Pergamon

0038-1101(93)E0128-N

Solid-State Electronics Vol. 37, No. 9, pp. 1633–1647, 1994
Copyright © 1994 Elsevier Science Ltd
Printed in Great Britain. All rights reserved
0038-1101/94 $7.00 + 0.00

1D ANALYTICAL MODEL OF THE METAL–SEMICONDUCTOR CONTACT BEYOND THE WKB APPROXIMATION

ANDREAS SCHENK

Swiss Federal Institute of Technology, Integrated Systems Laboratory, Gloriastrasse 35, CH-8092 Zürich, Switzerland

(Received 1 June 1993; in revised form 18 November 1993)

Abstract—A 1D analytical model of the metal–semiconductor contact is developed in the framework of emission theory which exploits an improved description of the transmission probability. The usual WKB approximation is substituted by an interpolation scheme where the eigensolutions of a given potential barrier are expressed by Airy functions and then mapped to Gaussians in order to enable analytical integration. The method is demonstrated for a parabolic barrier where the eigenfunctions are known and direct comparison between the model and an "exact" reference is possible. Contact currents are given in a fully analytical form and agree well with those from the reference model over the whole range of doping concentrations from Schottky to Ohmic. Boundary conditions for device simulation are derived by merging the calculated emission current to the drift-diffusion current at a certain distance from the metal–semiconductor interface. Experimental data of Ti on intermediately doped n-Si are used to compare with the model and to discuss the influence of the most important parameters.

NOTATION

A_c, A_M	effective Richardson constants [$AK^{-2}cm^{-2}$]
Ai, Ai′, Bi, Bi′	Airy functions and their derivatives
b	numerical fit parameter
E	absolute energy [eV]
E_c	energy of the bulk conduction band edge [eV]
$E_{F,M}, E_{F,S}$	Fermi energies of metal and semiconductor, resp. [eV]
E_{max}	energy limit for Gaussian approximation [eV]
E_λ	normalization energy [eV]
$E^+, E_1^-, E_2^-, E_1^d, E_2^d$	energy levels of maximum current flow [eV]
$\mathrm{Erf}(x_1, x_2)$	generalized error function
$\mathrm{erf}(x)$	error function
$\mathrm{erfc}(x)$	complementary error function
f_M, f_S	Fermi distribution functions of metal and semiconductor, resp.
$\hbar$	reduced Planck's constant
$\mathrm{Im}(Z)$	imaginary part of Z
i	imaginary unit
j	current density [$A\,cm^{-2}$]
$j^+, j_1^-, j_2^-, j_1^d, j_2^d, j_3$	partial current densities [$A\,cm^{-2}$]
k	Boltzmann constant
$\mathbf{k}_M, \mathbf{k}_S$	momentum in metal and semiconductor, resp.
m_M, m_c, m_{eff}	effective electron masses in metal and semiconductor, resp.
m_0	electron rest mass
n	electron density [cm^{-3}]
N_c	effective density of states in conduction band [cm^{-3}]
N_D	donor concentration [cm^{-3}]
q	elementary charge
$\mathrm{Re}(Z)$	real part of Z
$S(x)$	action integral
T	absolute temperature [K]
$\mathscr{T}$	transmission probability
t_0	coordinate of the first maximum of the Airy function
U, V, U', V'	parabolic cylinder functions and their derivatives
U_0, V_0, U_0', V_0'	values of U, V, U', V' at the end of the space charge region
$U_{\zeta_B}, V_{\zeta_B}, U_{\zeta_B}', V_{\zeta_B}'$	values of U, V, U', V' at the metal–semiconductor interface
U_{appl}	external voltage [V]
u	normalized external voltage qU_{appl}/kT
$v_{M,x}, v_{S,x}$	x-component of the electron velocity in metal and semiconductor, resp.
x_B	barrier width [cm]
x_T	extension of the physical contact [cm]
$Y(S)$	function of the action $(3S/2)^{2/3}$
$\Gamma(x)$	Gamma function
$\gamma(E)$	equals Y'^{-1}
δ	real number
ϵ	normalized energy E $(=E/E_\lambda)$
ϵ_0, ϵ_s	vacuum and semiconductor permittivity, resp.
ϵ_c	normalized conduction band edge in the bulk $(=E_c/E_\lambda)$
ϵ_{max}	normalized energy E_{max} $(=E_{max}/E_\lambda)$
$\epsilon^+, \epsilon_1^-, \epsilon_2^-, \epsilon_1^d, \epsilon_2^d$	normalized energies $E^+, E_1^-, E_2^-, E_1^d, E_2^d$
$\epsilon_{F,M}$	normalized Fermi energy of metal, $(=E_{F,M}/E_\lambda)$
η	energy normalized by kT $(=E/kT)$
η_c	E_c normalized by kT $(=E_c/kT)$
Θ, Θ_{max}	angle between x-axis and momentum vector in the metal and its limit
$\Theta(x)$	step function
$\vartheta(E), \vartheta_1(E)$	arguments of the error functions
κ_M, κ_S	normalized x-components of the momenta
λ	normalization length [cm]

