Nonexponential corrections to the Zener tunneling rate

Ping Ao*
Department of Physics, University of Illinois at Urbana-Champaign, 1110 West Green Street, Urbana, Illinois 61801

Jørgen Rammer
Institutt for Fysikk, Universitetet i Trondheim-NTH, N-7034 Trondheim, Norway
(Received 31 July 1991)

In artificial structures, such as superlattices, arbitrary relationships between band-structure parameters can be achieved, enforcing the need for a reinvestigation of the interband transition of a crystal electron originally studied by Zener. We report the existence of nonexponential corrections to the Zener tunneling rate and demonstrate their importance in parameter regimes relevant for quantum-transport properties of superlattices, such as Bloch oscillations. Oscillations in the I-V characteristic of a superlattice are predicted with periods determined by the band-structure parameters.

Recent advancement in microfabrication of superlattices[1] has reached the regime where a direct observation of the long-predicted Bloch oscillations should be possible. However, as we shall demonstrate in the following, interband transitions, which can be the dominant breakdown mechanism for the oscillations, can occur in superlattices with a substantially altered rate in comparison to the generic situation in semiconductors. Zener tunneling[2] was originally contemplated in the context of bulk insulators, which, as we show, in the present context are characterized by their large bandwidths compared to their energy gaps. In superlattices small bandwidths occur and we shall therefore reconsider the Zener tunneling theory in order to be able to deal with arbitrary bandwidths. As the present analysis reveals, in addition to the breakdown feature of the interband transition rate, oscillatory behavior will occur as a function of the applied field. This feature of the interband transition rate is manifested both in transport and optical properties of superlattices. Furthermore, the finite-bandwidth effect we shall discuss is relevant to Bloch oscillations in crystals,[3] Josephson and normal junctions,[4] and has implications both for the dynamics of an electron in a metallic ring driven by a changing magnetic flux[5] and interband transitions in an optical band structure.[6]

The energy eigenvalues of extended electron states in a crystal form energy bands separated by energy gaps. When an external electric force is applied to the electron, it gives rise to coupling between the energy bands and causes interband transitions. The occurrence of Bloch oscillations corresponds to the neglect of such interband transitions corresponding to treating the band index as a conserved quantity. The interband transition rate was first calculated more than half a century ago by Zener,[2] who viewed the transition as tunneling through the gap region, and the phenomenon has since been referred to as Zener tunneling. The Zener tunneling rate has subsequently been calculated using quite differing methods.[3] Although the results differ in details depending on assumed band structure, the Zener tunneling probability is essentially of the form $\exp(-F_0/F)$, where F is the applied force and F_0 a combination of the energy gap, lattice constant, and the effective mass of the electron, as standard textbook results show.[7] In this Brief Report, we demonstrate that the interband transition rate exhibits oscillatory behavior as a function of the inverse field strength $1/F$, with periods determined by the lattice constant, energy gap, and bandwidth, when proper account of a finite bandwidth is taken.

The Hamiltonian of an electron in a crystal experiencing an applied electric force F is

$$H=-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2}+v(x)-Fx \ . \tag{1}$$

Here, m is the free-electron mass and $v(x)$ is the periodic potential with periodicity d. A one-dimensional crystal is considered here, in accordance with assuming the applied electric field along a reciprocal-lattice vector. Assuming that the electron at some initial time (taken to be at time $t=0$) is in a given Bloch state, we ask for the probability that the electron is to be found in a different band after one Block period T, $T=h/Fd$. The corresponding transition amplitude from band n' to band n in one Bloch period is given by[8]

$$A_{nn'}=\frac{F}{i\hbar}\int_0^T dt\, X_{nn'}[k(t)] \times\exp\left[\frac{i}{\hbar}\int_0^t dt'\{\varepsilon_n[k(t')]-\varepsilon_{n'}[k(t')]\}\right] \tag{2}$$

corresponding to the Houston expansion.[3] Here, $X_{nn'}(k)=-i\int dx\, u_n^*(k,x)(\partial/\partial k)u_{n'}(k,x)$ is the force-induced interband coupling, with $u_n(k,x)$ being the spatially periodic part of the Bloch function, and $\varepsilon_n(k)$ the corresponding energy eigenvalue. The temporal relation between the crystal momentum and the force is $k(t)=Ft/\hbar$. For weak applied force F, the crystal momentum of the electron changes slowly, and Eq. (2) for the transition amplitude is the corresponding adiabat-

ic expression. One expects that in the case of a strong force F, the adiabatic expansion is not valid for considering the long-time behavior. A study of the dynamics of an electron in this case can be found elsewhere.[9]

