Time-dependent resonant tunneling of wave packets in the tight-binding model

J. A. Støvneng* and E. H. Hauge

Institutt for Fysikk, Norges Tekniske Høgskole, Universitetet i Trondheim, N-7034 Trondheim-NTH, Norway

(Received 6 May 1991)

We consider the dynamics of resonant tunneling through a double barrier within a simple one-dimensional tight-binding model. High-frequency experiments on double-barrier semiconductor structures have motivated theoretical studies of the various time scales that exist in a resonant-tunneling process. Numerical results involving wave packets have shown that the transient process of *building up* a resonant state in a double-barrier well is qualitatively different from the *exponential decay* out of the well. These results are confirmed in the present paper. We investigate how the transient buildup depends on barrier and initial-state parameters, and find the perhaps surprising result that, for opaque barriers, the particle density inside the well evolves in time essentially *independently* from barrier width and height. In our opinion, this shows that one cannot use the concept of a classical velocity for particles moving through classically forbidden regions of space. Also, we find that the buildup time mainly reflects the spatial extent of the incoming wave packet. Our conclusion is, therefore, that the limiting factor for a maximum operating frequency is either due to properties of the incoming electrons, in particular, their spatial extent, or due to inelastic-scattering or external-circuit effects, not incorporated in our simple model.

I. INTRODUCTION

The physics of resonant tunneling has been known since the early days of quantum mechanics. The phenomenon is treated in every introductory textbook,[1] and is usually discussed with reference to a one-dimensional double-barrier potential structure (DBS).

Such systems can be manufactured today, utilizing epitaxial growth techniques like molecular-beam epitaxy (MBE) or metallo-organic chemical-vapor deposition (MOCVD). One of the most common semiconductor structures consists of $Al_xGa_{1-x}As$ barrier layers with GaAs in the well and the contacts. Because of the difference in band gaps between these materials, an electron moving in the conduction band of GaAs will effectively come under the influence of the scattering potential of Fig. 1 when approaching the $Al_xGa_{1-x}As$ layers.

The resonant behavior in the transmission coefficient makes the DBS interesting for device applications.[2] In particular, the DBS is a strong candidate for very-high-speed oscillator components. In 1983, Sollner *et al.*[3] reported resonant current response at a frequency of 2.5 THz in a DBS consisting of 50-Å $Al_{0.3}Ga_{0.7}As$ barriers separated by a 50-Å GaAs well. This indicates that the charge transport mechanism is at least as fast as $\tau \sim 1/f = 4 \times 10^{-13}$ s. Clearly, this experiment raises questions of a fundamental nature: What is the frequency limit of such a DBS? What time scales exist in a resonant-tunneling process, and which is the decisive one for the maximum operating frequency?

The resonant-state lifetime τ_d is a well-established quantity in this context.[4] The transmission coefficient in Fig. 1 has a Lorentzian line shape close to the resonances E_1 and E_2, with widths Γ_1 and Γ_2, respectively. An elec-

tron with energy $E_{1(2)}$ initially placed inside the quantum well will then escape at a rate proportional to $\exp(-\Gamma_{1(2)}t/\hbar)$ thus having a lifetime $\tau_{d1(d2)} = \hbar/\Gamma_{1(2)}$. Clearly, the resonant-state lifetime represents the relevant time scale in several types of experiment. For example, when electrons are photoexcited from the valence band to a resonant state in the quantum-well conduction band, time-resolved photoluminescence spectroscopy gives de-
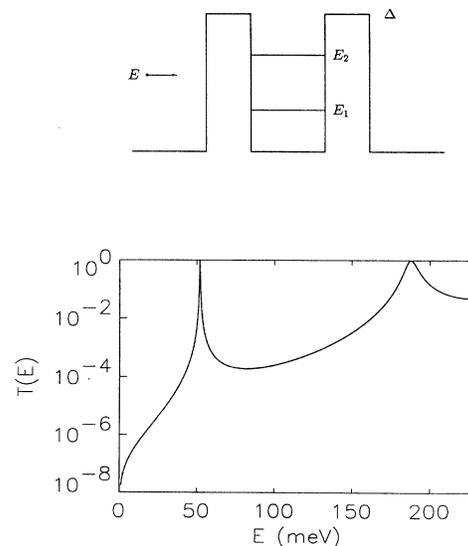

FIG. 1. A symmetric double barrier with two resonant levels below the barrier edge. An electron with energy E scatters off the structure. The transmission spectrum shown corresponds to a GaAs-$Al_{0.3}Ga_{0.7}As$ structure with two 50-Å barriers enclosing a 70-Å well. The barrier height Δ is 230 meV.

     ©1991 The American Physical Society

cay times in good agreement with calculated lifetimes.[5] Another example is the time delay for an electron tunneling through a DBS on resonance. All calculations give a result that essentially equals the resonant-state lifetime.[6]

However, when the DBS is placed in an ac circuit, as was the case in Sollner's experiment, it is not clear that the decay time τ_d is the relevant time scale. In that and other experiments[3,7] the metastable state in the well has initially a higher energy than all the available charge carriers and is therefore completely empty. When a voltage is applied across the DBS, the metastable state is pulled below the Fermi level of, say, the left contact, thus enabling electrons to tunnel resonantly through the DBS. As argued by Ricco and Azbel,[8] under such conditions a transient time is required for the electrons to *enter* the state in the well. In order to have a current response reflecting the resonant features of the DBS under high-frequency conditions, the electrons should at least be given time to enter the resonant state on a time scale set by the period of oscillation. In Ref. 8 it is suggested that this transient time is of the same order as the resonant-state lifetime. The argument is simply that the time required to fill the resonant state should be the same as that required to empty it. Let us see what this proposal implies for the limiting frequency in the Sollner experiment. Assuming a barrier height $\Delta = 230$ meV, there is only one resonant level in the well with energy $E_{res} \sim 60$ meV. We take the barrier and well effective masses to be $m_b^* = 0.09 m_e$ and $m_w^* = 0.067 m_e$, respectively, where m_e is the bare electron mass. The decay time is then[4] $\tau_d = \hbar/\Gamma \simeq w v_{res}^{-1} \exp(2\kappa_{res} b) \simeq 5$ ps. In this expression $v_{res} = \hbar k_{res}/m_w^*$ is the electron velocity on resonance, $\kappa_{res} = [2m_b^*(\Delta - E_{res})]^{1/2}/\hbar$, and w and b are the well and barrier width, respectively, in this case $w = b = 50$ Å. Assuming τ_d to be the speed-limiting factor, one finds a maximum frequency of about 200 GHz. Clearly, this is not compatible with the 2.5 THz observed in the experiment.

The observation above shows that the process of entering the resonant state may take place on a time scale, hereafter denoted the buildup time τ_b, considerably shorter than the decay time τ_d. In order to study the buildup time in a simple model, Jauho[9] and Guo et al.[10] solved the time-dependent Schrödinger equation for an electron, described by a wave packet, tunneling through a DBS. Their conclusions can be summarized as follows: Jauho finds a τ_b substantially shorter than τ_d. Guo et al. find that τ_b can be shorter than, equal to, or longer than τ_d, depending on the barrier parameters. For parameters corresponding to the Sollner experiment, $\tau_b < \tau_d$ is found by both groups.

The following questions arise naturally: What determines, quite generally, the buildup time τ_b? How does it depend on the barrier parameters, and to what extent is it determined by the incoming wave packet? References 9 and 10 give only partial answers to these questions. The purpose of the present paper is to clarify these points as far as possible. Like Refs. 9 and 10 we start from the premise that only an explicitly time-dependent solution of the problem can give reliable information on the dynam-

ics of the resonant-tunneling process. Unlike those authors, who worked with continuum effective-mass theory, we shall base our discussion on a simple tight-binding (TB) Hamiltonian.

An advantage of this model is the built-in discretization of real space, making numerical implementation quite straightforward. In addition, the finite bandwidth of the TB model is in better correspondence with reality than is the infinite bandwidth of the effective-mass approximation. When interpreting numerical results, however, one must be careful to sort out effects caused by the finite bandwidth and the presence of bound states from the transient buildup and decay in which we are primarily interested. To prevent confusion, we have, therefore, included a discussion of these extraneous effects, illustrated with numerical examples.

In the limit of a bandwidth large compared with the barrier heights and the energy spread of the wave packet, the TB model reproduces the effective-mass results. This limit involves wave packets with a large spatial extent. With such wave packets, the observed time dependence of tunneling to a large extent reflects the shape of the wave packet in real space. In an attempt to avoid such effects, we shall mainly study wave packets *narrow* in real space.

Our model is an extremely idealized one. Effects of inelastic scattering, caused by coupling to phonons and other electrons, have been disregarded. In addition, transport is taken to be strictly one dimensional (1D), which assumes perfect translational invariance in the lateral dimensions. Finally, effects of the external circuit have not been taken into account. In a real experiment the assumptions underlying such idealizations are violated to a higher or lesser degree, and this may significantly affect the temporal behavior of resonant tunneling. Nevertheless, we believe that thorough investigations of simple models should precede more realistic, and therefore more complex, calculations.

