Generalization of the WKB Theory for Tunneling between Two Different One-Dimensional Periodic Potentials*

T. E. Feuchtwang
Department of Physics, The Pennsylvania State University, University Park, Pennsylvania 16802
(Received 10 April 1969; revised manuscript received 15 January 1970)

A general analysis of tunneling in a one-dimensional heterojunction is given. The analysis is given in terms of a generalization of the WKB theory specifically developed to apply to potentials which, as x tends to $\pm\infty$, tend to periodic functions of position, rather than to constants. The theory is particularly suited for narrow junctions. The tunneling probability is shown to factor into a bulk factor which is proportional to the product of the group velocities in the periodic potentials on both sides of the junction, and into a barrier factor. The latter depends primarily on solutions of the barrier Hamiltonian, and only for a thick junction may it be approximated by a simple exponential.

I. INTRODUCTION

We consider the tunneling between two different one-dimensional periodic potentials which are joined smoothly. We assume that the transition from one (perturbed) periodic potential to the other is primarily localized in a relatively narrow region. The potential in this "junction" may differ significantly from either of the periodic potentials. That is, we assume that the continuously differentiable potential $V(x)$ can be decomposed into three terms,

$$V(x)=\theta(-x+X_l)V_l(x)+\theta(x-X_r)V_r(x)+V'(x)\ . \quad (1.1)$$

Here, $V'(x)$ is appreciable only in the junction proper, $X_l<x<X_r$, and vanishes like $|x|^{-(2+\eta)}$ as $|x|$ tends to infinity.[1] $V_{l,r}(x)$ are two distinct periodic potentials,

$$V_{l,r}(x)=V_{l,r}(x+na_{l,r})\ , \quad (1.2)$$

whose mean values are

$$\overline{V}_{l,r}=a_{l,r}^{-1}\int_0^{a_{l,r}}V_{l,r}(x)\,dx\ . \quad (1.3)$$

In our analysis we require an approximation of the linearly independent solutions of the Hamiltonian outside the junction. The solutions which we derive, in Sec. II, represent a proper generalization of the ordinary WKB functions,

$$\psi_{1,2}(x)=[Q(x)]^{-1/2}\exp[\pm i\int^x Q(x_1)\,dx_1]\ , \quad (1.4)$$

$$Q(x)=[E-V(x)]^{1/2}\ . \quad (1.5)$$

Such a generalization is necessary because our potential violates the conditions under which the ordinary WKB approximation is valid.[2,3] One possible way to overcome this difficulty is to estimate the rapidly varying periodic components of the potential by means of the effective Hamiltonian theory, and to apply the ordinary WKB approximation to the effective Hamiltonian $E(-i\nabla)+V'(x)$.[4,5] However, the effective Hamiltonian theory was developed for problems involving a single periodic potential: The theory leads to conceptual and mathematical difficulties when it is applied to problems such as ours, involving several periodic potentials.[6] We therefore chose to develop a generalization of the WKB approximation which is explicitly valid when

$$\lim_{x\to\pm\infty}V(x)=V_{r,l}(x)=V_{r,l}(x+na_{r,l})\ ,$$

and consequently does not have to involve the effective Hamiltonian theory.

Our theory is based on an adaptation of the technique of "variation of parameters" for the solution of inhomogenous differential equations.[7] The procedure is outlined in Sec. I A below. In the course of this outline, we also introduce the notation to be used in the remainder of this paper. In Sec. III, we discuss the joining relations appropriate for our generalization of the WKB theory. The entire theory is then applied, in Sec. IV, to calculate the transmission coefficient for a heterojunction. It is shown that this quantity is always proportional to the product of the group velocities in the two periodic potentials, divided by the product of the extended wave vectors which characterize the two asymptotic states. This result contradicts the conclusion of Harrison[5] about the energy dependence of the transmission coefficient within a (generalized) WKB theory. Our results are summarized and interpreted in Sec. V.

A. WKB Approximation from a "Variation of Parameters" Point of View

The solutions of

$$[\mathcal{H}_0 - E - (\hbar^2/2m)\epsilon(x)]\,\Psi = 0 \tag{1.6}$$

can be represented in terms of two linearly independent solutions $\psi_{1,2}(x)$ of

$$(\mathcal{H}_0 - E)\psi = 0 \ , \tag{1.7}$$

and a pair of "envelope functions" $a_{1,2}(x)$ such that

$$\begin{pmatrix} \Psi(x) \\ \frac{d}{dx}\Psi(x) \end{pmatrix} = \begin{pmatrix} \psi_1(x) & \psi_2(x) \\ \frac{d}{dx}\psi_1(x) & \frac{d}{dx}\psi_2(x) \end{pmatrix} \begin{pmatrix} a_1(x) \\ a_2(x) \end{pmatrix} . \tag{1.8}$$

