Negative-mass transport in semiconductor diodes

A. Dyson and B. K. Ridley
Department of Electronic Systems Engineering, University of Essex, Colchester, United Kingdom
(Received 21 June 2005; published 7 November 2005)

An analysis of high-field electron transport in a semiconductor diode is presented using a simple kinetic model. A cosine band structure is assumed for the conduction band and the diode is taken to be short enough for the rate of energy-relaxation collisions to be negligible. The drift velocity of the electrons is determined by momentum-relaxation collisions independent of the applied electric field and the band structure, and transfer or tunneling to other valleys or bands is ignored. The analysis is applied to GaN, and conditions for the appearance of a negative differential resistance (NDR) associated with the negative-mass states in the band are derived for the regime of diode lengths and fields in which Bloch oscillations do not occur. It is shown that NDR in this regime occurs over a limited range of diode lengths and fields. Lengths and fields corresponding to breakdown conditions are explicitly avoided.

DOI: 10.1103/PhysRevB.72.193301 PACS number(s): 73.63.−b, 72.20.Ht, 85.30.De, 85.30.Kk

I. INTRODUCTION

Interest in high-field transport in semiconductors, in which carriers are excited into regions of the band structure where their effective mass is negative, goes back to Krömer's proposal for a negative-mass amplifier in the 1950s (Refs. 1 and 2). The amplifier he proposed was based on the valence-band structure of Ge, but conventional hot-electron techniques failed to excite a sufficiently large population of holes into the negative-mass states, so the idea of a negative-mass amplifier was dropped. With the advent of new growth techniques that produced the semiconductor superlattice, the idea of realizing a negative differential resistance (NDR) associated with negative-mass states was revived by Esaki and Tsu.[3] Since then, superlattice NDR has been observed[4–6] and this has attracted a great deal of attention.[7–10] Much interest has also been shown in the possibility of generating Bloch oscillations (see Refs. 11–13, and references therein). Bloch oscillations in the bulk are achievable only via a combination of weak scattering and high electric fields, a combination that has generated interest in the phenomenon of the intracollisional field effect.[14–16]

Recently, attention has begun to focus on the possibility of negative-mass NDR and Bloch oscillations in bulk material. Since a high field is an essential ingredient, the only feasible semiconductors to consider are those with a large band gap and high breakdown field. One of the most promising materials is GaN and its alloys with Al. Their technological usefulness for light-emitting diodes (LEDs) and high-power microwave field-effect transistors (FETs) will guarantee steady efforts to improve the crystalline quality of bulk layers and heterostructures. It has recently been proposed that negative-mass NDR is achievable in ultrashort GaN diodes, provided that injection of electrons at the cathode is directly into a high-energy state of the conduction band.[17] The condition calls for essentially ballistic transport. But even in the case of long diodes in which the transit time is very long, and therefore in which collisions dominate, it has been shown that negative-mass NDR in GaN and in AlN is achievable in sufficiently high fields,[18–22] using the same criterion as Esaki and Tsu.[3] These predictions call for fields of order 1 MV/cm and higher, so even though NDR may occur in long diodes, a limitation on length is set by the breakdown voltage (some 4–5 V in GaN).

The prediction from the analysis of ballistic transport[17] is that the NDR can provide a source of teraherz radiation. If so, an AlGaN/GaN structure could provide a convenient, battery operated source of teraherz radiation. It is therefore of some practical, as well as theoretical, interest to analyze further the effect of collisions and diode length on negative-mass NDR.

The introduction of collisions immediately highlights the necessity to consider two distinct regimes, which we label A and B. In regime A the combination of field strength and diode length is such that Bloch oscillations cannot occur. In this case collisions can be regarded in the usual way as producing transitions between states described by Bloch functions. In regime B Bloch oscillations lead to the transformation of band states into Wannier–Stark states. In this case electron flow occurs only as a consequence of collisions as a hopping mechanism, and via Zener tunneling to higher bands. The work described here focuses entirely on regime A. As a consequence, the parameter's diode length and transit time emerge as factors as important as electric field. The inclusion of diode length and transit time makes the present analysis significantly different from the Esaki–Tsu approach. As a result we find that in regime A, NDR occurs only over a limited range of diode lengths and fields. The corresponding voltages are well below those which would lead to breakdown.

In Sec. II, we prepare the ground by describing our band-structure model and defining the regimes that are involved. In Sec. III, our scattering model is presented and in Sec. IV the results of our analysis are presented.

