Nonresonant and resonant Zener tunneling

S. Glutsch*

Friedrich-Schiller-Universität Jena, Institut für Festkörpertheorie und Theoretische Optik, Max-Wien-Platz 1, 07743 Jena, Germany

(Received 14 April 2003; revised manuscript received 16 March 2004; published 23 June 2004)

The Zener tunneling rate is derived by quantum-mechanical perturbation theory, using the Kane functions as base functions. This result is able to describe both nonresonant and resonant tunneling and is in line with recent experimental findings on Zener tunneling in superlattices. In contrast, the result published in the literature, which is the state of art up until now, does not adequately describe resonant Zener tunneling and is not an improvement over the semiclassical theory. The reason for that is a misconception about the initial condition.

DOI: 10.1103/PhysRevB.69.235317 PACS number(s): 73.21.Cd, 72.20.Ht, 78.67.Pt

I. INTRODUCTION

The motion of an electron with mass m and charge $-e$ in one dimension in a lattice-periodic potential $U(z)=U(z+a)$, subjected to an electric field F is governed by the Hamiltonian

$$\hat{H}=-\frac{\hbar^2}{2m}\frac{d^2}{dz^2}+U(z)+eFz. \tag{1}$$

This problem has drawn the attention of generations of physicists in the last 75 years.[1] In his famous paper from 1934, Zener calculated the tunneling rate of an electron from the valence to the conduction band due to an external electric field.[2] His calculation was based upon a semiclassical approach, where energies in the gap region are assigned to a complex Bloch wave number. This concept of branch-point tunneling, which has been introduced previously by Landau,[3] is known as Landau-Zener tunneling.

In 1940, Houston derived an approximate solution of the time-dependent Schrödinger equation, where the wave function is given by a Bloch function with time-dependent wave number $k(t)=k(0)-eFt/\hbar$.[4] Because the points k and $k+2\pi/a$ in reciprocal space are equivalent, the electron performs oscillations with the period $T=2\pi\hbar/|eFa|$, so-called Bloch oscillations. In the 1950s and 1960s, approximate eigenfunctions of the Hamiltonian (1) were derived in the framework of the crystal-momentum representation.[5–10] These so-called Kane functions give rise to ladders of equally spaced eigenvalues $E_{jl}=E_{j0}+eFal$; $l\in\mathbb{Z}$ for each band j, known as Wannier-Stark ladders. It can be shown that the Houston functions and the crystal-momentum representation are equivalent by Fourier or gauge transform.[11–13] The Zener tunneling rate has been calculated by quantum-mechanical perturbation theory on the basis of both the Houston and the Kane functions, leading to the same result.[4,6–8,13,14] It was claimed by some authors that the quantum-mechanical result would approach the semiclassical result in the limit $F\rightarrow 0$.[15,16] As perturbation theory is also limited to small values of F, it is not clear in which way the quantum-mechanical result is an improvement over the semiclassical result.

It was shown only in 1977 that the spectrum of a Bloch electron in an electric field is entirely continuous,[17] which is a prerequisite of Zener tunneling. Then the Wannier-Stark ladders become resonances, i.e., nonstationary states with a long lifetime. In case that $U(z)$ is an analytic function, the resonances can be defined as complex eigenvalues of an analytically continued Hamiltonian.[18] The imaginary part of the resonances, which governs the lifetime of wave packets,[19] approaches Zener's expression for the tunneling rate in the limit of small F.[20]

In recent years, the topic of Zener tunneling resurfaced in the context of superlattices[21–29] and accelerating optical potentials.[30–32] The experimental data and numerical calculations show resonances which cannot be explained by Zener's semiclassical theory. These resonances are related to interactions ("anticrossings" or "avoided crossings") of Wannier-Stark ladders and are often called "resonant Zener tunneling." This can also be explained in the real-space picture, where the effective tunneling length of the potential $U(z)+eFz$ is proportional to $1/F$, but also shows oscillations due to the periodicity of $U(z)$. Anticrossings of Wannier-Stark ladders have been observed previously in the optical spectra of superlattices without a direct connection made to Zener tunneling.[33,34] In fact, Zener tunneling requires an entirely continuous spectrum and, therefore, cannot be explained on the basis of discrete Wannier-Stark ladders.

In many cases, the experimental data are in good agreement with first-principle or model calculations.[22,26,27,30,31] However, the relationship between nonresonant tunneling, as described by the original Zener formula, and resonant tunneling remains largely unclear. In particular, it is not clear, if the Houston-Kane formalism—apparently an improvement over the semiclassical theory—can explain the effect of resonant tunneling. The literature is remarkably silent about this point. To our knowledge, only one publication by Ao and Rammer deals with this question.[35] These authors found that the result of perturbation theory contains $1/F$ oscillations, which are not found in Zener's semiclassical result.

In the present paper we derive a formula for the tunneling rate which correctly takes into account nonresonant and resonant Zener tunneling. We show that the result from the literature, derived in the framework of Houston or Kane functions, is inconsistent, considerably underestimates the effect of resonant tunneling, and cannot be considered an improvement over the semiclassical theory. Explicit results are given for a shallow superlattice, which has recently been studied experimentally and theoretically. In Sec. II we introduce important definitions and notations, and give a brief review on the semiclassical theory and the Kane functions. The tunneling rate is derived in Sec. III on the basis of Kane functions and it is shown that the assumptions of the Houston-Kane theory are inconsistent. In Sec. IV we show explicit results for a shallow superlattice and make a comparison with experimental data. A summary is given in Sec. V.

II. PRELIMINARIES

In this section we introduce the basic definitions, notations, and conventions used in Sec. III. Furthermore, we specify the result from the semiclassical theory as derived by Zener and others.

A. Bloch, Kane, and Houston functions

For $F=0$, the eigenfunctions of the Hamiltonian (1) can be written as Bloch functions,

$$\psi_{jk}(z) = \frac{1}{\sqrt{2\pi}} e^{ikz}\, u_{jk}(z); \quad u_{jk}(z) = u_{jk}(z+a), \tag{2}$$

where $j \in \mathbb{N}$ is the band index and $k \in (-\pi/a, +\pi/a]$ is the Bloch wave number. The spectrum consists of energy bands $E_j(k)$. We assume that U is piecewise continuous, $U(z) = U(-z)$, and all gaps are open. For this case, a number of analytical properties have been derived in the famous paper by Kohn.[36] In the extended zone scheme, $k \in \mathbb{R}$, it holds that $E_j(k) = E_j(k+2\pi/a)$. Furthermore, the bands $E_j(k)$ correspond to different branches of one multi-valued complex analytic function $E(k)$, the so called "energy function." The phase of the Bloch functions can be chosen such that $\psi_{jk}(z) = \psi^*_{j,-k}(z)$; $\psi_{j,k+2\pi/a}(z) = \psi_{jk}(z)$ and that $\psi_{jk}(z)$ is a complex analytic function of k. A recipe for the numerical calculation, which ensures analyticity of the Bloch functions, was given by the present author.[37] We assume the Bloch functions to be normalized as

$$\langle \psi_{jk} | \psi_{j'k'} \rangle = \int_{-\infty}^{+\infty} dz \varphi^*_{jk}(z) \varphi_{j'k'}(z) = \delta_{jj'} \underline{\delta}(k-k'),$$

where $\underline{\delta}$ is the periodic delta function with period $2\pi/a$.