The functions $a_{1,2}(x)$ satisfy a pair of coupled first-order differential equations, whose general solution can be written in terms of their fundamental solution matrix $\vec{\mathrm{T}}(x,x')$ and the values of $a_{1,2}(x)$ at the arbitrary fixed point $x = x'$,

$$\begin{pmatrix} a_1(x) \\ a_2(x) \end{pmatrix} = \vec{\mathrm{T}}(x,x') \begin{pmatrix} a_1(x') \\ a_2(x') \end{pmatrix} . \tag{1.9}$$

The T matrix $\vec{\mathrm{T}}(x,x')$ satisfies the integral equation

$$\vec{\mathrm{T}}(x,x') = \vec{\mathrm{I}} + \int_{x'}^{x} \vec{\mathrm{M}}(x_1)\vec{\mathrm{T}}(x_1,x')\,dx_1 \ , \tag{1.10}$$

where

$$M = \epsilon(x)\psi_1(x)\psi_2(x)[W]^{-1} \times \begin{pmatrix} 1 & \psi_2(x)[\psi_1(x)]^{-1} \\ -\psi_1(x)[\psi_2(x)]^{-1} & -1 \end{pmatrix} , \tag{1.11}$$

and W is the x-independent Wronskian determinant of $\psi_1(x)$ and $\psi_2(x)$. Combining Eqs. (1.8) and (1.9), we obtain the matrix equation

$$\vec{\Psi}(x) = \vec{\mathrm{W}}(\psi_1(x),\psi_2(x))\vec{\mathrm{T}}(x,x')\vec{\mathrm{a}}(x') \ , \tag{1.12}$$

where $\vec{\mathrm{W}}$ is the Wronskian matrix of ψ_1 and ψ_2. Froman and Froman[8] have shown that the WKB theory can be formulated as a special case of the preceding analysis, in which the model Hamiltonian is chosen so that

$$|T_{ij}(\infty,x) - \delta_{ij}| \le A_{ij}\{\exp[B\delta(\infty,x)] - 1\} \quad \text{if } x > X_r \tag{1.13}$$

and

$$|T_{ij}(x,-\infty) - \delta_{ij}| \le C_{ij}\{\exp[D\delta(x,-\infty)] - 1\} \quad \text{if } x \le X_l \ . \tag{1.14}$$

Here A_{ij}, B, C_{ij}, and D are appropriate positive constants,

$$\delta(x,x') = \int_{x'}^{x} |q(x_1)\epsilon(x_1)|\,dx_1 \ , \tag{1.15}$$

where $q(x)$ is an appropriate local wave vector, and

$$\lim_{X_r < x \to \infty} \delta(\infty,x) = 0 = \lim_{X_l > x \to \infty} \delta(x,-\infty) \ . \tag{1.16}$$

Equations (1.13)–(1.16) specify bounds on $\vec{\mathrm{T}}$ which imply that for $x > X_r$ or $x < X_l$ this matrix may be approximated by a unit matrix. The approximation involves an error whose leading term is proportional to the "error integrals," $\delta(\infty,x)$ and $\delta(x,-\infty)$. That is, over these intervals the linearly independent solutions of $\mathcal{H} - E$ may be approximated by ψ_1 and ψ_2. The error involved in this approximation can be estimated with the help of Eqs. (1.13)–(1.15). This error increases as x tends to $X_l(X_r)$ from below (above), and $X_{l,r}$ are the upper (lower) boundaries of the intervals over which the error is acceptable. As a rule, the interval $X_l < x < X_r$ contains one or more points at which $\epsilon(x)$ diverges. In the neighborhood of these classical turning points, the approximation fails completely.[9] Although the integral equation (1.10) is singular at the turning points, the matrix $\vec{\mathrm{T}}(X_r,X_l)$ exists. Consequently Eq. (1.12) can still be used to represent $\Psi(x > X_r)$ in terms of $\Psi(x < X_l)$. The last two assertions can be proved by the following argument: Let the matrix $\vec{\mathrm{R}}(x,x')$ be the fundamental solution matrix of the Schrödinger equation

$$\begin{pmatrix} \frac{d}{dx}\Psi \\ \frac{d^2}{dx^2}\Psi \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\frac{2m}{\hbar^2}[E - V(x)] & 0 \end{pmatrix} \begin{pmatrix} \Psi \\ \frac{d}{dx}\Psi \end{pmatrix} , \tag{1.17}$$