For an arbitrary periodic potential $v(x)$, the integral in Eq. (2) is intractable, but we shall, by treating exactly a relevant model Hamiltonian, be able to obtain the general behavior of the transition amplitude. We consider, therefore, a two-band model that contains the essential physics and permits the integration in Eq. (2) to be performed. The two-band model is described by the Hamiltonian

$$H=\sum_{n=1,2}\sum_{p=-\infty}^{\infty}[W_n(a^{\dagger}_{n,p}a_{n,p+1}+a^{\dagger}_{n,p+1}a_{n,p}) +(\varepsilon_n-Fdp)a^{\dagger}_{n,p}a_{n,p}] +FX\sum_{p=-\infty}^{\infty}(a^{\dagger}_{1,p}a_{2,p}+a^{\dagger}_{2,p}a_{1,p}), \tag{3}$$

where the first term describes the hopping of the electron in the presence of the applied field and the second term describes the interband coupling. In the above site representation, the two-band model corresponds to two orbitals at each site. For the orbitals $n=1$ and 2, $W_n=(-1)^nW/4$ are the hopping constants corresponding to tunneling between neighboring sites (labeled by integer p), and $\varepsilon_n=(-1)^n\varepsilon$ are the on-site energies. The force-induced coupling between the two orbitals is FX and is for convenience chosen at the same site. Within this model, dispersion relations for the bands $n=1,2$ are simple harmonics $\varepsilon_n(k)=(-1)^n[\varepsilon+(W/2)\cos(kd)]$ with bandwidth W, and the matrix elements $X_{21}=X^*_{12}=X$ are real. Therefore, according to Eq. (2), the interband transition amplitude in one Bloch period is

$$A=-iX\int_0^{2\pi/d}dk\,\exp\left[\frac{i}{Fd}[2\varepsilon kd+W\sin(kd)]\right]. \tag{4}$$

The transition amplitude A for the chosen model can be expressed in terms of a known function, viz, the Anger function $\hat{\mathbf{J}}_\nu(z)$,[10] $A=(-2\pi iX/d)e^{i\pi\nu}\hat{\mathbf{J}}_\nu(z)$, where $\nu=2\varepsilon/Fd$ and $z=W/Fd$. Despite the compact form, the field dependence enters in a complicated fashion. In order to make contact with known results and to systematically study corrections to the Zener expression, we first consider the low-field limit where the voltage drop across the unit cell is smaller than the bandwidth, $Fd<W$. In the low-field regime we obtain, by using the asymptotic expansion of the Anger function,[10], for the transition amplitude (from now on we drop the irrelevant phase factor)

$$A=\frac{2\pi X}{d}\left[J_\nu(z)+\frac{1+\nu/\pi}{\pi z}\sin(\pi\nu)\right], \tag{5}$$

where $J_\nu(z)$ is the Bessel function. If we further assume that Fd/Δ is not in far excess of $(\Delta/W)^{1/2}$ (a further restriction only in the limit where the energy gap $\Delta=2\varepsilon-W$, is much smaller than the bandwidth), we can use the tangent approximation for the Bessel function,[10] and obtain

$$A=\frac{2\pi X}{d}\left[\left[\frac{Fd}{2\pi\Delta(1+2W/\Delta)^{1/2}}\right]^{1/2}\exp\left\{-\left[\frac{\Delta}{Fd}\right]\left[\left[1+\frac{W}{\Delta}\right]\operatorname{arccosh}\left[1+\frac{\Delta}{W}\right]-\left[1+2\frac{W}{\Delta}\right]^{1/2}\right]\right\} +\frac{(2+\Delta/W)Fd}{\pi W}\sin\left[\frac{\pi(W+\Delta)}{Fd}\right]\right]. \tag{6}$$

Expression Eq. (6) is valid for arbitrary relationship between gap energy and bandwidth. The limiting cases that we now discuss are appropriate for different cases of material parameters.

