Sektion Physik der Humboldt-Universität zu Berlin[1])

Calculation of Interband Tunneling in Inhomogeneous Fields

By

A. Schenk, M. Stahl, and H.-J. Wünsche

A new approach to the interband tunneling in semiconductor junctions is developed. It generalizes the traditional WKB and EMA theories, which fail in the limits of strong fields and inhomogeneous fields, respectively. The given expression for the energetic tunneling rate remains valid in both limits. In the strong field case it tends to the parabolic band EMA result, whereas for low fields with medium inhomogeneity any band structure model is possible. A Kane model is used to calculate tunneling and total $I-U$-characteristics for narrow gap diodes. The breakdown behaviour yields information about the electrically active doping profile.

Es wird ein neuer Zugang zum Interband-Tunneln in pn-Übergängen von Halbleitern entwickelt. Er verallgemeinert die herkömmlichen WKB- und EMA-Theorien, die in den Grenzfällen des starken bzw. inhomogenen elektrischen Feldes versagen. Der angegebene Ausdruck für die energetische Tunnel-Generationsrate bleibt in beiden Grenzfällen gültig. Im Falle des starken Feldes geht er in das EMA-Resultat für parabolische Bandstruktur über, während für schwache Felder mit mittlerer Inhomogenität ein beliebiges Bandstrukturmodell verwendet werden kann. Ein Kane-Modell wird zur Berechnung von Tunnel- und Gesamtstrom–Spannungscharakteristiken für Dioden aus schmallückigem Material verwendet. Das Durchbruchverhalten liefert Informationen über das elektrisch aktive Dotierungsprofil.

1. Introduction

The investigation of $I-U$ characteristics of narrow gap diodes shows a soft reverse breakdown occurring at relatively low bias and the "breakdown voltage" progressively decreasing as temperature is decreased [1, 7], which is attributed to interband tunneling. In narrow gap material, due to the small depletion layer, the electric field is essentially inhomogeneous and consequently the use of standard tunneling formulas with constant or "averaged" field strengths should fail. In recent years, much work has been devoted to overcome this shortage [2 to 7]. In most cases the WKB tunneling probability was calculated using a certain form of the potential barrier, e.g., a parabolic one-sided abrupt junction potential [2, 3, 4, 7]. Such a model originates from the highly asymmetrical doping profiles of n⁺p-HgCdTe photovoltaic detectors, where the reduction of the performance at reverse bias operation can be explained by leakage currents from interband or deep level-to-band tunneling processes [6]. But there is evidence, that Hg diffusion smooths out the electrical active profile [8], making an abrupt junction model questionable. Therefore, an expression for the tunneling current density valid for a wide class of junction potentials is desirable.

In this paper we describe a method for the analytical treatment of interband tunneling in semiconductor junctions which allows to take into account the inhomogeneous field of the space-charge region in a more general form. The only restriction to the

---

[1]) PSF 1297, Invalidenstr. 110, DDR-1040 Berlin, GDR.

potential is a demand on its "medium" variation, such that it can be linearized over a portion Δl_t of the tunneling length l_t, defined by

$$\frac{\Delta l_t}{l_t} = 2\left(\frac{\hbar\theta}{E_g}\right)^{3/4} \tag{1}$$

(E_g gap energy, $\hbar\theta$ electrooptical energy, see Section 2). The condition $E_g \gg \hbar\theta$ governs any WKB approach to the tunneling problem and can only be avoided in EMA calculations with constant electric field. Although we choose an EMA framework, we feel our model — field strength nearly constant over Δl_t defined by (1) — interpolating between both the WKB and the EMA approximations, conserving the advantages of both methods. It should be applicable to all conventional p–n junctions. We denote it by "model of medium inhomogeneity".

In Section 2 a short review of previous results will be given. We describe the model in detail in Section 3 and give the result in form of an energetic tunneling generation rate. In Section 4 we compare the tunneling I–U characteristics with those from previous models and present total I–U characteristics of a model diode.

2. The Model of Weakly Inhomogeneous Fields

The interband tunneling is illustrated schematically in Fig. 1. As a result of the isoenergetic transition of an electron from the valence band to the conduction band an electron and a hole are generated at different places x_c and x_v (classical turning points). Both charge carriers have covered the tunneling distance

$$l_t = |x_c - x_v| . \tag{1a}$$

There are two major methods of evaluating the corresponding tunneling probability:

1. The barrier penetration probability is calculated using the WKB approximation, assuming a certain shape for the barrier.
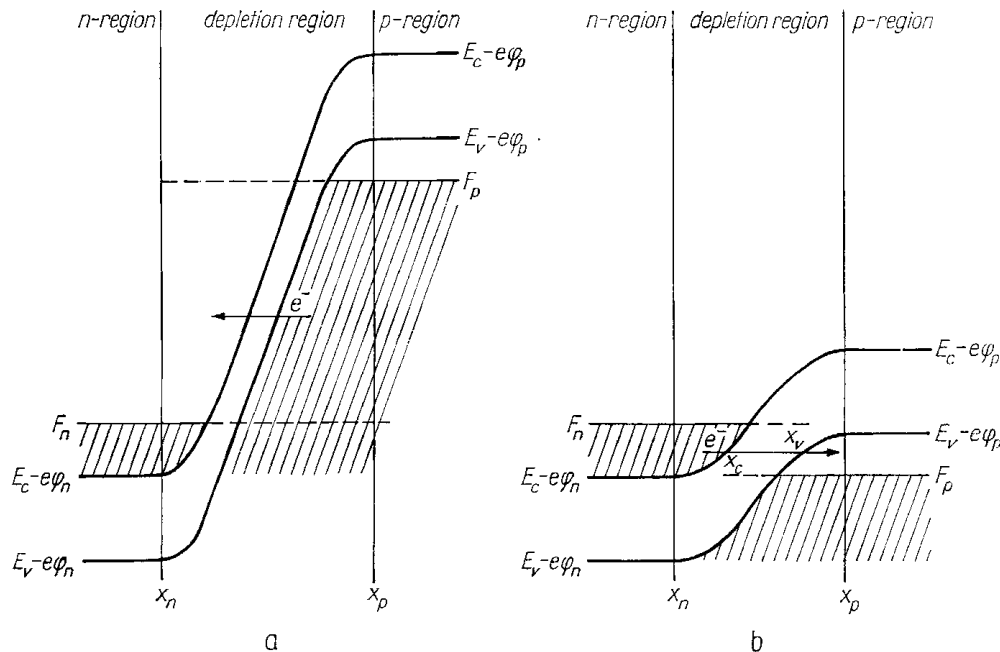

