

Electron transfer between regions of quasi-two-dimensional and three-dimensional dynamics in semiconductor microstructures

Alfred M. Kriman

Naval Research Laboratory, Washington D.C. 20375

P. Paul Ruden*

*Naval Research Laboratory, Washington D.C. 20375
and North Carolina State University, Raleigh, North Carolina 27695*

(Received 1 July 1985)

A theoretical study is carried out on the transfer of electrons at the edge of a microstructure, where electrons may cross between a region of quasi-two-dimensional dynamics and a "contact region" of three-dimensional dynamics. Working in the one-electron, effective-mass approximation, a formal solution is found for a model having a general profile for the confining potential in the two-dimensional region. The general form of the threshold behavior is found for ejection probabilities and injection cross sections. Numerical results are obtained for the case of a parabolic confining potential.

I. INTRODUCTION

Semiconductor microstructures, such as metal-oxide-semiconductor field-effect transistors (MOSFET's), $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ heterojunctions, and ultrathin p - n - p doping layers, have provided researchers with a tool to study quasi-two-dimensional electronic systems.^{1,2} The electrons are dynamically two dimensional in the sense that their motion is unrestricted in two spatial dimensions but quantized by a confining potential in the third dimension. Each electronic band of the bulk system is split into a number of two-dimensional subbands corresponding to quantum levels of the confining potential.

The properties of electrons in such two-dimensional structures have been studied extensively by transport and magnetotransport measurements.¹ In these experiments the microstructure is usually in the form of a channel which is terminated at opposite edges by uniformly doped regions in the semiconductor that serve as contacts. In the contact regions the electrons' dynamical behavior is bulklike. At the source end, a net current of carriers is injected from the contact region into the channel; at the drain a net current is ejected from the channel. Both injection and ejection involve a change in the dynamic behavior of the electron between two and three dimensional. In this paper we study these processes in detail.

A correct treatment of ejection from or injection into the channel must be explicitly quantum mechanical. For example, an electron approaching the (channel-drain) interface from the channel side has unit classical probability of being ejected into the drain. Quantum mechanically, the ejection probability is, in general, less than unity; there is a finite probability of reflection at the interface. The ejection probability depends on the incident subband, and the carrier may be reflected into a different subband of the channel. We will call the determination of the various reflection and ejection probabilities the ejection problem.

In the complementary injection problem, in which an electron in the source approaches the (source-channel) interface, the classical theory predicts a single cross section for injection into the channel. One quantum-mechanical effect is to raise the threshold energy: the classical threshold is determined by the minimum of the confining potential; the quantum threshold is raised above this by the zero-point energy of the ground state. Furthermore, one may define cross sections for injection into individual subbands, and the total cross section exhibits structure reflecting the presence of the subbands.

In present state-of-the-art semiconductor microstructures, the transition from the (three-dimensional) contact region to the (two-dimensional) channel is gradual in the sense that the mean free path of the electron is short compared to the length over which the confining potential changes from zero to its full value. Therefore, quantum-mechanical effects which require a wave coherence length larger than the transition region have not, to our knowledge, been observed. However, current trends toward increased sophistication in crystal-growth and processing capabilities lead us to expect that the transition between the channel and the contact regions will become sufficiently sharp in the near future for the effects discussed here to become observable.³

In the next section a fairly general model is described which we will use to study the ejection and injection problems. In Sec. III this model is solved formally. The connection is made with observable quantities, and some exact properties of the solutions are found, including the behavior at energies near subband minima.

In Sec. IV we treat the specific case of a parabolic confining potential. Accurate numerical solutions are found; because the problem can be solved in dimensionless form, the results found describe the behavior for all choices of effective mass and well size. Section V is a discussion and summary.

An alternative formalism, equivalent to that described in Sec. III, can be developed along the lines of conventional scattering theory. This approach is described in the Appendix.

II. MODEL

We will work in the single-particle approximation. The charge carrier will be treated in the effective-mass approximation, with a single effective mass throughout. The restriction to a single mass is most appropriate for doping structures in a homogeneous host material. It is also applicable to heterostructures if the band offset is sufficiently large to confine the carriers effectively to one material.

A single electronic band will be considered. The present model will thus be more immediately applicable to the conduction band than to the valence band, which for typical semiconductors consists of overlapping light- and heavy-hole bands.

Scattering will be neglected. Thus, our treatment will be valid only in a region within a mean free path of the interface. The transition probabilities and cross sections to be determined in succeeding sections then serve to match currents across the interface; the distributions of charge carriers on either side are determined by a Boltzmann-equation treatment far from the interface.⁴

In the context of the stated approximations, the confinement of the electron is described as arising from an effective potential. We will define a potential sufficiently general to include any quantum-well profile, but having an abrupt transition between the channel and contact regions (see Fig. 1). Calling z the direction perpendicular to the channel (the confining potential varies with z), and placing the interface along $x=0$, we write

$$H = \frac{-\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + H^\perp, \quad (1)$$

$$H^\perp = \begin{cases} H^I = \frac{-\hbar^2}{2m} \frac{\partial^2}{\partial z^2} + V^I(z) & \text{for } x < 0, \\ H^{II} = \frac{-\hbar^2}{2m} \frac{\partial^2}{\partial z^2} + V^{II}(z) & \text{for } x > 0. \end{cases} \quad (2)$$

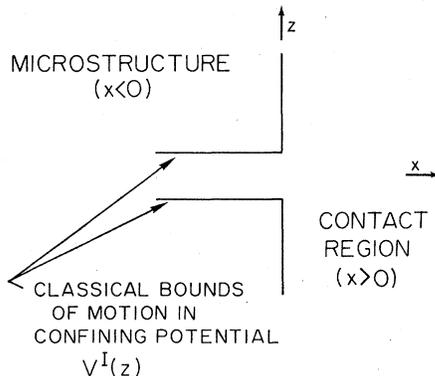


FIG. 1. Schematic of the interface region.