The Bloch functions (2) can be used as base functions for solving the eigenvalue problem of the Wannier-Stark Hamiltonian (1). This is done by the ansatz

$$\varphi(z) = \sum_j \sqrt{\frac{a}{2\pi}} \int_{-\pi/a}^{+\pi/a} dk \tilde{\varphi}_j(k) \psi_{jk}(z).$$

To determine the unknown functions $\tilde{\varphi}_j(k)$, we have to calculate the matrix elements of the perturbation eFz. Employing the explicit form of the Bloch functions (2) we find[38]

$$\int_{-\infty}^{+\infty} dz\ \psi^*_{jk}(z) eFz \psi_{j'k'}(z) = ieF\ \delta_{jj'} \frac{d}{dk} \underline{\delta}(k-k') + eF \underline{\delta}(k-k') Z_{jj'}(k),$$

where

$$Z_{jj'}(k) = \frac{1}{a} \int_{-a/2}^{+a/2} dz\ u^*_{jk}(z) i \frac{\partial u_{j'k}(z)}{\partial k} = Z^*_{j'j}(k).$$

For symmetric potentials and the phase choice introduced above the diagonal matrix elements are either $Z_{jj}(k) \equiv 0$ or $Z_{jj}(k) \equiv \pm a/2$ (Kohn's cases A and B, see Ref. 36). If a single band is considered, the origin of the z axis can always be chosen such that $Z_{jj}(k) \equiv 0$.

Now the eigenvalue problem for the $\tilde{\varphi}_j(k)$ reads

$$\left[E_j(k) + ieF \frac{d}{dk} \right] \tilde{\varphi}_j(k) + eF \sum_{j'} Z_{jj'}(k) \tilde{\varphi}_{j'}(k) = E \tilde{\varphi}_j(k),$$

with the boundary condition $\tilde{\varphi}_j(k) = \tilde{\varphi}_j(k+2\pi/a)$. The solution is trivial if the coupling between different bands is neglected and it holds that

$$E_{jl} = \frac{a}{2\pi} \int_{-\pi/a}^{+\pi/a} dk' \overline{E}_j(k') + eFal; \quad l \in \mathbb{Z}$$

$$\tilde{\varphi}_{jl}(k) = \exp\left\{ \frac{1}{ieF} \int_0^k dk' [E_{jl} - \overline{E}_j(k')] \right\}, \tag{3}$$

where

$$\overline{E}_j(k) = E_j(k) + eFZ_{jj}(k).$$

The resulting functions

$$\varphi_{jl}(z) = \sqrt{\frac{a}{2\pi}} \int_{-\pi/a}^{+\pi/a} dk\ \tilde{\varphi}_{jl}(k) \psi_{jk}(z), \tag{4}$$

which are the approximate eigenfunctions of the Hamiltonian (1), are known as Kane functions.[8,10,12,38] They are orthonormal, i.e., $\langle \varphi_{jl} | \varphi_{j'l'} \rangle = \delta_{jj'} \delta_{ll'}$, and complete.

It is interesting to see the connection between Kane and Houston functions. For fixed band index j, the Kane functions φ_{jl} are a complete system in the subspace of the jth band. The expansion of a Bloch function with wave-number k_0 in terms of Kane functions gives

$$\psi_{jk_0}(z) = \sum_l \langle \varphi_{jl} | \psi_{jk_0} \rangle \varphi_{jl}(z) = \sqrt{\frac{a}{2\pi}} \sum_l \tilde{\varphi}^*_{jl}(k_0) \varphi_{jl}(z).$$

The time dependence of the Kane functions in the absence of tunneling, i.e., when the nondiagonal matrix elements $Z_{jj'}(j \neq j')$ are neglected, is

$$\varphi_{jl}(z,t) = e^{E_{jl} t / i\hbar} \varphi_{jl}(z).$$

Then the time-dependent Bloch functions

$$\psi_{jk_0}(z,t) = \exp\left\{ \frac{1}{i\hbar} \int_0^t dt' \overline{E}[k(t')] \right\} \psi_{j,k(t)}(z) \tag{5}$$

with

$$k(t) = k_0 - eFt/\hbar$$

are identical to the Houston functions.[4] The replacement of the $E_j(k)$ by $\overline{E}_j(k)$ was made later.[7,11,15,16]

B. Semiclassical Zener theory

In his famous paper from 1934, Zener calculated the tunneling rate of a Bloch electron in the presence of an electric field.[2] The main idea was to assign the gap region to complex values of the Bloch wave number, where the imaginary part is responsible for a tunneling probability lower than unity. By deforming the integral on the complex energy plane, the tunneling probability can be expressed as[20,32]

$$p = \exp\left[\frac{1}{\mathrm{i}|eF|}\oint dk\ E(k)\right], \tag{6}$$

where the integration is carried out on a closed contour, cycling in clockwise direction both conjugated complex branch points, which connect the two energy bands. The tunneling rate is then given by the tunneling probability p, divided by the period of the Bloch oscillations $T=2\pi\hbar/|eFa|$.

In the case of nearly free electrons, the above integral can be carried out explicitly and the tunneling probability from the first to the second band is[2,32]

$$w = \frac{|eF|a}{2\pi\hbar}\exp\left(-\frac{m\ a\ E_g^{*2}}{4\hbar^2|eF|}\right) \tag{7}$$

with $E_g^*=E_2(k_0)-E_1(k_0)$ being the first gap, located at $k_0 = \pm\pi/a$. The function (7) monotonously increases with F and does not show any oscillations. The behavior in the gap region is parabolic

$$E_2(k) - E_1(k) = E_g^* + \frac{\hbar^2(k-k_0)^2}{2m^*} + \cdots,$$

and the tunneling rate as function of the effective-mass parameters E_g^* and m^* is[39]

$$w = \frac{|eF|a}{2\pi\hbar}\exp\left(-\frac{\sqrt{m^*}\ \pi E_g^{*3/2}}{2\hbar|eF|}\right). \tag{8}$$

This formula was also derived by Franz without the assumption of nearly free electrons[40] and by Eilenberger for tightly bound electrons.[16] It is also found true for the crossing of any two bands, in the approximation of nearly free electrons. Hence, it is reasonable to consider Eq. (8) as the most general form of the semiclassical result.