which we can rewrite as

$$\frac{d}{dx}\vec{\Psi} = \vec{\mathrm{A}}(x)\vec{\Psi}(x) \ . \tag{1.18}$$

Then, the regularity of the matrix $\vec{\mathrm{A}}(x)$ implies that

$$\vec{\Psi}(X_r) = \vec{\mathrm{R}}(X_r,X_l)\vec{\Psi}(X_l) \ . \tag{1.19}$$

Combining Eqs. (1.12)–(1.14) and (1.19) we now have

$$\vec{\mathrm{W}}(\psi_1(X_r),\psi_2(X_r))\vec{\mathrm{a}}(X_r) = \vec{\mathrm{R}}(X_r,X_l)\vec{\mathrm{W}}(\psi_1(X_l),\psi_2(X_l))\vec{\mathrm{a}}(X_l) \ . \tag{1.20}$$

Hence,

$$\vec{\mathrm{a}}(X_r) = \{[\vec{\mathrm{W}}(\psi_1(X_r)\psi_2(X_r))]^{-1} \times \vec{\mathrm{R}}(X_r,X_l)\vec{\mathrm{W}}(\psi_1(X_l),\psi_2(X_l))\}\vec{\mathrm{a}}(X_l) \ , \equiv \vec{\mathrm{T}}(X_r,X_l)\vec{\mathrm{a}}(X_l) \ . \tag{1.21}$$

Evidently the matrix $\vec{T}(X_r, X_l)$ defined by the expression in the curly brackets represents a somewhat unusual formulation of the joining relations of the WKB theory. Froman and Froman used special analytic techniques to approximate $\vec{T}(X_r, X_l)$ without reference to the matrix $\vec{R}$; we shall not be able to follow such a procedure.

II. GENERALIZED WKB FUNCTIONS

In adapting the scheme described above to our problem, we first have to find model Hamiltonians for which Eqs. (1.13) and (1.14) are likely to be valid. Here we shall be guided by physical intuition and analogy to the ordinary WKB theory: (i) The model Hamiltonians should tend to the periodic Hamiltonians characteristic of the two bulk regions. (ii) The solutions of the model Hamiltonians should be essentially Bloch functions whose constant wave vector is replaced by a local position-dependent wave vector. The band-bending model suggests that this wave vector should be defined by the implicit equation

$$E - \bar{V}_{l,r} - V'(x) - E_{l,r;\nu}(q_{l,r}(x)) = 0 \ . \tag{2.1}$$

Here, $E_{l,r;\nu}(q)$ is the νth eigenvalue of the differential equation satisfied by the periodic factor of the Bloch function $b(x; q_{l,r}, E) = e^{iq_{l,r}x}u(x; q_{l,r}, \nu)$:

$$\left[\frac{d^2}{dx^2} + 2iq_{l,r}\frac{d}{dx} + \frac{2m}{\hbar^2} \times \left(E - V_{l,r}(x) - \frac{(\hbar q_{l,r})^2}{2m}\right)\right] u(x; q_{l,r}, \nu) = 0 \ . \tag{2.2}$$

[Note that in Eq. (2.2) energies are referred to $\bar{V}_{l,r}$.] The concept of a local (position-dependent) wave vector $q(x)$ enters the ordinary WKB theory as the physical interpretation of a Liouville transformation of the independent variable,[10]

$$x \to w(x) = \int^x q(x_1)\, dx_1 \ , \tag{2.3}$$

where

$$q(x) = Q(x) = [E - V(x)]^{1/2} \tag{2.4}$$

and

$$\lim q(x) = q_{r,l} \quad \text{as } x \to \pm\infty \ . \tag{2.5}$$

Hence,

$$\lim w(x) = q_{r,l}x + w_{r,l} \quad \text{as } x \to \pm\infty \ . \tag{2.6}$$

These observations combine with Eq. (1.4) to suggest for the generalized WKB functions, which are adapted to our problem, the form

$$\psi_{1,2;r}(x) = [q_r(x)]^{-1/2} e^{\pm iw(x)} u\{[w(x) - w_r]q_r^{-1};\ \pm q_r, \nu\} \ , \tag{2.7}$$

$$w(x) = \int_{X_r}^x q_r(x_1)\, dx_1 \ , \tag{2.8}$$

if $x > X_r$, and an analogous form if $x < X_l$. Here $q_{r,l}(x)$ are those solutions of Eq. (2.1) which represent the extended (local) wave vectors, rather than the conventional reduced wave vectors. This assures that $q_{r,l}(x)$ has a positive lower bound $(\nu - 1)\pi(a_{r,l})^{-1}$. The term u satisfies Eq. (2.2) in the variable $[w(x) - w_r]q_r^{-1}$.