Fig. 1. Schematical illustration of interband tunneling for a) reverse and b) forward bias

2. The transition probability between valence and conduction band states is calculated approximating the interband matrix element. The most usual way here is the effective mass approximation (EMA).

Since our calculation is based on the EMA, we should comment here only upon the latter method.

Up to now, EMA is restricted to the case of weakly inhomogeneous fields [12], i.e., the field strength $F(x)$ changes only little over l_t

$$|F(x_c) - F(x_v)| \ll F(x_{c,v}) .$$ (2)

In this case, the electron and hole tunneling generation rates $G_n(x)$ and $G_p(x)$ will also be nearly the same and can approximated by their values in a homogeneous field of the local strength $F(x)$,

$$G_n(x) = G_p(x) = G_{hom}(F(x)) ,$$ (3)

with [10, 12]

$$G_{hom}(F) = \frac{(e \, |F|)^3}{8\hbar^2 E_g \theta} \left[Ai^2 \left(\frac{E_g}{\hbar\theta} \right) - \frac{E_g}{\hbar\theta} Ai'^2 \left(\frac{E_g}{\hbar\theta} \right) \right].$$ (4)

Here,

$$\theta = \left(\frac{e^2 F^2}{2\mu\hbar} \right)^{1/3} \quad \text{and} \quad \mu = \frac{m_c m_v}{m_c + m_v}$$ (5)

are the electrooptical energy and the reduced effective mass, respectively. The tunneling current density is obtained integrating the generation rate over the space-charge region (SCR),

$$j_t = -e \int_{SCR} G_{hom}(F) \, dx .$$ (6)

Equations (4) to (6) will be called the local field approximation. At this point it turns out that at least one numerical integration will be left in evaluating the tunneling current.

For large arguments $E_g \gg \hbar\theta$ (WKB limitation) the asymptotic forms of the Airy functions can be used [13] and instead of (4) one gets

$$G_{hom}(F) = \frac{(e \, |F|)^3}{64\pi\hbar E_g^2} \exp \left\{ -\frac{F_c}{|F|} \right\}$$ (7)

with the critical field strength

$$F_c = \frac{4 \sqrt{2\mu E_g^3}}{3e\hbar} .$$ (8)

Equation (7), however, is restricted to the case $F \ll F_c$ ($\hbar\theta \ll E_g$). In $Hg_{0.8}Cd_{0.2}Te$, e.g., we have $E_g \approx 0.1$ eV and $m_{c,v} \approx 0.007m_0$. The critical field strength becomes $F_c \approx 1.3 \times 10^5$ V/cm and the critical current density is about 10^6 A/cm². In highly doped, abrupt junctions the condition $F \ll F_c$ does not hold in the whole space-charge layer and hence, neither (7) nor any other WKB method are applicable.

Let us illustrate now the limitations of the local field approximation. In deriving (6) it has been assumed that a joint tunneling rate exists for both electrons and holes up to the boundaries of the space-charge layer. However, at the boundaries x_p and x_n there is a region with the extent of a tunneling length l_t, where only holes but no electrons are generated and vice versa, only electrons but no holes, respectively. It is clear from Fig. 1, that these regions will overlap and, consequently, cover the whole

space-charge layer, if $-eU_{ext} \leqq 2E_g - eU_{bi}$. Therefore, a sensible definition of a joint tunneling generation rate is possible only for

$$e(U_{bi} - U_{ext}) \geqq 2E_g \tag{9}$$

(U_{ext} negative for reverse bias). This limitation will be overcome by the model of medium inhomogeneity to be presented in the next section.

3. The Model of Medium Inhomogeneity

Our calculation is based on a two-band model. In the case of narrow gap materials only the light holes will contribute to the transition rate because of the large effective mass difference between heavy and light holes of about two orders of magnitude. Nevertheless, the complexity of the valence band structure at the Γ-point could give some modifications of the hole tunneling mass to be used.

We choose parabolic bands to find an analytical interpolation between the quasi-classical forms of the wave functions. Whereas these asymptotic solutions are not restricted to a particular dispersion relation $E(k)$, a closed solution in the neighbourhood of the classical turning points is only possible in the parabolic band approximation. Reaching the WKB limit the tunneling current is determined by the correct wave functions including the nonparabolic band structure. Since we are not able to find these solutions analytically, in the case of large field strengths our expression will tend to the parabolic band EMA results.

The effect of relaxing the parabolic band approximation will be discussed in the end of this section.