It was claimed that in the limit $F\to 0$ the result of perturbation theory would go over into the semiclassical result (8)[15,16] and in the mathematical theory, the imaginary part of the resonance for small F is found to be asymptotically equal to expression (6).[20] The results between perturbation theory and semiclassical theory were found to differ by a factor of $(\pi/3)^2 \approx 1.079$. We shall come back to this $\pi/3$ problem in the next section.

III. TUNNELING RATE

In this section we derive the tunneling rate by means of quantum-mechanical perturbation theory and make a comparison with the result published in the literature. The perturbational approach is more rigorous than the original Zener theory, but also makes assumptions which cannot be rigorously justified within the theory. As in the semiclassical theory, the transition probability $p(T)=[n(0)-n(T)]/n(0)$ is defined as the relative decrease of the occupation number during one Bloch cycle and the tunneling rate is given by $w=p(T)/T$.

To calculate the tunneling rate from the first to the second band, it is assumed that tunneling from the second to the third band is much more effective than from the first to the second band, etc., so that downward tunneling is negligible in comparison to upward tunneling. Then, from the point of view of the theory, the excited states can be considered empty. Interestingly, the increase of tunneling probability with increasing band index, which holds true for all realistic potentials $U(z)$, can be employed to prove the continuous spectrum of the Wannier-Stark Hamiltonian (1).[41] In turn, a continuous spectrum is necessary for Zener tunneling, otherwise there would be closed subspaces of Wannier-Stark ladders with stationary states. Because much of the Zener theory was developed before the continuous nature of the spectrum was established, these assumptions and approximations are not always clearly stated.

As a consequence of the above assumptions, the decrease of the occupation number is linearly exponential with the number of Bloch cycles and tunneling during different Bloch cycles is uncorrelated. The first condition is in contradiction to the time-reversal symmetry. Niu and Raizen found a nonexponential behavior of $n(t)$ near $t=0$, which goes over into an exponential decay at larger times.[42] Holthaus found that the ablation of daughter wave packets during different Bloch cycles is not uncorrelated.[32] The above-mentioned difficulties are avoided in the mathematical theory, where the tunneling rate is defined as the imaginary part of the resonances. This approach, which can also be used for numerical calculations,[29,31] is restricted to analytic potentials. Alternatively, one can numerically calculate the density of states with no restriction in the number of subbands. Then the tunneling rate is manifested as a natural linewidth of the Wannier-Stark resonances.[25,26]

As base functions for the following time-dependent perturbation theory we use the Kane functions (4), which are the eigenfunctions in the absence of band-to-band coupling. The diagonal matrix elements $(j=j')$ of the Hamiltonian (1) are given by Wannier-Stark ladders

$$H_{jl,jl'} = \langle\varphi_{jl}|\hat{H}|\varphi_{jl'}\rangle = E_{jl}\delta_{ll'}. \tag{9}$$

For the nondiagonal matrix elements $(j\neq j')$ we find

$$\begin{aligned} H_{jl,j'l'} &= \int_{-\infty}^{+\infty} dz\ \varphi_{jl}^*(z)eFz\ \varphi_{j'l'}(z) \\ &= \frac{a}{2\pi}\int_{-\pi/a}^{+\pi/a} dk\ \tilde{\varphi}_{jl}^*(k)eFZ_{jj'}(k)\tilde{\varphi}_{j'l'}(k) \\ &= \frac{eFa}{2\pi}\int_{-\pi/a}^{+\pi/a} dk\ Z_{jj'}(k)\exp\left\{\frac{1}{\mathrm{i}eF}\int_0^k dk'[\overline{E}_j(k') \right. \\ &\quad \left. - \overline{E}_{j'}(k') - E_{jl} + E_{j'l'}]\right\}. \end{aligned} \tag{10}$$

Without loss of generality, we consider tunneling between two bands $j=1$ and $j'=2$. The von Neumann equation for the density matrix takes the form

$$\mathrm{i}\hbar\frac{d}{dt}N(t)=[H,N(t)];\quad H=\left(\begin{array}{c|c}H_{11}&H_{12}\\\hline H_{21}&H_{22}\end{array}\right);\quad N=\left(\begin{array}{c|c}N_{11}&N_{12}\\\hline N_{21}&N_{22}\end{array}\right),\tag{11}$$

where, for convenience, we introduced the block notations $H_{jj'}=(H_{jl,j'l'})$ and $N_{jj'}=(N_{jl,j'l'})$. As we neglect downward scattering from the second to the first miniband, at $t=0$, we assume the first Wannier-Stark ladder to be completely filled and the second Wannier-Stark ladder to be completely empty so that the initial condition for the density matrix is

$$N(t=0)=\mathrm{Diag}(1,0).\tag{12}$$

This is in line with the translational symmetry of the problem, because no Wannier-Stark level stands out against the others.

Now, we apply Dirac's perturbation theory, where the unperturbed operator is given by the diagonal part $H^{(0)}=\mathrm{Diag}(H_{11},H_{22})$ and the perturbation by the nondiagonal part $H^{(i)}=H-H^{(0)}$. The correspondence between Heisenberg and Dirac picture is given by

$$\tilde{H}_{jl,j'l'}(t)=e^{-E_{jl}t/\mathrm{i}\hbar}\,H^{(i)}_{jl,j'l'}e^{+E_{j'l'}t/\mathrm{i}\hbar};$$

$$\tilde{N}_{jl,j'l'}(t)=e^{-E_{jl}t/\mathrm{i}\hbar}N_{jl,j'l'}(t)e^{+E_{j'l'}t/\mathrm{i}\hbar}.\tag{13}$$

Importantly, it holds that $[H^{(0)},N(0)]=0$. The quantities with the tilde obey the same equation of motion (11) and the initial condition is $\tilde{N}(0)=N(0)$. The first nonvanishing contribution to $\tilde{N}_{22}$ is found second order in time and it holds that

$$\tilde{N}^{(2)}_{22}(t)=\frac{1}{\hbar^2}\int_0^t dt_1\int_0^t dt_2\,\tilde{H}_{21}(t_1)\tilde{N}_{11}(0)\tilde{H}_{12}(t_2).\tag{14}$$