Our conjecture is borne out by an analysis based on the model Hamiltonians $\mathcal{H}_{0;r,l}$:

$$\frac{2ma_r^2}{\hbar^2}(E - \mathcal{H}_{0;r}) = \left(\frac{q_r}{q_r(x/a_r)}\right)^2 \frac{d^2}{d(x/a_r)^2} + \frac{2ma_r^2}{\hbar^2}\left(E - V_r\left\{\left[w\left(\frac{x}{a_r}\right) - w_r\right](q_r a_r)^{-1} a_r\right\}\right) - \left(\frac{q_r}{q_r(x/a_r)}\right)^2 q_r^{1/2}\left(\frac{x}{a_r}\right)\frac{d^2}{d(x/a_r)^2}\left[q_r^{-1/2}\left(\frac{x}{a_r}\right)\right] \tag{2.9}$$

if $x > X_r$. If $x < X_l$, an analogous expression holds for $\mathcal{H}_{0;l}$.

Before we proceed with our analysis, we shall briefly examine the significance of the factor $[q_r(x)]^{-1/2}$ in Eqs. (1.4) and (2.7). This factor is often interpreted merely as a normalization of the WKB function in Eq. (1.4) so that it carries unit flux. Consequently, it is argued that $[Q(x)]^{-1/2}$ should be identified with the inverse square root of the "local velocity" $(\hbar^{-1}dE/dQ)^{-1/2}$. In the ordinary WKB approximation this point of view is of no particular consequence. But, whenever $Q(x)$ is not a constant multiple of the local group velocity, the preceding argument is quite misleading: The Liouville transformation of the independent variable, defined by Eq. (2.8), has to be coupled to a transformation of the wave function $\Psi(x) = [q(x)]^{-1/2} \times \theta(w(x))$ to assure that the differential equation for $\theta(w)$ has the same form as that for $\Psi(x)$. More precisely, the Liouville transformation introduces a term proportional to $(d/dw)\,\theta$ into the differential equation for $\theta(w)$. This term is eliminated by the above transformation of the wave function, whereas, the substitution $\Psi(x) = (\hbar^{-1}dE/dq)^{-1/2}\,\theta(w(x))$ fails to achieve this. The occurrence of a first derivative in the equation for $\theta(w)$ means that either the model Hamiltonian or the error term $\epsilon(x)$ has to include this additional term. In the first case it is trivial to check that the solution of the model Hamiltonian will be proportional to the term $[(\hbar^{-1}dE/dq) \times q^{-1}]^{-1/2}$, indicating that our initial substitution should have been $\Psi(x) = [q(x)]^{-1/2}\theta$.

In the second case the same model Hamiltonian can be chosen for both substitutions. Let the solu-

tion of the model Hamiltonian be $\theta(w)$. The generally larger error term $\epsilon(x)$ implies that the approximation of $\Psi(x)$ by $\phi(x) = (\hbar^{-1} dE/dq)^{-1/2} \theta(w(x))$ will generally be considerably poorer than that by $\psi(x) = [q(x)]^{-1/2} \theta(w(x))$.

Finally, we note that the former, poorer, substitution will essentially reproduce Harrison's results.[5]

In Eq. (2.7) we introduced the dimensionless space variables $x/a_{r,l}$ and wave vectors, $q_{r,l}(x/a_{r,l})a_{r,l}$. This enables us to express the errors in terms of small dimensionless quantities. Making the corresponding changes in Eq. (2.6), it is easily verified that $\psi_{1,2;r}$ are a pair of linearly independent eigenfunctions of $\mathcal{H}_{0;r}$. The wave vector $q_r(x)$ defined by Eq. (2.1) satisfies Eqs. (2.4) and (2.5), because $V'(x)$ vanishes at infinity. Consequently, $\mathcal{H}_{0;r}$ tends to the periodic Hamiltonian of the bulk material on the right-hand side. Thus, Eqs. (2.7)–(2.9) conform to our intuitive requirements (i) and (ii).

On any interval on which $[q_r/q_r(x/a_r)]^2$ is positive and bounded, the solutions of

$$(E - \mathcal{H})\Psi = 0 \tag{2.10}$$

and of

$$\left[E - \mathcal{H}_{0;r} + \frac{\hbar^2}{2ma_r^2}\epsilon_r\left(\frac{x}{a_r}\right)\right]\Psi = \left(\frac{q_r}{q_r(x/a_r)}\right)^2 (E - \mathcal{H})\Psi = 0 \tag{2.11}$$

are identical.