3.1 Wave functions

The standard envelope method, applied to the Schrödinger equation

$$[H_0 - e\varphi(x)] \, \Phi(r) = E\Phi(r) \,, \tag{10}$$

where H_0 represents the crystal Hamiltonian, and $-e\varphi(x)$ is the space-charge potential, gives

$$\Phi^\mu_{E,k_\perp}(r) = \sqrt{\frac{2}{\Omega}} \; e^{ik_\perp r_\perp} \, u_\mu(r) \, f^\mu_{E,k_\perp}(x) \tag{11}$$

(Ω crystal volume, μ band index). The envelope functions $f^\mu_{E,k_\perp}(x)$ are subject to the one-dimensional Schrödinger equations

$$\left\{ \frac{d^2}{dx^2} - [\varkappa^\mu_{E,k_\perp}(x)]^2 \right\} f^\mu_{E,k_\perp}(x) = 0 \,. \tag{12}$$

In (12) the quasi-classical momenta $\varkappa$ are determined from the relation

$$[\varkappa^\mu_{E,k_\perp}(x)]^2 = k^2_\perp \pm \frac{2m_\mu}{\hbar^2} \left(E_\mu - E - e\varphi(x) \right) \,. \tag{13}$$

Here and in the following the upper sign refers to the conduction band ($\mu = c$) and the lower to the light hole band ($\mu = v$). The other symbols are: E_μ band edge energies for zero potential φ, m_μ band edge effective masses, $k_\perp$ wave vector parallel to the junction plane.

Obviously, the zeros of $\varkappa^\mu_{E,k_\perp}(x)$ are the classical turning points, denoted by $x^\mu_{E,k_\perp}$. In general, there is no closed solution of (12). In order to find an approximate form of the envelope function, we neglect the potential drop outside the space-charge region,

which is possible for the current densities of interest. In this approximation, we can immediately write down the solutions outside the depletion layer,

$$f^{\mu}_{E,\mathbf{k}_{\perp}}(x) = \sin\left(k_{1\mu}x + \delta_{\mu}\right) \tag{14}$$

with the constant wave numbers

$$k_{1c}(E, \mathbf{k}_{\perp}) = \sqrt{\frac{2m_c}{\hbar^2}\left(E - E_c(\mathbf{k}_{\perp}) + e\varphi_n\right)}, \tag{15a}$$

$$k_{1v}(E, \mathbf{k}_{\perp}) = \sqrt{\frac{2m_v}{\hbar^2}\left(E_v(\mathbf{k}_{\perp}) - E - e\varphi_p\right)}, \tag{15b}$$

where $\varphi_{n,p}$ denote the boundary values of the electrostatic potential. Far away from classical turning points the WKB approximation holds [14],

$$f^{\mu}_{E,\mathbf{k}_{\perp}}(x) = \sqrt{\frac{k_{1\mu}(E, \mathbf{k}_{\perp})}{|\varkappa^{\mu}_{E,\mathbf{k}_{\perp}}(x)|}} \begin{cases} \sin\left(|S_{E,\mathbf{k}_{\perp}}(x)| + \dfrac{\pi}{4}\right), & \tag{16a} \\[2ex] \dfrac{1}{2}\exp\left(-|S_{E,\mathbf{k}_{\perp}}(x)|\right) & \tag{16b} \end{cases}$$

with the action (in units of $\hbar$)

$$S^{\mu}_{E,\mathbf{k}_{\perp}}(x) = \int_{x^{\mu}_{E,\mathbf{k}_{\perp}}}^{x} \mathrm{d}x'\, \varkappa^{\mu}_{E,\mathbf{k}_{\perp}}(x'). \tag{17}$$

The prefactor of (16) has been determined in matching the WKB solution (16a) for the classically allowed region to the plane wave solution (14). The decaying function (16b) holds for the classically forbidden region.

Now we develop the potential in the vicinity of the classical turning points

$$\varphi(x) = \varphi(x^{\mu}_{E,\mathbf{k}_{\perp}}) + F(x^{\mu}_{E,\mathbf{k}_{\perp}})\left(x - x^{\mu}_{E,\mathbf{k}_{\perp}}\right), \tag{18}$$

and insert (18) into the differential equation (12), getting an Airy equation,

$$\left[\frac{\mathrm{d}^2}{\mathrm{d}x^2} - q^3(x - x^{\mu}_{E,\mathbf{k}_{\perp}})\right] f(x^{\mu}_{E,\mathbf{k}_{\perp}}) = 0, \tag{19}$$

with

$$q = \mp e\left(\frac{2m_{\mu}F(x^{\mu}_{E,\mathbf{k}_{\perp}})}{e^2\hbar^2}\right)^{1/3}. \tag{20}$$

Taking into account only the first fundamental solution of (19), the envelope function near the classical turning points reads

$$f^{\mu}_{E,\mathbf{k}_{\perp}}(x) = \sqrt{\pi\,\frac{k_{1\mu}(E, \mathbf{k}_{\perp})}{|q|}}\,\mathrm{Ai}\left[q(x - x^{\mu}_{E,\mathbf{k}_{\perp}})\right]. \tag{21}$$

In order to use (21) as an interpolation between the quasiclassical functions (16), one has to assume that 1. the potential can be approximately linearized within just the branch where the WKB approach fails, and 2. the contribution of the second solution Bi $\left(q(x - x^{\mu}_{E,\mathbf{k}_{\perp}})\right)$ of (19) can be neglected. We investigated the Schrödinger equation (12) numerically for various potential shapes and found that both assumptions are justified as long as the condition $E_g \gg \hbar\theta$ is fulfilled.

The prefactor of (21) has been determined to ensure that the solution turns into its asymptotic form (16) in the case of a constant electric field, where (19) is not restricted to the neighbourhood of the classical turning points.