With $\tilde{N}_{11}(0)=N_{11}(0)=1$ and Eq. (13), the explicit expressions in the Heisenberg picture are

$$N^{(2)}_{2l,2l'}(t)=\frac{1}{\hbar^2}\int_0^t dt_1\int_0^t dt_2\sum_{l''=-\infty}^{+\infty}e^{+E_{2l}(t-t_1)/\mathrm{i}\hbar}\times H_{2l,1l''}e^{+E_{1l''}(t_1-t_2)/\mathrm{i}\hbar}H_{1l'',2l'}\,e^{+E_{2l'}(t_2-t)/\mathrm{i}\hbar}.$$

The elements of $N^{(2)}_{22}$ depend only on $l-l'$ because of the translational symmetry. The tunneling probability in lowest order is given by $p(t)=N^{(2)}_{2l,2l}(t)=N^{(2)}_{20,20}(t)=\tilde{N}^{(2)}_{20,20}(t)$ and it follows that

$$p(t)=\sum_{l=-\infty}^{+\infty}\left|\frac{1}{\mathrm{i}\hbar}\int_0^t dt'\,e^{(E_{20}-E_{1l})t'/\mathrm{i}\hbar}\right|^2|H_{20,1l}|^2=\sum_{l=-\infty}^{+\infty}\left(\frac{t}{\hbar}\right)^2\mathrm{sinc}^2\frac{(E_{20}-E_{1l})t}{2\,\hbar}|H_{20,1l}|^2,$$

where sinc $x=\sin(x)/x$. With the explicit form of the interaction matrix elements (10), the tunneling rate can be written as

$$w=\frac{p(T)}{T}=\frac{|eF|a}{2\pi\hbar}\int_{-\pi/a}^{+\pi/a}dk_1\int_{-\pi/a}^{+\pi/a}dk_2\,f^*(k_1)g\times(k_1-k_2,b)f(k_2),\tag{15}$$

where

$$f(k)=Z_{12}(k)\exp\left\{\frac{1}{\mathrm{i}eF}\int_0^k dk'[\overline{E}_1(k')-\overline{E}_2(k')]\right\},$$

$$g(k,b)=\sum_{l=-\infty}^{+\infty}\mathrm{sinc}^2\left[(b-al)\frac{\pi}{a}\right]e^{+\mathrm{i}(b-al)k};\quad b=\frac{E_{20}-E_{10}}{eF}.\tag{16}$$

The latter sum can be carried out explicitly (Appendix B) and it holds that

$$g(k,b)=1-\frac{|k|a}{2\pi}\left(1-\cos\frac{2\pi b}{a}\right)+\mathrm{i}\frac{ka}{2\pi}\sin\frac{2\pi b}{a};\quad -\frac{2\pi}{a}\leqslant k\leqslant+\frac{2\pi}{a}.\tag{17}$$

The result (15) is the same for any integration limits $(k_0,k_0+2\pi/a)$, because the shift of the origin in the definition of the Kane functions (3) results only in a phase factor, which is canceled in all measurable quantities. This is also expected from the periodicity of the Bloch functions in reciprocal space. The function $g(k,b)$ in Eq. (15) is a source of $1/F$ oscillations with period $ea/(E_{20}-E_{10})$, which is the result of anticrossings between the first and second Wannier-Stark ladders.

We mention that the result does not depend on a particular initial condition and would be the same for any $N(0)$ that is diagonal in the band indices, if the tunneling probability is defined as $p=[N_{1l,1l}(0)-N^{(2)}_{1l,1l}(t)]/N_{1l,1l}(0)$. In particular, for a pure state $\hat{N}(0)=|\varphi_{1l}\rangle\langle\varphi_{1l}|$, the solution of the von Neumann equation is equivalent to the solution of a time-dependent Schrödinger equation with an initial state $|\varphi_{1l}\rangle$. This is a result of second-order perturbation theory, where the second band is considered always empty.

Let us now compare our result to the literature. In the context of the crystal-momentum approximation, a tunneling rate of the form

$$w=\frac{|eF|a}{2\pi\hbar}\left|\int dkZ_{12}(k)\exp\left\{\frac{1}{\mathrm{i}eF}\int_0^k dk'[\overline{E}_1(k')-\overline{E}_2(k')]\right\}\right|^2 \tag{18}$$

has been derived by many authors.[4,6–8,13,14] In some cases, the integration is carried out from $-\pi/a$ to $+\pi/a$, in other cases from 0 to $2\pi/a$, and some authors do not specify integral limits. One would expect that the result is independent on the limits of the outer integral, except that the domain length has to be $2\pi/a$, but this is not the case, as we shall see in the next section. Formula (18) is very similar to our result (15), except that $g(k,b)$ is set equal to unity. However, this function ensures independence of the integration limits. Furthermore, having identified $g(k,b)$ as a major source of $1/F$ oscillations, the question arises, in which way formula (18) is able to describe resonant tunneling and goes beyond the semiclassical approach.

To find out which misconceptions lead to the expression (18), we shall rederive Eq. (15) in a slightly different way, which will also lead to a much simpler form of the present result. In the Dirac picture, as a result of the translation symmetry, all submatrices matrices of $\tilde{H}$ and $\tilde{N}$ are Toeplitz matrices, i.e., their elements depend only on the difference of the indices $l-l'$. A product of such matrices can be conveniently calculated by Fourier transform (Appendix A). With the Fourier-transformed Hamiltonian,

$$\begin{aligned}\tilde{H}_{12}(q,t)&=\tilde{H}^*_{21}(q,t)=\sum_{l=-\infty}^{+\infty}\mathrm{e}^{-\mathrm{i}qa(l-l')}\,\tilde{H}_{1l,2l'}(t)\\&=e\,F\,\mathrm{e}^{\frac{(E_{10}-E_{20})q}{\mathrm{i}eF}}\,Z_{12}\left(q-\frac{eFt}{\hbar}\right)\\&\quad\times\exp\left\{\frac{1}{\mathrm{i}eF}\int_0^{q-eFt/\hbar}dq'[\overline{E}_1(q')-\overline{E}_2(q')]\right\},\end{aligned}$$

Equation (14) goes over into

$$\begin{aligned}\tilde{N}_{22}(q,t)&=\frac{1}{\hbar^2}\int_0^t dt_1\int_0^t dt_2\,\tilde{H}_{21}(q,t)\tilde{N}_{11}(q,0)\tilde{H}_{12}(q,t)\\&=\left|\frac{eF}{\hbar}\int_0^t dt'Z_{12}\left(q-\frac{eFt'}{\hbar}\right)\right.\\&\quad\times\exp\left\{\frac{1}{\mathrm{i}eF}\int_0^{q-eFt'/\hbar}dq'[\overline{E}_1(q')\right.\\&\quad\left.\left.-\overline{E}_2(q')]\right\}\right|^2\tilde{N}_{11}(q,0).\end{aligned}$$