If we substitute Eq. (2.7) into Eq. (2.11) we obtain an explicit expression for the dimensionless quantities $\epsilon_{r,l}$,

$$\begin{aligned}\epsilon_r\left(\frac{x}{a_r}\right) = {} & \frac{2ma_r^2}{\hbar^2}\left(\frac{q_r}{q_r(x/a_r)}\right)^2 \\ & \times \left[E - V'\left(\frac{x}{a_r}\,a_r\right) - V_r\left(\frac{x}{a_r}\,a_r\right)\right] \\ & - \left(E - V_r\left\{\left[w\left(\frac{x}{a_r}\right) - w_r\right](q_r a_r)^{-1} a_r\right\}\right) \\ & + \left(\frac{q_r}{q_r(x/a_r)}\right)^2 q_r^{1/2}\left(\frac{x}{a_r}\right)\frac{d^2}{d(x/a_r)^2}\left[q_r^{-1/2}\left(\frac{x}{a_r}\right)\right]\end{aligned} \tag{2.12}$$

if $x > X_r$. A similar expression holds for $\epsilon_l(x/a_l)$ if $x < X_l$. Equations (2.4) and (2.5) imply that

$$\lim_{x\to\infty} \epsilon_r \frac{x}{a_r} = 0 = \lim_{x\to-\infty} \epsilon_l \frac{x}{a_l}\,. \tag{2.13}$$

However, this is not sufficient to prove that Eq. (2.6) represents the required generalization of the WKB functions. To prove this assertion we substitute Eqs. (2.6) and (2.12) into the general equation (1.10). It is then easy to show that the T matrix for our problem satisfies an inequality such as Eq. (1.13). [Each element in the sequence of successive iterations of Eq. (1.10) can be bounded by an element of a convergent sequence.] Finally, an examination of the error integral Eq. (1.15) indicates that Eqs. (2.1), (2.3), and (2.12) are consistent with Eq. (1.16), if and only if

$$\lim_{x\to\infty} x^{(2+\eta)} V'(x) \leq C < \infty\,.$$

That is, if $V'(x)$ vanishes too slowly at infinity, the eigenfunctions of $\mathcal{H}$ cannot be approximated arbitrarily well by the Bloch functions for $V_{r,l}(x)$. Hence, although $\psi_{1,2}$ tend to those Bloch functions, the upper bound of the error does not tend to zero.

III. JOINING RELATIONS

Our discussion in Sec. II excluded the interval $X_l < x < X_r$, where the relatively rapid transition from one (perturbed) periodic potential $V_l(x)$ to another (perturbed) periodic potential $V_r(x)$ occurs. In this interval $V'(x)$ is so strong that it is no longer meaningful to view $V(x)$ as a perturbed periodic potential. It follows that as this region is approached $V'(x)$ tends to exhibit an appreciable variation over intervals of the order of interatomic distances. That is, both $V'(x)$ and $dV'/dx \approx \Delta V(x)/a$ tend to become appreciable. No WKB-type approximation should be expected to be valid under such conditions. A qualitative analysis of Eqs. (1.15), (2.1), and (2.12) leads to an intuitive picture of the deterioration of our generalization of the WKB approximation as the transition region is approached. From Eq. (2.1) we have

$$\begin{aligned}\frac{1}{q_{r,l}(x)}\,\frac{d}{dx}\,(q_{r,l}(x)) &= -\frac{1}{q_{r,l}(x)}\,\frac{dV'}{dx}\left(\frac{dE_{\nu;r,l}}{dq}\bigg|_{q=q_{r,l}(x)}\right)^{-1} \\ &= -\frac{dV'}{dx}\,[\hbar q_{r,l}(x)\, v_{g;r,l}(x)]^{-1} \\ &\approx -\frac{dV'}{dx}\,[E - \bar{V}_{r,l} - V'(x)]^{-1}\,.\end{aligned} \tag{3.1}$$

Here, $v_{g;r,l}(x)$ is the local group velocity. Equation (3.1) indicates that as the transition region is approached, the rapid variation of $V'(x)$ induces a correspondingly rapid variation in the local wave vector. And when the variation of $V'(x)$ over a period $\Delta V(x)$ approaches the local kinetic energy, then both $a^2\epsilon(x) = \epsilon(x/a)$ and the error integrals $\delta(\infty, x)$ and $\delta(x, -\infty)$ tend to large values. Often the

approximation fails, however, not primarily because of the increase in the absolute value of dV'/dx but rather because of the gradual decrease in the magnitude of the local group velocity. Just as in the ordinary WKB theory, our approximation fails at the classical turning points at which the local group velocity vanishes: At these points both $\epsilon(x)$ and the error integral $\delta(\infty, x)$ and $\delta(-x, \infty)$ diverge.[9]