The basic idea of this paper is to unite the piecewise valid solutions (14), (16), and (21) in the envelope function

$$f^{\mu}_{E,\mathbf{k}_{\perp}}(x) = \sqrt{\pi \left| \frac{k_{1\mu}(E,\mathbf{k}_{\perp})}{\varkappa^{\mu}_{E,\mathbf{k}_{\perp}}(x)} \right|} \left[\frac{3}{2} S^{\mu}_{E,\mathbf{k}_{\perp}}(x) \right]^{1/6} \mathrm{Ai}\left(\left[\frac{3}{2} S^{\mu}_{E,\mathbf{k}_{\perp}}(x) \right]^{2/3} \right). \tag{22}$$

Using the properties of the Airy function one can easily verify, that (22) is smooth everywhere and will give the correct asymptotic forms (16) as well as (21) in the vicinity of the classical turning points. In evaluating the normalization constant according to

$$\langle \Phi^{\mu}_{E,\mathbf{k}_{\perp}}(\mathbf{r}) \mid \Phi^{\mu'}_{E',\mathbf{k}'_{\perp}}(\mathbf{r}) \rangle = \delta_{\mu\mu'} \delta_{\mathbf{k}_{\perp},\mathbf{k}'_{\perp}} \delta_{k_{1\mu},k'_{1\mu}}, \tag{23}$$

we took into consideration that the contribution of the space-charge region to the normalization integral remains finite and therefore negligible in the limit $\Omega \to \infty$.

3.2 Transition probability

The interband coupling due to the electrostatic potential of the space-charge layer is treated as perturbation, inducing tunneling transitions. Applying the usual "golden rule", the probability for a transition from a valence band state with E, $\mathbf{k}_{\perp}$ to a conduction band state with E', $\mathbf{k}'_{\perp}$ is given by

$$w(\mathrm{v}, E, \mathbf{k}_{\perp}; \mathrm{c}, E', \mathbf{k}'_{\perp}) = \frac{2\pi}{\hbar} |M_{\mathrm{vc}}(E, \mathbf{k}_{\perp}; E', \mathbf{k}'_{\perp})|^2 \delta(E - E') \tag{24}$$

with the matrix element

$$M_{\mathrm{vc}}(E, \mathbf{k}_{\perp}; E', \mathbf{k}'_{\perp}) = \langle \Phi^{\mathrm{v}}_{E,\mathbf{k}_{\perp}}(\mathbf{r}) \mid e\varphi(x) \mid \Phi^{\mathrm{c}}_{E',\mathbf{k}'_{\perp}}(\mathbf{r}) \rangle. \tag{25}$$

We use the EMA states (11), integrate over the crystal and neglect, as is usual in EMA, the $\mathbf{k}$-dependence of the Bloch matrix element,

$$x_{\mathrm{cv}}(\mathbf{k}) = \frac{i}{\Omega_0} \int_{\Omega_0} \mathrm{d}^3 r \, u^*_{\mathrm{v}\mathbf{k}}(\mathbf{r}) \frac{\partial}{\partial k_x} u_{\mathrm{c}\mathbf{k}}(\mathbf{r}). \tag{26}$$

Since the potential drop across a Wigner-Seitz cell (volume Ω_0) is tiny, we find for the transition matrix element

$$M_{\mathrm{vc}}(E, \mathbf{k}_{\perp}; E', \mathbf{k}'_{\perp}) = \frac{2e}{\Omega^{1/3}} x_{\mathrm{cv}}(0) \, \delta_{\mathbf{k}_{\perp}\mathbf{k}'_{\perp}} \int_{\Omega^{1/3}} \mathrm{d}x \, f^{\mathrm{v}*}_{E,\mathbf{k}_{\perp}}(x) \, F(x) \, f^{\mathrm{c}}_{E,\mathbf{k}_{\perp}}(x). \tag{27}$$

The overlap of the envelope functions within the gap results in a sharp peak at the position x_0, which one can determine from

$$|f^{\mathrm{v}*}_{E,\mathbf{k}_{\perp}}(x_0) f^{\mathrm{c}}_{E,\mathbf{k}_{\perp}}(x_0)| = \max |f^{\mathrm{v}*}_{E,\mathbf{k}_{\perp}}(x) f^{\mathrm{c}}_{E,\mathbf{k}_{\perp}}(x)|. \tag{28}$$

With increasing distance from x_0 the integrand in (27) will rapidly fall off, because one of the envelopes is strongly damped. Therefore, the main contribution to the matrix element (27) stems from the immediate neighbourhood of the point x_0 with an extent denoted by Δl_{t}. The outlined behaviour of the integrand in (27) is not essentially changed by the factor $F(x)$, the variation of which is comparatively much smaller.

Assuming equal effective masses m_{c} and m_{v} one can estimate the halfwidth of the resulting peak produced by the overlap of both envelope functions within the gap,

$$\Delta l_{\mathrm{t}} = 2l_{\mathrm{t}} \left(\frac{eF\hbar}{E_{\mathrm{g}} \sqrt{2\mu E_{\mathrm{g}}}} \right)^{1/2} = 2l_{\mathrm{t}} \left(\frac{\hbar\theta}{E_{\mathrm{g}}} \right)^{3/4}. \tag{29}$$
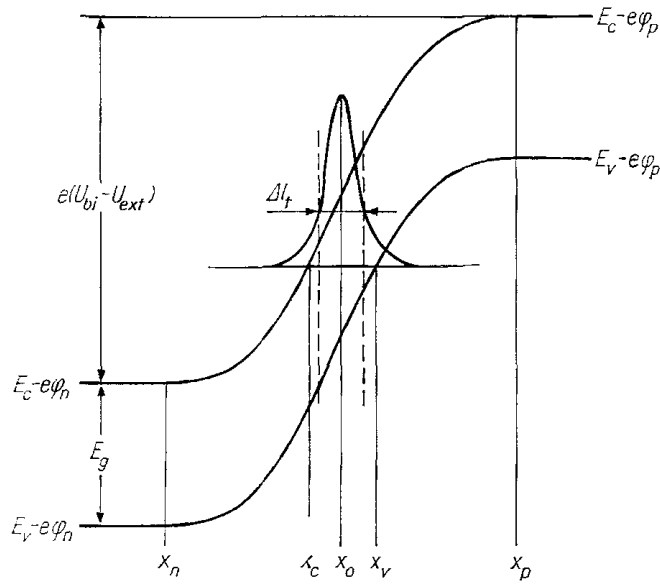
Fig. 2. Illustration of the overlap peak of the envelopes within the space-charge layer at maximum field (worst case)