A surprisingly simple conclusion from our work is the following: For opaque barriers (i.e., for most barriers of practical interest) the buildup time is found to be essentially *independent* of the barrier parameters. The buildup process is almost exclusively determined by the shape of the incoming wave packet in our simple model. Thus, factors limiting the frequency of oscillators based on resonant tunneling must be sought, either among those that determine the shape of the incoming wave packets or among effects neglected in our simple elastic one-electron picture.

We have organized the paper as follows. In Sec. II we define the tight-binding Hamiltonian and show how to introduce a double-barrier potential in the 1D lattice. A recursive Green-function method is used to derive the transmission and reflection coefficients for the DBS, and to illustrate the appearance of bound states associated with the scattering potential. In Sec. III we derive an exact expression in terms of the Green function for the time evolution of an arbitrary initial state. It is shown how to construct an initial state that has the minimum uncertainty Gaussian wave packet as its continuum limit. In

Sec. IV we present numerical results that serve to illustrate some of the basic features of the model. Examples of these are the exponential decay of resonant states, interference between resonant and bound states, and effects of a finite bandwidth. Section V is devoted to a study of the buildup time τ_b, and we investigate its dependence on barrier and wave-packet parameters. Our conclusions are collected in Sec. VI.

II. STATIONARY PROPERTIES OF THE TIGHT-BINDING MODEL

The tight-binding model is described in every introductory textbook on solid-state physics, and the reader is referred to the literature[11] for an extensive discussion of the motivation behind the model and its physical ingredients. We will describe the noninteracting conduction-band electrons with the simplest form of a 1D nearest-neighbor TB Hamiltonian,

$$H = \sum_i [\,|i\rangle \varepsilon_i \langle i| + u(\,|i\rangle\langle i+1| + |i\rangle\langle i-1|)\,]\ . \quad (2.1)$$

Here the (Wannier) states $|i\rangle$ form a complete, $\sum_i |i\rangle\langle i| = 1$, orthonormal, $\langle i|j\rangle = \delta_{ij}$, set of atomiclike orbitals centered at the sites i, which form a regular lattice with interatomic spacing a. The nearest-neighbor hopping elements u are, for simplicity, taken to be real quantities, independent of position. In the translationally invariant case, $\varepsilon_i = \varepsilon_0$ and $H = H_0$. Here H_0 corresponds to the free-particle Hamiltonian with eigenfunctions of the Bloch type,

$$|k\rangle = \sum_j e^{ikja}|j\rangle\ . \quad (2.2)$$

We reserve the notation $|k\rangle$ with the letter k for *momentum* eigenstates, whereas other letters will be used to denote the localized orbitals of (2.1). The dispersion relation associated with these Bloch states reads

$$E(k) = \varepsilon_0 + 2u\cos(ka)\ , \quad (2.3)$$

i.e., the model has an energy band, extending from $\varepsilon_0 - |2u|$ to $\varepsilon_0 + |2u|$, with the first Brillouin zone spanning the k interval $[-\pi/a, \pi/a]$. The choice of $\varepsilon_0 = -2u$ for the trivial parameter ε_0 ensures that $k = 0$ corresponds to the bottom of the band. Furthermore, the choice $u = -\hbar^2/2m^*a^2$ sees to it that the results of the present model close to the bottom of the band will coincide with those of effective-mass theory.

A double-barrier structure is included in the model in a very natural way. The atoms in the barrier layers are given a higher value of the local energy ε_n, the addition Δ_n corresponding to the barrier height in position $x = na$. An arbitrary scattering potential is shown in Fig. 2. We shall mainly be concerned with the symmetric DBS where the two barriers have equal width $b = N_b a$ and equal, constant height Δ. In an $Al_{0.3}Ga_{0.7}As$-$GaAs$ structure Δ will be about 230 meV, which is the conduction-band offset between $GaAs$ and $Al_{0.3}Ga_{0.7}As$. Typically, the electron effective mass in the barriers is different from that in the well and in the contacts. (We shall refer to the homogeneous regions outside the DBS
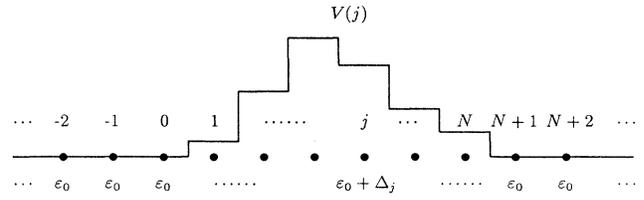

FIG. 2. An arbitrary scattering potential in a one-dimensional lattice. The potential profile $V(j) = \varepsilon_0 + \Delta_j$ extends across N lattice constants.

as "contacts" or "leads.") This could easily be modeled by using a different hopping energy in the barriers. The quantitative effects of ignoring the space dependence of m^* may be quite large for certain quantities, like the exponential decay time τ_d,[12] but qualitatively our results remain unchanged. Thus, we shall assume a constant m^* and use the same value of u throughout the lattice.

With reference to Fig. 2, the stationary scattering problem is solved by finding the scattering eigenstates with positive k,

$$\psi_k(j) = \begin{cases} e^{ikja} + r(k)e^{-ikja}, & j \leq 0 \\ t(k)e^{ikja}, & j \geq N+1\ , \end{cases} \quad (2.4)$$

and those with negative k,

$$\tilde{\psi}_k(j) = \begin{cases} e^{ikja} + \tilde{r}(k)e^{-ikja}, & j \geq N+1 \\ \tilde{t}(k)e^{ikja}, & j \leq 0\ . \end{cases} \quad (2.5)$$

These wave functions describe plane-wave electrons approaching the scattering potential from $x = -\infty$ and $x = +\infty$, respectively, with transmission and reflection amplitudes $t(k)$ and $r(k)$ for $k > 0$, and $\tilde{t}(k)$ and $\tilde{r}(k)$ for $k < 0$. In (2.4) and (2.5) there is no voltage drop between the two contacts, which have equal and constant potential ε_0. A voltage drop V across the structure would simply require that ε_0 be replaced by $\varepsilon_0 - eV$ in the low-voltage contact. Clearly, arbitrary depletion and accumulation layers may also be taken into account (at least in a quasistatic approximation), but such effects will not be considered here.

In order to derive expressions for the transmission and reflection amplitudes, we shall use a recursive Green-function technique. This method has been applied in several different contexts, and examples are the study of the conductivity of a disordered linear chain[13] and transport through quantum wires with geometrical scattering centers.[14] The relation between the Green function G and the S matrix for scattering in leads has been given by Fisher and Lee,[15] and was generalized to an arbitrary scattering structure by Stone and Szafer.[16] For clarity, and to introduce a notation convenient for our purposes, we go through the necessary steps in the Appendix. The Green function[17] is defined through

$$G(z)(z - H) \equiv 1\ , \quad (2.6)$$

where z is an energy variable. In the energy representation (2.6) becomes

$$G(z) = \sum_\alpha \frac{|\alpha\rangle\langle\alpha|}{z - \varepsilon_\alpha} , \tag{2.7}$$

where $|\alpha\rangle$ is the complete set of energy eigenstates, with $H|\alpha\rangle = \varepsilon_\alpha |\alpha\rangle$. The real space matrix elements of $G(z)$ are defined as

$$G_{ij}(z) \equiv \langle i|G(z)|j\rangle . \tag{2.8}$$

Here $|i\rangle$ and $|j\rangle$ are the states of Eq. (2.1). The transmission and reflection amplitudes are now given by

$$t(k) = -2iu \sin(ka)e^{-i(N+1)ka}G_{N+1,0}(k) ,$$
$$r(k) = -2iu \sin(ka)G_{00}(k) - 1 , \tag{2.9}$$

with similar expressions for $\tilde{t}(k)$ and $\tilde{r}(k)$. Here $G_{ij}(k)$ means $G_{ij}(z)$ evaluated in $z = E(k)$. Note that Eq. (2.9) is based on using the *retarded* Green function, by which it is meant that $(z - H)$ should be understood as $\lim_{\eta \to 0^+}(z - H + i\eta)$.

In the case of free particles, the eigenstates $|k\rangle$ of Eq. (2.2) constitute a complete, orthogonal set. Since the localized orbitals $|j\rangle$ are taken to be normalized to unity, one has the following δ-function normalization for the plane waves:

$$\langle k'|k\rangle = 2\pi\delta(k - k') . \tag{2.10}$$

In the presence of a scattering potential there are, in addition to the scattering states $|\psi_k\rangle$ and $|\tilde{\psi}_k\rangle$, always one or more bound states $|b\rangle$ localized in the vicinity of the scattering center. This is a well-known result for a 1D *attractive* potential, but in the present model there are also bound states when the potential is *repulsive*. Close to the upper edge of the continuum band, the effective mass is negative, and for the same reason as a bound state is formed *below* the band in an attractive potential, a repulsive potential will form a bound state *above* the band. From Eq. (2.7) it is clear that bound states will appear at energies corresponding to simple poles in the Green function. The presence of bound states will have implications on the discussion of the time dependence of tunneling. We will come back to this later.