In the following we assume that the number of the classical turning points is even, and that all of them are contained in the interval $X_l < x < X_r$ which also includes the transition region which was discussed above.[11] That is, all singularities of $\epsilon_r(x/a_r)$ and of $\epsilon_l(x/a_l)$ are included in $X_l < x < X_r$; furthermore, the smallness of these quantities implies that we may set

$$\vec{T}(\infty, x) \approx \vec{I} \text{ if } x > X_r$$
and
$$\vec{T}(x, -\infty) \approx \vec{I} \text{ if } x < X_l \ . \tag{3.2}$$

It follows from Eqs. (3.2) that for $x > X_r$, $\psi_{1,2;r}$ are a pair of (approximate) linearly independent eigenfunctions of $\mathcal{H}$. But the exact eigenfunction of $\mathcal{H}$, $\Psi_i(x)$, which for $x > X_r$ is approximated by $\psi_{i;r}$, exists also for $x < X_l$. The approximation of $\Psi_i(x)$, when $x < X_l$ in terms of $\psi_{1;l}$ and $\psi_{2;l}$ is indicated by Eq. (1.21). Using a matrix notation these remarks may be summarized by the equation

$$(\Psi_1(x), \Psi_2(x)) = \begin{cases} (\psi_{1;r}(x), \psi_{2;r}(x)), & x > X_r \\ (\psi_{1;l}(x), \psi_{2;l}(x))\vec{T}^{-1}(X_r, X_l), & x < X_l \end{cases} . \tag{3.3}$$

Here, the matrix $\vec{T}(X_r, X_l)$ is defined by the second equation in Eq. (1.21). [Note that the quantities $\psi_{1,2}(X_r)$ in Eq. (1.21) are to be interpreted as $\psi_{1,2;r}(X_r)$ and, similarly, $\psi_{1,2}(X_l)$ is to be replaced by $\psi_{1,2;l}(X_l)$.]

The joining relations, represented by the T matrix $\vec{T}(X_r, X_l)$ depend explicitly on the fundamental solution matrix $\vec{R}(x, x')$ of the full Schrödinger equation in the junction $X_l < x < X_r$. Depending on the nature of the problem, a variety of approximation techniques can be applied to determine $\vec{R}$. In this paper we are concerned primarily with narrow junctions, only a few atomic layers wide; this enables us to consider approximations which are ordinarily impractical, in particular, the iterative solution of the integral equation

$$\vec{R}(X_r, X_l) = \vec{I} + \int_{X_l}^{X_r} \vec{A}(x_1)\vec{R}(x_1, X_l)\, dx_1 \ , \tag{3.4}$$

where $\vec{A}$ is defined by Eqs. (1.17) and (1.18). Evidently, the matrix $\vec{A}$ may also be approximated by another matrix leading to a more tractable problem. This is indeed the procedure followed in the ordinary WKB theory, by one of the standard derivations of the joining relations across a single turning point: Here the exact solution of the Schrödinger equation with a linear potential is used as an approximation which is valid in the neighborhood of the turning point, where the WKB approximation fails.

IV. TRANSMISSION FACTOR FOR A NARROW JUNCTION

The transmission factor S is defined as the ratio of the flux traveling outward on the right-hand side to the incoming flux on the left-hand side, when the solution on the right-hand side is a purely outgoing wave. If we assume that both on the right-hand side and on the left-hand side the group velocities are positive, then it follows from Eq. (3.3) that

$$S = \frac{v_{g;r}}{a_r q_r} \frac{a_l q_l}{v_{g;l}} \frac{1}{|(T^{-1})_{11}|^2} \ , \tag{4.1}$$

where $v_{g;l,r} = dE_{l,r;\nu}/dq$ is the group velocity in the periodic potential $V_{l,r}(x)$, and

$$(T^{-1})_{11} = [T(X_r, X_l)]^{-1}_{11} \ . \tag{4.2}$$

Using Eqs. (1.21) and (2.6), we can express $(T^{-1})_{11}$ in terms of the WKB functions and the elements of $\vec{R}$. We now recall that

$$\det[W(\psi_{1;l}(x), \psi_{2;l}(x))] = \det[W(\psi_{1;l}(x), \psi^*_{1;l}(x))] = \text{const}, \quad x < X_l$$

and hence can be evaluated in the limit $x \to -\infty$, where it reduces to $(2mi/\hbar)\, v_{g;l}/a_l q_l$. Thus, we obtain

$$S = \frac{v_{g;r}}{a_r q_r} \frac{v_{g;l}}{a_l q_l} \left(\frac{2m}{\hbar}\right)^2 S_b \ , \tag{4.3}$$