The situation is illustrated in Fig. 2 for a reverse bias beyond the breakdown voltage using typical parameters of a HgCdTe photodiode in the step junction model and looking at the maximum field strength. — The feature of the electric field to alter strongly in regions, where its absolute value is small, reduces Δl_t there according to (29). For that reason

$$\left| 1 - \frac{F(x_0 + \Delta l_t)}{F(x_0)} \right| \ll 1 \qquad (30)$$

can be assumed, and $F(x)$ may be replaced by its value at x_0 in (27). Therefore, we are left with the problem to evaluate the overlap integral

$$\Gamma_{\mathrm{vc}}(E, \boldsymbol{k}_\perp) = \int_{\Omega^{1/3}} \mathrm{d}x\, f_{E,\boldsymbol{k}_\perp}^{\mathrm{v}*}(x)\, f_{E,\boldsymbol{k}_\perp}^{\mathrm{c}}(x) \qquad (31)$$

using the envelope wave functions (22). We again take advantage of (30), i.e. the model of medium inhomogeneity, and treat the slowly varying factors in f as constants with $x = x_0$. Furthermore, the arguments of the Airy functions are developed up to first order in the deviation of x from the point of maximum overlap x_0. The remaining integral over a product of Airy functions with linear arguments can be solved [15], yielding

$$\Gamma_{\mathrm{vc}}(E, \boldsymbol{k}_\perp) = \frac{3\pi \big(k_{1\mathrm{c}}(E, \boldsymbol{k}_\perp)\, k_{1\mathrm{v}}(E, \boldsymbol{k}_\perp)\, |S_{E,\boldsymbol{k}_\perp}^{\mathrm{v}}(x_0)\, S_{E,\boldsymbol{k}_\perp}^{\mathrm{c}}(x_0)|\big)^{1/2}}{2\varkappa_{E,\boldsymbol{k}_\perp}^{\mathrm{v}}(x_0)\, \varkappa_{E,\boldsymbol{k}_\perp}^{\mathrm{c}}(x_0)\, [\frac{3}{2}\, \Delta S_{E,\boldsymbol{k}_\perp}(x_0)]^{1/3}} \times$$
$$\times\, \mathrm{Ai}\left(\left[\frac{3}{2}\, \Delta S_{E,\boldsymbol{k}_\perp}(x_0)\right]^{2/3}\right) \qquad (32)$$

with

$$\Delta S_{E,\boldsymbol{k}_\perp}(x_0) = |S_{E,\boldsymbol{k}_\perp}^{\mathrm{c}}(x_0) - S_{E,\boldsymbol{k}_\perp}^{\mathrm{v}}(x_0)| \,.$$

In order to find the point x_0 as a function of energy E and transverse momentum part $\boldsymbol{k}_\perp$, we assume x_0 to be such that both envelopes can be described by their asymptotic forms (16b) there. This assumption is fulfilled, if $m_c \approx m_v$, and provided that again $E_g \gg \hbar\theta$ holds. The point x_0 then is determined from the condition that the sum of

the absolute values of the actions (17) becomes minimum. In this way an implicit relation $x_0 = x_0(E, \mathbf{k}_\perp)$ is obtained,

$$(m_\mathrm{c} + m_\mathrm{v})\, \varphi\big(x_0(E, \mathbf{k}_\perp)\big) = m_\mathrm{v}\varphi(x^\mathrm{v}_{E,\mathbf{k}_\perp}) + m_\mathrm{c}\varphi(x^\mathrm{c}_{E,\mathbf{k}_\perp}) \tag{33}$$

or

$$[\varkappa^\mathrm{v}_{E,\mathbf{k}_\perp}(x_0)]^2 = [\varkappa^\mathrm{c}_{E,\mathbf{k}_\perp}(x_0)]^2 = \mathbf{k}^2_\perp + \frac{2\mu E_\mathrm{g}}{\hbar^2}. \tag{34}$$

Equation (33) removes the uncertainty of the correct local field strength to be used in the final integration over the space-charge region, which is always present, if the tunneling rate is calculated from (4).

With (24), (27), and (32) we find the transition probability to be

$$w(\mathrm{v}, E, \mathbf{k}_\perp; \mathrm{c}, E', \mathbf{k}'_\perp) = \frac{3\pi^3\hbar^3}{\Omega^{2/3}}\, |eF(x_0)\, x_\mathrm{cv}(0)|^2\, \frac{k_{1\mathrm{v}}(E, \mathbf{k}_\perp)\, k_{1\mathrm{c}}(E, \mathbf{k}_\perp)}{\mu^2 \left(E_\mathrm{g} + \dfrac{\hbar^2 k^2_\perp}{2\mu}\right)^2}\, S^\mathrm{red}_{E,\mathbf{k}_\perp}(x_0) \times$$

$$\times \left[\frac{3}{2}\, \Delta S_{E,\mathbf{k}_\perp}(x_0)\right]^{1/3} \mathrm{Ai}^2\left(\left[\frac{3}{2}\, \Delta S_{E,\mathbf{k}_\perp}(x_0)\right]^{2/3}\right) \times$$

$$\times\, \delta_{\mathbf{k}_\perp, \mathbf{k}'_\perp}\, \delta(E - E')\,, \tag{35}$$

where S^red denotes the reduced action

$$S^\mathrm{red}_{E,\mathbf{k}_\perp}(x_0) = \frac{|S^\mathrm{c}_{E,\mathbf{k}_\perp}(x_0)\, S^\mathrm{v}_{E,\mathbf{k}_\perp}(x_0)|}{\Delta S_{E,\mathbf{k}_\perp}(x_0)}\,. \tag{36}$$