In the commonly prevailing situation in semiconductors where the bandwidth is much larger than the band gap, $W>\Delta$, we obtain

$$A=X\sqrt{\pi F}\left[\frac{m^*}{\Delta\hbar^2}\right]^{1/4}\exp\left[-\frac{2}{3F}\left[\frac{m^*}{\hbar^2}\right]^{1/2}\Delta^{3/2}\right] +\frac{2FX}{W}\sin\left[\frac{\pi W}{Fd}\right], \tag{7}$$

where, to compare with Zener's result, we have rewritten the first term in the transition amplitude in one Bloch period in terms of the effective mass $m^*=2\hbar^2/Wd^2$. The first term is Zener's result, which is the contribution near the energy gap to the integral of Eq. (2). It exhibits the breakdown feature and is, at small fields, exponentially small. The second term is the result of a proper account of the finite time to complete the transition, and although this time is very long in the limit of small applied force, this term can be the dominant one. The amplitude of the second term will not be linear in the applied force F for general forms of force-induced interband coupling and energy dispersion relations, but the powerlike and oscillatory behavior remains. We note that the difference between the result obtained here and Zener's result vanishes when the bandwidth W becomes large, as a consequence of the arbitrary large energy differences enabling the phase factor in Eq. (2) to render the finite-time restriction irrelevant, that is, there will be no contribution away from the gap region because of the fast oscillation of the phase factor, and we are left with the first term in Eq. (7), the stationary phase contribution. The second term of Eq. (7) is seen to exhibit oscillatory behavior in the inverse field strength $1/F$, with a period equal to $2d/W$.

In the limit where the bandwidth is smaller than the band gap, $Fd<W<\Delta$, a situation achievable in superlat-

tices, we obtain

$$A=\frac{2\pi X}{d}\left[\left[\frac{Fd}{2\pi\Delta}\right]^{1/2}\exp\left[-\frac{\Delta^2}{FdW}\right]+\frac{Fd}{\pi\Delta}\left[\frac{\Delta}{W}\right]^2\sin\left[\frac{\pi\Delta}{Fd}\right]\right]. \tag{8}$$

We note that the nonexponential term dominates in this parameter regime, and the transition amplitude exhibits oscillatory behavior in the inverse field strength $1/F$, with a period equal to $2d/\Delta$.

In the high-field limit where the bandwidth is small compared to the voltage drop across the unit cell, $W<Fd$, a situation achievable in superlattices, we note from the series representation of the Anger function,[10] that the transition amplitude exhibits oscillations in the inverse field strength $1/F$, with a period equal to $4d/(\Delta+W)$. For example, in the case where the energy gap is of the same order of magnitude as the bandwidth, only the lowest-order term in W/Fd need be retained and we obtain for the transition amplitude

$$A=\frac{2\pi X}{d}\frac{\cos[\pi(\Delta+W)/2Fd]}{\Gamma(1+(\Delta+W)/2Fd)\Gamma(1-(\Delta+W)/2Fd)}, \tag{9}$$

where Γ denotes the gamma function. We note that the breakdown feature of the transition amplitude is absent in this limit.

The original point of view of Zener, that the interband transition can be viewed as tunneling in real space, is heuristic as it introduces a space-dependent band structure. Such an approach is incapable of treating the effect of the finite-time restriction on the interband transition. In the following we shall therefore calculate the interband transition rate assuming the initial state to be not a Bloch state but rather the complementary, a state where the particle at the initial time is localized on one of the lower orbitals. For such a localized initial state the interband transition probability can be found by using the fact that the Hamiltonian Eq. (3) is an exactly solvable one if the last, interband, term is absent. The transition probability from a definite orbital 1 site (chosen to be $p=0$) to an arbitrary orbital 2 site in time span t is

$$P(t)=\sum_{p=-\infty}^{\infty}|\langle p,2|U(t)|1,0\rangle|^2, \tag{10}$$

where $U(t)$ is the evolution operator corresponding to the Hamiltonian given in Eq. (3). The intraband propagator can be expressed in terms of the Bessel function[11]