Motion in the y direction is trivial and the kinetic energy for motion in that direction is left out of the Hamiltonian.

The normalized eigenstates of H^I will be labeled by a lower-case Latin character. Only m and n will be used: $H^I|n\rangle = E_n|n\rangle$. Generally, the quantum well will be chosen to lie in region I ($x < 0$), so the lowest-energy eigenstates $|n\rangle$ form a discrete set.

Normalized eigenstates of H^{II} will be labeled by a lower-case Greek character; μ and ν will be used: $H^{II}|\nu\rangle = E_\nu|\nu\rangle$. In the approximation that motion in region II ($x > 0$) is unbounded, the states $|\nu\rangle$ form a continuous set. Except where a particular basis has been chosen, however, we shall use the notation of a discrete set for normalization of and sums over the states ν .

Both sets $\{|\mu\rangle\}$ and $\{|m\rangle\}$ are chosen to be complete orthonormal bases. Since the total Hamiltonian H , restricted to a single region ($x < 0$ or $x > 0$), is separable, a complete basis of product states can be found for each region. For total energy E (excluding kinetic energy in the y direction), these are $|n, \sigma\rangle$, $\sigma = \pm 1$ in region I, defined (using the coordinate representation for the x direction only) by

$$\langle x | n, \sigma \rangle \equiv \frac{1}{\sqrt{2\pi}} \exp(-i\sigma k_n x) \Theta(-x) |n\rangle. \quad (3)$$

Here, Θ is the unit step function: $\Theta(y) = 1$ for $y > 0$, $\Theta(y) = 0$ for $y < 0$. Similarly, a basis in region II is given by

$$\langle x | \nu, \sigma \rangle \equiv \frac{1}{\sqrt{2\pi}} \exp(+i\sigma k_\nu x) \Theta(+x) |\nu\rangle. \quad (4)$$

The wave vectors in Eqs. (3) and (4) have the values

$$k_n \equiv \left[\frac{2m}{\hbar^2} (E - E_n) \right]^{1/2}, \quad k_\nu \equiv \left[\frac{2m}{\hbar^2} (E - E_\nu) \right]^{1/2}, \quad (5)$$

where the "positive" square root is taken (i.e., $k = r$ or ir , where r is real and non-negative).

With the given convention for the square root, states with $\sigma = +1$ either have momentum in the direction away from the interface at $x=0$ or decay exponentially with increasing distance from the interface. Conversely, $\sigma = -1$ corresponds to motion or decay toward the interface. The physical eigenstates are bounded at infinity, so for imaginary wave vector the states with $\sigma = -1$ are excluded. It should be noted that for convenience in writing superpositions, the states $|n, \sigma\rangle$ and $|\nu, \sigma\rangle$ have been defined by (3) and (4) for all x , although each state is zero in one or the other region.

We will take an approach similar to that used in solving muffin-tin models in band structure and similar to the R -matrix approach in nuclear and atomic scattering. The solutions $|\phi\rangle$ of the Schrödinger equation with Hamiltonian H are most naturally described as superpositions of the $|n, \sigma\rangle$ in region I and of $|\nu, \sigma\rangle$ in region II. A solution valid throughout space is determined by satisfying the continuity conditions

$$\langle x | \phi \rangle_{x=0^-} = \langle x | \phi \rangle_{x=0^+} \quad (6)$$

and

$$\left[\frac{d}{dx} \langle x | \phi \rangle \right]_{x=0^-} = \left[\frac{d}{dx} \langle x | \phi \rangle \right]_{x=0^+} . \quad (7)$$

III. FORMAL THEORY

A. Wave functions

For the ejection problem we seek energy eigenstates consisting of a carrier incident on the interface in the state $|n, -\rangle$ plus a weighted sum over the outgoing states ($|m, +\rangle$, $|\mu, +\rangle$) into which the carrier scatters. We write

$$\begin{aligned} |\psi_n^{(+)}\rangle &= |n, -\rangle - |n, +\rangle + \sum_m |m, +\rangle F_{mn} k_n \\ &+ \sum_\mu |\mu, +\rangle F_{\mu n} k_n , \end{aligned} \quad (8)$$

with expansion coefficients F_{mn} , describing reflected waves, and coefficients $F_{\mu n}$, describing ejected waves, to be determined by the conditions (6) and (7). Most of the coefficients F_{mn} and $F_{\mu n}$ are amplitudes for nonpropagating waves (wave vectors k_m, k_μ are imaginary). These amplitudes do not describe any net stationary current, but they are needed to satisfy the continuity conditions.