3.3 Energetic tunneling rate

Expression (35) defines the probability of an elementary tunneling process connected with the generation of an electron–hole pair. The tunneling current density is given by the total number of generated electrons per cm² and per second,

$$j_\mathrm{t} = -\frac{e}{\Omega^{2/3}}\, \frac{\mathrm{d}N}{\mathrm{d}t}\,. \tag{37}$$

The temporal change of the total number N is obtained integrating the transition probability (35) over all initial and final states. The summation includes the Fermi-Dirac functions as well as the one-dimensional densities of states,

$$\varrho_\mu(E, \mathbf{k}_\perp) = \int \mathrm{d}k_{1\mu}\, \delta\big(E - E(\mathbf{k}_\perp, k_{1\mu})\big)\,, \tag{38}$$

which follow from (15) in the form

$$\varrho_\mathrm{c}(E, \mathbf{k}_\perp) = \frac{m_\mathrm{c}}{\hbar^2 k_{1\mathrm{c}}(E, \mathbf{k}_\perp)}\, \theta[E - E_\mathrm{c}(\mathbf{k}_\perp) + e\varphi_\mathrm{n}]\,, \tag{38a}$$

$$\varrho_\mathrm{v}(E, \mathbf{k}_\perp) = \frac{m_\mathrm{v}}{\hbar^2 k_{1\mathrm{v}}(E, \mathbf{k}_\perp)}\, \theta[E_\mathrm{v}(\mathbf{k}_\perp) - E - e\varphi_\mathrm{p}]\,. \tag{38b}$$

Utilizing the δ-functions in (35) we obtain

$$\frac{\mathrm{d}N}{\mathrm{d}t} = \frac{\Omega^{2/3}}{\pi^3\hbar}\, |eF(x_0)\, x_\mathrm{vc}(0)|^2 \int \mathrm{d}E \int \mathrm{d}^2k_\perp [f_\mathrm{v}(E) - f_\mathrm{c}(E)] \times$$

$$\times\, \varrho_\mathrm{v}(E, \mathbf{k}_\perp)\, \varrho_\mathrm{c}(E, \mathbf{k}_\perp)\, |\Gamma_\mathrm{vc}(E, \mathbf{k}_\perp)|^2\,. \tag{39}$$

Using this, (37) can be written as

$$j_t = -e \int G(E) \left[f_v(E) - f_c(E) \right] \mathrm{d}E \tag{40}$$

with

$$G(E) = \frac{|eF(x_0)\, x_{cv}(0)|^2}{\pi^3 \hbar} \int \mathrm{d}^2 k_\perp \, \varrho_v(E, k_\perp) \, \varrho_c(E, k_\perp) \, |\Gamma_{vc}(E, k_\perp)|^2 \ . \tag{41}$$

The latter quantity is the number of electrons tunneling per energy interval and per time unit in case of $f_v = 1$ and $f_c = 0$. We call it energetic tunneling rate. An analytical result for $G(E)$ is possible, if the spherical $k_\perp$-integration is carried out approximately. We may proceed in this way, because only small $|k_\perp|$ contribute to the integral in (41) due to the fact that the tunneling probability will rapidly decrease with increasing tunneling length l_t. Therefore, we set $|k_\perp| = 0$ in all slowly varying factors, develop the arguments of the Airy functions, and extend the integration to infinity. The resulting integral over Ai^2 can be solved [15]. Furthermore, the expressions for the actions simplify to

$$S_\mu(x_0) = S_{E,0}^\mu(x_0) = \frac{\sqrt{2m_\mu}}{\hbar} \int\limits_{x_\mu}^{x_0} \mathrm{d}x' \sqrt{e\, |\varphi(x') - \varphi(x_\mu)|} \ , \tag{42}$$

where x_μ denotes the classical turning points $x_\mu = x_{E,0}^\mu$ now, which follow from

$$E = E_\mu - e\varphi(x_\mu) \ . \tag{43}$$

The point x_0, where the maximum overlap of the envelope wave functions takes place, is the solution of the reduced implicit relation

$$(m_c + m_v)\, \varphi(x_0) = m_v \varphi(x_v) + m_c \varphi(x_c) \ . \tag{44}$$

Finally we end up in the following expression for the energetic tunneling rate:

$$G(E) = \frac{3\, |eF(x_0)|^3\, m_v m_c S_{red}(x_0)}{128 \pi \mu^2 E_g^3 \sqrt{2\mu E_g}} \left[f_v(E) - f_c(E) \right] \theta(E_v - E - e\varphi_p) \times$$

$$\times \ \theta(E - E_c + e\varphi_n)\, M(x_0) \tag{45}$$

with

$$M(x_0) = 8\pi \left[\tfrac{3}{2} \Delta S(x_0)^{2/3} \right] \times$$

$$\times \ \{ \mathrm{Ai}^2 \left([\tfrac{2}{3} \Delta S(x_0)]^{2/3} \right) - [\tfrac{3}{2} \Delta S(x_0)]^{2/3}\, \mathrm{Ai}'^2 \left([\tfrac{3}{2} \Delta S(x_0)]^{2/3} \right) \} \ . \tag{46}$$

In deriving (45) we made use of

$$|x_{cv}(0)|^2 = \frac{\hbar^2}{4\mu E_g}, \tag{47}$$

attainable from a Kane two-band model [16].

In evaluating the tunneling current density one has to pass through the following steps: The point x_0 and the turning points $x_{c,v}$ have to be determined for a given potential $\varphi(x)$, which satisfies the medium inhomogeneity condition (30), from (43) and (44), respectively. After that the actions (and their difference) are computable using (42), which has to be done numerically in general, unless simple shapes like linear or parabolic ones are considered. The final integration over E yields the tunneling current density according to (40). The step functions in (45) confine the energy interval to those energy levels for which tunneling is possible, and the difference of the Fermi functions attends to the correct current direction. For zero external voltage the total

tunneling current vanishes. $M(x_0)$, defined by (46), is the predominant factor in the tunneling rate. Provided that $E_g \gg \hbar\theta$, it turns into the well-known WKB exponential,

$$M(x_0) \to \exp\left[-2\,\Delta S(x_0)\right].\tag{48}$$

For a homogeneous field, (6) together with (7) is reproduced.