$$\langle p,n|U_0(t)|n',p'\rangle=\delta_{n,n'}[(-1)^n i]^{p-p'}\exp\left[i\left[\frac{Fd}{2\hbar}(p-p')t-(-1)^n\frac{\varepsilon}{\hbar}t\right]\right]J_{p-p'}\left[\frac{W}{Fd}\sin\left[\frac{Fd}{2\hbar}t\right]\right], \tag{11}$$

and is apart from a phase factor periodic in time with the Bloch period. In the following we shall calculate the interband transition probability assuming that we only need to take into consideration one interband hop. The transition probability within one Bloch period, $P(T)$, can then be given in terms of the Bessel function,

$$P(T)=\left[\frac{2X}{d}\right]^2\left[\pi^2\cos(\nu\pi)J_{-\nu}(z)J_{\nu}(z)-\pi\sin(\nu\pi)\frac{\partial}{\partial\nu}[J_{-\nu}(z)J_{\nu}(z)]\right]. \tag{12}$$

Again, using the tangent approximation for the Bessel function, we can obtain an explicit expression similar to Eq. (6). For definiteness, we only state the result for the case where the bandwidth is larger than the energy gap, which in turn is much larger than the voltage drop across the unit cell, $W>\Delta>Fd$. The transition probability in this case is

$$P(T)=\left[\frac{2\pi X}{d}\right]^2\left\{\frac{Fd^2}{4\pi\sqrt{\Delta}}\left[\frac{m^*}{\hbar^2}\right]^{1/2}\exp\left[-\frac{4}{3F}\left[\frac{m^*}{\hbar^2}\right]^{1/2}\Delta^{3/2}\right]+\frac{F^2d^3}{4\pi^2\Delta^{3/2}}\left[\frac{m^*}{\hbar^2}\right]^{1/2}\sin^2\left[\frac{\pi W}{Fd}\right]\right\}. \tag{13}$$

The first term in Eq. (13) is the Zener result and the same as we found for the Bloch state. The second, nonexponential correction term is of a form similar to the one we found for the Bloch state. However, the prefactor is relatively larger by the order of magnitude $(W/\Delta)^{3/2}$, as the effective weight in the gap region diminishes for the Bloch state compared to the localized initial state when the bandwidth increases. The effect of the finite-time restriction is thus dependent on the initial state.

Regarding material parameters and the possibility of observing the effects of the oscillatory behavior of the transition rate, the following situation prevails. In semiconductors and insulators, the bandwidth is much larger than the energy gap. In this case the interband transition rate is close to the exponential form as follows from Eq. (7), and the I-V characteristic can exhibit the breakdown feature, except for possible additional oscillations at low fields. However, at low fields one has to take into account the influence of the environment such as phonons, since the oscillatory behavior is sensitive to this influence. Generally speaking, the neglect of inelastic-scattering processes requires that the inelastic-scattering time is larger than the Bloch period. This requirement, together with the smallness of the oscillatory terms, makes their observation difficult in materials provided by nature. However, owing to present fabrication technology, these difficulties can be overcome using artificial structures. For instance, in superlattices, the energy gap Δ and the energy bandwidth W can be manufactured to be of the same order of magnitude, say of the order of 10 meV for

a superlattice constant d of order 10 nm, by tuning the single-layer structural parameters. Therefore, the finite-bandwidth effect on the interband transition rate should be observable in such systems at electric-field strengths of the order of 10^4 V/cm. The conductance (per electron unit cell) is given by

$$G=\frac{e^2}{h}P\ , \tag{14}$$

where P is the transition probability for the Bloch state, $P=|A|^2$. The conductance due to interband transitions will therefore exhibit the typical oscillatory behavior depicted in Fig. 1, and depending on the relative magnitude of the two terms in Eq. (6), a variety of behaviors of the conductance can occur.