The sum over m in (8) includes a term with $m=n$; an extra term $-|n, +\rangle$ is included explicitly in (8) to simplify the equations which determine the F coefficients. There is also a physical motivation for the form chosen. If one defines an "unperturbed" version of the problem by placing an infinite potential barrier at the interface, then one may define $|n, -\rangle - |n, +\rangle$ to be the "unperturbed" eigenstates in region I. The (large) perturbation of removing the barrier (to restore contact between the regions) induces scattering out of the unperturbed states. The F coefficients in (8) describe the rate of scattering out of the state $|n, -\rangle - |n, +\rangle$. In this picture, $-|n, +\rangle$ is the part of the unscattered wave "beyond" or after the scattering region $x=0$. The wave $|n, +\rangle F_{nn} k_n$ is the analog here of the forward-scattered wave in ordinary potential scattering. These ideas are made more precise in the Appendix, where an S -matrix approach is developed.

We now apply the continuity conditions (6) and (7) to the state described by (8), obtaining

$$\sum_m |m\rangle F_{mn} = \sum_\mu |\mu\rangle F_{\mu n} \quad (9)$$

and

$$2|n\rangle = \sum_m k_m |m\rangle F_{mn} + \sum_\mu k_\mu |\mu\rangle F_{\mu n} , \quad (10)$$

respectively.

The state $|\psi_n^{(+)}\rangle$ is defined only for those n with real wave vector k_n ($E > E_n$); F coefficients for other n have not been defined. There is nothing to prevent our defining the latter, however, by an extension of (9) and (10) to all n . We can then define a linear operator

$$F \equiv \sum_{m,n} |m\rangle F_{mn} \langle n| , \quad (11)$$

with matrix elements $F_{mn} = \langle m | F | n \rangle$. Equation (9) may now be interpreted as a transformation rule, demonstrating that the $F_{\mu n} = \langle \mu | F | n \rangle$ are different matrix elements of the same operator.

If we define operators

$$k^I \equiv \sum_m |m\rangle k_m \langle m| , \quad (12)$$

$$k^{II} \equiv \sum_\mu |\mu\rangle k_\mu \langle \mu| , \quad (13)$$

and

$$K \equiv \frac{1}{2}(k^I + k^{II}) , \quad (14)$$

then Eq. (10) becomes $1 = KF$, and the solution (8) is completely specified by

$$F = K^{-1} . \quad (15)$$

For the ejection problem we can define, in analogy with (8) for the injection problem, a state

$$|\psi_\nu^{(+)}\rangle = |\nu, -\rangle - |\nu, +\rangle + \sum_M |M, +\rangle F_{M\nu} k_\nu , \quad (16)$$

which satisfies the boundary condition that the incident component of the wave function at infinity is $|\nu, -\rangle$. In Eq. (16) and below, capital letters (M, N) are used to label energy eigenstates and their eigenvalues for either H^I or H^{II} . A sum over M will mean, unless otherwise indicated, a sum over all m and all μ .

The same procedure used to find the F_{Mn} shows that the $F_{M\nu}$ are matrix elements $\langle M | F | \nu \rangle$ of the operator F given by Eq. (15).

B. Formal properties of the solutions

1. Reciprocity relations

The full model Hamiltonian H , as well as H^I and H^{II} , is invariant under time reversal, which gives rise to reciprocity relations among the matrix elements of F . These are most easily derived by choosing the wave functions $\langle z | M \rangle$ to be real. In this case the transformation matrix $\langle m | \mu \rangle$ is orthogonal, making K_{mn} , and hence F_{mn} , symmetric. All reciprocity relations can be derived by means of basis transformations from the symmetry of F in a real basis.

For any pair of time-reversal-invariant states M and N ,

$$|F_{MN}|^2 = |F_{NM}|^2 , \quad (17)$$

because such states are proportional to real states. It is noted that, for a one-dimensional Hamiltonian, all states bound at positive or negative infinity are nondegenerate and thus time-reversal-invariant. In particular, this implies that relation (17) holds for any pair of discrete states.

Of special interest is the case of a constant potential in region II, for which it is convenient to choose momentum eigenstates

$$\langle z | \mu \rangle = \frac{1}{\sqrt{2\pi}} \exp(i\mu z) , \quad (18)$$

with transverse (i.e., z -component) momentum $\hbar\mu$. The states $|\mu\rangle$ and $|\mu\rangle$ go into each other under time reversal. Thus,

$$|F_{\mu\nu}|^2 = |F_{-\nu, -\mu}|^2 \quad (19)$$

and

$$|F_{\mu N}|^2 = |F_{N, -\mu}|^2 \quad (20)$$

for a state $|N\rangle$ invariant under time reversal.

Reflection across the x axis also reverses the transverse momentum, so for a symmetric confining potential [$V^I(z) = V^I(-z)$] symmetry under the simultaneous operations of time reversal and reflection extends the validity of (17) to include momentum eigenstates among the possible M, N .

The eigenstates of a one-dimensional Hamiltonian are, at most, doubly degenerate (degeneracy occurs when the energy exceeds the potential at both positive and negative infinity). Thus, the most general reciprocity relation, valid for any choice of basis, is

$$\sum (|F_{NN'}|^2 - |F_{N'N}|^2) = 0, \quad (21)$$

where the sum runs over either one or both state(s) $|N\rangle$ in a region with the single energy E , and the one or two states $|N'\rangle$ (in the same or different region) with energy E' .