In the case of narrow gap diodes the correct dispersion relation is actually the two-band $k \times p$ approximation valid near $\Gamma = 0$ in the Brillouin zone. The generalization of our interpolation procedure to a Kane model is straightforward, if $\hbar\theta \ll E_g$ holds for the field strength. Instead of the quasi-classical momenta (13) we have

$$[\varkappa^\mu_{E,\,\mathbf{k}_\perp}(x)]^2 = \frac{m_\mu}{\hbar^2}\left[\frac{2}{E_g}\left(E - E_0 + e\varphi(x)\right)^2 - \frac{E_g}{2} - \frac{\hbar^2\mathbf{k}_\perp^2}{m_\mu}\right],\tag{49}$$

where E_0 denotes the midgap energy level. Using (49) in evaluating the actions yields the tunneling current for the Kane model, as long as the field strength does not exceed the WKB limit, since the neighbourhood of the turning points is unimportant in this case. Consequently Kane's formula [16] is reproduced with (48), assuming a constant field strength, up to a small difference in the pre-exponential factor, which has its origin in our EMA treatment.

The solution for the envelope wave functions in the neighbourhood of the classical turning points, where the WKB approximation breaks down, is found in form of its integral representation only, linearizing the potential there. To continue the formalism one had to find the inverted function of the action and to analyse all integrals numerically. For the sake of this difficulties we use the Airy-like interpolation functions (22) again, which have the correct asymptotic behaviour also in the case of a Kane model. When the WKB approximation breaks down, i.e. when $E_g < \hbar\theta$, according to (1), Δl_t becomes larger than the tunneling distance l_t. But at the same time the potential is linearized by the reverse bias over a broader range of the space-charge layer, covering Δl_t. Thus, for large reverse bias our result tends to the well-known EMA expression [9 to 11] involving the case $\hbar\theta > E_g$, but in the parabolic band approximation.

4. Numerical Example

We used the common abrupt junction model to demonstrate the change in the tunneling characteristics, if the different levels of description are applied, which have been outlined in this paper. In Fig. 3 the local field approximation (6) and the model of medium inhomogeneity ((40), (45)) are compared using the doping concentrations N_D and N_A as parameters. Additionally, we have regarded the model of an averaged constant field $|F| = |U_{bi} - U_{ext}|/W$, where U_{bi} denotes the built-in voltage and W the space-charge region width. In this case the tunneling current density becomes simply

$$j_t = j_0(F)\exp\left(-F_c/|F|\right)\tag{50}$$

with

$$j_0(F) = -\frac{e^4 W\,|F|^3}{64\pi\hbar E_g^2}.\tag{51}$$

The use of (50) ignores the spatial dependence of the electric field in the whole space-charge region. This approximation is suitable only for p⁺–i–n⁺ structures, where approximately the whole reverse bias drop is across the intrinsic layer i.

Fig. 3 shows that the simple model of an constant field strongly underestimates the tunneling current in all cases. For $N_D = N_A = 10^{16}$ cm^{-3} the tunneling branch is not distinguished from the voltage axis.
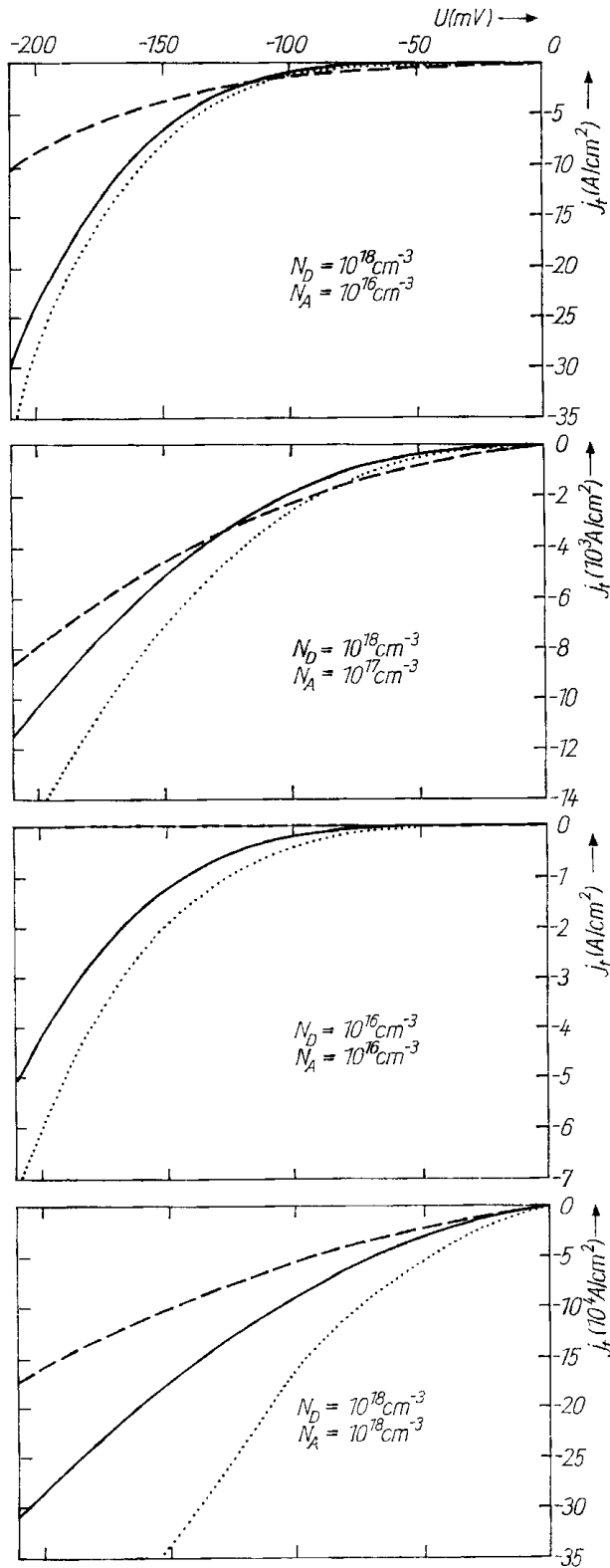
The comparison of the medium inhomogeneity model with the local field approximation shows the latter to give the best results at higher reverse voltages, when the influence of the field inhomogeneity decreases. The deviation then is primarily due to the incorrect integration interval and the asymptotic representation of the Airy functions. On the other hand, the difference between both models increases for low reverse voltages, when the inhomogeneity of the electric field becomes increasingly important. Then the local field approximation overestimates the tunneling current remarkably. At zero voltage the differential resistance can differ by orders of magnitude.