The transition probability during one Bloch cycle is given by

$$p(T)=N^{(2)}_{20,20}(T)=\tilde{N}^{(2)}_{20,20}(T)=\frac{a}{2\pi}\int_{-\pi/a}^{+\pi/a}dq\tilde{N}^{(2)}_{22}(q,T)$$

and with the initial condition $\tilde{N}_{11}(q,0)=N_{11}(q,0)=1$ we find for the tunneling rate

$$w=\frac{|eF|a}{2\pi\hbar}\frac{a}{2\pi}\int_{-\pi/a}^{+\pi/a}dq\left|\int_{q-\pi/a}^{q+\pi/a}dk\,f(k)\right|^2 \tag{19}$$

with $f(k)$ defined in Eq. (16). It can be shown (Appendix C) that this result is equivalent to the previous formulation (15).

The result (19) is closely related to the expression used in the literature (18). The spurious dependence of the tunneling probability on the integration limits is fixed by taking the average over all intervals of length $2\pi/a$.

Let us now discuss the assumption that lead to the result in the literature. In order to obtain formula (18), the initial condition has to be chosen such that $\tilde{N}_{11}(q,t=0)=N_{11}(q,t=0)=2\pi/a\underline{\delta}(q-q_0)$, which is compatible with the symmetry of the problem, but $H^{(0)}$ does not commute with $N^{(0)}$. This is precisely the initial condition for the Houston functions. The popular view that Bloch oscillations are a prerequisite for tunneling, in other words, that the electron needs some run-up before jumping, is the reason why the existing works on Zener tunneling only rederive Houston's result from 1940.

Let us come back to the $\pi/3$ problem, mentioned in Sec. II. It can be shown that the correct prefactor is unity, if all orders of the perturbation series are taken into account.[43] The reason is that the passage through the forbidden gap does not occur infinitely slowly. A systematic treatment of nonadiabatic transitions is possible by introducing a sequence of superadiabatic bases, where, with increasing order, the prefactor reduces from $\pi/3$ to unity.[44,45]

Finally, we discuss the influence of the result on the number of Bloch periods. The function f, defined in Eq. (16), fulfills $f(k+2\pi/a)=e^{+\mathrm{i}2\pi b/a}f(k)$. Therefore, the tunneling rate for N periods is

$$w_N=\frac{p(NT)}{N\,T}=\frac{\sin^2\frac{N\pi b}{a}}{N\sin^2\frac{\pi b}{a}}w,$$

where w is the tunneling rate for one Bloch cycle (19). The same prefactor also appears in formula (18) when N Bloch cycles are considered.[13] The "spectrometer function" has $N-1$ zeros between the ladder crossings $b/a\in\mathbb{Z}$ and for $N\to\infty$ goes over into $\underline{\delta}(b/a)$ with period 1, which is Fermi's golden rule. Therefore, the occurrence of nonresonant tunneling, which means that w is nonzero for all $F\neq 0$, crucially depends on the restriction to one Bloch cycle or, equivalently, the assumption that tunneling during different Bloch cycles is uncorrelated. As the memory is located in the nondiagonal elements of the density matrix, this means that their dephasing time (not to be confused with the lifetime of the Wannier-Stark states, which is equal to the inverse tunneling rate) needs to be well below T so that the next Bloch cycle starts again with a diagonal N. So far, all perturbational approaches of Zener tunneling implicitly make this assumption, as the dephasing on the account of higher Wannier-Stark ladders is not included in the two-band model.

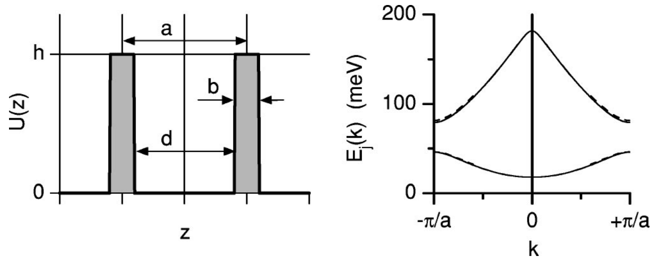

FIG. 1. Left: Sketch of the superlattice potential. Right: First and second miniband (solid line) and approximation of nearly free electrons (dashed line).

IV. NUMERICAL EXAMPLE

Superlattices are ideal objects to study Zener tunneling. The parameters of the potential $U(z)$ can be controlled by the sample geometry and the material combination and the tunneling rate can be directly measured as an increase of the linewidth of the optical transitions.

Here we present explicit results for a $GaAs/Ga_{0.92}Al_{0.08}As$ superlattice, which was studied experimentally and theoretically in Refs. 25–28. The electron confinement is modeled by a rectangular potential $U(z)$, which is shown in the left part of Fig. 1. The parameters are $d=76$ Å; $b=39$ Å; $a=115$ Å, and $h=63.2$ meV. The electron effective mass is $m_e=0.067\,m_0$, with $m_0=9.109\times10^{-31}$ kg being the electron rest mass.

The eigenvalue problem of $\hat{H}$ for $F=0$ was solved numerically using third-order (i.e., sectionally quadratic) B splines with a step size of $a/230=0.5$ Å. The Brillouin zone was sampled by 400 points and convergence of all results was ensured. The right part of Fig. 1 shows the dispersion of the first two minibands $E_{1,2}(k)$. The approximation of nearly free electrons (dashed line) is very accurate already for the first two minibands and is nearly indistinguishable from the numerical solution for higher minibands (not shown). Therefore, the conduction-electron states can be reasonably described in the approximation of nearly free electrons and it makes no difference if the semiclassical tunneling rate is calculated by formula (7) or by formula (8). For this choice of the origin, both the first and the second miniband belong to case A so that $\overline{E}_1(k)=E_1(k)$ and $\overline{E}_2(k)=E_2(k)$.

The tunneling rate is shown in Fig. 2, where the quantity $\hbar w$, which has the dimension of an energy, is plotted versus the electric field F. The result of the present paper (solid line) is compared with the result of the literature for integration limits $(0,2\pi/a)$ (dashed line) and $(-\pi/a,+\pi/a)$ (dash-dotted line), and with the semiclassical result, given by expression (8) (dotted line).