2. Optical theorems

The solution (15) for F yields the identity

$$F + F^\dagger = F^\dagger(K^\dagger + K)F. \quad (22)$$

The N, N' matrix element of (22) can be expressed [using Eqs. (12)–(14)] as

$$F_{N'N} + F_{NN'}^* = \sum'_M F_{MN'}^* F_{MN} k_M, \quad (23)$$

where the prime on the sum over M indicates that the sum is over all states with real wave vector k_M ($E > E_M$). For $N' = N$, this becomes

$$2\text{Re}(F_{NN}) = \sum'_M |F_{MN}|^2 k_M. \quad (24)$$

In the terminology of formal scattering theory, (23) and (24) are the generalized optical theorem and the optical theorem, respectively [see remarks following Eq. (A.12)].

C. Kinematics

Having found wave functions appropriate to the ejection and injection problems (Sec. III A), we now seek to relate experimentally accessible quantities to the matrix elements of F which describe these wave functions. For this section we will consider a confining potential in region I that has only bound states [$V^I(z) \rightarrow \infty$ as $|z| \rightarrow \infty$] and a potential in region II that is identically zero.

The different dynamics of the two regions lead to descriptions of the phenomena in terms of two different kinds of quantities. In region I the eigenstates $|m\rangle$ describing transverse motion may be normalized to con-

tain one carrier, so ejection from a particular subband is described in terms of probabilities of reflection into various allowed subbands and in terms of an angular probability density for ejection (generically, transition probabilities). In region II, eigenstates $|\mu\rangle$ must be normalized not to a number but to a (lineal) density of carriers. Therefore, the injection problem is concerned with cross sections (having units of length) for injection into various subbands, and a differential cross section for scattering.

We consider the ejection problem first. The current density of the ingoing and outgoing parts of $|\psi_m^{(+)}\rangle$ is integrated over a surface at $x < 0$ to find the incident and reflected fluxes. The ratio of these is the probability

$$R_{mn} = \frac{k_n}{k_m} |\delta_{mn} - F_{nm} k_m|^2 \quad (25)$$

that a carrier in the m th subband will be reflected back into the n th subband (when k_n is real). By the reciprocity relation (17), the off-diagonal (out-of-subband) reflection probabilities are symmetric: $R_{mn} = R_{nm}$.

For a general potential in region II, expressions similar to (25) can be found for the transmission probabilities. As we are considering only the case of zero potential ($V^{\text{II}} = 0$), we adopt a basis of momentum eigenstates [Eq. (18)]. Using polar coordinates in region II ($x = r \cos\theta$, $z = r \sin\theta$, $|\theta| \leq \pi/2$, $r \geq 0$), a stationary-phase analysis of $|\psi_m^{(+)}\rangle$ at large positive x yields

$$\psi_m^{(+)} \sim \left[\frac{Q}{2\pi r} \right]^{1/2} f_m(\theta) \cos\theta \exp[i(Qr + \pi/4)], \quad (26)$$

where $Q = (2mE/\hbar^2)^{1/2}$ is the total wave vector of a carrier in region II, and $f_m(\theta) = F_{\mu m}$ for $\mu = Q \sin\theta$. Thus, a carrier approaching the interface in the m th subband has a probability

$$T_m(\theta) d\theta = k_m Q^2 \cos^2\theta |f_m(\theta)|^2 d\theta \quad (27)$$

of being ejected into a small angular range $d\theta$ about the angle θ .

The conservation of probability current can be expressed as

$$\sum'_n R_{mn} + \int_{-\pi/2}^{\pi/2} T_m(\theta) d\theta = 1 \quad (28)$$

(the prime indicates that the sum is over states with real wave vector k_n). This also follows from the optical theorem (24).

We now consider the injection problem. Injection cross sections are defined as ratios of injected flux to incident-flux density. The cross section for injection into the m th subband from a carrier incident with total momentum $\hbar Q$ from the θ direction is

$$\Lambda_m(\theta) = 2\pi Q \text{Re}(k_m) \cos^2\theta |f'_m(\theta)|^2, \quad (29)$$

where $f'_m(\theta) = F_{m, -Q \sin\theta}$. The reciprocity relation (20) implies that $|f'_m(\theta)|^2 = |f_m(\theta)|^2$, so that

$$\Lambda_m(\theta) = \frac{2\pi}{Q} T_m(\theta), \quad (30)$$

and the angle-averaged cross section for injection into a

subband m ($\bar{\Lambda}_m$) is $2/Q$ times the total probability of ejection from the m th subband. As the total ejection probability is bounded above by 1, this imposes a "unitarity limit" of $2/Q$ on $\bar{\Lambda}_m$. Because the classical ejection probability is 1, we shall call this the "quasiclassical" upper bound.

In general, a carrier in the contact region has a nonzero cross section for being reflected nonspecularly (scattered) from the interface. It is noted that for the particular model chosen here, in which the potential change at $x=0$ is abrupt, the classical prediction is of zero differential scattering cross section.