Fig. 3. Comparison of the tunneling characteristics calculated with three different models: Averaged constant field (— — —), local field approximation (········), and model of medium inhomogeneity (————). The abrupt junction model and the following parameters have been used: $T = 80$ K, E_g ($x = 0.2$) = 0.084 eV, $m_c = m_v = 0.007 m_0$, $\varepsilon_s = 17.6$
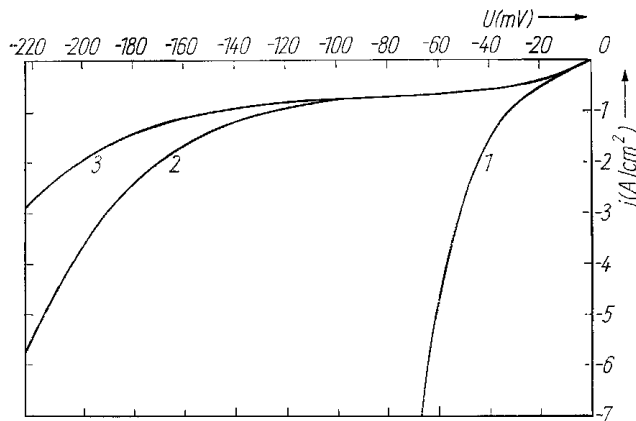
Fig. 4. Current–voltage characteristics for three different gradients of the linear doping profile. (1) ∞ (abrupt), (2) 2.0×10^{17} cm^{-3}/µm, (3) $1.54 \times \times 10^{17}$ cm^{-3}/µm. Parameters: $N_A = 3.2 \times 10^{16}$ cm^{-3}, $N_D = 10^{18}$ cm^{-3}, $T = 100$ K, $E_g = 0.1059$ eV, $n_i = = 1.81 \times 10^{14}$ cm^{-3} (intrinsic density), $\tau_{SHR} = 10^{-10}$ s, $m_c = m_v = 0.007 m_0$, $m_{hh} = 0.44 m_0$

Fig. 4 shows total I–U characteristics, calculated with the model of medium in-homogeneity ((40), (45)) for three different doping profiles: abrupt junction (1), linearly graded doping concentrations with 2.0×10^{17} cm^{-3}/µm (2), and $1.54 \times \times 10^{17}$ cm^{-3}/µm (3) gradients. The recombination current comprises Auger, Shockley-Read-Hall (SHR), and surface recombination currents with dominating SHR current ($\tau_{SHR} = 0.1$ ns). It is clearly seen that the breakdown shifts toward higher reverse voltages and becomes flatter, if the doping gradient decreases. This conclusion from the breakdown behaviour to the electrically active doping profile is helpful in characterizing HgCdTe photodiodes.

Acknowledgement

References

[1] W. W. ANDERSON, Appl. Phys. Letters 41, 1080 (1982).
[2] J. L. MOLL, Physics of Semiconductors, New York, McGraw-Hill Publ. Co., New York 1964 (p. 164).
[3] W. W. ANDERSON, Infrared Phys. 14, 147 (1977).
[4] W. W. ANDERSON, Infrared Phys. 20, 353 (1980).
[5] W. A. BECK and N. E. BYER, IEEE Trans. Electron Devices 31, 292 (1984).
[6] W. W. ANDERSON and H. J. HOFFMAN, J. appl. Phys. 53, 9130 (1982).
[7] Y. NEMIROVSKY and I. BLOOM, Infrared Phys. 27, 143 (1987).
[8] L. O. BUBULAC, J. Crystal Growth 86, 723 (1988).
[9] C. B. DUKE, Tunneling in Solids, Academic Press, Inc., New York 1969 (p. 78).
[10] K. HENNEBERGER and R. ENDERLEIN, phys. stat. sol. (b) 72, 547 (1975).
[11] R. ENDERLEIN and K. PEUKER, phys. stat. sol. (b) 48, 231 (1971).
[12] K.-H. HASLER and H.-J. WÜNSCHE, Wiss. Z. Humboldt-Universität zu Berlin, math.-nat. R. 35, 2 (1986).
[13] M. ABRAMOWITZ and A. STEGUN, Pocketbook of Mathematical Functions, Verlag Harri Deutsch, Thun, Frankfurt (Main) 1984 (p. 164).
[14] L. D. LANDAU and E. M. LIFSCHITZ, Lehrbuch der theoretischen Physik Vol. 3, Akademie-Verlag, Berlin 1965 (p. 47).
[15] D. E. ASPNES, Phys. Rev. 147, 554 (1966).
[16] E. O. KANE, J. Phys. Chem. Solids 23, 173 (1962).