λ_D	Debye length [cm]
μ_n	electron mobility [cm^2/V s]
ν, ν_{max}	equals $\sin(\Theta)$, $\sin(\Theta_{max})$
ξ_B	normalized barrier width x_B $(=x_B/\lambda)$
Φ_B	barrier height [eV]
Φ_N	energetic distance between E_c and $E_{F,S}$ [eV]
Φ_T	energy level of negligible spectral current density [eV]
φ_B	barrier height normalized by kT $(=\Phi_B/kT)$
φ_n	electron quasi Fermi potential [V]
Ψ	electrostatic potential [V]
$\psi(x)$	wave function

1. INTRODUCTION

The metal–semiconductor interface is one of the most challenging problems in the field of solid-state theory and device physics. This is because of a variety of physical effects, e.g. the influence of interface states on barrier height[1,2], the effect of interfacial layers (dipole, oxide or contamination)[3–6], inelastic scattering events[7,8], recombination, trapping[9,10] and trap-assisted tunneling[11], vertical and lateral potential fluctuations[12], barrier height fluctuations[13], interface roughness[14], band-state mixing (making the tunneling mass an uncertain quantity)[15], hot carrier effects, realistic image forces, and some others which make the theoretical modeling a complicated task. Simplified contact models, e.g. suitable for device simulation, have to neglect most of all these effects.

A step towards a more physical model instead of the commonly used boundary conditions for ideal Ohmic contacts (neutrality, equilibrium) and ideal rectifying (Schottky) contacts (thermionic emission neglecting tunneling) was made recently by Schroeder[16]. He developed the well established formula of thermionic field emission (TFE) to an analytical expression, using a simple WKB transmission probability of the parabolic potential barrier and separating the contact region (emission) from the remaining device, where the current is governed by drift-diffusion (Crowell and Sze[7], Chang and Sze[12]). Supplemented by a proper boundary condition for the Poisson equation such a model is able to describe the transition from Schottky to Ohmic contacts as the doping level increases. A shortage of that treatment is the use of the WKB approximation which breaks down if the de Broglie wavelength becomes comparable to the barrier width (Ohmic contact or strong reverse bias) and/or if the current flows predominantly in the vicinity of the maximum of the barrier (Schottky contact).

It is the aim of the paper to show that the WKB approximation can be substituted by a better approach, which preserves the accuracy of an "exact" reference model over the entire doping range. The "exact" reference model is chosen as follows: parabolic potential barrier (constant doping level in the barrier region, Schottky approximation, no image effect, no interfacial layer, etc.), independence of the transmission probability on transverse momentum, and constant effective mass for all energies above and below the band edge of majority carriers. For such a model the exact transmission coefficient is well known, and derived results may be compared with those of our analytical approach.

In the next section we recapitulate the 1D theory of the metal–semiconductor contact and define the basic quantities. Section 3 outlines the interpolation method for the transmittance which removes the drawbacks of the WKB approximation. Additionally, the transmittance will be written in a form suitable for analytical integration. The analytical model of the contact current is derived in Section 4 and jV-characteristics are compared with those from the "exact" reference model there. In Section 5 Dirichlet boundary conditions for device simulation are calculated. A certain fraction of the barrier is defined as that region where no Fermi level exists and which is shrinked to the contact boundary in numerical simulation. Matching the emission current to the drift-diffusion current at the end of this region yields an explicit expression for the boundary value of the quasi Fermi potential. The theoretical results are compared with data obtained from a Kelvin structure of titanium on intermediately doped n-silicon in Section 6. We discuss the effect of the most important physical parameters there. Conclusions are given in the last section.