D. Threshold behavior

As the total energy E increases through the minimum of a subband m , a new "channel" (in the sense of scattering theory) becomes available: For $E > E_m$, the initial and/or final state may be in the m th subband. The energy E_m is called a threshold.

Some quantities, such as R_{mn} and $\Lambda_m(\theta=0)$, are undefined or identically zero below E_m and must exhibit some nonanalytic threshold behavior. The matrix elements of F are constant to lowest order in $E - E_m$, so in Eq. (29) the dominant threshold behavior arises from the factor $\text{Re}(k_m) = k_m \theta(E - E_m)$, and $\Lambda_m(\theta=0)$ increases as $(E - E_m)^{1/2}$ above threshold. Similarly, the angle-averaged injection cross section $\bar{\Lambda}_m$ and the probabilities for ejection and out-of-subband reflection involving the m th subband all grow asymptotically as the one-half power of $E - E_m$ above threshold.

This threshold behavior is controlled by explicit factors of k_m in the expressions for cross sections and transition (i.e., ejection, reflection) probabilities. Thus, it is associated only with processes having the initial state or final state, or both, in the new subband; we call these threshold effects kinematic.

There is also threshold structure associated with singular variation of F . By expanding Eq. (15) in powers of the small quantity k_m , one finds

$$F_{NN'} = F_{NN'}^{\text{th}} - (2m/\hbar^2)^{1/2} F_{Nm}^{\text{th}} F_{mN'}^{\text{th}} (E - E_m)^{1/2} + O(E - E_m), \quad (31)$$

with the square-root convention of Eq. (5), and $F^{\text{th}} = F(E = E_m)$. The first-order correction in (31) gives rise to an asymmetric cusp in the probabilities (or cross sections) for all processes which involve initial and final states both coupled to the m th subband. (That is, F_{Nm} and $F_{mN'}$ both nonzero; this is assumed to be the case unless forbidden by symmetry.)

In terms of a perturbation theory (such as one associated with the S matrix of the Appendix), the cusp in F results from intermediate states in the m th subband; we call the associated threshold effects dynamic.

IV. NUMERICAL RESULTS

A. Truncation approximation

We shall find it convenient to solve for F in the discrete representation (corresponding to the basis of states in the

confined region), in which

$$K_{mn} = \frac{1}{2} \left[k_m \delta_{mn} + \sum_{\mu} \langle m | \mu \rangle k_{\mu} \langle \mu | n \rangle \right]. \quad (32)$$

If K_{mn} were diagonal, then its inverse, F_{mn} , would also be diagonal, with its diagonal elements the inverses of the corresponding elements of K . Equivalently, one may say that inversion of K commutes with projection of the matrix into a space spanned by an individual eigenvector.

In practice, the magnitudes of the elements K_{mn} tend to decrease with increasing distance from the diagonal $m = n$, so that there is only a small matrix element directly connecting distant levels. This suggests a natural approximation scheme based on the assumption that inversion commutes approximately with projection into a moderately large subspace.

For the example discussed in the following section, this idea was implemented by defining a finite square K_{mn} matrix for index values up to some chosen truncation level N_T : $m, n \leq N_T$. The finite matrix is readily inverted to give an approximation to F_{mn} . This truncation approximation is particularly useful because the reflection probabilities are described completely by the finite set of F_{mn} corresponding to real k_m, k_n ("small" m, n).

The truncation approximation can be improved systematically by increasing the truncation level. Observation of the convergence of F as the truncation level is increased can be used to estimate the accuracy of the procedure.

We take all matrix elements F_{mn} not explicitly calculated at some truncation level (i.e., for m or $n > N_T$) to be identically zero. With this choice, other matrix representations of the operator F are defined consistently in the truncation approximation by the usual transformations from the purely discrete representation. For example,

$$F_{\mu n} = \sum_{m \leq N_T} \langle \mu | m \rangle F_{mn}. \quad (33)$$

With this convention, it can be shown that the operator identity (22) holds, so the truncation approximation satisfies the optical theorems.

The reciprocity relations (Sec. III B 1) also hold identically in the truncation approximation. These were proven for the exact F to follow from certain symmetry properties of F , and it is easily confirmed that the same symmetries hold in the present case.

B. Example of parabolic confining potential

As a concrete example, we now choose the confining potential to be parabolic and let the potential in region II be zero:

$$V^I(z) = \frac{m\omega^2}{2} z^2, \quad (34a)$$

$$V^{II}(z) = 0. \quad (34b)$$

In a semiconductor microstructure one may think of the parabolic potential as that produced by the space charge of a thin layer of ionized donors sandwiched between layers of ionized acceptors (p - n - p structure).^{5,6}

We choose $\hbar\omega$ to be the energy unit, so H^I has energy levels $E_n = n + \frac{1}{2}$, $n=0,1,2,\dots$. With the length unit chosen to be $(\hbar/m\omega)^{1/2}$, the Hamiltonian takes the form

$$H = -\frac{1}{2} \left[\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} \right] + \frac{1}{2} z^2 \Theta(-x), \quad (35)$$

which contains no adjustable parameters.