The *resonant states* $|\psi_r\rangle$ will be of particular importance in our discussion. They have features in common with the localized bound states in the sense that the amplitude of the wave function $\psi_r(j)$ is enhanced in the well of the DBS. On the other hand, $|\psi_r\rangle$ does not go to zero, but reduces, except for a phase factor, to the unperturbed Bloch wave $|k\rangle$ on the far side of the DBS. In the context of 1D tunneling the resonant behavior manifests itself through peaks in the transmission coefficient $T(k) = |t(k)|^2$. As is well known, even a single barrier exhibits resonances when the electron is transmitted with energy higher than the barrier potential Δ. If the barrier is modeled with one atom only, there is no resonant structure in $T(k)$. A single, isolated atom gives rise to a single state, and when coupled to the two contacts, this appears as a bound state above the continuum band. A barrier with more atoms, however, creates several locally enhanced states, some of which may be degenerate with the continuum band and appear as resonant states.[18] In

general, a barrier consisting of N_b atoms will give rise to N_b states in the energy range $(\varepsilon_0 + \Delta - |2u|, \varepsilon_0 + \Delta + |2u|)$. The value of Δ determines how many of these will be resonant states, degenerate with the continuum band $(\varepsilon_0 - |2u|, \varepsilon_0 + |2u|)$, and how many will be bound states with energies $z_b > \varepsilon_0 + |2u|$.

In the single-barrier case the transmission peaks always occur *above* the barrier edge. This is no longer true when we introduce a second barrier. An isolated well with W atoms has W energy states in the range $(\varepsilon_0 - |2u|, \varepsilon_0 + |2u|)$ (assuming $\varepsilon_j = \varepsilon_0$ is the same in the well as in the contacts). When the well is coupled to the contacts through potential barriers, these states, all degenerate with the continuum band, give rise to W resonances. Some of them may lie *below* the barrier edge, and through these states an electron will tunnel resonantly (see Fig. 3).

The limiting symmetric case of very high barriers of width $b = N_b a$ enclosing a well consisting of a single atom will be used later in the discussion of the time dependence of the trapped wave function inside the well. In this case, with Δ larger than the bandwidth B, there is
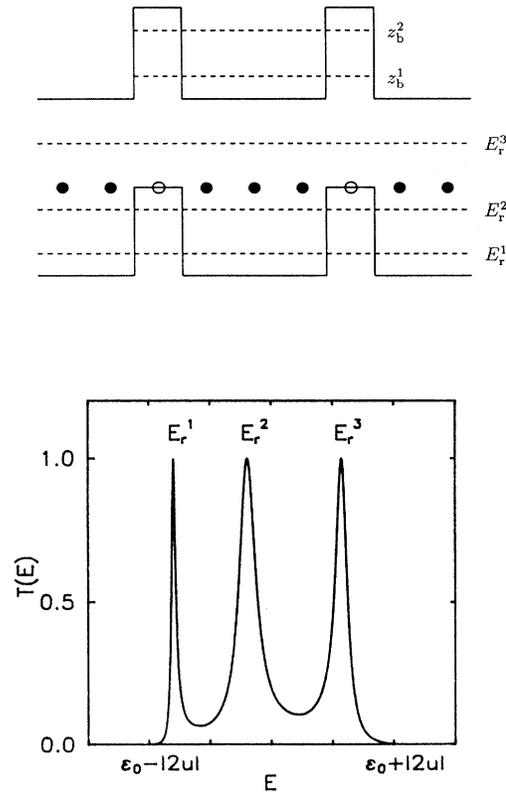

FIG. 3. A double barrier with three atoms in the well and one in each barrier. The barrier height is $\Delta = |2.0u|$. There are five locally enhanced states. Two of them appear as truly bound states, with energies z_b^1 and z_b^2. There are three resonant levels, and the two lowest lie below the barrier edge. Through these levels electrons can tunnel resonantly. Since the structure is symmetric, there is full transmission, $T = 1$, on resonance.

J. A. STØVNENG AND E. H. HAUGE

one resonance, and the number of bound states equals the total number $2N_b$ of barrier atoms. In the simplest case, with one atom in each barrier, the bound-state energies are

$$z_b^1 = \varepsilon_0 + \Delta + u^2/\Delta \; ,$$

(2.11)

$$z_b^2 = \varepsilon_0 + \Delta + 3u^2/\Delta - 8u^4/\Delta^3 + 44u^6/\Delta^5 + \cdots \; ,$$

whereas the resonant level energy and width are

$$E_r = \varepsilon_0 - 2u^2/\Delta + 4u^4/\Delta^3 + \cdots \; ,$$

$$\Gamma_r = 4|u|^3/\Delta^2 + \cdots \; .$$

(2.12)

As expected, the resonance is narrow and lies close to ε_0, the energy of an isolated atom. Increasing the barrier thickness to two atoms yields

$$E_r = \varepsilon_0 - 2u^2/\Delta + 2u^4/\Delta^3 + \cdots \; ,$$

$$\Gamma_r = 4|u|^5/\Delta^4 + \cdots \; .$$

(2.13)

The position of the resonance is almost unchanged, but the width is reduced in a manner consistent with exponential dependence on barrier thickness. Both results are consistent with earlier calculations within the WKB approximation.[19]

III. PROPAGATION OF WAVE PACKETS IN THE 1d TIGHT-BINDING CHAIN

In this section we shall study how the time evolution of an initial state $|\Psi(t=0)\rangle$ is given by the Green function. For all times $t \geq 0$ the state $|\Psi(t)\rangle$ must be a solution of the Schrödinger equation

$$\left[i\frac{\partial}{\partial t} - H \right] |\Psi(t)\rangle = 0 \; .$$

(3.1)

Here, the Hamiltonian H is independent of time, and we choose units in which $\hbar = 1$. We perform a Laplace transformation on (3.1),

$$\int_0^\infty dt \, e^{-st} \left[i\frac{\partial}{\partial t} - H \right] |\Psi(t)\rangle = 0 \; ,$$

(3.2)

since we are interested in $|\Psi(t)\rangle$ for positive times only. Apply partial integration on the first term in (3.2). Then, with the definition

$$|\Phi(s)\rangle \equiv \int_0^\infty dt \, e^{-st} |\Psi(t)\rangle \; ,$$

(3.3)

and changing variable $s \rightarrow -iz$, we obtain

$$|\Phi(-iz)\rangle = (z - H)^{-1} i |\Psi(0)\rangle = iG(z)|\Psi(0)\rangle \; .$$

(3.4)

The time-dependent state $|\Psi(t)\rangle$ is the inverse Laplace transform of $|\Phi\rangle$, given in terms of the Mellin inversion integral.[20] In the present notation one has

$$|\Psi(t)\rangle = \int_{-\infty+i\gamma}^{\infty+i\gamma} \frac{dz}{2\pi} e^{-izt} iG(z)|\Psi(0)\rangle \; ,$$

(3.5)

where the integration line lies above all the singularities of $G(z)$ in the complex z plane. We have seen that $G(z)$

is analytic, except for a branch cut along the real z axis, corresponding to the continuum band of H, and possibly one or more simple poles on the real z axis, corresponding to the discrete eigenvalues of the bound states of H. Thus, we may choose any real $\gamma > 0$ in Eq. (3.5). Furthermore, since $t \geq 0$, we may close the integration contour in the lower half-plane. Then, by Cauchy's theorem, it may be deformed to any contour enclosing the branch cut and the poles of $G(z)$. An example, suitable for numerical solution of the integral, is shown in Fig. 4.

The initial state $|\Psi(0)\rangle$ is usually constructed as a minimum-uncertainty Gaussian wave packet. In a continuum model it has the form[21]

$$\Psi(x;t=0) \sim (\Delta x)^{-1/2} \exp\left[-\left[\frac{x - \langle x(0)\rangle}{2\Delta x} \right]^2 + i\langle k\rangle x \right] \; ,$$

(3.6)

where Δx is a measure of the real space extent of the initial wave packet, and $\langle x(0)\rangle$ and $\langle k\rangle$ denote the mean position and wave number, respectively. For this initial state one has $\Delta x \Delta k = \frac{1}{2}$.