where

$$S_b^{-1} = \Big| R_{22}\, \psi_{1;r}(X_r) \frac{d}{dx} \psi_{2;l}(X_l) - R_{11} \frac{d}{dx} \psi_{1;r}(X_r)\, \psi_{2;l}(X_l) - R_{12} \frac{d}{dx} \psi_{1,r}(X_r) \times \frac{d}{dx} \psi_{2,l}(X_l) + R_{21}\, \psi_{1,r}(X_r)\, \psi_{2;l}(X_l) \Big|^2 \tag{4.4}$$

and

$$R_{ij} = R_{ij}(X_r, X_l) \ . \tag{4.5}$$

There are two points which should be emphasized in the interpretation of Eqs. (4.3) and (4.4): First, if $\vec{R}$ is indeed the exact fundamental solution (matrix) on $X_l < x < X_r$ then the only approximation involved is the use of Eq. (3.2). That is, the errors are proportional to (the dimensionless quantity)

$\delta(\infty X_r)+\delta(X_l, -\infty)$ and can be controlled by the choice of X_r and X_l. Second, the transition probability always factors into a bulk factor which is *proportional to the product of the group velocities in the two periodic potentials*, and a "barrier" factor S_b. This second factor depends on the complete solution of $\mathcal{H}$ in the junction. It may be approximated either by approximating the Hamiltonian in the junction or by using an approximate solution to the full Hamiltonian. In either case an independent estimate of the error is obtainable by standard techniques applied to Eq. (3.4).

V. CONCLUSIONS

The tunneling between two different one-dimensional periodic potentials was analyzed in terms of a generalization of the WKB approximation specifically designed to apply to problems in which, as x tends to $\pm\infty$, the potential tends to periodic functions of position rather than to constants. The approximation was used to calculate the probability for tunneling from one periodic potential to the other across the junction. The transition probability was found to be proportional to the product of the group velocities in the two periodic potentials, and inversely proportional to the product of the extended wave vectors which characterize the two states involved in the transition.[12] Our derivation suggests that this result is exact in the sense that it should hold also for the exact solution. Our result refutes the commonly accepted assertion of Harrison that the WKB approximation excludes such a dependence of the tunneling probability on the group velocities.[5] Our tunneling probability clearly exhibits the correct dependence on the energy relative to extrema of the band structure on the right and on the left. This obviates the need for such *ad hoc* corrections to the WKB tunneling probability as proposed by Conley and Mahan.[13]

The second factor in the transmission coefficient for the junction was expressed in terms of an explicit solution of the Schrödinger equation in the junction. We were concerned with the case of a narrow junction, and indicated how such a solution might be determined. In the opposite limit, the junction is so wide that one might approximate it by a homogenous bulk material (i.e., a periodic potential) joined by a pair of (narrow) transition regions to the two periodic potentials on the right-hand side and left-hand side. If these transition regions were neglected one could adapt Kane's analysis of the internal field emission[14] to determine the required (approximate) fundamental solution in the junction. It is easily seen that in this case the familiar exponential tunneling probability of internal field emission would enter our result. However, it would be multiplied by two different preexponentials: first, the product of the group velocities discussed above, and second, the terms resulting from the approximate evaluation of the factor S_b defined in Eq. (4.4).

At this point it might be instructive to compare our work with the analysis of Harrison in some more detail. Harrison assumes that the effective-mass function, which modulates the cell-periodic function u, can be approximated by an appropriate WKB function, which he writes in the form

$$\psi \propto (\hbar^{-1} dE/dq)^{-1/2} \exp[i \int^x q(x_1)\, dx_1] .$$

This should be compared to our amplitude function

$$[q(x)]^{-1/2} \exp[i \int^x q(x_1)\, dx_1] .$$

The formal justification given by Harrison for his choice of a WKB function is sketchy and involves several implicit assumptions which are not generally valid. The most crucial assumption is the reality of the functions $\alpha(x)$ and $\beta(x)$ which he defines by the requirement that $\alpha(x)\, d\phi/dx$ and $\beta(x)\phi(x)$ are to be continuous. It is easily seen that α and β cannot generally be real. Physically, Harrison identifies (improperly) the wave vector in the preexponential factor with the particle velocity, while keeping the conventional interpretation of the wave vector in the exponential. There is no point in repeating our discussion, following Eq. (2.9), concerning the significance of the preexponential factor in the WKB function. It is easy to verify that Harrison's amplitude function will in general lead to a much poorer approximation of the solutions of the full Schrödinger equation than our WKB function. The notable exception to this rule is the case where $E(q) = (\hbar q)^2/2m^*$, with a constant (q independent) effective mass m^*. In this case both Harrison's and our amplitude functions have the same form. Furthermore, if the energy of an electron in the periodic potentials on the right-hand side and on the left-hand side is strictly quadratic in the wave vector, then our transmission coefficient defined by Eq. (4.3) reduces to a constant (energy independent) multiple of S_b. If, in addition, we approximate S_b by the simple exponential tunneling probability for internal field emission, which was discussed above, then our tunneling coefficient reduces to the expression quoted by Harrison.