The present result (solid line) shows an overall increase of the tunneling rate with the electric field, but also exhibits pronounced oscillations, which are the result of interacting Wannier-Stark ladders. The relative role of these oscillations increases for decreasing field and $w(F)$ reveals an infinite number of maxima and minima.

The result (18) strongly depends on the integration limits, which is widely unnoticed in the literature. An obvious choice would be $(0,2\pi/a)$ so that the electron jumps in the

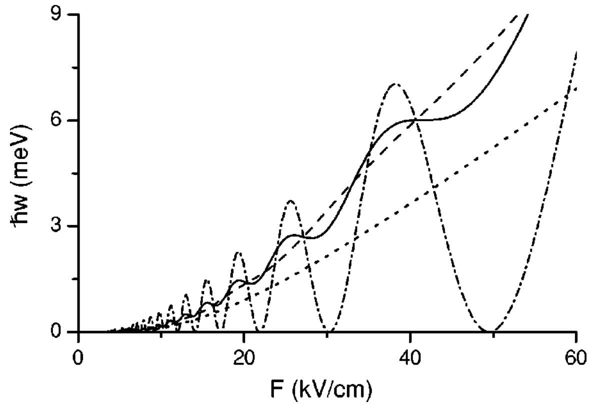

FIG. 2. Theoretical tunneling rate $\hbar w$ versus electric field F. Solid line: present result (19). Dashed line: formula (18) with integration limits $(0,2\pi/a)$. Dash-dotted line: formula (18) with integration limits $(-\pi/a,+\pi/a)$. Dotted line: semiclassical result (8).

middle of the period. In this case (dashed line), the tunneling rate is nearly a monotonous function of the field. It reveals very small oscillations which are negligible, compared to the oscillations observed in the present result (solid line). It also closely resembles the semiclassical result (dotted line). For large F, the tunneling rate is by about 50% larger than the semiclassical result. Taking into account the factor $\pi^2/9$, the difference reduces to about 40%. It is not clear if the semiclassical result is too small and the result of perturbation theory is too large, or both, because both approaches become inaccurate for large fields. A numerical calculation by Holthaus also gives a tunneling rate which is larger than the semiclassical result.[32] For small F, the result (18) with integration limits $(0,2\pi/a)$ closely approaches the semiclassical result, multiplied by $\pi^2/9$. This has also been found in analytical calculations by Franz[15] and Eilenberger.[16] However, both functions are not asymptotically equal, because the role of oscillations increases for the dashed curve.

In the opposite case, with integration limits $(-\pi/a,+\pi/a)$ (dash-dotted line), the tunneling rate has an infinite number of zeros, which appear between the level crossings. By averaging over all possible integration intervals (19), the present result (solid line) combines features of both the dashed curve and the dash-dotted curve and, therefore, describes nonresonant and resonant tunneling. All three results (solid line, dashed line, dash-dotted line) are equal at those points where $\overline{E}_{20}-\overline{E}_{10}=eFal, l\in\mathbb{Z}$, or, equivalently, where b/a is an integer. In this case, the integral (18) is the same for all integration limits $(q_0,q_0+2\pi/a)$.

The explicit results shown in Fig. 2 demonstrate that the perturbational result from the literature (18) with appropriate integration limits symmetric around the gap cannot be considered an improvement over the semiclassical result (8) and any physical conclusions drawn from the difference between the semiclassical result (8) and the expression (18) are spurious. Thus the state of art in the field is a result from 1940, which after all does not go beyond the semiclassical result from 1934! The result derived in this paper (19), which is not based upon the assumption of Bloch oscillations, consistently takes into account resonant tunneling, which is mani-

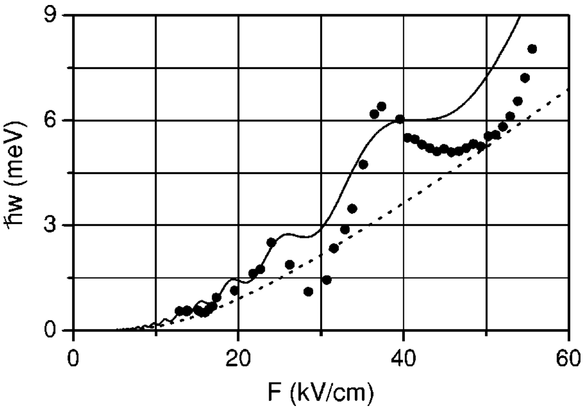

FIG. 3. Tunneling rate $\hbar w$ versus electric field F. Solid line: present result (19). Dotted line: semiclassical result (8). Circles: experimental data (from Ref. 27).

fested in pronounced oscillations of the tunneling rate.

Let us now make a comparison with recent experimental results.[27]. The Zener tunneling leads to an increase of the linewidth with increasing field in the optical absorption. For exponential decay of the wave function, the resulting absorption line is a Lorentzian with a half width at half maximum given by the decay constant. Then the tunneling rate, which describes exponential decay of the probability, should be compared to the full width at half maximum (FWHM) of the absorption lines. Experimentally, it is observed that the line shape is not exactly Lorentzian, but falls off more rapidly. This is in line with the fact that, contrary to the assumptions of Zener theory, $p(t)$ is a smooth function at $t=0$.

The comparison between theory and experiment is shown in Fig. 3. The present result (solid line) and the semiclassical result (dashed line) are compared with the linewidth (FWHM) of the optical transitions (black circles; cf. Fig. 2 of Ref. 27). The broadening at $F=0$, which is due to other decay mechanisms, was subtracted from the experimental linewidth. Although the theory does not use any adjustable parameters, we observe good agreement between theory and experiment. The overall increase of the linewidth is described by the semiclassical result. The experimental linewidth shows strong oscillations, which are due to interaction of Wannier-Stark ladders (cf. Fig. 1 of Ref. 26). For intermediate and large fields, the positions of the maxima and minima agree with the prediction of the theory (solid line). The differences for small fields result from Coulomb interaction, which leads to excitonic Stark ladders.[46] With regard to the oscillations, the theoretical result is in good qualitative agreement with the experimental data, compared to the result of the literature (dashed curve in Fig. 2), which underestimates the oscillations by orders of magnitude. However, the amplitude of the oscillations in the experiment is systematically larger than predicted by theory. This is due to a principal limitation of the theory, which assumes tunneling during different Bloch cycles to be uncorrelated, while the experimental results suggest a weak correlation between tunneling during subsequent Bloch cycles.

For the superlattice under consideration there exist also all-numerical results for the optical density of states.[25] These results are found in better agreement of the the experimental data than the result of perturbation theory.[26] The reason is that the full numerical calculation takes into account all minibands and does not rely on assumptions about the dephasing of nondiagonal elements. An even better agreement between theory and experiment is achieved by calculating the absorption including Coulomb interaction.[27] However, some discrepancies remain, which can be addressed to the neglect of band mixing in the calculation.