The tight-binding model is not continuous in real space, so we cannot use the form (3.6) directly. Instead, we shall construct the initial state on the 1D chain as a linear combination of the states $|j\rangle$,

$$|\Psi(0)\rangle = \sum_j c_j |j\rangle \; .$$

(3.7)

The complex coefficients $c_j = |c_j|\exp(i\alpha_j)$ determine the shape of the wave packet through their absolute values, and its mean group velocity $\langle v\rangle$ through their phases. We choose the coefficients so that a minimum-uncertainty Gaussian wave packet is reproduced in the continuum limit. Such an initial state, occupying the S sites from $j = -S+1$ to $j = 0$, has expansion coefficients

$$c_j^s = \left[\frac{(S-1)!}{(S-1+j)!(-j)!2^{S-1}} \right]^{1/2} \exp(i\alpha j) \; .$$

(3.8)

This choice yields a wave packet normalized to unity, and the linear dependence of the phase on position ensures an initial state with no cross-correlation between velocity and position, i.e.,
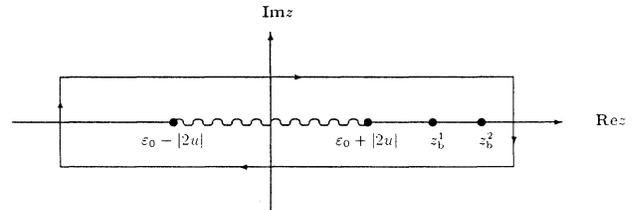

FIG. 4. Integration contour for the integral in Eq. (3.5). The contour encloses the branch cut between $\varepsilon_0 - |2u|$ and $\varepsilon_0 + |2u|$, corresponding to the continuum band, and all bound-state energies, in this example z_b^1 and z_b^2.

$$\tfrac{1}{2}\langle(\Delta v\,\Delta x+\Delta x\,\Delta v)\rangle=0 \ . \tag{3.9}$$

The group velocity is

$$v=\frac{\partial E(k)}{\partial k}=-2ua\,\sin(ka) \ , \tag{3.10}$$

and the mean velocity of the packet follows from α as

$$\langle v\rangle=-2uaF(S)\sin\alpha \ . \tag{3.11}$$

The value of $F(S)$ lies between 0.5 and 1.0, and is close to 1.0 except for very small S. This is just as expected, since the Fourier transform of $|\Psi(0)\rangle$ peaks at $ka=\alpha$. A broad wave packet (large S) has a sharp k distribution, and the mean velocity will be close to the value given by (3.10), $\langle v\rangle\simeq-2ua\sin\alpha$. Obviously, $F(S)<1.0$, since $v_{\mathrm{max}}=-2ua$.

Equation (3.5) is an exact expression for the time evolution of a given initial state $|\Psi(0)\rangle$. It is computationally quite efficient, given the procedure in the Appendix for calculating the Green function. However, we will also consider the standard approach of expanding $|\Psi(0)\rangle$ directly in the eigenstates of the system. In our case the complete, orthonormal set consists of the scattering states of Eqs. (2.4) and (2.5) and the bound states $|b\rangle$. Thus,

$$\begin{aligned}|\Psi(0)\rangle=&\int_0^{\pi/a}\frac{dk}{2\pi}d(k)|\psi_k\rangle\\&+\int_{-\pi/a}^0\frac{dk}{2\pi}\tilde{d}(k)|\tilde{\psi}_k\rangle+\sum_b d_b|b\rangle \ ,\end{aligned} \tag{3.12}$$

where $d(k)$, $\tilde{d}(k)$, and d_b are the expansion coefficients, given by

$$\begin{aligned}d(k)&=\sum_j c_j[e^{-ikja}+r^*(k)e^{ikja}] \ ,\\ \tilde{d}(k)&=\sum_j c_j\tilde{t}^*(k)e^{-ikja} \ , \tag{3.13}\\ d_b&=\sum_j c_j\langle b|j\rangle=\sum_j c_j b_j^* \ .\end{aligned}$$

Here we have assumed that the particle is initially on the left side of the scattering potential. The time evolution of $|\Psi(0)\rangle$ is then

$$\begin{aligned}|\Psi(t)\rangle=&\int_0^{\pi/a}\frac{dk}{2\pi}d(k)|\psi_k\rangle e^{-iE(k)t}\\&+\int_{-\pi/a}^0\frac{dk}{2\pi}\tilde{d}(k)|\tilde{\psi}_k\rangle e^{-iE(k)t}\\&+\sum_b d_b|b\rangle e^{-iz_b t} \ . \tag{3.14}\end{aligned}$$

As mentioned in Sec. II, the presence of bound states may influence the time-dependent tunneling process. Assume that we want to study how the trapped wave function in the well of a DBS evolves in time. If the particle was initially very far away from the DBS, the overlap between $|\Psi(0)\rangle$ and the bound states is very small, and $\Psi_w(t)\equiv\sum_{j\in\mathrm{well}}\langle j|\Psi(t)\rangle$ is accurately described in terms of the scattering states of the system. However, if $|\Psi(0)\rangle$ is a narrow wave packet, initially localized close to the

DBS, a substantial fraction of $\Psi_w(t)$ may be due to the bound states. This will be illustrated with numerical examples in the next section.

Note that in order to have a complete set of scattering states, one must solve *both* stationary scattering problems, i.e., with particles coming from left ($k>0$) and right ($k<0$). The expansion coefficients are then calculated as in (3.13). However, one can show that there is a simpler but equivalent form of (3.14), namely,

$$|\Psi(t)\rangle=\int_{-\pi/a}^{\pi/a}\frac{dk}{2\pi}\phi(k)|\psi_k\rangle e^{-iE(k)t}+\sum_b d_b|b\rangle e^{-iz_b t} \ , \tag{3.15}$$

where the expansion coefficients $\phi(k)$ are the Fourier components of the initial state, and the positive k eigenstates $|\psi_k\rangle$ are *analytically continued* to negative values of k. The equivalence between (3.14) and (3.15) is a consequence of particle conservation and time-reversal symmetry, which lead to the relations (A23) between the S-matrix elements.

Under circumstances where contributions from bound states may be neglected, the most efficient approach is usually to evaluate the integral in (3.15). However, the exact expression in (3.5) may still be preferable, especially when the resonances are extremely narrow. In that case the numerical integral is easier to perform along a contour with Im$z\neq0$, since that yields a less sharply peaked integrand.

IV. SOME NUMERICAL ILLUSTRATIONS

The main application of the formalism presented up to now will be a study of the transient behavior of resonantly tunneling wave packets. Before we get to that subject, we will present numerical examples in order to highlight some implications of the discussion in Secs. II and III.

A. Effects of a finite bandwidth

The finite bandwidth of the tight-binding model $B=|4u|$ has consequences for the time evolution of an initial state $|\Psi(0)\rangle$. Let us look at free-particle motion, for which analytic results are easily obtained. In this case there are no bound states, and the integral in (3.15) is an expansion in the plane waves $|k\rangle$ of Eq. (2.2). If we have an arbitrary initial state $|\Psi(0)\rangle=\sum_j c_j|j\rangle$, the amplitude at a given site m will evolve as

$$\begin{aligned}\Psi_m(t)&=\int_{-\pi/a}^{\pi/a}\frac{dk}{2\pi}\phi(k)\langle m|k\rangle e^{-iE(k)t}\\&=e^{-i\varepsilon_0 t}\sum_j c_j\int_{-\pi/a}^{\pi/a}\frac{dk}{2\pi}e^{ik(m-j)a-2iut\cos(ka)}\\&=\frac{1}{a}e^{-i\varepsilon_0 t}\sum_j c_j i^{|m-j|}J_{|m-j|}(|2u|t) \ , \tag{4.1}\end{aligned}$$

where $J_n(x)$ is the Bessel function of nth order. In Fig. 5 we have plotted $|\Psi_2(t)|^2$ for wave packets of initial widths $\Delta x=0.5a$ and $4.0a$. The narrow packet oscillates strongly in time with period $T\sim2\pi/B$, reflecting the interference of only two Bessel functions in (4.1). The width in k
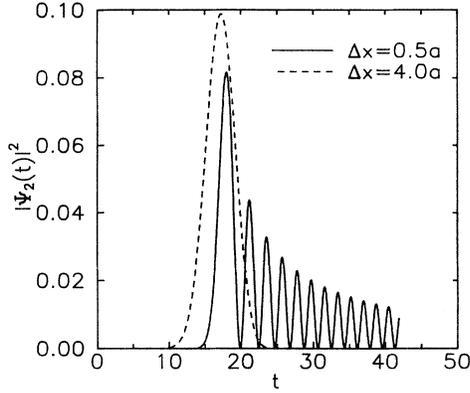
FIG. 5. The time evolution (in units of $|u|^{-1}$) of freely propagating wave packets. We have plotted the particle density (i.e., the absolute square of the wave function) on site 2. The hopping amplitude is $u = -1.0$. At $t = 0$ the wave packets were centered around site -32. The solid and dashed lines represent packets initially occupying 2 and 65 sites, respectively, having initial spatial widths $\Delta x = 0.5a$ and $4.0a$.

space of $|\Psi(0)\rangle$, $\Delta k \sim 1/a$, is of the same order as the total bandwidth $B_k = 2\pi/a$, and large deviations from an effective-mass approach (with infinite B) are observed. The broad packet, on the other hand, has $\Delta k \sim 1/8a \ll B_k$, and the resulting $|\Psi_2(t)|^2$ is a smooth function, simply reflecting the shape of $|\Psi(0)\rangle$ in real space. At this point one should note that in order to have $\Delta x \gtrsim 10\text{--}20$ lattice constants, one needs an initial state occupying more than 1000 sites. As a result, the numerics becomes exceedingly time consuming. However, our choice of initial state reduces to a Gaussian for broad packets. Under such circumstances one may use the exact Gaussian form of (3.6) as a good approximation and save considerable amounts of computer time.

B. Effects of bound states and resonances

As discussed in Sec. II, one or more potential barriers will have localized states and resonances associated with them. We shall now see how these states influence the propagating wave packet.

