term for the current is derived and interpreted, and limiting cases are discussed.

II. MODEL

The microscopic model, described in detail in Refs. 24 and 25, is repeated here for convenience. The electrons are described by a single-particle Hamiltonian,

$$H_0=\frac{1}{2m^*}(\mathbf{p}-e\,\mathbf{A})^2+V(\mathbf{r}) \tag{2.1}$$

[$V(\mathbf{r})$ includes the electrostatic potential and conduction-band discontinuities], and interact with a phonon bath maintained in thermodynamic equilibrium,

$$H_R=\sum_{\mathbf{q}}\hbar\omega_{\mathbf{q}}(a_{\mathbf{q}}^{\dagger}a_{\mathbf{q}}+\tfrac{1}{2}) , \tag{2.2}$$

through the electron-phonon interaction

$$H'=\frac{1}{\sqrt{V}}U\sum_{\mathbf{q}}e^{i\mathbf{q}\cdot\mathbf{r}}(a_{\mathbf{q}}e^{-i\omega_0 t}+a_{-\mathbf{q}}^{+}e^{i\omega_0 t}) , \tag{2.3}$$

where $a_{\mathbf{q}}$ is the phonon operator, U is a fixed strength, V is the volume, and $\hbar\omega_0=36$ meV is the optical-phonon energy. The interaction is local in space. This model is similar to that used by Anda and Flores,[17] where it is written in tight-binding notation, and Wingreen, Jacobsen, and Wilkins.[10] The choice of the strength U^2 and the relation between U^2 and the dimensionless coupling constant g is described in Appendix C.

The electron-phonon interaction is treated in the self-consistent first Born approximation. Within this approximation, the self-energies are local. The resulting electron (hole) outscattering rates, $1/\tau_{n(p)}$ are given by the expressions

$$\frac{1}{\tau_n(E)}=\frac{2\pi}{\hbar}U^2\{N_B(\hbar\omega_0)N_0(E+\hbar\omega_0)[1-f(E+\hbar\omega_0)]+[N_B(\hbar\omega_0)+1]N_0(E-\hbar\omega_0)[1-f(E-\hbar\omega_0)]\} , \tag{2.4a}$$

$$\frac{1}{\tau_p(E)}=\frac{2\pi}{\hbar}U^2\{N_B(\hbar\omega_0)N_0(E-\hbar\omega_0)f(E-\hbar\omega_0)+[N_B(\hbar\omega_0)+1]N_0(E+\hbar\omega_0)f(E+\hbar\omega_0)\} , \tag{2.4b}$$

where $N_B(\hbar\omega_0)$ is the Bose-Einstein factor, N_0 is the local density of states, and U is a constant describing the strength of the phonon coupling. At low temperatures (assuming $\hbar\omega_0 \gg k_B T$), (2.4a) and (2.4b) become

$$\frac{1}{\tau_n(E)}=\frac{2\pi}{\hbar}U^2N_0(E-\hbar\omega_0)[1-f(E-\hbar\omega_0)] \tag{2.5a}$$

and

$$\frac{1}{\tau_p(E)}=\frac{2\pi}{\hbar}U^2N_0(E+\hbar\omega_0)f(E+\hbar\omega_0) . \tag{2.5b}$$

The total scattering rate $1/\tau_\phi$ is given by

$$1/\tau_\phi(E)=1/\tau_n(E)+1/\tau_p(E) . \tag{2.6}$$

In (2.4)–(2.6), the position coordinate has been suppressed for clarity. The scattering rates are related to the self-energies in the Keldysh notation by $\Sigma^{<}(z,z';E)=i\hbar/\tau_p(z;E)\delta(z-z')$, $\Sigma^{>}(z,z';E)=-i\hbar/\tau_n(z;E)\delta(z-z')$, and $\mathrm{Im}\Sigma^{R}(z,z';E)=-\hbar/2\tau_\phi(z;E)\delta(z-z')$.

For simplicity, we have considered a strictly one-dimensional picture as in Refs. 10, 11, 14, 16, and 17. In real three-dimensional structures, phonons provide coupling between electron states with different transverse energies. We believe that the main effect of this transverse momentum transfer is to broaden the phonon peak feature in the I-V characteristic. This can be understood as follows. In a one-dimensional treatment, the scattering rate $1/\tau_{n(p)}$ is enhanced only if the resonant energy E_r is exactly one phonon energy below the incident energy E ($E=E_r+\hbar\omega_0$). But in a three-dimensional treatment, at low temperature, the scattering rate for an electron incident with energy E (and zero transverse energy)

is given by [cf. Eq. (2.5a)]

$$\frac{1}{\tau_n(E)}=\frac{2\pi}{\hbar}\sum_{\mathbf{q}}|U_{\mathbf{q}}|^2N_0\left[E-\hbar\omega_0-\frac{\hbar^2\mathbf{q}^2}{2m^*}\right]\times\left[1-f\left[E-\hbar\omega_0-\frac{\hbar^2\mathbf{q}^2}{2m^*}\right]\right], \tag{2.7}$$

where $U_{\mathbf{q}}$ is the coupling strength for phonons with transverse momentum $\hbar\mathbf{q}$. The scattering rate is thus enhanced over a range of energies:

$$E_r+\hbar\omega_0<E<E_r+\hbar\omega_0+\frac{\hbar^2Q^2}{2m^*}$$

(where Q is the maximum value of the phonon wave vector such that $U_{\mathbf{q}}$ is negligible for $q>Q$), instead of at a single energy $E=E_r+\hbar\omega_0$. This leads to a broadening of the phonon peak. Indeed, if $U_{\mathbf{q}}$ were independent of $\mathbf{q}$, Q would be effectively infinite and the phonon peak would transform into a phonon step. Calculations by Turley and Teitsworth[20,22] show that the dominant scattering for narrow wells is due to the inner symmetric interface phonon mode whose coupling $U_{\mathbf{q}}$ does decrease with increasing q, but only as $q^{-1/2}$.[22,26] Consequently, the phonon peak in real three-dimensional structures will be broader than the one-dimensional results presented in this paper. This is in agreement with the experiment.[2-6]

III. FORMALISM

The formalism is described in detail in Refs. 24 and 25. Here we will merely state the results. Using the self-consistent Born approximation for the local phonon model, the Keldysh equations can be written as[27]

$$I(z;E)=\frac{e}{h}\int dz'T(z,z';E)[f(z;E)-f^{\tau}(z';E)]\ . \tag{3.1}$$

We write all of our expressions for the current and the electron density for a single spin channel; a sum over spins is implied. The occupation factor $f(z;E)$ and the scattering factor $f^{\tau}(z';E)$ in (3.1) are defined as follows:

$$f(z;E)\equiv n(z;E)/N_0(z;E)\ , \tag{3.2}$$

$$f^{\tau}(z';E)\equiv\frac{1/\tau_p(z';E)}{1/\tau_{\phi}(z';E)}\ , \tag{3.3}$$

where n is the electron density per unit energy, $n(z;E)=-iG^{<}(z,z;E)/2\pi$, and N_0 is the local density of states which is related to the spectral function A, $N_0(z;E)=A(z,z;E)/2\pi=-\mathrm{Im}G^R(z,z;E)/\pi$. G^R is obtained from the equation (in matrix notation) $(E-H_0-\Sigma^R)G^R=1$. The kernel $T(z,z';E)$ is calculated from the relation

$$T(z,z';E)=\frac{\hbar^2|G^R(z,z';E)|^2}{\tau_{\phi}(z;E)\tau_{\phi}(z';E)}\ . \tag{3.4}$$

Note that $I(z;E)$ in (3.1) is the external current which is nonzero only inside the contacts. Equation (3.1) is first solved for $f(z;E)$ and $f^{\tau}(z;E)$ inside the device subject to the boundary condition: $f(z;E)=f^{\tau}(z;E)=f_0(E-\mu_i)$ if $z\in$ contact "i," where f_0 is the Fermi function and μ_i is the electrochemical potential at contact "i." Once f and f^{τ} are known everywhere, Eq. (3.1) can be used to calculate $I(z;E)$ inside the contacts; integrating $I(z;E)$ over contact "i" yields the terminal current I_i.