II. MODEL SYSTEM

We need to describe how the motion of an electron is affected by (1) an electrical field, (2) band structure, and (3) scattering. The quantum-mechanical result of adding an electric field of arbitrary strength to the crystal Hamiltonian results in the familiar acceleration law for the time dependence of the electron wave vector

$$\dot{\mathbf{k}} = e\mathbf{F}/\hbar, \tag{1}$$

where $\mathbf{k}$ is the electron wave vector and $\mathbf{F}$ is the field. This is a rigorous quantum-theoretic result.[13] In a constant, uniform electric field and in the absence of collisions the electrons travel through the Brillouin zone with constant rate. If the field is directed along a principal crystallographic direction in which the lattice constant is a, Bragg reflection causes the electron to oscillate within the range $-\pi/a < k < \pi/a$ with an angular frequency

$$\omega_B = eFa/\hbar. \tag{2}$$

How these Bloch oscillations relate to velocity depend upon the band structure which, in general, has no simple analytical form. However, in large band-gap semiconductors, where the influence of the valence band in particular is weak, the band along the field direction can be approximately described in terms of a simple cosine function,

$$E = E_0(1 - \cos ka). \tag{3}$$

This is quite a good approximation for GaN (Ref. 18). The velocity in the direction of the field is

$$v = v_m \sin ka,$$

$$v_m = E_0 a/\hbar = \hbar/m_0^* a, \tag{4}$$

where m_0^* is the band-edge effective mass. If the electron is injected at the cathode into the state k_0 at $t=0$,

$$v(t) = v_m \sin(\omega_B t + \alpha),$$

$$\alpha = k_0 a. \tag{5}$$

In general, the injected electron will have a perpendicular component of $\mathbf{k}$, but this can usually be neglected if the injection is via thermionic emission over or tunneling through a barrier.

Our interest is in the situation in which the field is large and the electron has high energy, both as a consequence of being injected over or through a barrier and being accelerated by the field. Under these circumstances the relevant scattering mechanism in reasonably pure, and unflawed material can be taken to be associated with optical phonons. Intervalley scattering via the deformation potential interaction in many semiconductors, GaN included, can lead to NDR via the electron-transfer mechanism. In GaN the lowest valley is at least 1.4 eV above the conduction-band edge and, moreover, the Bloch overlap function is small[18] so that the intervalley scattering rate is relatively small compared with the intravalley rate associated with polar-optical-phonon scattering. We feel justified, therefore, in neglecting intervalley scattering and its effect on transport. In the presence of high fields Zener tunneling into higher lying bands becomes possible, but we will assume that this process can be neglected except at extremely high fields. In short, we ignore the presence of bands and valleys other than the conduction band in which the electron remains.

Polar-optical-phonon scattering is an inelastic process. In wurtzite GaN it is also somewhat anisotropic, but for simplicity we will ignore that. At room temperature in GaN, scattering will be principally via the emission of a phonon whose energy is about 90 meV. We can distinguish three regimes on the basis of the three time constants associated with intravalley scattering, namely, the scattering time τ_s, the momentum-relaxation time τ, and the energy-relaxation time τ_E. For polar-optical-phonon scattering these time constants are roughly $\tau_s = 10$ fs, $\tau = 30$ fs, and $\tau_E = 500$ fs in the high-energy regime.[22] The energy dependences of the scattering and energy-relaxation rates are very weak; that of the momentum-relaxation time is somewhat more marked but since it passes through a weak minimum towards high energies it will be sufficient to assume a constant average value. We can then identify the three regimes as (1) Bloch oscillations, $\omega_B \tau_s \geqslant 1$; (2) quasiballistic, $\omega_B \tau \geqslant 1$; and (3) lucky drift, $\omega_B \tau_E \geqslant 1$.

Entry into the Bloch-oscillation regime triggers a radical change in the physics of the transport. Scattering becomes no longer a process describing a transition from one well-defined $\mathbf{k}$ state to another but, instead, a transition between Wannier–Stark quantized states. The velocity averaged over an oscillation is zero, so conduction depends on the scattering-induced hopping between Wannier–Stark states down the potential gradient. As already mentioned, we will not discuss this regime here. Nevertheless, an approach to this regime at fields of order 1 MV/cm may have an effect on the scattering process via the intracollisional field effect.[14] However, this effect depends on how sensitive the scattering rate is to the electron energy gained from the field during the scattering process. In the case of polar-optical-phonon scattering the rate is not very sensitive to energy at large energies, so the intracollisional field effect is expected to be weak in this case. Future work in this area is required to assess the effect on the magnitudes of time constants already quoted.