In Fig. 6 we have plotted $|\Psi_2(t)|^2$ when a single barrier is located on site 1, and the initial state occupies two sites to the left of the barrier. When the wave packet starts close to the barrier (solid line), the overlap with the bound state is large, and a substantial fraction of the particle density is trapped in this state forever. However, if one takes the initial state and the bound state farther apart (dashed line), $|\Psi_2(t)|^2$ is totally dominated by the contribution from the scattering states. One can show that the amplitude of the bound states decreases exponentially with the distance R from the scattering center.[22] Thus, one may neglect the presence of bound states when the overlap integrals $\langle \Psi_R(0)|b\rangle \sim \exp(-R/R_b)$ are small compared with the scattering-states contribution.
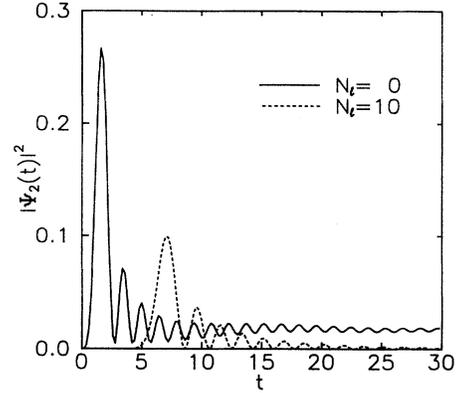

FIG. 6. Time evolution of the particle density on site 2 when a single barrier of height $\Delta = |1.25u|$ is located on site 1. The hopping amplitude is $u = -1.0$. The initial wave packet occupies the two sites $-1, 0$ (solid line) and $-11, -10$ (dashed line). N_l is the number of sites between barrier and initial packet.

Here R_b is the localization length of $|b\rangle$. When, as in the next section, only the scattering states are of interest, one must locate the particle initially a distance larger than R_b away from the barriers.

The final example is a double barrier of height $\Delta = |4.0u|$ located on sites 1 and 3; see the insets of Fig. 7. In Fig. 7(a) we observe the temporal evolution inside
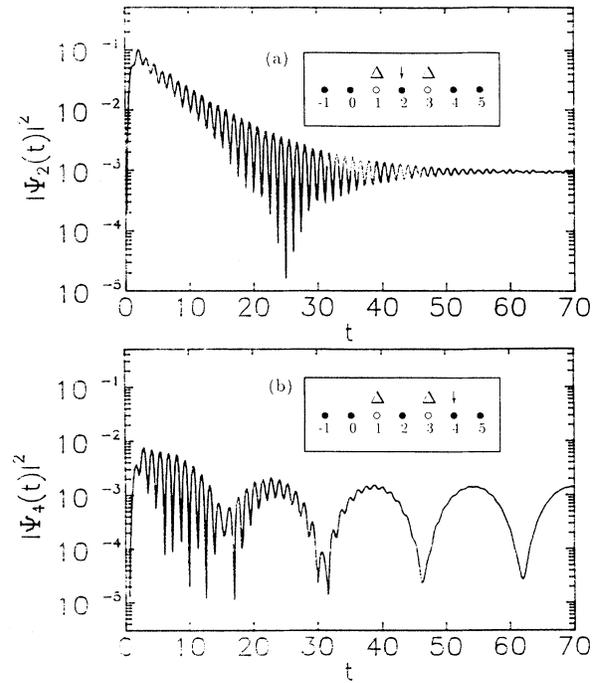

FIG. 7. Time evolution of tunneling through a double barrier. The barriers have height $\Delta = |4.0u|$ and are located on sites 1 and 3. The initial wave packet occupies the two sites -1 and 0. The hopping amplitude is $u = -1.0$. (a) Particle density inside the well (site 2). (b) Transmitted particle density (site 4).

the well, whereas in Fig. 7(b) we look at the transmitted wave on site 4. In this case we have a resonance with energy eigenvalue $E_r = \varepsilon_0 - |0.45u|$ and width $\Gamma = |0.20u|$, as well as two bound states at $z_b^1 = \varepsilon_0 + |4.25u|$ and $z_b^2 = \varepsilon_0 + |4.65u|$ [cf. Eqs. (2.11) and (2.12)]. In both figures there is an exponential decay $\exp(-t/\tau_d)$, with $\tau_d = 1/\Gamma = 5.0|u|^{-1}$. Since τ_d is so short, the exponential behavior is barely observable. The rapid oscillations correspond to interference between the resonance and a bound state. They start as soon as the resonance is built up, and exhibit the strongest oscillations for $t \sim 25|u|^{-1}$ in Fig. 7(a), when the amplitudes of the two states are approximately equal. Eventually, the resonant state is empty, and the rapid oscillation disappears. Then, for $t \gtrsim 20|u|^{-1}$ in Fig. 7(b), the interference between the two bound states comes into play. Note, however, that $|\Psi_2(t)|^2$ in Fig. 7(a) simply goes to a constant value, whereas $|\Psi_4(t)|^2$ in Fig. 7(b) oscillates with period $T = 2\pi/(z_b^2 - z_b^1) \simeq 15.7|u|^{-1}$. This can be understood as follows: Because one of the bound states has odd parity, it has vanishing amplitude on the well site. Therefore, $|\Psi_2(t)|^2$ only depends on the even-parity bound state and decays to a constant value, as in Fig. 6. On site 4, however, both bound states have nonzero amplitude, and $|\Psi_4(t)|^2$ oscillates with frequency according to the energy splitting of the two levels.

V. TRANSIENT BEHAVIOR OF RESONANT TUNNELING

This section is devoted to a study of the buildup time τ_b, which is the time it takes to establish the resonant state in double-barrier tunneling. We shall, somewhat arbitrarily, define τ_b as the time required for $|\Psi_w(t)|^2$ to increase from 1% to 100% of its maximum value. We have chosen this definition simply because a function f starts to become discernible from zero around the value $0.01 f_{\max}$ in a linear plot on a normal-size paper. As mentioned in the Introduction, the buildup time is believed to be important for determining the frequency limit of resonant-tunneling devices.

The exponential decay law receives most of the attention in textbook treatments of time-dependent resonant tunneling. However, the transient process *is* sometimes mentioned, and even explicitly stated to occur within a time much shorter than the decay time.[23] This should be contrasted with the conjecture in Ref. 8 that τ_b and τ_d are *both* connected to the energy width of the resonance via the uncertainty relation as $\tau_b \simeq \tau_d \simeq \hbar/\Gamma$. The latter is clearly based on the assumption that only wave-packet components with energies in a range of order Γ around the resonant energy participate in the buildup process. We will now show that this is, in fact, *not* the case. We start with a qualitative argument that will eventually be supported by numerical calculations.

Assume that we have a symmetric double barrier with very high barriers such that $\exp(-\kappa b) \ll 1$. The wave number κ measures the difference between barrier height and particle energy (see Sec. I), and b is the thickness of each barrier. Close to resonance the value of the wave

function inside the well is $|\psi_r| \sim \exp(\kappa b)$, whereas far from resonance it is $|\psi| \sim \exp(-\kappa b)$. We assume a width $\Delta k = O(1)$ of the initial wave packet. However, the resonance width in k space is $\Gamma_k \sim \exp(-2\kappa b)$. This implies that the time-dependent state inside the well $\Psi_w(t)$ is made up of a resonant and an off-resonant part, each contributing with weight $|\int dk\, \psi| \sim \exp(-\kappa b)$. Clearly, one cannot ignore the off-resonant contribution.

In Fig. 8 we illustrate the dramatic effect of only including states close to resonance. The solid curve is the exact result for $|\Psi_w(t)|^2$, where we have integrated over the whole Brillouin zone. Bound states can be ignored in this example, since the overlap integrals $\langle \Psi(0)|b \rangle$ are very small. The long-dashed line results from integrating over a k interval ($k_{res} - 2.5\Gamma_k, k_{res} + 2.5\Gamma_k$). The buildup time, in reality very short in this case, seems to be of the same order as the exponential decay time. The short-dashed line includes a ten-times-larger k interval, extending over 50 resonance widths, but the deviations from the exact result are still substantial. Note, however, that all three curves merge after a long time. This example shows that the exponential decay is indeed described by the resonance properties of the DBS, whereas the transient behavior, in general, is not. It is described by a complicated interplay between resonant *and* off-resonant wave-packet components.