Before proceeding, a few comments will be made concerning Eq. (3.1). Equation (3.1) is derived rigorously from the Keldysh formalism. For points lying within the device [$I(z;E)=0$], it follows directly from the integral equation $G^{<}=G^R\Sigma^{<}G^A$. For points lying in the contacts (which are large reservoirs in local equilibrium in zero magnetic field), $I(z;E)\neq 0$ and Eq. (3.1) allow us to calculate the external current as described above. We have shown both analytically and numerically that this is equivalent to a direct evaluation of the current from the relation

$$\mathbf{J}(z;E)=\left[\frac{-e\hbar}{4\pi m^*}(\nabla-\nabla')G^{<}(z,z';E)\right]_{z=z'}-\frac{e^2}{m^*}\mathbf{A}(z)n(z;E)\ .$$

The present formulation incorporates the external current into the same equation as that used to calculate the occupation factor $f(z;E)$. This has two advantages. First, it clearly shows that current is conserved even in arbitrary multiterminal geometries because of the "sum rule" obeyed by the kernel.[24]

$$\int dz'T(z,z'E)=\int dz'T(z',z;E)=\frac{hN_0(z;E)}{\tau_{\phi}(z;E)}\ . \tag{3.5}$$

Using (3.5), we obtain, from (3.1),

$$\int dz\int dE\,I(z;E)=\int dz\int dE\left[\frac{en(z;E)}{\tau_{\phi}(z;E)}-\frac{eN_0(z;E)}{\tau_p(z;E)}\right]=\int dz\int dE\left[\frac{en(z;E)}{\tau_n(z;E)}-\frac{ep(z;E)}{\tau_p(z;E)}\right]\ . \tag{3.6}$$

The integrand represents $\nabla\cdot\mathbf{J}(z;E)$ and as such disappears on integrating over E [this can be shown explicitly by substituting the expressions for τ_p and τ_n (Ref. 28)]. Thus the sum of all the terminal currents is assured to be zero. Second, it expresses the Keldysh equations in a form analogous to those obtained from a Laudauer-Büttiker formalism.[29,30] If we simply extend a multiprobe Büttiker formula[31] to a continuous distribution of probes,[32-35] we obtain

$$I(z;E)=\frac{e}{h}\int dz'T(z,z';E)[f(z;E)-f(z';E)]\ , \tag{3.7}$$

which is similar to Eq. (3.1). The only difference is that the $f^{\tau}(z';E)$ in Eq. (3.1) has been replaced by $f(z';E)$ in (3.7). The relation between f^{τ} and f can be expressed in the form (suppressing the arguments z,E for simplicity)

$$f^{\tau}=f\frac{N_0/\tau_p}{n/\tau_{\phi}}\Longrightarrow f\frac{\text{Inscattering rate}}{\text{Outscattering rate}}\ .$$

If the outscattering and inscattering are balanced in each energy channel, then $f^\tau = f$ and Eq. (3.1) reduces to (3.7). But inelastic processes can cause "vertical flow" from one energy to another so that, in general, $f^\tau \neq f$ within the device.

Alternatively, the effect of inelastic processes on the current equation can be seen by writing Eq. (3.1) as (3.7), plus a correction factor due to "vertical flow." This form of Eq. (3.1) is

$$I(z;E)=\frac{e}{h}\int dz' T(z,z';E)[f(z;E)-f(z';E)] - \int dz' T(z,z';E)\frac{\tau_\phi(z';E)}{hN_0(z';E)}\nabla\cdot\mathbf{J}(z';E) . \quad (3.8)$$

The correction term is clearly zero if current is conserved at each energy. This form for the current with the inelastic term explicitly displayed is similar to that of Eq. (5) of Hershfield, Davies, and Wilkins.[36]

IV. ANALYSIS

The numerical results discussed in this paper are obtained from a direct solution of Eq. (3.1). To obtain simple analytical expressions (for the terminal current and the occupation of the resonant level) we consider a model structure [Fig. 1(a)] consisting of three regions: the emitter contact E, the collector contact C, and the well W, where, in each region, the occupation factor f is assumed to be independent of position. Transmission coefficients between the regions are defined by integrating $T(z,z';E)$ over the regions. For example,

$$T_{E,C}=\int_{z\in E}dz\int_{z'\in C}dz' T(z,z';E) . \quad (4.1)$$

Since we are not considering magnetic fields, the transmission coefficients are symmetric.

With the above definitions, the current per unit energy in the emitter contact is obtained from (3.1):

$$I_E=\frac{e}{h}[T_{E,C}(f_E-f_C)+T_{E,W}(f_E-f_W^\tau)] . \quad (4.2)$$

In the well, $I(z;E)$ is zero and (3.1) gives a general expression for the occupation of the well.

$$f_W=\frac{T_{E,W}f_E+T_{C,W}f_C+T_{W,W}f_W^\tau}{T_{E,W}+T_{C,W}+T_{W,W}} . \quad (4.3)$$

Note that all the quantities appearing in Eqs. (4.2) and (4.3) are functions of energy, E. In particular, we are interested in two values of energy, the energy E_i of the incident electrons and the energy E_r of the resonant level.

As discussed earlier, the main peak and the phonon peak behave very differently as the collector transmissivity is reduced. This difference arises because the function $f^\tau(z;E)$ is very different in the two cases. At the bias corresponding to the main peak $E_i \simeq E_r$, and there is effectively only one energy channel; under these conditions, $f^\tau = f$ as discussed earlier. For this case, Eqs. (4.2) and (4.3) yield

$$I_E=\frac{e}{h}\left[T_{E,C}(f_E-f_C)+\frac{T_{E,W}T_{C,W}}{T_{E,W}+T_{C,W}}(f_E-f_C)\right] \quad \textit{main peak} \quad (4.4)$$

and

$$f_W=\frac{T_{E,W}f_E+T_{C,W}f_C}{T_{E,W}+T_{C,W}} \quad \textit{main peak} . \quad (4.5)$$

The coherent component of the current is given by the first term in (4.4) and the sequential component of the current is given by the second term. Note that the sequential component depends on $T_{C,W}$.

In contrast, at the phonon peak $E_i = E_r + \hbar\omega_0$ and, in the low-temperature limit, it is apparent from Eqs. (2.5) that, $f_W^\tau(E_i)\simeq 0$, since $1/\tau_p(E_i)\simeq 0$, while $f_W^\tau(E_r)\simeq 1$, since $1/\tau_n(E_r)\simeq 0$. For low temperature and high bias, $f_C \simeq 0$ and $f_E \simeq 1$. Hence, from (4.2),

$$I_E(E_i)=\frac{e}{h}[T_{E,C}+T_{E,W}] \quad \textit{phonon peak} . \quad (4.6)$$

Also, noting that $T_{E,W}(E_r)=0$, from (4.3) we obtain

$$f_W(E_r)=\left[\frac{T_{C,W}}{T_{W,W}}+1\right]^{-1} \quad \textit{phonon peak} . \quad (4.7)$$

The coherent component of the current is given by the first term in (4.6) and the current in the inelastic channel is given by the second term. Note that this inelastic component is independent of $T_{C,W}$, unlike the sequential component in (4.4).

To obtain useful analytical expressions for I_E and f_W, we must evaluate the transmission coefficients. We have done so in two ways (the appendixes contain details of the derivations). In the first, we calculate the Green functions in (3.4) working in the continuous (coordinate) representation for a structure similar to that of Fig. 1, except that the potential drop is approximated as steplike occurring at the two barriers. In the second, we use a one-dimensional tight-binding model with the central site weakly coupled to the emitter (collector) lead. This model is similar to the one investigated by Hershfield, Davies, and Wilkins with the Hubbard U repulsion replaced by the electron-phonon interaction.[36] For both methods, we assume that the scattering times $\tau_{n(p)}$ and τ_ϕ are independent of position within any one region.