In conclusion, we have shown that Zener's original view regarding the interband transition as a tunneling process through a forbidden gap region is only reliable in the limit of large bandwidth, as a proper account of the finite time to complete the transition leads to corrections to the original Zener result, which has a different analytic dependence on the field strength. The considered two-band tight-binding model allowed us to perform a complete analytical treatment of the interband transition, and since the level-crossing problem is ubiquitous in physics our analysis has implications for the wide range of phenomena mentioned in the Introduction. In particular, we note that the sophistication of today's fabrication techniques makes it possible to construct superlattices with arbitrary relationship between the band gap and bandwidth. We have shown that in such a situation there can be substantial corrections to the usual Zener tunneling rate, and have derived the interband transition rate for arbitrary relationships between the band gap and bandwidth. In parameter regimes achievable in superlattices we predict oscillations in the I-V characteristic with periods determined by the band-structure parameters and inversely proportional to the field strength, thus allowing for a simple experimental method to determine band-structure parameters.

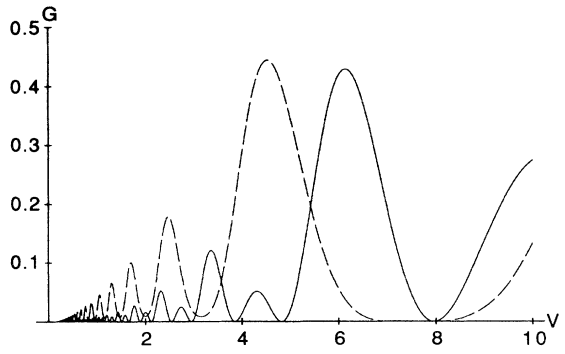

FIG. 1. The conductance [in units of e^2/h times the scale factor $(2\pi X/d)^{-2}$] for a superlattice consisting of 100 unit cells as a function of the voltage V (measured in volts), calculated according to Eqs. (6) and (14). The dashed curve corresponds to the values $\Delta=10$ meV and a ratio $\Delta/W=0.1$, and the solid curve to $\Delta=50$ meV and a ratio $\Delta/W=0.5$.

It is a pleasure to thank A. J. Leggett, Q. Niu, and A. Zawadowski for helpful discussions. This research was supported in part by the John D. and Catherine T. MacArthur Foundation at the University of Illinois under Grant No. 0-6-40129 and by the National Science Foundation under Grant No. DMR-86-12860. One of us (J.R.) acknowledges support by the Norwegian Research Council for Sciences and the Humanities and Nordisk Ministerråd.

*Present address: Department of Physics, FM-15, University of Washington, Seattle, WA 98195.

[1]*Proceedings of the International Symposium on Nano-Structure Physics and Fabrication,* edited by M. A. Reed and W. P. Kirk (Academic, New York, 1989).

[2]C. Zener, Proc. R. Soc. London Ser. A **145**, 523 (1934).

[3]W. V. Houston, Phys. Rev. **57**, 184 (1940); L. V. Keldysh, Zh. Eksp. Teor. Fiz. **33**, 994 (1958) [Sov. Phys. JETP **6**, 763 (1958)]; W. Z. Franz, Z. Naturforsch. **14A**, 415 (1959); E. O. Kane, J. Phys. Chem. Solids **12**, 181 (1959); G. Eilenberger, Z. Phys. **164**, 59 (1961); L. Fritsche, Phys. Status Solidi **13**, 467 (1966).

[4]G. Schön and A. D. Zaikin, Phys. Rep. **198**, 237 (1990).

[5]M. Büttiker, Y. Imry, and R. Landauer, Phys. Lett. **96A**, 365 (1983).

[6]D. Lenstra and W. van Haeringen, in *Coherence and Quantum Optics VI,* edited by J. H. Eberly, L. Mendel, and E. Wolf (Plenum, New York, 1989).

[7]J. M. Ziman, *Principles of the Theory of Solids* (Cambridge University Press, Cambridge, 1972); J. Callaway, *Quantum Theory of the Solid State* (Academic, New York, 1976); S. V. Vonsovsky and M. I. Katsnelson, *Quantum Solid-State Physics* (Springer-Verlag, Berlin, 1989).

[8]J. B. Krieger and G. J. Iafrate, Phys. Rev. B **33**, 5494 (1986).

[9]P. Ao, Phys. Rev. B **41**, 3998 (1990).

[10]I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products* (Academic, New York, 1980).

[11]H. Fukuyama, R. A. Bari, and H. C. Fogedby, Phys. Rev. B **8**, 5579 (1973).