In the previous example we found a buildup time much shorter than the resonant-state lifetime, consistent with the prediction in Ref. 23 and the numerical work of Refs. 9 and 10. However, in Ref. 10 it was shown that τ_d could become smaller than τ_b when the barriers were made sufficiently thin. We shall argue that this is a conse-
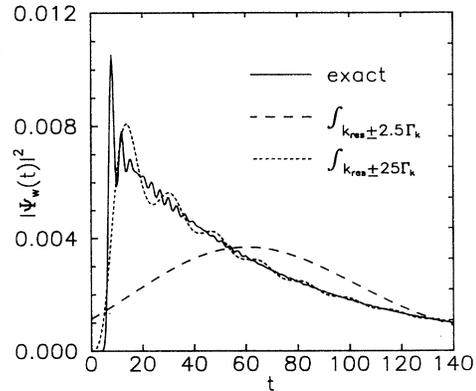

FIG. 8. Time evolution of the particle density inside the well (on site 2) of a double barrier located on sites 1 and 3. The hopping amplitude is $u = -1.0$. The initial wave packet occupies the two sites -11 and -10, making the overlap integrals $\langle \Psi(0)|b \rangle$ between the initial state and bound states negligible. The solid line is the exact result, where we have included all the scattering states in the first Brillouin zone. The long- and short-dashed lines are results of integrating over k intervals $k_{res} \pm 2.5\Gamma_k$ and $k_{res} \pm 25\Gamma_k$, respectively. Here k_{res} is the wave number on resonance, and Γ_k is the width of the resonance in k space. In this case $\Gamma_k \simeq 1.2 \times 10^{-3} B_k$, where $B_k = 2\pi/a$ is the total width of the first Brillouin zone.

quence of the spatial extent of the initial wave packet. When the packet approaches the DBS, it moves more or less like a free particle with mean velocity $\langle v \rangle$. (There is self-interference in front of the barrier causing deviations from free-particle motion; see Ref. 6.) It seems rather obvious, then, that the width Δx of the wave packet will provide a lower bound on the time required to establish the resonance, simply because different parts of the packet arrive at different times. In fact, we would expect $\tau_b \gtrsim \Delta x / \langle v \rangle$, assuming there are negligible position-velocity cross-correlations in the wave packet [see Eq. (3.9)]. This is exactly what we observe in Fig. 9, where we have plotted τ_b as a function of the initial width $\Delta x = [(S-1)/4]^{1/2}a$. The initial state occupies S sites and hits the DBS with mean velocity on resonance. The buildup time increases linearly with Δx. Note the deviation from linear behavior for $\Delta x \gtrsim 6a$. In this region the decay out of the well has become significant before the whole packet has arrived, and as a result, $|\Psi_w(t)|^2$ reaches its maximum somewhat earlier than expected from a linear dependence of τ_b on Δx.

When the barriers are made thinner, and the wave-packet width is kept fixed, the major effect is to reduce the decay time. In addition, τ_b will decrease, since a smaller τ_d makes the decay process start earlier, as explained above. However, the latter is a much smaller effect than the decrease in decay time, since τ_d depends exponentially on barrier thickness. Thus, one expects to find $\tau_d > \tau_b$ for thick barriers, but as the barrier thickness is reduced, and the resonance width Γ eventually becomes comparable to the energy spread of the wave packet, one should find $\tau_d < \tau_b$. This is exactly what was observed in Ref. 10.

It is important to realize that from the discussion above one cannot draw any conclusions about the speed of the tunneling process into the well; i.e., we cannot determine the single-barrier "tunneling time," a quantity that has been extensively discussed in the literature.[6] We have excluded the possibility that this process is of the
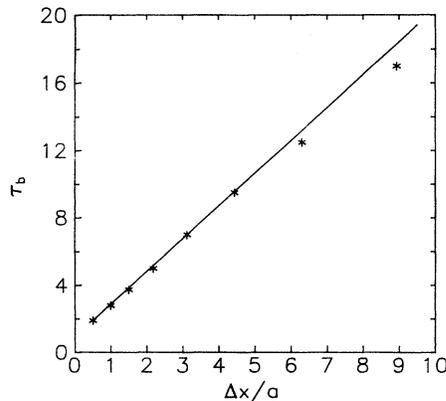
same nature as the decay out of the resonant state. However, it may well take place on a very short time scale, since the lower bound on τ_b found above only reflects wave-packet motion *outside* the DBS. In an attempt to discuss the time scale for the tunneling process itself, we shall investigate propagation of wave packets that are *narrow* in real space. As we have seen, broad packets will have a large τ_b, which makes it difficult to resolve small variations as one varies the barrier parameters. Clearly, these variations are best resolved when the initial state occupies only a few sites. Furthermore, we shall be using high barriers to eliminate the possibility of the particle passing above them without actually tunneling. Also, we use a narrow well, on a single site, in order to have a simple system with one resonance only. Finally, at this stage we want to look at propagation via scattering states only. This is accomplished by starting the wave packet sufficiently far away from the DBS, to make the contribution from bound states negligible.

The present setup is clearly not very realistic for typical transport in semiconductors, and certainly not for the structure in Fig. 1. In that case the barriers are low, and the transport takes place near the bottom of the GaAs conduction band. Representative wave packets are believed to have a width comparable to the electron mean-free path, typically around 500 Å in GaAs.[12] In contrast, our parameters will correspond to barriers higher than the bandwidth, and wave packets initially localized on a few sites, thus having an energy spectrum extending across the whole band. However, one might follow Hyldgaard and Jauho[24] and let the model describe transport within a miniband of an imperfect superlattice, where the DBS represents two barriers that are higher than those of the superlattice. In such a structure the bandwidth, typically between 10 and 100 meV, could easily become smaller than the height of the two particular barriers, and it is certainly possible to imagine the initial state as localized to within a few superlattice wells.

Next we shall see how the buildup time depends on barrier thickness. We use a wave packet initially localized on two sites, and starting a distance of $N_l = 20$ sites from the left edge of the DBS. The barriers have equal height $\Delta = 16|u|$. In Fig. 10 we have plotted $|\Psi_w(t)|^2$ for barrier widths $N_b = 1$, 10, and 20 sites. The results require a couple of comments. First, one observes oscillations in the probability density. This is the finite-bandwidth effect discussed in Sec. IV A. Due to a wide momentum distribution, the wave packet has developed rather strong cross-correlations when it hits the DBS, i.e., the high-velocity components arrive first. Thus, the first oscillation "period" in $|\Psi_w(t)|^2$ (for $10|u|^{-1} \lesssim t \lesssim 16|u|^{-1}$), corresponding to the leading "pulse" of the packet (see Fig. 11), had a mean velocity outside the barrier somewhat larger than that of the full packet. Second, one finds an almost constant buildup time, $\tau_b \simeq 3.7|u|^{-1}$, independent of barrier thickness. This is not surprising, since the *shape* of the wave packet immediately before impact (Fig. 11) is also almost independent of N_b. As discussed earlier, this shape essentially determines the value of τ_b. However, we find it

FIG. 9. The buildup time τ_b as function of initial wave packet width Δx. The double barrier is located on sites 1 and 3 and has height $\Delta = |4.0u|$. We use $u = -1.0$. The initial packet is located close to the left edge of the DBS.
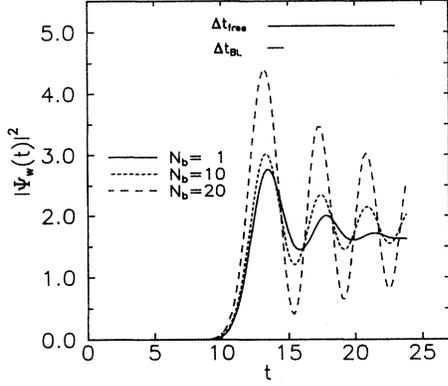
FIG. 10. Time evolution of the particle density inside the well of a DBS for different barrier widths [and normalized with respect to the single-barrier tunneling factor $(\Delta/u)^{2N_b}$]. As usual, $u = -1.0$. The well consists of a single atom. The initial wave packet occupies two sites at a distance of $N_l = 20$ sites from the left barrier. The barrier height is $\Delta = |16u|$. Solid line: $N_b = 1$. Short-dashed line: $N_b = 10$. Long-dashed line: $N_b = 20$. The horizontal bars indicate the expected delays of the peak when increasing N_b from 1 to 20, assuming the single-barrier tunneling process can be described with free-particle motion (long bar) or "Büttiker-Landauer velocity" (short bar).

FIG. 11. The wave packet at $t = 0$ and 9, with parameters as in Fig. 10. The corresponding curves for $|\Psi(t=9)|^2$ with $N_b = 1$ or 20 are indistinguishable from the present example, where $N_b = 10$.

surprising that $|\Psi_w(t)|^2$ reaches its maximum value at practically the *same instant* for every value of N_b. In other words, the time required for the wave packet to travel a certain distance, partly in a classically allowed region and partly in a classically forbidden one, is, within numerical uncertainty, independent of the classically forbidden distance traveled. Consequently, we cannot ascribe a finite velocity to the tunneling process, which seems to take place in essentially zero time. (Since we use a nonrelativistic model, there is no paradox in finding infinite speed.) This is certainly not a result that appeals to one's intuition, and it illustrates how classical ways of thinking may fail completely when one deals with purely quantum-mechanical phenomena.

Possible candidates for a finite "tunneling velocity" might be $v_{\text{free}} = -2ua \sin(ka)$, corresponding to free-particle motion, and v_{BL}, corresponding to the so-called traversal time proposed by Büttiker and Landauer.[25] In the effective-mass approximation one has $v_{\text{BL}} = \hbar\kappa/m^*$ (see Sec. I for a definition of κ). In the TB model, with $\hbar = 1$, one has $v_{\text{BL}} = \partial(\Delta - E)/\partial\kappa \simeq \Delta a$ (close to resonance). In Fig. 10 we have indicated with horizontal bars the expected delays of the particle density peak when the barrier thickness is increased from $N_b = 1$ to $N_b = 20$
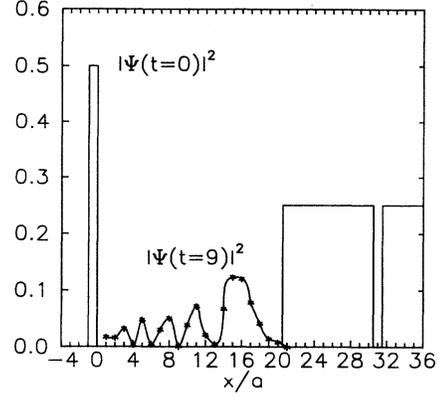
sites, the long bar corresponding to free-particle velocity and the short one corresponding to "Büttiker-Landauer velocity." Clearly, both candidates disagree qualitatively with the observed behavior. In fact, the maximum of $|\Psi_w(t)|^2$ occurs somewhat *earlier* as N_b is increased.