From the calculation in the continuous coordinate representation, we find

$$T_{E,C}=\frac{\Gamma_E\Gamma_C}{\Gamma}A , \quad (4.8)$$

$$T_{E,W}=\frac{\Gamma_E\hbar/\tau_\phi}{\Gamma}AF_{E,W} , \quad (4.9)$$

$$T_{C,W}=\frac{\Gamma_C\hbar/\tau_\phi}{\Gamma}AF_{C,W} . \quad (4.10)$$

Also, from the sum rule [Eq. (3.5)],

$$T_{E,W}+T_{C,W}+T_{W,W}=\frac{\hbar}{\tau_\phi}A , \quad (4.11)$$

where $\Gamma=\Gamma_C+\Gamma_E+\hbar/\tau_\phi$, and $1/\tau_\phi$ is the scattering rate in the well. The emitter (collector) tunneling rate $\Gamma_{E(C)}/\hbar$ is related to the transmission probability $T_{E(C)}$ of the emitter (collector) barrier by the relation $\Gamma_{E(C)}=\hbar v T_{E(C)}$, where $v=v/2d$ is the attempt frequency in the well (v is the velocity and d is the well width). The spectral function in the well, $A(E)$, defined by $\int_{\text{well}} dz\ A(z,z;E)$, is

$$A(E)=\frac{\Gamma}{[4(\hbar v)^2\sin^2(\theta/2)+\frac{1}{4}\Gamma^2]}\ , \tag{4.12}$$

where θ is the round-trip phase shift in the well; $\theta=2kd+\phi_E+\phi_C$, where $\phi_{E(C)}$ is the phase of the reflection amplitude of the emitter (collector) barrier. The factors $F_{E,W}$ and $F_{C,W}$ depend on the propagation and reflection phase shifts. They are of order one and will be neglected in the following discussion. In this case our equations (4.2) and (4.3) can be cast in a form similar to that found by Hershfield, Davies, and Wilkins.[36,37]

The tight-binding calculation yields the same result as Eqs. (4.8)–(4.11) (without the factors $F_{E(C),W}$), except that the spectral function has the form

$$A(E)=\frac{\Gamma}{[(E-E_r)^2+\frac{1}{4}\Gamma^2]}. \tag{4.13}$$

Close to resonance, $4(\hbar v)^2\sin^2(\theta/2)\simeq(E-E_r)^2$, so that the two expressions are equivalent, but for the off-resonant current, using Eq. (4.12) for A gives much better quantitative agreement with the numerical results.

A. Main peak

With the above results for the transmission coefficients, Eq. (4.4) for the current per unit energy at the main peak becomes

$$I_E(E_i)=\frac{e}{h}\frac{\Gamma_E\Gamma_C}{\Gamma_E+\Gamma_C}A(f_E-f_C)\ . \tag{4.14}$$

Integrating over energy, assuming $\Gamma_{E(C)}$ and $f_{E(C)}$ are slowly varying compared to A, results in the usual expression for the current at the main peak:

$$I=\frac{e}{\hbar}\frac{\Gamma_E\Gamma_C}{\Gamma_E+\Gamma_C}(f_E-f_C)\ , \tag{4.15}$$

where the quantities are evaluated at the resonant energy. If we were to use (4.14) to evaluate the off-resonant current, we would obtain an expression identical to that found by Büttiker.[38] From (4.5), the occupation of the resonance becomes

$$f_W(E_r)=\frac{\Gamma_E f_E+\Gamma_C f_C}{\Gamma_E+\Gamma_C}\ . \tag{4.16}$$

B. Phonon peak

First we consider the general expression for the current and analyze the role of "backflow." Next we write the high bias, low-temperature limits of the equation for the current and the occupation of the resonance, and analyze the effect of the collector barrier asymmetry.

1. "Backflow"

The general expression for the emitter current per unit energy [Eq. (4.2)], yields

$$I_E(E_i)=\frac{e}{h}\frac{A}{\Gamma}\left[\Gamma_E\Gamma_C(f_E-f_C)+\Gamma_E\hbar/\tau_\phi\left[f_E-\frac{\hbar/\tau_p}{\hbar/\tau_\phi}\right]\right]\ . \tag{4.17}$$

The first term in the brackets is the coherent current and the second term is the inelastic current. Consider the inelastic current:

$$I_E^{\text{inelastic}}(E_i)=\frac{e}{h}\frac{A}{\Gamma}\Gamma_E(f_E\hbar/\tau_\phi-\hbar/\tau_p)$$
$$=\frac{e}{h}\frac{A}{\Gamma}\Gamma_E[f_E\hbar/\tau_n-(1-f_E)\hbar/\tau_p]\ , \tag{4.18}$$

where we have used (2.6). To give a physical interpretation to the two terms within the brackets of (4.18), we use the expression for the occupation of the well at the incident energy. From Eq. (4.3), we obtain

$$f_W=\frac{\Gamma_E f_E+\Gamma_C f_C+\hbar/\tau_p}{\Gamma}\ , \tag{4.19}$$

where all of the terms in (4.17)–(4.19) are evaluated at the incident energy, E_i. The electron density per unit energy in the well is, therefore,

$$n_W(E_i)=\frac{A}{2\pi}f_W=\frac{1}{2\pi}\frac{A}{\Gamma}[\Gamma_E f_E+\Gamma_C f_C+\hbar/\tau_p]\ . \tag{4.20}$$

The first two terms in the brackets represent the contribution to $n_W(E_i)$ due to injection from the emitter and collector contacts, respectively, at energy E_i. The last term represents the contribution to $n_W(E_i)$ due to "backflow" to the incident energy from the resonant energy E_r. Thus we can write $n_W(E_i)$ as the sum of three contributions:

$$n_W(E_i)=n_{W\leftarrow E}+n_{W\leftarrow C}+n_{W(E_i)\leftarrow W(E_r)}\ , \tag{4.21a}$$

where

$$n_{W\leftarrow E(C)}=\frac{1}{2\pi}\frac{A}{\Gamma}\Gamma_{E(C)}f_{E(C)} \tag{4.21b}$$

and

$$n_{W(E_i)\leftarrow W(E_r)}=\frac{1}{2\pi}\frac{A}{\Gamma}\hbar/\tau_p\ . \tag{4.21c}$$

The interpretation of the first two terms is verified by injecting a plane wave from the emitter and calculating the resulting electron density per unit energy in the well (details of the calculation are shown at the ends of Appendixes A and B). The result is precisely Eq. (4.21b).

This interpretation is now applied to the inelastic current equation (4.18). We rewrite (4.18) in the transparent form

$$I_E^{\text{inelastic}}(E_i)=e\left[\frac{n_{W\leftarrow E}}{\tau_n}-n_{W(E_i)\leftarrow W(E_r)}\frac{\Gamma_E}{\hbar}(1-f_E)\right] . \tag{4.22}$$

The first term is the rate at which electrons injected from the emitter are scattered down to the resonance. This is the term calculated in the approach of Chevoir and Vinter.[12] The second term is the rate at which electrons are scattered up to the incident energy from the resonant energy and back out to the emitter. This is the correction term due to "backflow."

Field theories of many-particle systems treat electrons and holes in the conduction band on an equal footing. This suggests an interesting interpretation for the "backflow" term. The hole density per unit energy in the well due to injection of holes from the emitter is [cf. Eq. (4.21b)]

$$p_{W\leftarrow E}=\frac{1}{2\pi}\frac{A}{\Gamma}\Gamma_E(1-f_E) . \tag{4.23}$$

Therefore, the "backflow" term is $p_{W\leftarrow E}/\tau_p$ and the equation for the inelastic current takes the symmetric form

$$I_E^{\text{inelastic}}(E_i)=e\left[\frac{n_{W\leftarrow E}}{\tau_n}-\frac{p_{W\leftarrow E}}{\tau_p}\right] . \tag{4.24}$$

In this view, the "backflow" term is the outscattering rate in the well of holes injected from the emitter.