III. COLLISION-DOMINATED TRANSPORT

We consider the situation in which the field is too low for well-defined Wannier–Stark states to be established. We exploit the large difference between the rates for energy and momentum relaxation and take the collisions to be essentially elastic. (Exploitation of the large difference in energy- and momentum-relaxation rates has previously been the inspiration of the lucky-drift theory of impact ionization.[23]) We take the momentum-relaxation time to be the time between collisions that reduce the velocity in the direction of the field to zero. We further assume that the duration of a collision is infinitesimally small (a common assumption in transport theory), so no matter how short the diode is, there is the possibility that an electron may make an infinite number of collisions.

We consider the electron making zero, one, two, etc., momentum-relaxing collisions and obtain the velocity $v_j(t)$, where j is the number of collisions, weighted by the probability of having j collisions. Thus, the weighted velocity after zero collisions is

$$v_0(t) = v_m e^{-t/\tau} \sin(\omega_B t + \alpha). \tag{6a}$$

The probability of an electron avoiding collision for a time t_1 but then having a collision in the subsequent interval dt is $\exp(-t_1/\tau)dt/\tau$. The probability of having no further colli-

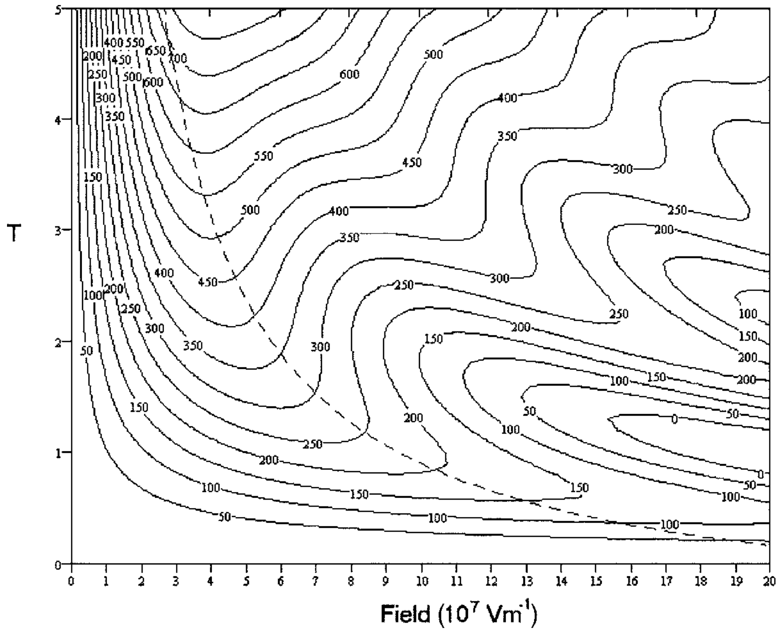

FIG. 1. Transit time dependence on field for various diode lengths (in angstroms) with cold injection $\alpha=0$. The contour $\omega_B T=\pi$ is the dashed line. The transit time is normalized to the momentum-relaxation time.

sion is $\exp\{-(t-t_1)/\tau\}$. Summing t_1 from 0 to t gives

$$v_1(t) = v_m \int_0^t e^{-(t-t_1)/\tau} \sin\{\omega_B(t-t_1)\} e^{-t_1/\tau} dt_1/\tau$$
$$= v_m e^{-t/\tau} \int_0^t \sin\{\omega_B(t-t_1)\} dt_1/\tau. \tag{6b}$$

A collision at t_1 followed by a second at t_2 leads to

$$v_2(t) = v_m e^{-t/\tau} \int_0^t (t_2/\tau) \sin\{\omega_B(t-t_2)\} dt_2/\tau, \tag{6c}$$

and so on. The average velocity is then given by

$$v(t) = \sum_{j=0}^{\infty} v_j(t) = v_m \left\{ \frac{\omega_B\tau}{1+(\omega_B\tau)^2} + e^{-t/\tau}\left[\sin\{\omega_B t+\alpha\} - \frac{1}{1+(\omega_B\tau)^2}(\sin\omega_B t + \omega_B\tau\cos\omega_B t)\right]\right\}. \tag{7}$$