The total Hamiltonian H has reflection symmetry about $z=0$, so parity is conserved. Since eigenstates $|n\rangle$ of H^I have parity $(-1)^n$, F_{mn} is zero unless n and m are both even or both odd. Physically, this means that a carrier in the channel can be reflected from the interface only into a subband of the same parity. It also means that dynamical threshold effects at $E=E_m$ will not be observed in reflections into states of parity opposite that of state m .

From a computational point of view, reflection symmetry allows one to apply the truncation approximation separately to the even and odd states, avoiding the calculation of matrix elements that are identically zero and diminishing both computation time and storage space by a factor of between $\frac{1}{2}$ and $\frac{1}{4}$.

Another computational advantage is associated with the symmetry of the confining potential under an interchange of momentum and coordinate: The wave functions have the same form in the momentum as in the coordinate representation. Thus, the transformation matrix $\langle \mu | m \rangle$ to harmonic-oscillator eigenstates of momentum μ , needed to determine k_{mn}^{II} , does not require a separate computation.

We have computed the amplitude matrix F in the truncation approximation, using eight even and eight odd states. For the energies investigated ($0 < E \lesssim 12$), the values of the matrix elements F_{mn} converge quickly with increasing truncation level. For the energies plotted in the curves below, errors in computation of probabilities are believed to be smaller than 10^{-4} .

Reflection probabilities are plotted in Fig. 2 for energies between 0.5 and 4.0. The diagonal reflection probabilities (R_{nn}) fall off quickly with energy from the threshold value of unity; off-diagonal reflection probabilities in-

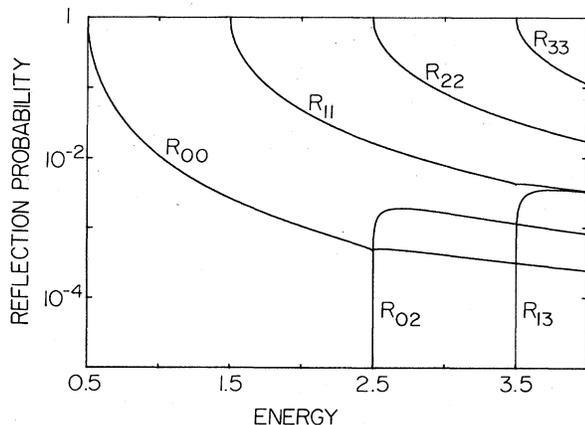


FIG. 2. Reflection probabilities involving the lowest few subbands. Energy is in units of the subband spacing.

crease quickly at threshold. Both effects follow the predictions of kinematic threshold behavior described in Sec. III D.

For comparison, we consider a one-dimensional (1D) analog of the present model, in which an electron moves in a potential $V(x) = V_0 \Theta(-x)$, with $V_0 > 0$. This problem is easily solved; the probability of reflection for a carrier incident from $x < 0$ (region I) is given by

$$R_{1D} = \tanh^2 \left[\frac{1}{4} \ln \left| 1 - \frac{V_0}{E} \right| \right]. \quad (36)$$

This one-dimensional reflection probability has the same form of threshold behavior as the three-dimensional result.

For an estimate of the reflection probability from subband n , we choose $V_0 = E_n$. For the range of energies studied here, R_{1D} was always greater than the total reflection probability, typically by a factor of about 2–3. This approximate tracking of one- and two-dimensional behavior covers a range of 4–5 orders of magnitude in reflection probabilities and a range in energy of from 0 to 12.

The one-dimensional model clearly does not have a process analogous to off-diagonal (out-of-subband) reflection. It is striking that the off-diagonal reflection probability can be quite large. A carrier in the zeroth band, with energy greater than $\sim 1.01 E_2$, is actually more likely to be reflected back in the second subband than in the original subband. The ratio of off-diagonal to diagonal reflection probabilities falls off with increasing subband index and also (slowly) with energy.

The total angle-averaged injection cross section ($\bar{\Lambda} = \sum_m \bar{\Lambda}_m$) is plotted as curve (a) in Fig. 3 for energies between 0.0 and 5.0. For comparison, the corresponding classical prediction (for the same Hamiltonian) is also plotted [curve (b)]. The classical cross section $\bar{\Lambda}_{cl}$ is found to be $(2E)^{1/2}$ by noting that the classical condition for injection is that the kinetic energy associated with motion in

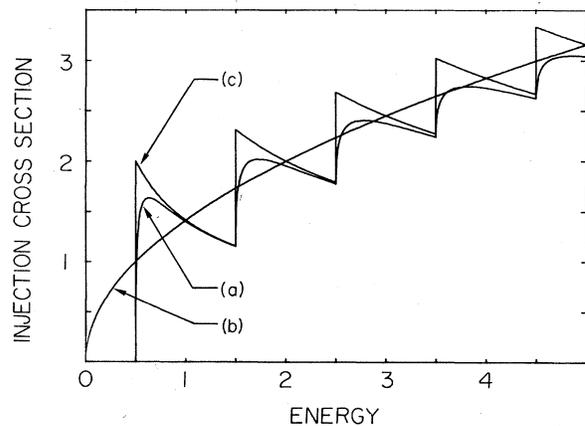


FIG. 3. (a) Total cross section for injection, averaged over angle of incidence, in dimensionless length units as described in the text. (b) Classical value of same. (c) Quasiclassical upper bound.

the negative x direction be greater than the increase in potential energy (V^1) at the point of impact on the interface.