There are three cases in which "backflow" is negligible. "Backflow" is proportional to the hole outscattering rate at the incident energy which, using (2.4b), is

$$\frac{p_{W\leftarrow E}}{\tau_p(E_i)}\simeq\frac{2\pi}{\hbar}U^2N_B(\hbar\omega_0)N_0(E_r)f(E_r)p_{W\leftarrow E} .$$

If $\hbar\omega_0\gg k_BT$, then $N_B(\hbar\omega_0)\simeq 0$. Thus, "backflow" is negligible at low temperatures. Second, if the occupation of the resonance is very small, $f(E_r)\simeq 0$, then $1/\tau_p(E_i)\simeq 0$ and there is no "backflow." Finally, since $p_{W\leftarrow E}\propto 1-f_E(E_i)$, "backflow" is suppressed if the injection energy lies $3k_BT$ or more below the Fermi energy of the emitter. For all three cases, the equation for the inelastic current reduces to the form used by Chevoir and Vinter:[12]

$$I_E^{\text{inelastic}}=e\frac{n_{W\leftarrow E}}{\tau_n} . \tag{4.25}$$

2. *Low-temperature limit*

In the low-temperature, high-bias limit, at the bias corresponding to the phonon peak, Eq. (4.6) yields

$$I_E(E_i)=\frac{e}{h}\frac{A}{\Gamma}[\Gamma_E\Gamma_C+\Gamma_E\hbar/\tau_\phi]$$
$$\simeq\frac{e}{h}\frac{A}{\Gamma}[\Gamma_E\Gamma_C+\Gamma_E\hbar/\tau_n] . \tag{4.26}$$

All of the quantities in (4.26) are evaluated at the energy E_i of the incident carriers. The second term in the brackets is the current in the inelastic channel. Since E_i is $\hbar\omega_0$ above the resonant energy,

$$A(E_i)/\Gamma\simeq[4(\hbar v)^2\sin^2\theta/2]^{-1} \tag{4.27}$$

and the current in the inelastic channel has no explicit dependence on Γ_C. The inelastic current depends only on the collector barrier through the effect of the factor of $[1-f(E_r)]$ on the value of $1/\tau_n(E_i)$ [see Eq. (2.5a)].

The total inelastic current is

$$\int dE_i\frac{e}{h}\frac{A(E_i)}{\Gamma(E_i)}\Gamma_E(E_i)\tilde{U}^2A(E_i-\hbar\omega_0)$$
$$\times[1-f_W(E_i-\hbar\omega_0)], \tag{4.28}$$

where we have used (2.5a), we write $\hbar/\tau_n(E_i)$ in (4.28) as

$$\hbar/\tau_n(E_i)=\tilde{U}^2A(E_i-\hbar\omega_0)[1-f_W(E_i-\hbar\omega_0)] . \tag{4.29}$$

$A(E)$ is given by (4.12) and $\tilde{U}^2=U^2/V_\square$, where $V_\square$ is the volume of the quantum box between the barriers. $V_\square$ accounts for the magnitude of the wave function, i.e., $2\pi N_0(z;E)\simeq A(E)/V_\square$, where z lies in the well. Since $A(E)$ is strongly peaked at the resonant energy E_r, the major contribution to the integral in (4.28) will come from values of $E_i\simeq E_r+\hbar\omega_0$. If we assume that $f_W(E_r)$ is slowly varying compared to $A(E_r)$ [this assumption is confirmed by Eq. (4.35)], we can move all of the quantities outside of the integral except $A(E_i-\hbar\omega_0)$, whose integral gives a factor of 2π. Thus the emitter current in the inelastic channel is

$$I_E=\frac{e}{\hbar}\frac{A(E_r+\hbar\omega_0)}{\Gamma(E_r+\hbar\omega_0)}\Gamma_E(E_r+\hbar\omega_0)\tilde{U}^2[1-f_W(E_r)] . \tag{4.30}$$

We define an effective phonon coupling constant g' as

$$g'=\frac{\tilde{U}^2}{4(\hbar v)^2\sin^2\frac{\theta(E_i)}{2}}=g\frac{(\hbar\omega_0)^2}{4(\hbar v)^2\sin^2\frac{\theta(E_i)}{2}} , \tag{4.31}$$

with the usual dimensionless phonon coupling constant $g=\tilde{U}^2/(\hbar\omega_0)^2$ from Eq. (C9). Then, using (4.27), the total emitter current in the inelastic channel becomes

$$I_E=\frac{e}{\hbar}g'\Gamma_E(E_i)[1-f_W(E_r)] , \tag{4.32}$$

where $E_i=E_r+\hbar\omega_0$. Equation (4.32) shows clearly that the inelastic current due to optical-phonon emission depends on the collector barrier only through the occupation of the resonance $f_W(E_r)$.

The occupation factor $f_W(E_r)$ is obtained from (4.7):

$$f_W(E_r)=\left[\frac{\Gamma_C}{\hbar/\tau_\phi}+1\right]^{-1}$$
$$\simeq\left[\frac{\Gamma_C}{\hbar/\tau_p}+1\right]^{-1} . \tag{4.33}$$

All of the quantities in (4.33) are evaluated at the energy E_r of the resonant level. Although Eq. (4.33) does not ex-

plicitly contain a factor of Γ_E, the resonance cannot be filled if the emitter barrier transmissivity becomes very small. The effect of the emitter barrier transmissivity on the occupation of the resonance is contained in the scattering rate $1/\tau_p(E_r)$ through the occupation factor at the incident energy, $f(E_i)$. From (2.5b), $\hbar/\tau_p(E_r) = \tilde{U}^2 A(E_i) f_W(E_i)$, where $A(E_i)$ is given by (4.12). Using (4.19) and (4.27) and remembering that $\hbar/\tau_p(E_i) \simeq 0$, we see that

$$\hbar/\tau_p(E_r) \simeq \frac{\tilde{U}^2 \Gamma_E(E_i)}{4(\hbar v)^2 \sin^2 \dfrac{\theta(E_i)}{2}} = g' \Gamma_E(E_i) \ . \tag{4.34}$$

Using (4.34), (4.33) becomes

$$f_W(E_r) = \left[\frac{\Gamma_C(E_r)}{g' \Gamma_E(E_i)} + 1 \right]^{-1} \ . \tag{4.35}$$

Thus, filling of the resonance at the phonon-peak bias occurs for $\Gamma_C \lesssim g' \Gamma_E$.

Substituting the expression for the occupation of the resonance (4.35) into the expression for the current (4.32), we obtain our final expression for the inelastic component of the phonon-peak current:

$$I_E = \frac{e}{\hbar} \frac{g' \Gamma_E \Gamma_C}{g' \Gamma_E + \Gamma_C} \ . \tag{4.36}$$

The form of Eq. (4.36) for the phonon-peak current is identical to the expression for the low-temperature main-peak current [see Eq. (4.15)], with the tunneling rate through the emitter barrier Γ_E scaled by the coupling constant g'. Our final low-temperature results show that for weakly transmitting barriers at the phonon-peak bias, the coupling of the emitter at the incident energy to the well at the resonant energy is the usual coupling through the emitter barrier Γ_E scaled by the phonon coupling constant g'.

V. COMPARISON OF ANALYTICAL AND NUMERICAL RESULTS

For the analytical calculations in this paper, we have used the values of the scattering rates obtained from the numerical simulations. However, we have checked that a self-consistent solution of (4.2) and (4.3) with (2.5a) and (2.5b) yields results in agreement with the full numerical solution. Figures 2(a) and 2(b) show the current I and the occupation factor of the resonance $f_W(E_r)$ for a series of devices at a temperature of 4.2 K with different collector barrier heights [Fig. 1(a)], both at the phonon peak and at the main peak. It is apparent that there is good agreement between the numerical results computed from Eq. (3.1) and the results obtained from the analytical expressions in Eqs. (4.15), (4.16), (4.26), and (4.33). Figure 2(c) shows $\hbar/\tau_\phi(E_r)$ and $\hbar/\tau_\phi(E_i)$ calculated numerically from Eqs. (2.4) at a temperature of 4.2 K. Note that, at the phonon-peak bias, the occupation factor $f_W(E_r)$ in the well becomes significant when Γ_c becomes smaller than $\hbar/\tau_\phi(E_r)$, as predicted by Eq. (4.33). At this point, $\hbar/\tau_\phi(E_i)$ decreases and the phonon peak is suppressed. However, due to the low density of states at E_i, $\tau_\phi(E_r)$ is very long [see Eq. (2.5b)] and Γ_C has to be quite small before filling of the well affects the phonon-peak current.