With $\alpha=0$, this result agrees exactly with the Esaki–Tsu prescription for NDR,

$$v(t) = \frac{eF}{\hbar^2}\int_0^t e^{-t/\tau}\frac{d^2E}{dk^2}dt. \tag{8}$$

Putting $t=\infty$, corresponding to an infinitely long diode, yields the Esaki–Tsu result

$$v(\infty) = v_m \frac{\omega_B\tau}{1+(\omega_B\tau)^2}, \tag{9}$$

which predicts the appearance of NDR when $\omega_B\tau > 1$. However, in our case, the diode is of finite length and t must be restricted to the transit time T, which we assume is much shorter than the energy-relaxation time. Thus, we must retain Eq. (7) in its entirety. With $t=T$ Eq. (7) can be expressed in normalized notation as follows:

$$v(s) = v_m \left\{ \frac{\beta}{1+\beta^2} + e^{-s}\left[\sin\{\beta s+\alpha\} - \frac{1}{1+\beta^2}(\sin\beta s + \beta\cos\beta s)\right]\right\}, \tag{10}$$

where $s=T/\tau$ and $\beta=\omega_B\tau$. For a diode of length L the transit time for a given field is obtained self-consistently from

$$L = v(s)T. \tag{11}$$

In ultrashort diodes in which the motion is essentially collision free, the situation is different. In this case the velocity is $v_m \sin(\omega_B t+\alpha)$ and time dependent throughout. There is no constant time-averaged velocity so the current is the sum of drift and displacement components. This was the situation analyzed in Ref. 17, where conditions for the appearance of NDR were obtained. Here, we focus on the situation in diodes that are too long for purely ballistic motion to be dominant and assume the essential features of the Esaki–Tsu model. Thus the differential mobility is

$$\mu = \frac{dv(s)}{dF}. \tag{12}$$

With n equal to the electron density, the differential conductivity is $en\mu$. At low energies $\mu=e\tau/m^*$, corresponding to a conductivity of $\sigma_0=e^2n\tau/m^*$, where m^* is the band-edge effective mass.

IV. RESULTS AND DISCUSSION

The velocity, in addition to having an explicit dependence on field, has an implicit dependence through the transit time. The dependence of transit time on field for various diode lengths is shown in Fig. 1 for the case of cold injection: $\alpha=0$. Knowing this we can calculate the conductivity as a function of field, and this is shown in Fig. 2, where the appearance of a negative differential conductivity is evident. At low fields the transit time reduces with increasing field, responding to the increase in average velocity. Towards high

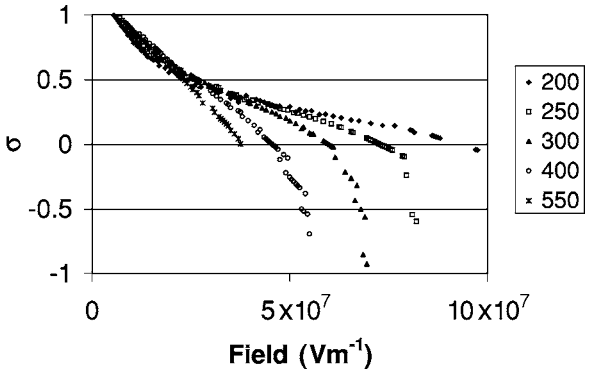

FIG. 2. Conductivity as a function of field for various diode lengths (in angstroms). The normalization is $\sigma_0 = e^2 n\tau / m^*$.

fields the average velocity reduces and hence the transit time increases. At very high fields in short diodes the influence of the onset of Bloch oscillations can be seen. The conductivity plots were deliberately limited to the Bloch-oscillation-free regime ($\omega_B T \leqslant \pi$), illustrated by the region in Fig. 1 to the left of the contour $\omega_B T = \pi$, since our kinetic theory of the effect of collisions is limited to that regime. The appearance of Bloch-oscillation-free NDR is therefore limited to diode lengths such that $\omega_B T \leqslant \pi$ and at fields where T increases, viz. 200 Å$<L<$550 Å (Fig. 2).

For hot injection, the fields have to be limited so that $\omega_B T + \alpha \leqslant \pi$ which, in turn, lowers the upper limit to the diode length set to avoid Bloch oscillations. Thus, for $\alpha = \pi/8, \pi/4$, we have found the limits to be 500 and 350 Å, respectively.