It is even possible to show analytically that the time evolution of the particle density in the well is (approximately) independent of barrier thickness. Using Eq. (3.15) and ignoring bound states, one has

$$\Psi_w(t) = \sum_m c_m \int_{-\pi/a}^{\pi/a} \frac{dk}{2\pi} e^{-ikma}[-2iu \sin(ka)]$$
$$\times G_{N_b + N_l + 1, 0} e^{-i[\varepsilon_0 + 2u \cos(ka)]t}$$

$$(5.1)$$

for an arbitrary initial state $|\Psi(0)\rangle = \sum_m c_m |m\rangle$ separated by N_l sites from the left edge of the DBS. The well consists of a single site, and (A18) was used to express $\psi_k(N_b + N_l + 1)$ in terms of the Green function. Next we use the results of the Appendix to find an expression for $G_{N_b + N_l + 1, 0}$ in the limit of very large barrier height Δ. To leading order one obtains

$$G_{N_b + N_l + 1, 0} = \frac{1}{u}\left[\frac{E(k) - \varepsilon_0}{u} + \frac{2u}{\Delta}\right]^{-1}$$
$$\times \left[-\frac{u}{\Delta}\right]^{N_b} e^{ik(N_l + 1)a}, \qquad (5.2)$$

which yields

$$\Psi_w(t) = \left[-\frac{u}{\Delta}\right]^{N_b} \sum_m c_m \int_{-\pi/a}^{\pi/a} \frac{dk}{2\pi}(e^{ik(N_l - m)a} - e^{ik(N_l - m + 2)a})\left[2\cos(ka) + \frac{2u}{\Delta}\right]^{-1} e^{-i[\varepsilon_0 + 2u \cos(ka)]t} \qquad (5.3)$$

for the wave function. The particle density $|\Psi_w(t)|^2$ depends on the barrier width only through the time-independent scaling factor $(u/\Delta)^{2N_b}$. Except for this factor, which corresponds exactly to the well-known term $\exp(-2\kappa b)$ in the continuum model, the temporal evolution of Ψ_w is also independent of Δ when $|\Delta/u| \gg 1$.

Let us summarize the results of the present section. The transient buildup of resonantly tunneling wave packets is of a nature different from the exponential decay out of the resonant state. The main reason is that in order to describe the buildup process, one needs to take both resonant and off-resonant scattering states into account. The exponential decay, on the other hand, is well described by states near resonance only. Thus, the decay time is linked to the resonance energy width through the uncertainty relation as $\tau_d = \hbar/\Gamma$. For wide wave packets the buildup time mainly reflects the spatial extent of the packet. A study of narrow packets yields a transient time independent of barrier thickness and height for opaque barriers, indicating that a finite "tunneling velocity" cannot be defined.

VI. CONCLUDING REMARKS

In this paper we have discussed stationary as well as time-dependent properties of resonant tunneling within a simple tight-binding formalism. The tight-binding model is more satisfying than the effective-mass approximation in that it incorporates a finite bandwidth, but reduces to an effective-mass description in the limit of very large bandwidth. However, when one investigates time-dependent properties, it is important to be aware of the specific features of the model, as was shown in Sec. IV. For explicit solutions we have applied a recursive Green-function technique, which turns out to be an effective tool in the present context, especially for numerical implementation, but also for analytic results in simple examples and certain asymptotic limits (see Secs. II, IV, and V).

Our main focus has been on the transient behavior of resonant tunneling. Contrary to earlier predictions, the process of building up resonance conditions is shown to be qualitatively different from the decay of the resonant state. Consequently, it takes place on a time scale different from the decay time τ_d. This result is in agreement with the numerical calculations of Refs. 9 and 10. We have shown how the buildup time τ_b is directly related to the spatial width Δx of the wave packet. This dependence of τ_b on Δx dominates the picture to an extent that we are forced to go to extremely small Δx, and correspondingly high barriers, in order to pinpoint a possible dependence of τ_b on the barrier thickness. Our conclusion, based on both analytic and numerical evidence, is that, to leading order in inverse barrier height, τ_b is *independent* of the thickness of the barrier. It makes no sense to ascribe a velocity to the electron tunneling through the initial barrier into the resonant state. Predictions based on the classical velocity v_{free} or the Büttiker-Landauer velocity v_{BL} are clearly contradicted by our results. This statement should not be read as an entry in the debate on the relative merits of various proposals for a general tunneling time. In our opinion, rather than to debate this general question, it is more fruitful to focus on the dynamics of tunneling in very specific models of some relevance to experiment. The present effort represents a first step in this direction.

Since the buildup time may be substantially shorter than the decay time, there is no contradiction between our results and the high-frequency experiments by Sollner *et al.* However, we cannot draw any conclusions on a maximum operating frequency, since the situation in a real sample is much more complex than in our simple one-electron model. Our results indicate that the limiting factor is either due to properties of the incoming electrons, in particular their spatial extent, or due to inelastic-scattering or external-circuit effects not incorporated in our simple model.

APPENDIX

We want to find the transmission and reflection amplitudes for the scattering potential in Fig. 2. The stationary Schrödinger equation is

$$(z-H)|\psi_z\rangle = 0 , \tag{A1}$$

where z is the eigenenergy and $|\psi_z\rangle$ is the corresponding energy eigenstate. Let us investigate the scattering states $|\psi_k\rangle$ of Eq. (2.4), i.e., for $k > 0$. We start by writing the total scattering state as a "source" term

$$|S_k\rangle = \sum_{j \leq -1} e^{ikja}|j\rangle , \tag{A2}$$

plus a remainder $|\varphi_k\rangle$. From (2.6) and (A1) we have

$$|\varphi_k\rangle = G(k)\{-[E(k)-H]|S_k\rangle\} , \tag{A3}$$

where $G(k)$ is the retarded Green function, and $E(k)$ is the eigenenergy of $|\psi_k\rangle$. With Eqs. (2.1) and (A2) and the orthonormality condition for the states $|j\rangle$, the term in curly brackets becomes

$$-[E(k)-H]|S_k\rangle = \sum_{j \leq -1} [-E(k)+\varepsilon_0$$
$$+u(e^{-ika}+e^{ika})]e^{ikja}|j\rangle$$
$$+ue^{-ika}|0\rangle - u|-1\rangle . \tag{A4}$$

From Eq. (2.3), $E(k) = \varepsilon_0 + 2u\cos(ka)$, and the sum in (A4) vanishes. Thus,

$$\varphi_k(j) \equiv \langle j|\varphi_k\rangle = ue^{-ika}G_{j0} - uG_{j,-1} , \tag{A5}$$

where we have used the definition (2.6), and, for simplicity, have suppressed the argument k in the Green-function

matrix elements.

In order to proceed from here, we use the Dyson equation. With $H = H^0 + H'$, the Dyson equation expresses the Green function G of the full Hamiltonian H in terms of the perturbation H' and the Green function G^0 of the unperturbed Hamiltonian H^0:

$$G = G^0 + G^0 H' G = G^0 + G H' G^0 . \tag{A6}$$

This exact result follows directly from the definition $(z - H^0)G^0 = 1$.

If we let the perturbation be the matrix elements of H between two sites, l and $l + 1$,

$$H' = |l\rangle u \langle l + 1| + |l + 1\rangle u \langle l| , \tag{A7}$$

G^0 and H^0 describe two separate semi-infinite chains $\{j \leq l\}$ and $\{j \geq l + 1\}$. Since there is no coupling between the two chains, any off-diagonal element G^0_{mn} with $m \leq l$ and $n \geq l + 1$ (or vice versa) must be zero. Furthermore, the diagonal element of G^0 at the end site of a semi-infinite chain requires special attention. With the "cut" (A7) (removing the terms $|l\rangle u \langle l + 1|$ and $|l + 1\rangle u \langle l|$ from H literally means that we cut the chain in two pieces) we define

$$\Gamma_l^- \equiv G^0_{ll} ,$$
$$\Gamma_{l+1}^+ \equiv G^0_{l+1,l+1} , \tag{A8}$$

where "+" denotes a chain extending (here, from site $l + 1$) in the positive direction and "−" denotes a chain extending (here, from site l) in the negative direction. The matrix elements Γ^+ and Γ^- will be called "surface" Green functions (SGF's).