V. SUMMARY AND CONCLUSIONS

In this paper we have derived the Zener tunneling rate by means of quantum-mechanical perturbation theory using the Kane functions as base functions. The tunneling rate increases with field, like the semiclassical result, but also shows pronounced $1/F$ oscillations, which are the result of interacting Wannier-Stark levels. Therefore, the formula derived in this paper, Eq. (19), is capable of describing both nonresonant and resonant tunneling.

In contrast, the result derived in the context of the crystal-momentum representation (18), which is the state of the art until now, does not adequately describe resonant tunneling, contrary to statements in the literature. The tunneling rate as a function of the electric field is very similar to the semiclassical result (8) in the whole field range. For small fields, both functions are nearly identical, even though they are not asymptotically equal. The result (18) does not contain new qualitative features and cannot be considered as an improvement over the semiclassical result. The inadequacy of the previous result (18) to describe resonant tunneling and a spurious dependence on the integration interval result from the fact that the initial wave function is a Bloch function, which is not an eigenstate of the unperturbed Hamiltonian. In the present approach, these problems are fixed by choosing a Kane function as initial wave function or, equivalently, an initial density matrix which is diagonal in the basis of Kane functions.

The formula derived in this paper Eq. (19) reproduces the evolution of the line broadening in the optical absorption of a biased superlattice. The oscillations observed in the experiment are somewhat stronger than in theory, which indicates that the transitions during different Bloch cycles are not completely uncorrelated.

ACKNOWLEDGMENTS

The author wishes to thank F. Bechstedt and J.-M. Wagner for comments on the manuscript and B. Rosam for interesting discussions and the experimental data.

APPENDIX A: FOURIER SERIES

Let v be a complex-valued function of real variable k and $v(k+2\pi/a)$; $a>0$. Then the Fourier series of v is given by

$$v(k)=\sum_{l=-\infty}^{+\infty} v_l e^{-\mathrm{i}kal}; \quad v_l=\frac{a}{2\pi}\int_{-\pi/a}^{+\pi/a} dk\, e^{\mathrm{i}kal} v(k). \quad \text{(A1)}$$

Many rules for Fourier series are in complete analogy to the rules for the Fourier transform, for example

$$\frac{a}{2\pi}\int_{-\pi/a}^{+\pi/a} dk'\ v(k-k')w(k') = \sum_{l=-\infty}^{+\infty} v_l w_l e^{-\mathrm{i}kal} \quad \text{(A2)}$$

and

$$\frac{a}{2\pi}\int_{-\pi/a}^{+\pi/a} dk\ v^*(k)w(k) = \sum_{l=-\infty}^{+\infty} v_l^* w_l, \quad \text{(A3)}$$

can be considered as the convolution theorem and Parseval's theorem for Fourier series.

APPENDIX B: PROOF OF EQ. (17)

We wish to calculate the function

$$g(k,b) = \sum_{l=-\infty}^{+\infty} \mathrm{sinc}^2\left[(b-al)\frac{\pi}{a}\right] e^{\mathrm{i}(b-al)k} = e^{+\mathrm{i}kb} v(k). \quad \text{(B1)}$$

For fixed b we define a function v as

$$v(k) = \sum_{l=-\infty}^{+\infty} \mathrm{sinc}^2\left[(b-al)\frac{\pi}{a}\right] e^{-\mathrm{i}akl} \quad \text{(B2)}$$

and notice that

$$\sum_{l=-\infty}^{+\infty} \mathrm{sinc}\left[(b-al)\frac{\pi}{a}\right] e^{-\mathrm{i}akl} = w(k)$$

$$w(k) = e^{-\mathrm{i}kb} \quad \text{for} \quad k \in \left(-\frac{\pi}{a}, +\frac{\pi}{a}\right); \quad w\left(k+\frac{2\pi}{a}\right) = w(k),$$

which can be directly verified by calculating the Fourier coefficients. In order to calculate the Fourier series (B2), we use the convolution theorem (A2) and it follows that

$$v(k) = \frac{a}{2\pi}\int_{-\pi/a}^{+\pi/a} dk'\ w(k-k')w(k') = \begin{cases} \left(1+\frac{ka}{2\pi}\right)e^{-\mathrm{i}kb} - \frac{ka}{2\pi}\ e^{-\mathrm{i}(k+2\pi/a)b} & \text{for} \quad -\frac{\pi}{a} \leqslant k \leqslant 0 \\ \frac{ka}{2\pi} e^{-\mathrm{i}(k-2\pi/a)b} + \left(1-\frac{ka}{2\pi}\right)e^{-\mathrm{i}kb} & \text{for} \quad 0 \leqslant k \leqslant +\frac{\pi}{a}. \end{cases}$$

The function v is periodic with period $2\pi/a$ and continuous. The functional form of v is the same for $-2\pi/a \leqslant k \leqslant -\pi/a$ and $-\pi/a \leqslant k \leqslant 0$ and also for $0 \leqslant k \leqslant +\pi/a$ and $+\pi/a \leqslant k \leqslant +2\pi/a$. Multiplication with $e^{+\mathrm{i}kb}$ gives the result (17).

APPENDIX C: EQUIVALENCE OF EXPRESSIONS (15) AND (19)

The transition probability in Eq. (19) is equal to

$$p = \frac{a}{2\pi}\int_{-\pi/a}^{+\pi/a} dq \left| \int_{q-\pi/a}^{q+\pi/a} dk\ f(k) \right|^2 \quad \text{(C1)}$$

with f defined in Eq. (16). The function $f(k)e^{-\mathrm{i}bk} = u(k)$ is periodic. Therefore, f can be represented as

$$f(k) = u(k)e^{+\mathrm{i}bk} = \sum_{l=-\infty}^{+\infty} u_l e^{+\mathrm{i}(b-al)k}.$$

With

$$\frac{a}{2\pi}\int_{-\pi/a}^{+\pi/a} dq \int_{q-\pi/a}^{q+\pi/a} dk\ e^{-\mathrm{i}(b-al)k} \int_{q-\pi/a}^{q+\pi/a} dk'\ e^{+\mathrm{i}(b-al')k'} = \left(\frac{2\pi}{a}\right)^2 \mathrm{sinc}^2\left[(b-al)\frac{\pi}{a}\right]\delta_{ll'}$$

expression (C1) becomes

$$p = \left(\frac{2\pi}{a}\right)^2 \sum_{l=-\infty}^{+\infty} |u_l|^2\ \mathrm{sinc}^2\left[(b-al)\frac{\pi}{a}\right].$$

With the convolution theorem (A2) and Parseval's theorem (A3) it follows that

$$p = \int_{-\pi/a}^{+\pi/a} dk_1\ u^*(k_1) \int_{-\pi/a}^{+\pi/a} dk_2\ v(k_1-k_2)u(k_2)$$

with v defined in Eq. (B2). If u and v are substituted by f and g we obtain the transition probability in Eq. (15).