Given the simplifying assumptions that have underpinned this calculation, it is pertinent to consider the validity of its results. The choice of a cosine band structure is indeed a good fit to the band structure calculated by empirical pseudopotential methods. We have taken the conduction-band width to be 2.7 eV, but this is by no means generally accepted. A larger width would enhance the velocity but, otherwise, it would not affect our conclusions regarding the attainment of negative-mass NDR. More contentious is our neglect of transfer to the upper valley along the U direction. We have assumed that the valley is about 1.4 eV above the bottom of the conduction band, which is near the inflection point. Any large-scale transfer to the U valley would replace our negative-mass NDR with an electron-transfer NDR. This would have the effect, if the analogy with other III-V compounds is relevant, of reducing the frequency range of radiation that can be generated by the NDR, due to the slow return of electrons to the conduction band. The intervalley rates involved are by no means well known; they depend on the overlap integral of the cell-periodic components of the respective Bloch functions, which are known to be small. They also depend on the deformation constant associated with the intervalley phonons, which is not known with any accuracy, but usually assumed to be 1×10^9 eV/cm. Taking this to be the deformation potential, the single-valley effective density-of-states mass equal to 0.5 and the overlap integral to be 0.05, we get a rate assuming six valleys of order 2×10^{12}/s corresponding to a time constant of 500 fs. This is significantly longer than the intravalley rate, so neglecting transfer to the U valley is reasonably justified, though there remain uncertainties of overall density-of states mass and deformation potential. There are two higher-lying valleys: one at K, the other at L. Information about intervalley rates associated with these valleys is even scarcer than that for the U valley. If their contribution is of the same order, the total intervalley time constant would reduce to 150 fs, which is still five times longer than the intravalley momentum-relaxation time. But even were it to become comparable, the K and L valleys lie significantly far above the inflection point for negative-mass NDR to occur.

A more fundamental problem affecting both intravalley and intervalley scattering is the applicability of the Fermi golden rule for calculating rates in the presence of a high electric field. Another is the question of duration of collision, which assumes particular importance when a short diode is involved. Both of these problems in the particular context of short distances and short times are not new, yet they still need to be addressed in a way that is both transparent and rigorous.

ACKNOWLEDGMENTS

We are grateful to the Office of Naval Research for support under Grant No. (N00014-05-1-0192) and to Dr. Colin Wood and to Dr. Mark Taylor for their sponsorship.

[1] H. Krömer, Z. Phys. **134**, 435 (1953).
[2] H. Krömer, Phys. Rev. **109**, 1856 (1958).
[3] L. Esaki and R. Tsu, IBM J. Res. Dev. **14**, 81 (1970).
[4] A. Sibille *et al.*, Phys. Rev. Lett. **64**, 52 (1990).
[5] F. Beltram *et al.*, Phys. Rev. Lett. **64**, 3167 (1990).
[6] H. T. Grahn *et al.*, Phys. Rev. B **43**, R12094 (1991).
[7] R. Tsu and L. Esaki, Phys. Rev. B **43**, R5204 (1991).
[8] X. L. Lei *et al.*, J. Phys.: Condens. Matter **7**, 9811 (1995).
[9] A. A. Ignatov *et al.*, Mod. Phys. Lett. B **5**, 1087 (1991).
[10] R. R. Gerhardts, Phys. Rev. B **48**, R9178 (1993).
[11] A. A. Ignatov and V. I. Shashkin, Sov. Phys. JETP **66**, 526 (1987).
[12] N. Sawaki and T. Nishinaga, Inst. Phys. Conf. Ser. **43**, 323 (1979).
[13] F. Rossi, in *Theory of Transport Properties of Semiconductor Nanostructures*, edited by E. Schöll (Chapman and Hall, London, 1998), p. 283.
[14] J. R. Barker, Solid-State Electron. **21**, 267 (1978).
[15] K. K. Thornber, Solid-State Electron. **21**, 259 (1978).
[16] R. Bertoncini *et al.*, Solid-State Electron. **32**, 1167 (1989).
[17] B. K. Ridley *et al.*, J. Appl. Phys. **97**, 094503 (2005).
[18] C. Bulutay *et al.*, Appl. Phys. Lett. **77**, 2707 (2000).
[19] C. Bulutay *et al.*, Phys. Rev. B **62**, 15754 (2000).
[20] C. Bulutay *et al.*, J. Cryst. Growth **230**, 462 (2001).
[21] C. Bulutay *et al.*, Physica B **314**, 63 (2002).
[22] C. Bulutay *et al.*, Phys. Rev. B **68**, 115205 (2003).
[23] B. K. Ridley, J. Phys. C **16**, 3373 (1983).