It is interesting to note [Fig. 2(b)] that, while the filling of the resonant level at the main peak is close to unity for $\Gamma_E/\Gamma_C > 10$, it takes greater asymmetry, by a factor of $1/g'$, at the phonon peak for significant filling to occur. In highly asymmetric structures,[3–6] charging could affect the shape of the phonon peak just as it affects the main peak.

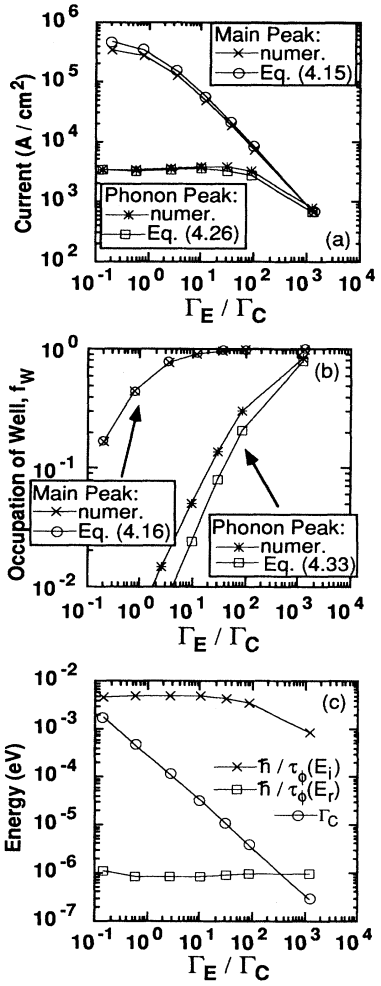

FIG. 2. Numerical results and analytical predictions are compared for (a) the magnitude of the current and (b) the occupation of the well, at the bias corresponding to the main current peak and at the bias corresponding to the phonon peak; (c) $\Gamma_C(E_i)$, $\hbar/\tau_\phi(E_i)$, and $\hbar/\tau_\phi(E_r)$ calculated numerically at the phonon-peak bias.

VI. CONCLUSIONS

In summary, starting from the Keldysh formalism, we obtain Eqs. (4.2) and (4.3) for the terminal current I and the occupation factor f_W in the well. The transmission coefficients appearing in Eqs. (4.2) and (4.3) are evaluated using the continuous (coordinate) representation [see Eqs. (4.8)–(4.11)]. Equations (4.2) and (4.3) are then used to calculate I and f_W at the main peak [Eqs. (4.15) and (4.16)] and at the phonon peak [Eqs. (4.26) and (4.33)]. The results agree well with detailed numerical calculations. The expression for the low-temperature total inelastic current at the phonon-peak bias Eq. (4.32) shows clearly that the inelastic current through a weakly coupled well only depends on the collector barrier through the occupation of the resonant level in the well. Substituting in the expression for the occupation of the resonant level [Eq. (4.35)] gives the final expression for the low-temperature inelastic current [Eq. (4.36)]. Equation (4.36) shows that the effective coupling between the incident energy in the emitter and the resonant energy in the well is $g'\Gamma_E$. Since $g'<1$, the effective emitter-well coupling is reduced. Until Γ_C is reduced to $\sim g'\Gamma_E$, the current is determined by $g'\Gamma_E$ alone. When $\Gamma_c \leq g'\Gamma_E$, charging of the well becomes significant at the phonon-peak bias. The high-temperature "backflow" term can be interpreted as the outscattering rate in the well of holes injected from the emitter. This term is suppressed if the injection energies lie several $k_B T$ below the Fermi level of the emitter, if the resonant level is empty or the temperature low.

ACKNOWLEDGMENTS

We thank P. Turley and S. Teitsworth for making available to us copies of their work and for valuable discussions. We acknowledge helpful discussions with M. P. Anantram. This work was supported by the Semiconductor Research Corporation under Contract No. 91-SJ-089.

APPENDIX A: DERIVATION OF THE TRANSMISSION COEFFICIENTS IN THE CONTINUOUS COORDINATE REPRESENTATION

The Green functions are calculated for the structure in Fig. 3. First we calculate $T_{E,C}$ as defined in Eq. (4.1). As described in Appendix C of Ref. 25, the integrations over the contacts result in factors of velocity,

$$T_{E,C}=\hbar^2|G^R(z_E,z_C;E)|^2 v_E v_C \ , \tag{A1}$$

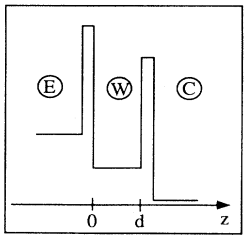

FIG. 3. The potential profile used for the analytical calculations of Appendix A.

where $v_{E(C)}$ is the velocity in the emitter (collector) contact. The quantity $G^R(z_E,z_C;E)$ can be calculated by matching wave functions and derivatives at the various interfaces. It can also be calculated more easily by summing multiple reflection paths (see Fig. 1.9 in Ref. 39). The result is

$$G^R(z_E,z_C;E)=\frac{-i}{\hbar v_E}\frac{t_E t_C e^{ikd}}{1-r_E r_C e^{i2kd}} \ , \tag{A2}$$

where $t_{E(C)}$ and $r_{E(C)}$ are, respectively, the transmission and reflection amplitudes through the emitter (collector) barriers. The complex wave vector in the well is

$$k=\frac{1}{\hbar}\left[2m^*\left[E-V+\frac{i\hbar}{2\tau_\phi}\right]\right]^{1/2}$$
$$\simeq\frac{1}{\hbar}\sqrt{2m^*(E-V)}+\frac{i}{2v_W\tau_\phi} \ ,$$

where $v_W=\sqrt{2(E-V)/m^*}$ and τ_ϕ is the scattering time in the well. Substituting (A2) into (A1) gives

$$T_{E,C}=\frac{v_C}{v_E}|t_E|^2|t_C|^2\frac{e^{-N}}{1+R_E R_C e^{-2N}-2\sqrt{R_E R_C}e^{-N}\cos\theta}$$
$$=\frac{T_E T_C e^{-N}}{(1-\sqrt{R_E R_C}e^{-N})^2+4\sqrt{R_E R_C}e^{-N}\sin^2\frac{\theta}{2}} \ , \tag{A3}$$

where $T_E=(v_W/v_E)|t_E|^2$, $T_C=(v_C/v_W)|t_C|^2$, $R_{E(C)}=|r_{E(C)}|^2$, and $N=d/v_W\tau_\phi=1/2\nu\tau_\phi$. Assuming that

$$T_E, T_C, N \ll 1 \ , \tag{A4}$$

which is always the case for the structures we consider,

$$T_{E,C}\simeq\frac{\Gamma_E\Gamma_C}{4(\hbar\nu)^2\sin^2\frac{\theta}{2}+\frac{1}{4}\left[\Gamma_E+\Gamma_C+\frac{\hbar}{\tau_\phi}\right]^2} \ . \tag{A5}$$

Using the form of the spectral function, Eq. (4.12), (A5) becomes (4.8).

To calculate the spectral function, we need $G^R(z,z;E)$, where z lies in the well. Again, the simplest approach is to sum multiply reflected paths. The result is

$$G^R(z,z;E)=\frac{-i}{\hbar v_W}\frac{(1+r_E e^{i2kz})(1+r_C e^{i2k(d-z)})}{1+r_E r_C e^{i2kd}} \ . \tag{A6}$$

Now we integrate over the well region and take twice the imaginary part to find

$$A=\frac{1}{\hbar\nu}\frac{1-R_E R_C e^{-2N}}{1+R_E R_C e^{-2N}-2\sqrt{R_E R_C}e^{-N}\cos\theta}+F_W \ , \tag{A7}$$

where F_W is a correction term. Using (A4), the first term of (A7) reduces to (4.12). The correction term is

$$F_W=\frac{1}{\hbar\nu}\mathrm{Re}\left\{\frac{(r_E+r_C)e^{ikd}}{1-r_Er_Ce^{i2kd}}\frac{\sin(kd)}{kd}\right\}. \quad \text{(A8)}$$

This term is small and can be shown to be identically equal to zero for $N=0$ and $R_{E(C)}=1$.