The quantum-mechanical cross section follows the general upward trend of the classical cross section, but has additional features associated with the subband structure. A part of this may be thought of as arising from a phase-space factor, which reflects the number of available injection states. The quasiclassical upper bound on $\bar{\Lambda}$, which is $2/E^{1/2}$ multiplied by the number of energetically accessible subbands, exhibits the general steplike behavior associated with the subband structure; it is plotted as curve (c) in Fig. 3. Because the number of subband minima with energy less than E is about equal to $E/\hbar\omega$, the quasiclassical bound is asymptotic to the classical cross section at high energies. The actual cross section comes close to the upper bound, reflecting the fact (cf. Fig. 2) that the ejection probability approaches unity quickly.

The initial rise in the cross section above each threshold is again a kinematic threshold effect. Dynamic threshold effects that can be distinguished from kinematic ones are again small.

V. SUMMARY AND DISCUSSION

We have solved formally a model for a sharp interface between a channel in which electrons are confined to move in two dimensions and a contact region in which their motion is unconfined. Expressions were found for injection cross sections and various transition probabilities. Choosing a parabolic confining potential, we have obtained numerical results for energies on the order of 0–12 subband spacings.

Rigorous qualitative results include the finding that, at energies just above a subband minimum E_m , cross sections and transition probabilities involving the m th subband exhibit kinematic threshold singularities proportional to $(E - E_m)^{1/2}$. There are useful reciprocity relations which lead, *inter alia*, to an upper bound on the injection cross section that does not depend on details of the confining potential.

Numerical results indicate that good electrical contact is established between the bulklike contact region and a subband in the channel at energies not far above the subband minimum. A measure of this is the approach of the ejection probability to unity, or equivalently the approach of the injection cross section to its quasiclassical upper bound. The subband structure is reflected in the steplike variation of the injection cross section as a function of energy.

Off-diagonal reflection (reflection back into the channel accompanied by a change of subband) is the dominant reflection mode at most energies for carriers approaching the interface in the lowest subband. The extent of off-diagonal reflection decreases with increasing subband index.

The numerical results obtained for the parabolic confining potential are expected to be qualitatively correct for nonparabolic potentials. If the potential is approximately parabolic, then quantitative results should be approximately correct if the subband spacing is used as the energy unit $\hbar\omega$ and if cross sections are scaled by the length

unit $(\hbar/m\omega)^{1/2}$.

The selection of a model with a sharp interface is most appropriate for situations in which the electron wavelength is long compared to the width of the actual interface. The present results should be qualitatively correct, moreover, when the wavelength is shorter.

The key approximation limiting the validity of this work is the neglect of scattering. It is necessary, for any of the present results to be valid, that the mean free path be long compared to the width of the transition between channel and contact regions.

ACKNOWLEDGMENTS

We thank F. Crowne and T. L. Reinecke for useful discussions. This work was supported in part by an Office of Naval Research contract. One of us (A.M.K.) acknowledges support from the National Research Council.

APPENDIX: S-MATRIX THEORY

In this appendix we show how the problem studied in the body of the paper may be treated by an S -matrix formalism resembling that conventional for the scattering of particles in free space. Rather than develop the whole formalism explicitly, we shall concentrate on exhibiting the connection with the F -matrix formalism of Sec. III, and quote freely some well-known results of formal scattering theory.⁷

To begin we generalize the model Hamiltonian [Eqs. (1) and (2)] by adding a delta-function potential barrier at the interface between the two regions. That is, we define

$$H_0 \equiv H + V_0 \delta(x), \quad (\text{A1})$$

which we call the *unperturbed*, or *uncontacted*, Hamiltonian. In this language the problem considered previously becomes the *perturbed* problem which is created by adding a perturbation $V = -V_0 \delta(x)$.

A particular basis of eigenstates for the unperturbed Hamiltonian H_0 may be found in the same manner as previously (Sec. III A) for the perturbed Hamiltonian H . Seeking states identical in form to (8) and (16), we write

$$|\psi_{Nk}\rangle = |N, -\rangle - |N, +\rangle + \sum_M |M, +\rangle F_{MN}^0 k_N, \quad (\text{A2})$$

where $k = k_N(E)$ is explicit to distinguish states of different energy.

The addition of a potential barrier leaves the continuity condition (6) for the wave function unchanged, so that the amplitudes F_{MN}^0 may again be interpreted as matrix elements of a single (energy-dependent) operator: $F_{MN}^0 = \langle M | F^0 | N \rangle$.