To conclude, one should be cautious in interpreting the energy dependence of the tunneling probability. First, in narrow junctions the probability is distinctly not a simple exponential. And even when the junction is wide, and an exponential factor enters the probability, there are

always energy-dependent preexponential factors which cannot generally be dismissed.

An extension of the present work to a three-dimensional setting is being considered.

*Research sponsored by the Air Force Office of Scientific Research, Office of Aerospace Research, U. S. Air Force, under AFOSR Grant No. AFOSR-69-1727.

[1]The significance of this restriction is explained at the end of Sec. II.

[2]A careful discussion of the ordinary WKB approximation is given by N. Froman and P. O. Froman, *The J WKB Approximation* (North-Holland, Amsterdam, 1965).

[3]See also R. Bellman, *Perturbation Techniques in Mathematics, Physics and Engineering* (Holt, Rinehart, and Winston, New York, 1966), pp. 80–88.

[4]A careful review of this approach is given by C. B. Duke, in *Solid State Physics*, edited by F. Seitz, D. Turnbull, and H. Ehrenreich (Academic, New York, 1969), Suppl. 10, pp. 23–29 and 36–80.

[5]One of the most thorough attempts to develop such an approach for heterojunctions and to provide a theoretical justification for the procedure is due to W. A. Harrison, Phys. Rev. 123, 85 (1961).

[6]These difficulties are exemplified by the controversy concerning the proper conditions to be imposed at points of discontinuity in the effective mass. [Compare, for instance, Harrison, Ref. 5 above, and J. W. Conley, C. W. Duke, G. D. Mahan, and J. J. Tiemann, Phys. Rev. 150, 466 (1966).]

[7]See E. L. Ince, *Ordinary Differential Equations* (Dover, New York, 1944), p. 122.

[8]Reference 2, p. 27.

[9]The characterization of the classical turning points at which the local velocity vanishes as singularities of $\epsilon(x)$ was noted by Froman. This point is further clarified by Eq. (3.1) below.

[10]See Ref. 3, p. 80.

[11]In principle, the turning points may be located well within the bulk region, far from the transition region. This however does not invalidate our assumption.

[12]Note that this bulk factor reduces to unity for a free particle.

[13]J. W. Conley and G. D. Mahan, Phys. Rev. 161, 681 (1967).

[14]E. O. Kane, J. Phys. Chem. Solids 12, 181 (1959).

PHYSICAL REVIEW B VOLUME 2, NUMBER 6 15 SEPTEMBER 1970

Ambipolar Transport of Electrons and Holes in Anisotropic Crystals*

J. F. Schetzina and J. P. McKelvey

Department of Physics, The Pennsylvania State University, University Park, Pennsylvania 16802

(Received 13 April 1970)

The problem of ambipolar diffusion and drift of electrons and holes, in which the diffusivity and mobility tensors for the two carrier species may be of different forms, is examined. Problems of this type arise, for example, in studying the diffusion and drift of excess carrier distributions in uniaxially stressed germanium and silicon as well as in certain naturally anisotropic substances. General ambipolar transport equations are obtained in situations where the quasineutrality approximation is justified. Solutions to these equations are quite easily obtained in certain cases where particular simplifying assumptions can be made. These solutions are explicitly obtained and the range of conditions under which they are applicable is outlined in detail. Certain other procedures have been employed to solve the general problem in cases when these conditions are not satisfied. Such methods are usually applicable only in cases involving rather special and restricted sample geometries. The transformation properties of the various terms in the ambipolar transport equation are discussed in various situations of physical interest and importance.

I. INTRODUCTION

The ambipolar transport behavior of excess carrier distributions in semiconductor crystals has been discussed by Herring,[1] Shockley,[2] and van Roosbroeck[3]; the most elegant and explicit treatment of the subject is contained in a subsequent article by the last author.[4] In this treatment the influences which the diffusing electrons and holes exert upon one another by virtue of their mutual electrostatic attraction are taken into account in an approximate way by assuming that the electrostatic forces are sufficient at all times to maintain an approximate state of electrical neutrality throughout the crystal. The electrons, whose mobility is greater, are thus envisioned as pulling the holes, whose mobility is less, along more rapidly than they would otherwise travel, and being