Finally, we calculate $T_{E,W}$, where

$$T_{E,W}=\frac{\hbar^2 v_E}{\tau_\phi}\int_W dz|G^R(z,z_E;E)|^2 . \quad \text{(A9)}$$

The Green function is

$$G^R(z,z_E;E)=\frac{-i}{\hbar v_E}t_E\frac{e^{ikz}+r_Ce^{ik(2d-z)}}{1-r_Er_Ce^{i2kd}} , \quad \text{(A10)}$$

where z lies in the well. Substituting (A10) into (A9) gives

$$T_{E,W}=\frac{1}{\hbar\nu}\frac{T_E\hbar/\tau_\phi}{1+R_ER_Ce^{-2N}-2\sqrt{R_ER_C}e^{-N}\cos\theta}F_{E,W} . \quad \text{(A11)}$$

In the limit (A4), (A11) reduces to (4.9). The correction factor is

$$F_{E,W}=\frac{(1-e^{-N})(1+R_Ce^{-N})}{2N}+\mathrm{Re}\left\{\sqrt{R_C}e^{-\theta_C}\frac{e^{i2kd}-1}{i2kd}\right\}. \quad \text{(A12)}$$

Again, using (A4),

$$F_{E,W}\simeq 1+\cos(\beta d+\phi_C)\frac{\sin\beta d}{\beta d}\simeq 1 , \quad \text{(A13)}$$

where $\beta=\mathrm{Re}k$. The expression for $T_{C,W}$ is obtained from the expression for $T_{E,W}$ by replacing the subscript E with the subscript C.

In all of our analytical calculations, we have ignored the factors F_W and $F_{E(C),W}$. The good agreement between the analytical results and the numerical results (see Fig. 2) indicates that this approximation is justified.

The calculation of the electron density in the well due to the injection of a plane wave from the emitter begins with the calculation of the resulting wave function in the well:

$$\psi(z)=t_E\frac{e^{ikz}+r_Ce^{ik(2d-z)}}{1-r_Er_Ce^{i2kd}} , \quad \text{(A14)}$$

where z lies in the well. Comparing with Eq. (A10), we obtain

$$\psi(z)=i\hbar v_E G^R(z,z_E;E) , \quad \text{(A15)}$$

where z_E is the point at the emitter-device interface. The corresponding electron density per unit energy in the well is

$$n(z;E)=\frac{1}{2\pi\hbar v_E(E)}|\psi(z;E)|^2 , \quad \text{(A16)}$$

where $1/2\pi\hbar v_E(E)$ is the density of states in the emitter. Integrating $n(z;E)$ over the well gives the final result:

$$n_{W\leftarrow E}=\frac{1}{2\pi}f_E\hbar v_E\int_W dz|G^R(z,z_E;E)|^2 =f_E\frac{1}{2\pi\hbar/\tau_\phi}T_{E,W} . \quad \text{(A17)}$$

In the limit of (A4), Eq. (A17) reduces to (4.21b).

APPENDIX B: TIGHT-BINDING CALCULATION OF TRANSMISSION COEFFICIENTS

We consider a one-dimensional tight-binding model with the electron-phonon interaction at the $n=0$ site (Fig. 4). The central site is weakly coupled to the emitter (collector) lead via the hopping matrix element $W_{E(C)}$. The hopping matrix elements in the leads are identical and equal to W. The site energies in the leads differ by the applied voltage $e|V|$.

The Hamiltonian in tight-binding form is

$$H=a\sum_j\{W_{j,j+1}|j\rangle\langle j+1|+\varepsilon_j|j\rangle\langle j|\} ,$$

where $|j\rangle$ is a state localized around site j, the symmetric nearest-neighbor coupling W_{jk} is as shown in Fig. 4, and the site energy is ε_j. Since the electron-phonon interaction is local, we include the self-energy Σ^R in the site energy so that ε_j contains an imaginary part $-i\hbar/2\tau_\phi$. We have chosen the normalization such that $\langle j|j\rangle=1/a$.

To calculate the transmission coefficients $T_{E,C}$, $T_{E,W}$, and $T_{W,W}$, we need the Green functions $G^R(-1,1;E)$, $G^R(-1,0;E)$, and $G^R(0,0;E)$. First we calculate $G^R(0,0;E)$ since it will give us the spectral function A. The three coupled equations centered at the sites -1, 0, and 1 are, respectively,

$$\begin{aligned} &WG_{-2}+(E-\varepsilon_L)G_{-1}+W_EG_0=0 ,\\ &W_EG_{-1}+(E-\varepsilon_0)G_0+W_CG_1=1/a ,\\ &W_CG_0+(E-\varepsilon_R)G_1+WG_2=0 , \end{aligned} \quad \text{(B1)}$$

where $G_i=G^R(i,0;E)$. Since the contacts are uniform, the Green function simply propagates as a plane wave so that $G_{\pm 2}=e^{ik_{\pm 1}a}G_{\pm 1}$. This relation is used to close the system of equations (B1). The Green function $G^R(i,j;E)$ for $i,j\in\{-1,0,1\}$ is thus

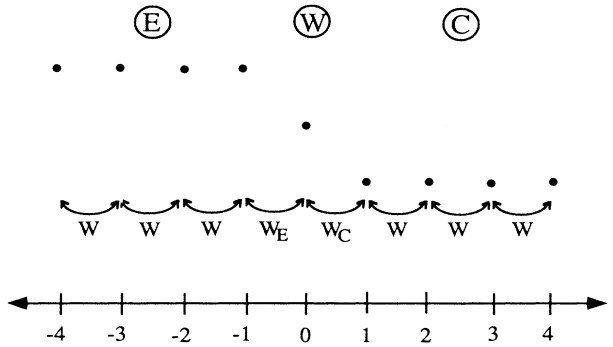

FIG. 4. The tight-binding chain. The coupling between the central site and the emitter (collector) lead is $W_{E(C)}$. The coupling between all other sites is W. The lattice spacing is a.

$$[G^R]=\frac{1}{a}\begin{bmatrix} -We^{-ik_E a} & W_E & 0 \\ W_E & E-\varepsilon_0 & W_C \\ 0 & W_C & -We^{-ik_C a} \end{bmatrix}^{-1}, \quad \text{(B2)}$$

where we have used the dispersion relations

$$E=\varepsilon_{E(C)}-2W\cos(k_{E(C)}a) . \quad \text{(B3)}$$

Inverting the matrix in (B2), we find

$$G^R(0,0;E)=\frac{1/a}{E-\tilde{\varepsilon}_0+i\frac{1}{2}(\Gamma_E+\Gamma_C+\hbar/\tau_\phi)} , \quad \text{(B4)}$$

$$G^R(-1,0;E)=\frac{W_E}{W}e^{ik_E a}G^R(0,0;E) , \quad \text{(B5)}$$

and

$$G^R(1,-1;E)=\frac{W_E}{W}e^{ik_E a}\frac{W_C}{W}e^{ik_c a}G^R(0,0;E) , \quad \text{(B6)}$$

where

$$\tilde{\varepsilon}_0=\left[\varepsilon_0+\frac{i\hbar}{2\tau_\phi}\right]-\frac{W_E^2}{W}\cos(k_E a)-\frac{W_C^2}{W}\cos(k_C a) \quad \text{(B7)}$$

(note that $\tilde{\varepsilon}_0$ is real),

$$\Gamma_{E(C)}=\left[\frac{W_{E(C)}}{W}\right]^2\frac{\hbar v_{E(C)}}{a} , \quad \text{(B8)}$$

and

$$v_{E(C)}=\frac{2aW}{\hbar}\sin(k_{E(C)}a) . \quad \text{(B9)}$$

The spectral function A used in definitions (4.8)–(4.11) is $A(0,0;E)$ multiplied by the lattice spacing a (the tight-binding equivalent of integrating over the central site):

$$A=\frac{\Gamma}{(E-\tilde{\varepsilon}_0)^2+\frac{1}{4}\Gamma^2}, \quad \text{(B10)}$$

where $\Gamma=\Gamma_C+\Gamma_E+\hbar/\tau_\phi$.