The delta-function barrier does, however, introduce a discontinuity in the wave-function derivative, so that the derivative matching condition (7) becomes

$$\left[\frac{d}{dx} \langle x | \phi \rangle \right]_{x=0^+} - \left[\frac{d}{dx} \langle x | \phi \rangle \right]_{x=0^-} = v_0 \langle x=0 | \phi \rangle, \quad (\text{A3})$$

where $v_0 \equiv 2mV_0/\hbar^2$. This leads to a modified equation for the amplitude operator F^0 :

$$F^0 = (K + iv_0/2)^{-1}. \quad (\text{A4})$$

Formal scattering theory proceeds by defining scattering solutions $|\psi_{Nk}^{(\pm)}\rangle$ of the full Hamiltonian that are related to unperturbed states of the same energy via the Lippmann-Schwinger equation:

$$|\psi_{Nk}^{(\pm)}\rangle = |\psi_{Nk}\rangle + (E - H_0 \pm i\eta)^{-1} V |\psi_{Nk}^{(\pm)}\rangle. \quad (\text{A5})$$

The Green's functions $(E - H_0 \pm i\eta)^{-1}$ (with $\eta \rightarrow 0^+$) are chosen so that the outgoing (ingoing) states $|\psi_{Nk}^{(+)}\rangle$ ($|\psi_{Nk}^{(-)}\rangle$) approach asymptotically the unperturbed state $|\psi_{Nk}\rangle$ in the remote past (future). In particular, this means that $|\psi_{Nk}^{(+)}\rangle$ satisfies the same incident-wave boundary condition as $|\psi_{Nk}\rangle$. This is the boundary condition that defines $|\psi_{Nk}^{(+)}\rangle$ of Eqs. (8) and (16), so this latter is identically the outgoing scattering state which solves (A5). In similar fashion, the incoming states are found to be

$$|\psi_{Nk}^{(-)}\rangle = |N, -\rangle - |N, +\rangle - \sum_M |M, -k_M^*\rangle F_{MN}^* k_N, \quad (\text{A6})$$

where $|M, k\rangle$ stands for $|M, \sigma\rangle$ with $k \equiv \sigma k_M$.

If we allow v_0 to become large, the unperturbed states $|\psi_{Mk}\rangle$ approach $|M, -k_M\rangle - |M, +k_M\rangle$, which are orthonormal. Scattering states have the same orthogonality relations as the unperturbed states, so

$$\langle \psi_{Mk}^{(+)} | \psi_{Nk'}^{(+)} \rangle = \langle \psi_{Mk}^{(-)} | \psi_{Nk'}^{(-)} \rangle = \delta_{MN} \delta(k - k'). \quad (\text{A7})$$

That is, outgoing and incoming scattering states separately form orthonormal bases. As these bases can be transformed continuously into a complete basis (that for v_0 infinite), and as there are no bound states, it is assumed that $\{|\psi_{Mk}^{(+)}\rangle\}, \{|\psi_{Mk}^{(-)}\rangle\}$ are complete bases.

The transformation matrix between these bases is called the S matrix:

$$S_{Mk, Nk'} \equiv \langle \psi_{Mk}^{(-)} | \psi_{Nk'}^{(+)} \rangle. \quad (\text{A8})$$

By construction, the S matrix is independent of the choice of v_0 of the unperturbed Hamiltonian. (For heuristic reasons, however, and because some important results follow from taking the limit, it is useful to think of $v_0 \rightarrow \infty$ as defining "the" unperturbed Hamiltonian.) As it is a transformation between complete orthonormal bases, S is unitary:

$$\int_0^\infty dk'' \sum_M S_{Mk'', Nk}^* S_{Mk'', N'k'} = \delta_{NN'} \delta(k - k'). \quad (\text{A9})$$

Because energy is conserved, the S matrix is proportional to a delta function in energy:

$$S_{Nk, Mk'} = \delta_{NM} \delta(k - k') - 2\pi i \delta(E - E') T_{NM}. \quad (\text{A10})$$

In this equation, T_{NM} is the transition matrix, which may be expressed as

$$T_{NM} = \langle \psi_{Nk}^{(-)} | V | \psi_{Mk} \rangle. \quad (\text{A11})$$

This does not depend on the choice of v_0 ; if v_0 is allowed to approach infinity, F_{NM}^0 is asymptotically $(2/iv_0)\delta_{NM}$, and one finds

$$T_{NM} = \frac{\hbar^2}{2\pi i m} F_{MN} k_M k_N. \quad (\text{A12})$$

Equations (A10) and (A12) establish the correspondence between the F -matrix and S -matrix formalisms.

The generalized optical theorem is the expression of the unitarity of the S matrix in terms of the transition matrix T ; using relation (A12), this takes the form of Eq. (23).

*Present address: Honeywell Physical Sciences Center, 10701 Lyndale Avenue South, Bloomington, MN 55420.

¹For a comprehensive review up to 1982, see T. Ando, A. B. Fowler, and F. Stern, *Rev. Mod. Phys.* **54**, 437 (1982).

²For some recent work, see Proceedings of the Fifth International Conference on Electronic Properties of Two-Dimensional Systems, Oxford, 1983 [*Surf. Sci.* **142**, Nos. 1-3 (1984)].

³C. E. C. Wood, *Progress, Problems, and Applications of Molecular-Beam Epitaxy*, Vol. II of *Physics of Thin Films*,

edited by G. Hass and M. H. Francombe (Academic, New York, 1980).

⁴B. E. Sernelius, K.-F. Berggren, M. Tomak, and C. McFadden, *J. Phys. C* **18**, 225 (1985).

⁵G. H. Döhler, *Surf. Sci.* **73**, 97 (1978).

⁶H. Künzel, G. H. Döhler, A. Fischer, and K. Ploog, *Appl. Phys. Lett.* **38**, 171 (1981).

⁷For a review, see Eugen Merzbacher, *Quantum Mechanics* (Wiley, New York, 1961).