In order to find the matrix elements of the Green function, we use the Dyson equation repeatedly. With the perturbation (A7) one has

$$G_{ll} = G^0_{ll} + G^0_{ll} u G_{l+1,l}$$
$$= G^0_{ll} + G^0_{ll} u G^0_{l+1,l+1} u G_{ll}, \tag{A9}$$

or, using (A8),

$$G_{ll}^{-1} = (\Gamma_l^-)^{-1} - u \Gamma_{l+1}^+ u,$$
$$G_{l+1,l} = \Gamma_{l+1}^+ u G_{ll} . \tag{A10}$$

Repeating the procedure, one readily finds

$$G_{l+m,l} = \left[\prod_{j=l+m}^{l+1} \Gamma_j^+ u \right] G_{ll}$$
$$= G_{l+m,l+m} \left[\prod_{j=l+m-1}^{l} u \Gamma_j^- \right] . \tag{A11}$$

The next step is to calculate the SGF's, and we start by deriving recursion relations for SGF's on neighboring sites. Take H' as in (A7) and let $H'' = H' + |l - 1\rangle u \langle l| + |l\rangle u \langle l - 1|$. Then G^0 corresponds to the two chains $\{j \leq l\}$ and $\{j \geq l + 1\}$, whereas G^{00} corresponds to the chains $\{j \leq l - 1\}$ and $\{j \geq l + 1\}$ plus an isolated site, $j = l$. The Dyson equation now yields a relation analogous to (A10),

$$(G^0_{ll})^{-1} = (G^{00}_{ll})^{-1} - u G^{00}_{l-1,l-1} u . \tag{A12}$$

Using the definition of G and $H^{00} = H - H''$, we have $(G^{00}_{ll})^{-1} = \langle l|(z - H^{00})|l\rangle = z - \varepsilon_l$. Furthermore, G^0_{ll} and $G^{00}_{l-1,l-1}$ are both diagonal elements of rightmost end sites of semi-infinite chains, so, from (A8), we have

$$(\Gamma_l^-)^{-1} = z - \varepsilon_l - u \Gamma_{l-1}^- u . \tag{A13}$$

By interchanging H' and H'' we find the analogous equation for Γ^+,

$$(\Gamma_l^+)^{-1} = z - \varepsilon_l - u \Gamma_{l+1}^+ u . \tag{A14}$$

An expression for G_{ll}, more symmetric than (A10), is then found by inserting (A13) in (A10),

$$G_{ll}^{-1} = z - \varepsilon_l - u \Gamma_{l-1}^- u - u \Gamma_{l+1}^+ u . \tag{A15}$$

The systems we want to study consist of *homogeneous* regions enclosing a potential profile. In particular, the linear chain has a potential that is equal and constant for $j \leq 0$ and $j \geq N + 1$; see Fig. 2. Obviously, all Γ_j^- are equal for $j \leq 0$, and all Γ_j^+ are equal for $j \geq N + 1$. Denote by Γ (not to be confused with the energy width of a resonance) the SGF of a homogeneous semi-infinite chain. Then, (A13) [or (A14)] is a quadratic equation for Γ, with solutions

$$\Gamma = \frac{1}{2u^2} \{ z - \varepsilon_0 \pm [(z - \varepsilon_0)^2 - 4u^2]^{1/2} \} . \tag{A16}$$

The proper sign in (A16) for the retarded Green function is that which yields a non-negative density of states. The surface density of states is given as $D_\Gamma(z) = (-1/\pi) \text{Im} \Gamma(z)$. For z located in the continuous spectrum of the Hamiltonian, namely $z = E(k) = \varepsilon_0 + 2u \cos(ka)$, we obtain

$$\Gamma = \frac{1}{u} e^{\pm ika} \quad (k \gtrless 0) . \tag{A17}$$

Here we have also assumed $u < 0$. Functions of z involving the square-root term in (A16) have a branch cut along the real z axis between $z = \varepsilon_0 - |2u|$ and $z = \varepsilon_0 + |2u|$. For real values of z not belonging to the branch cut, one must take $[(z - \varepsilon_0)^2 - 4u^2]^{1/2} = \pm |(z - \varepsilon_0)^2 - 4u^2|^{1/2}$ for $z \gtrless \varepsilon_0 \pm |2u|$.

Now we have all we need to derive the scattering amplitudes for the potential in Fig. 2. From (A5) we have

$$\psi_k(N + 1) = \varphi_k(N + 1) = -2iu \sin(ka) G_{N+1,0} , \tag{A18}$$

where we used $G_{N+1,-1} = G_{N+1,0} u \Gamma$ and (A17) for $k > 0$. Comparison of (A18) with (2.4) yields $t(k)$ in Eq. (2.9). To find the reflection amplitude, start from Eq. (A5) with $j \leq -1$. Using the Dyson equation repeatedly, one has

$$G_{j,-1} = G^0_{j,-1} + G_{j,0} u \Gamma ,$$
$$G^0_{j,-1} = \frac{1}{u} e^{-ikja} , \tag{A19}$$
$$G_{j,0} = e^{-ikja} G_{00} ,$$

and finally,

$$\psi_k(j) = S_k(j) + \varphi_k(j)$$

$$= e^{ikja} - [2iu \sin(ka)G_{00} + 1]e^{-ikja}. \tag{A20}$$

Again, comparing with (2.4), one obtains $r(k)$ in Eq. (2.9).

The inverse process, i.e., scattering of a plane wave coming from $x = +\infty$ with $k < 0$, is analyzed in a similar way, starting with a source term

$$|\widetilde{S}_k\rangle = \sum_{j \geq N+2} e^{ikja}|j\rangle . \tag{A21}$$

The resulting scattering amplitudes are

$$\widetilde{t}(k) = 2iu \sin(ka)e^{ik(N+1)a}G_{0,N+1} ,$$

$$\widetilde{r}(k) = [2iu \sin(ka)G_{N+1,N+1} - 1]e^{2ik(N+1)a} . \tag{A22}$$

The consequences of time-reversal symmetry and particle conservation lead to relations between the scattering amplitudes.[26] With $k > 0$ one has

$$\widetilde{t}(-k) = t(k) ,$$

$$\widetilde{r}(-k) = -r^*(k)t(k)/t^*(k) . \tag{A23}$$

*Present address: Nordita, Blegdamsvej 17, 2100 Copenhagen Ø, Denmark.

[1] See, for example, E. Merzbacher, *Quantum Mechanics* (Wiley, New York, 1970), Chaps. 6 and 7.

[2] For a review of device applications, see F. Capasso, K. Mohammed, and A. Y. Cho, IEEE J. Quantum Electron. **QE-22**, 1853 (1986).

[3] T. C. L. G. Sollner *et al.*, Appl. Phys. Lett. **43**, 588 (1983).

[4] E. Merzbacher, *Quantum Mechanics* (Ref. 1), Chap. 7.

[5] See, e.g., T. Matsusue *et al.*, Phys. Rev. B **42**, 5719 (1990).

[6] E. H. Hauge and J. A. Støvneng, Rev. Mod. Phys. **61**, 917 (1989).

[7] L. L. Chang, L. Esaki, and R. Tsu, Appl. Phys. Lett. **24**, 593 (1974).

[8] B. Ricco and M. Ya. Azbel, Phys. Rev. B **29**, 1970 (1984).

[9] A. P. Jauho, Acta Polytech. Scand. Electr. Eng. Ser. **58**, 192 (1987).

[10] H. Guo *et al.*, Appl. Phys. Lett. **53**, 131 (1988).

[11] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Saunders College, Philadelphia, 1976), Chap. 10.

[12] K. Diff *et al.*, in *Nanostructure Physics and Fabrication,* edited by M. A. Reed and W. P. Kirk (Academic, San Diego, 1989).

[13] D. J. Thouless and S. Kirkpatrick, J. Phys. C **14**, 235 (1981).

[14] F. Sols *et al.*, J. Appl. Phys. **66**, 3892 (1989).

[15] D. S. Fisher and P. A. Lee, Phys. Rev. B **23**, 6851 (1981).

[16] A. D. Stone and A. Szafer, IBM J. Res. Dev. **32**, 384 (1988).

[17] For a thorough and clear account of Green functions, see E. N. Economou, *Green's Functions in Quantum Physics* (Springer, Berlin, 1979).

[18] E. N. Economou, *Green's Functions in Quantum Physics* (Ref. 17), Chap. 7.

[19] See, for example, D. Bohm, *Quantum Theory* (Prentice-Hall, New York, 1951), Chap. 12.

[20] E. Butkov, *Mathematical Physics* (Addison-Wesley, Reading, MA, 1968), Chap. 5.

[21] E. Merzbacher, *Quantum Mechanics* (Ref. 1), Chap. 8.

[22] E. N. Economou, *Green's Functions in Quantum Physics* (Ref. 17), Chap. 6.

[23] D. Bohm, *Quantum Theory* (Ref. 19), p. 293.

[24] P. Hyldgaard and A. P. Jauho, J. Phys. Condens. Matter **2**, 8725 (1990).

[25] M. Büttiker and R. Landauer, Phys. Rev. Lett. **49**, 1739 (1982); Phys. Scr. **32**, 429 (1985); M. Büttiker, Phys. Rev. B **27**, 6178 (1983).

[26] M. Ya. Azbel, Phys. Rev. B **28**, 4106 (1983).