Using the definition of the transmission coefficients (4.1), and remembering that integration over a contact results in a factor of velocity and integration over the well results in a factor of a, relations (4.8)–(4.11) are obtained.

$$T_{E,C}=\hbar^2 v_E v_C|G^R(1,-1;E)|^2=\frac{\Gamma_E\Gamma_C}{\Gamma}A , \quad \text{(B11)}$$

$$T_{E,W}=av_E\frac{\hbar^2}{\tau_\phi}|G^R(-1,0;E)|^2=\frac{\Gamma_E\hbar/\tau_\phi}{\Gamma}A . \quad \text{(B12)}$$

Similarly,

$$T_{C,W}=\frac{\Gamma_C\hbar/\tau_\phi}{\Gamma}A . \quad \text{(B13)}$$

Finally,

$$T_{W,W}=a^2[\hbar/\tau_\phi]^2|G^R(0,0;E)|^2=\frac{[\hbar/\tau_\phi]^2}{\Gamma}A . \quad \text{(B14)}$$

Note that the sum $T_{E,W}+T_{C,W}+T_{W,W}$ does indeed satisfy the sum rule (4.11).

We now calculate the electron density in the well due to injection of a plane wave from the emitter. For sites $n\neq 0$, the wave function is written as

$$\psi_n=\begin{cases} e^{ik_E(n+1)a}+re^{-ik_E(n+1)a}, & n<0 \\ te^{ik_C(n-1)a}, & n>0 . \end{cases} \quad \text{(B15)}$$

The equations for ψ_i are identical to the equations for G_i, Eq. (B1), with $1/a$ replaced by 0 on the right-hand side. Substituting in the expressions for $\psi_{i=0}$ from (B15) into the three equations for ψ_i, $i\in\{-1,0,1\}$, using the dispersion relation (B3) results in

$$\psi_0=i\hbar v_E G^R(-1,0;E) , \quad \text{(B16)}$$

with v_E given by (B9) and $G^R(-1,0;E)$ given by (B5). The rest follows as in (A15)–(A17), with $T_{E,W}$ now defined by (B12).

APPENDIX C: CALCULATION OF SCATTERING STRENGTH U^2 AND THE COUPLING CONSTANT g

In the self-consistent first Born approximation (SCFBA), the self-energy $\Sigma^>$ is defined as

$$\Sigma^>(\mathbf{r},\mathbf{r}';E)=\int\frac{dE'}{2\pi\hbar}D^>(\mathbf{r},\mathbf{r}';E')G^>(\mathbf{r},\mathbf{r}';E-E') ,$$

where

$$D^>(\mathbf{r},\mathbf{r}';E)=\int d(t-t')e^{iE(t-t')/\hbar}\langle H'(\mathbf{r},t)H'(\mathbf{r}',t')\rangle .$$

The self-energy $\Sigma^>$, resulting from our interaction Hamiltonian (2.3), is

$$\Sigma^>(\mathbf{r}_1,\mathbf{r}_2;E)=U^2\delta^3(\mathbf{r}_1-\mathbf{r}_2)\{G^>(\mathbf{r}_1,\mathbf{r}_2;E-\hbar\omega_0)[N_B(\hbar\omega_0)+1]+G^>[\mathbf{r}_1,\mathbf{r}_2;E+\hbar\omega_0]N_B(\hbar\omega_0)\} . \quad \text{(C1)}$$

To obtain an order-of-magnitude estimate for the coupling strength U^2, we consider the interaction Hamiltonian for bulk polar optical phonons:

$$H'_{\mathrm{POP}}=\frac{1}{\sqrt{V}}\sum_{\mathbf{q}}\frac{M}{q}e^{i\mathbf{q}\cdot\mathbf{r}}(a_{\mathbf{q}}e^{-i\omega_0 t}+a^{+}_{-\mathbf{q}}e^{i\omega_0 t}) , \quad \text{(C2)}$$

where

$$M^2=\frac{2\pi e^2\hbar\omega_0}{4\pi\varepsilon_0}\left[\frac{1}{\kappa_\infty}-\frac{1}{\kappa_0}\right] . \quad \text{(C3)}$$

In (C3), $\hbar\omega_0=36$ meV, $\varepsilon_0=8.854\times10^{-12}$ F/m,

$\kappa_\infty = 10.92$, and $\kappa_0 = 12.90$. The self-energy $\Sigma^>$ resulting from (C2) in the SCFBA is

$$\Sigma^>(\mathbf{r}_1,\mathbf{r}_2;E) = \frac{\hbar\omega_0}{2}\frac{e^2}{4\pi\varepsilon_0}\left[\frac{1}{\kappa_\infty}-\frac{1}{\kappa_0}\right]\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|} \times\{G^>(\mathbf{r}_1,\mathbf{r}_2;E-\hbar\omega_0)[N_B(\hbar\omega_0)+1] + G^>(\mathbf{r}_1,\mathbf{r}_2;E+\hbar\omega_0)N_B(\hbar\omega_0)\} . \quad \text{(C4)}$$

We are concerned with the electron-phonon scattering within the quantum well. Within the well, the electrons are strongly confined. $G^>(\mathbf{r}_1,\mathbf{r}_2;E)$ provides a spatial cutoff within the confines of the quantum well to the nonlocal potential $1/|\mathbf{r}_1-\mathbf{r}_2|$. We are also considering a one-dimensional problem in which the quantum well is a three-dimensionally confined region, a quantum box.

To choose the strength of U^2 in (C1), we replace the nonlocal potential $1/|\mathbf{r}_1-\mathbf{r}_2|$ by a δ function with a strength given by the integral of $1/|\mathbf{r}_1-\mathbf{r}_2|$ over the quantum box, i.e.,

$$\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}\rightarrow\mu\delta^3(\mathbf{r}_1-\mathbf{r}_2) , \quad \text{(C5)}$$

where μ is defined by the relation

$$\int_\square d^3r_1\int_\square d^3r_2\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}=\mu\int_\square d^3r_1\int_\square d^3r_2\delta^3(\mathbf{r}_1-\mathbf{r}_2) . \quad \text{(C6)}$$

Substituting (C5) into (C4) and comparing (C4) to (C1), we find the value of U^2:

$$U^2=\frac{\hbar\omega_0}{2}\frac{e^2}{4\pi\varepsilon_0}\left[\frac{1}{\kappa_\infty}-\frac{1}{\kappa_0}\right]\mu . \quad \text{(C7)}$$

If one assumes that the spatial distribution of the electron density in the box is constant with a magnitude of $1/V_\square$, where $V_\square$ is the volume of the box, then the standard dimensionless coupling constant g defined by

$$g=\frac{1}{(\hbar\omega_0)^2}\frac{1}{V}\sum_{\mathbf{q}}\left|\frac{M\rho(\mathbf{q})}{q}\right|^2 , \quad \text{(C8)}$$

where $\rho(\mathbf{q})$ is the Fourier transform of the electron density in the well, is simply

$$g=\frac{1}{V_\square}\frac{U^2}{(\hbar\omega_0)^2} . \quad \text{(C9)}$$

This result is obtained by interchanging the order of the integration, performing first the integrals over $\mathbf{r}_1$ and $\mathbf{r}_2$ in (C6) and then the sum over $\mathbf{q}$ leading to (C4).

The cross-sectional dimensions used in the numerical program were 2×2 nm^2 with hard-wall boundaries. Since our numerical simulator always solves a three-dimensional problem, the narrow dimensions were used to ensure that there was no contribution to the spectral function from the tail of the second transverse subband. The length of the well in the longitudinal z direction was 7 nm. In the calculation of the strength (C6), the integral was performed over a sphere with a volume equal to the volume of the box ($2\times2\times7$ nm^3).