*Electronic address: stephan@ifto.physik.uni-jena.de

[1] F. Bloch, Z. Phys. **52**, 555 (1928).
[2] C. Zener, Proc. R. Soc. London, Ser. A **145**, 523 (1934).
[3] L. Landau, Phys. Z. Sowjetunion **1**, 88 (1932); **2**, 46 (1932).
[4] W. V. Houston, Phys. Rev. **57**, 184 (1940).
[5] E. N. Adams, J. Chem. Phys. **21**, 2013 (1953).
[6] E. N. Adams, Phys. Rev. **107**, 698 (1957).
[7] L. V. Keldysh, Zh. Eksp. Teor. Fiz. **33**, 994 (1957) [Sov. Phys. JETP **6**, 763 (1958)].
[8] E. O. Kane, J. Phys. Chem. Solids **12**, 181 (1959).
[9] G. H. Wannier, Phys. Rev. **117**, 432 (1960).
[10] E. O. Kane, J. Appl. Phys. **32**, 83 (1961).
[11] L. Fritsche, Phys. Status Solidi **13**, 487 (1966).
[12] D. E. Aspnes and N. Bottka, in *Semiconductors and Semimetals*, Vol. 9, edited by. R. K. Willardson and A. C. Beer (Academic, New York, 1972), p. 457.
[13] J. B. Krieger and G. J. Iafrate, Phys. Rev. B **33**, 5494 (1986); J. Zak, *ibid.* **38**, 6322 (1988); G. J. Iafrate and J. B. Krieger, *ibid.* **38**, 6324 (1988).
[14] J. B. Krieger and G. J. Iafrate, Phys. Rev. B **35**, 9644 (1987).
[15] W. Franz, Z. Naturforsch. A **14A**, 415 (1959).
[16] G. Eilenberger, Z. Phys. **164**, 59 (1961).
[17] J. E. Avron, J. Zak, A. Grossmann, and L. Gunther, J. Math. Phys. **18**, 918 (1977).
[18] I. W. Herbst and J. S. Howland, Commun. Math. Phys. **80**, 23 (1981).
[19] I. W. Herbst, Commun. Math. Phys. **75**, 197 (1980).
[20] V. Grecchi, M. Maioli, and A. Sacchetti, J. Phys. A **26**, L379 (1993); Commun. Math. Phys. **159**, 605 (1994).
[21] A. Di Carlo, P. Vogl, and W. Pötz, Phys. Rev. B **50**, 8358 (1994).
[22] K. Murayama, H. Nagasawa, S. Ozaki, M. Morifuji, C. Hamaguchi, A. Di Carlo, P. Vogl, G. Böhm, and G. Weimann, Superlattices Microstruct. **20**, 493 (1996); M. Morifuji, K. Murayama, C. Hamaguchi, A. Di Carlo, P. Vogl, G. Böhm, and M. Sexl, Phys. Status Solidi B **204**, 368 (1997).
[23] A. Sibille, J. F. Palmier, and F. Laurelle, Phys. Rev. Lett. **80**, 4506 (1998).
[24] M. Helm, W. Hilber, G. Strasser, R. De Meester, F. M. Peeters, and A. Wacker, Phys. Rev. Lett. **82**, 3120 (1999).
[25] S. Glutsch and F. Bechstedt, Phys. Rev. B **60**, 16 584 (1999).
[26] B. Rosam, D. Meinhold, F. Löser, V. G. Lyssenko, S. Glutsch, F. Bechstedt, F. Rossi, K. Köhler, and K. Leo, Phys. Rev. Lett. **86**, 1307 (2001).
[27] S. Glutsch, F. Bechstedt, B. Rosam, and K. Leo, Phys. Rev. B **63**, 085307 (2001).
[28] D. Meinhold, K. Leo, N. A. Fromer, D. S. Chemla, S. Glutsch, F. Bechstedt, and K. Köhler, Phys. Rev. B **65**, 161307 (2002).
[29] M. Glück, A. R. Kolovsky, H. J. Korsch, and F. Zimmer, Phys. Rev. B **65**, 115302 (2002).
[30] C. F. Bharucha, K. W. Madison, P. R. Morrow, S. R. Wilkinson, B. Sundaram, and M. G. Raizen, Phys. Rev. A **55**, R857 (1997).
[31] M. Glück, A. R. Kolovsky, and H. J. Korsch, Phys. Rev. Lett. **83**, 891 (1999).
[32] M. Holthaus, J. Opt. B: Quantum Semiclassical Opt. **2**, 589 (2000).
[33] H. Schneider, H. T. Grahn, K. v. Klitzing, and K. Ploog, Phys. Rev. Lett. **65**, 2720 (1990).
[34] R. Till, H. Kümmel, A. Philipp, G. Böhm, and G. Weimann, Superlattices Microstruct. **24**, 227 (1998).
[35] P. Ao and J. Rammer, Phys. Rev. B **44**, 11 494 (1991).
[36] W. Kohn, Phys. Rev. **115**, 809 (1959).
[37] S. Glutsch, J. Phys.: Condens. Matter **11**, 5533 (1999).
[38] J. Callaway, *Quantum Theory of the Solid State*, 2nd ed. (Academic Press, San Diego, 1991).
[39] K. B. McAfee, E. J. Ryder, W. Shockley, and M. Sparks, Phys. Rev. **83**, 650 (1951).
[40] W. Franz, in *Encyclopedia of Physics*, edited by S. Flügge, Vol. 17 (Springer, Berlin, 1956), p. 155.
[41] P. Ao, Phys. Rev. B **41**, 3998 (1990).
[42] Q. Niu and M. G. Raizen, Phys. Rev. Lett. **80**, 3491 (1998).
[43] J. P. Davis and P. Pechukas, J. Chem. Phys. **64**, 3129 (1976).
[44] M. V. Berry, Proc. R. Soc. London, Ser. A **429**, 61 (1990).
[45] M. V. Berry and R. Lim, J. Phys. A **26**, 4737 (1993).
[46] M. M. Dignam and J. E. Sipe, Phys. Rev. Lett. **64**, 1797 (1990); Phys. Rev. B **43**, 4097 (1991).