Using the physical constants listed after (C3), the coupling constant g calculated from (C9) is 0.18. Numerically, we found that this value for g resulted in unrealistically short scattering times $\tau_n(E_i)$ of ~1 fs. For this reason, the value actually used in the numerical simulations was considerably less, $g=1.84\times10^{-3}$, which was chosen to give more realistic scattering times $\tau_n(E_i)$ on the order of 0.1 ps.

APPENDIX D: NUMERICAL CALCULATION OF SCATTERING RATES

The devices we considered, with collector barriers ranging from 220 to 1000 meV, had resonances at the phonon-peak bias with intrinsic widths in energy, $\Gamma_i=\Gamma_C$, ranging from 0.5 meV to 0.1 μeV, respectively. For all of the devices except the most asymmetric structure, $\hbar/\tau_\phi(E_r)\lesssim\Gamma_C(E_r)$, so that $\Gamma(E_r)\simeq\Gamma_c(E_r)$. Since the magnitude of the spectral function at resonance is $4/\Gamma(E_r)$, the scattering rate at the incident energy, $\hbar/\tau_n(E_i)\propto A(E_r)$, would vary over three orders of magnitude between the symmetric structure and the most asymmetric structure. If the strength were chosen such that the scattering rate at the incident energy for the symmetric structure were $10^{12}(1/s)$, then the scattering rate at the incident energy for the most asymmetric structure would be $\sim5\times10^{15}(1/s)$, corresponding to an energy broadening $\hbar/\tau$ of 3.3 eV. Such a rate is unrealistic. To remove the dependence of the scattering rate at the incident energy on the intrinsic resonant width, we replaced the Einstein phonon spectrum with a normalized rectangular window function of finite but narrow width. The scattering rates were calculated by convolving the window function with the expressions in Eq. (2.4). The scattering rates become independent of the resonant widths if the width of the window function is chosen slightly larger than the width of the widest resonance. Unless such a procedure is used, a plot of $\hbar/\tau_\phi(E_i)$ as shown in Fig. 2(c) would not be meaningful. Since the current is obtained by integrating over the range of incident energies [see Eq. (4.28)], the current is independent of the window width Δ, provided $\Delta\ll\hbar\omega_0$, E_F, where E_F is the Fermi level in the emitter.

[1]*Resonant Tunneling in Semiconductors Physics and Applications*, edited by L. L. Chang, E. E. Mendez, and C. Tejedor (Plenum, New York, 1991).

[2]V. J. Goldman, C. D. Tsui, and J. E. Cunningham, Phys. Rev. B **36**, 7635 (1987).

[3]E. S. Alves, L. Eaves, M. Henni, O. H. Hughes, M. L. Leadbeater, F. W. Sheard, and G. A. Toombs, Electron. Lett. **24**, 1190 (1988).

[4]M. L. Leadbeater, E. S. Alves, L. Eaves, M. Henni, O. H. Hughes, F. W. Sheard, and G. A. Toombs, Semicond. Sci. Technol. **3**, 1060 (1988).

[5]M. S. Skolnick, D. G. Hayes, P. E. Simmonds, A. W. Higgs, G. W. Smith, H. J. Hutchinson, C. R. Whitehouse, L. Eaves, H. Henini, O. H. Hughes, M. L. Leadbeater, and D. P. Halliday,

Phys. Rev. B **41**, 10 754 (1990).
[6]P. J. Turley, C. R. Wallis, and S. W. Teitsworth (unpublished).
[7]M. L. Leadbeater, E. S. Alves, L. Eaves, M. Henni, O. H. Hughes, A. Celeste, J. C. Portal, G. Hill, and M. A. Pate, Phys. Rev. B **39**, 3438 (1989).
[8]G. S. Boebinger, A. F. J. Levi, S. Schmitt-Rink, A. Passner, L. N. Pfeiffer, and K. W. West, Phys. Rev. Lett. **65**, 235 (1990).
[9]J. G. Chen, C. H. Yang, M. J. Yang, and R. A. Wilson, Phys. Rev. B **43**, 4531 (1991).
[10]N. S. Wingreen, K. W. Jacobsen, and J. W. Wilkins, Phys. Rev. Lett. **61**, 1396 (1988); Phys. Rev. B **40**, 11 834 (1989).
[11]W. Cai, T. F. Zheng, P. Hu, B. Yudanin, and M. Lax, Phys. Rev. Lett. **63**, 418 (1989).
[12]F. Chevoir and B. Vinter, Appl. Phys. Lett. **55**, 1859 (1989).
[13]B. G. R. Rudberg, Semicond. Sci. Technol. **5**, 328 (1990).
[14]P. Hyldgaard and A. Jauho, J. Phys. Condens. Matter **2**, 8725 (1990).
[15]A. Jauho, Phys. Rev. B **41**, 12 327 (1990).
[16]J. A. Stovneng, E. H. Hauge, P. Lipavsky, and V. Spicka, Phys. Rev. B **44**, 13 595 (1991).
[17]E. V. Anda and F. Flores, J. Phys. Condens. Matter **3**, 9087 (1991).
[18]X. Wu and S. E. Ulloa, Phys. Rev. B **44**, 13 148 (1991).
[19]P. J. Turley and S. W. Teitsworth, Phys. Rev. B **44**, 3199 (1991).
[20]P. J. Turley and S. W. Teitsworth, Phys. Rev. B **44**, 8181 (1991).
[21]P. J. Turley and S. W. Teitsworth, Phys. Rev. B **44**, 12 959 (1991).
[22]P. J. Turley and S. W. Teitsworth, J. Appl. Phys. (to be published).
[23]L. V. Keldysh, Zh. Eksp. Teor. Fiz. **47**, 1515 (1964) [Sov. Phys. JETP **20**, 1018 (1965)].
[24]S. Datta, J. Phys. Condens. Matter **2**, 8023 (1990).
[25]R. Lake and S. Datta, Phys. Rev. B **45**, 6670 (1992).
[26]N. Mori and T. Ando, Phys. Rev. B **40**, 6175 (1989).
[27]For nonlocal scattering, see Eq. (3.2) of S. Datta, Phys. Rev. B **46**, 9493 (1992).
[28]See Eq. (22) of Ref. 25.
[29]R. Landauer, IBM J. Res. Dev. **32**, 306 (1988).
[30]M. Büttiker, Phys. Rev. Lett. **57**, 1761 (1986).
[31]M. Büttiker, Phys. Rev. B **33**, 3020 (1986).
[32]S. Datta, Phys. Rev. B **40**, 5830 (1989).
[33]J. L. D'Amato and H. M. Pastawski, Phys. Rev. B **41**, 7411 (1990).
[34]H. M. Pastawski, Phys. Rev. B **44**, 6329 (1991).
[35]S. Hershfield, Phys. Rev. B **43**, 11 586 (1991).
[36]S. Hershfield, J. H. Davies, and J. W. Wilkins, Phys. Rev. Lett. **67**, 3720 (1991).
[37]In our notation, Eq. (5) of Ref. 36 becomes

$$I(E)=\frac{e}{h}\frac{2\Gamma_E\Gamma_C}{\Gamma_C+\Gamma_C}A[f_E-f_C]+e\frac{\Gamma_C-\Gamma_E}{\Gamma_E+\Gamma_C}[p/\tau_p-n/\tau_n]\ ,$$

where the energy dependence of the quantities is implied.
[38]We set $\sin^2(\theta/2)$ equal to 1 since E_i is far from resonance. Assuming $(2\hbar\nu)^2 \gg \frac{1}{4}(\Gamma_E+\Gamma_C+\hbar/\tau_\phi)^2$ (which is always the case for the devices we consider), Eq. (4.14) becomes

$$I_E(E_i)=\frac{e}{h}\frac{1}{4}\frac{\Gamma}{\Gamma_E+\Gamma_C}T_E T_C(f_E-f_C)\ ,$$

which is identical to Eq. (23) of Büttiker in Ref. 1, p. 213.
[39]S. Datta, in *Quantum Phenomena, The Molecular Series on Solid State Devices,* edited by Robert F. Pierret and Gerold W. Neudeck (Addison-Wesley, Reading, MA, 1989), Vol. 8, p. 25.