2. EMISSION CURRENT THROUGH A PARABOLIC BARRIER

In the following we will discuss an n-type semiconductor with a parabolic conduction band. Figure 1 defines some basic quantities: barrier height Φ_B, barrier width $x_B = x_b(U_{appl})$, the two Fermi levels, and the energetic distance Φ_N between bulk quasi Fermi level $E_{F,S}$ and conduction band edge E_c in the bulk of the semiconductor. The energy zero is chosen such that $E_c = E_{F,M} + \Phi_N + qU_{appl}$ with the Fermi level $E_{F,M}$ in the metal. For the moment we assume emission to occur in the whole barrier, i.e. between $x = 0$ and $x = x_B$. In Section 5 we will restrict emission to a reasonable fraction of the barrier width. U_{appl} equals the "applied voltage" only for vanishing bulk resistance. This causes severe problems if one

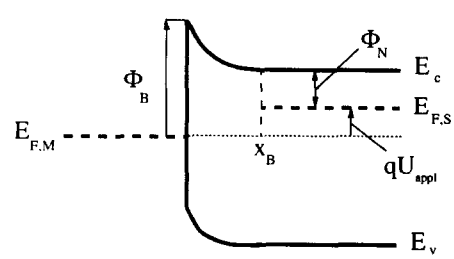

Fig. 1. Schematic band diagram of the metal–semiconductor interface.

wants to simulate more extended devices than a Schottky diode with the physical contact model.

Neglecting collisions inside the barrier region, transverse momentum and energy are conserved, and the emission current through the metal–semiconductor interface can be calculated from (Conley *et al.*[17]):

$$J = -\frac{q}{4\pi^3}\int \mathrm{d}^3\mathbf{k}_\mathrm{M} v_{\mathrm{M},x}\mathscr{T}(k_{\mathrm{M},x}, k_{\mathrm{S},x}) \times [f_\mathrm{M}(\mathbf{k}_\mathrm{M}) - f_\mathrm{S}(\mathbf{k}_\mathrm{S})], \quad (1)$$

(x-direction perpendicular to the interface). In eqn (1) the subscripts M and S refer to metal and semiconductor, respectively. $\mathscr{T}$ denotes the transmission probability and f the equilibrium distribution function. The transmission probability is obtained from the solutions of a 1D Schrödinger equation in the barrier region and on both sides of the interface. Therefore, both longitudinal momenta enter the expression of $\mathscr{T}(k_{\mathrm{M},x}, k_{\mathrm{S},x})$. However, the $k_{\mathrm{S},x}$-dependence can be expressed with the help of the conservation laws:

$$k_{\mathrm{S},x}^2 = \frac{m_\mathrm{c}}{m_\mathrm{M}} k_{\mathrm{M},x}^2 - \frac{2m_\mathrm{c}}{\hbar^2}E_\mathrm{c} - k_{\mathrm{M},\perp}^2\left(1 - \frac{m_\mathrm{c}}{m_\mathrm{M}}\right). \quad (2)$$

This relation clearly shows that due to the different effective masses on both sides of the barrier, $\mathscr{T}$ explicitly depends on the transverse momentum, i.e. $\mathscr{T}$ is a function of the two variables $k_{\mathrm{M},x}$ and $k_{\mathrm{M},\perp}$. In the following we will use the more convenient choice E (total energy) and Θ (angle between current axis and vector $\mathbf{k}_\mathrm{M}$).

For the evaluation of $\mathscr{T}(k_{\mathrm{M},x}, k_{\mathrm{M},\perp})$ we need the wave functions inside the barrier region, at least their WKB forms. Therefore, the potential shape must be given explicitly. Closed-form solutions of the Schrödinger equation exist in the case of a parabolic potential (uniform impurity density, Schottky approximation). We restrict ourselves to that case since it defines an "exact" reference for the accuracy of the analytical model. The method itself actually only demands the wave functions to be known far from the classical turning points as will be shown below, and hence is not restricted to a parabolic shape.

With above assumptions the electrostatic energy in the barrier region is given by (see Fig. 1):

$$E_\mathrm{c}(x) = E_\mathrm{c} + \frac{kT}{2\lambda_\mathrm{D}^2}(x - x_\mathrm{B})^2, \quad (3)$$

where λ_D denotes the Debye length $\lambda_\mathrm{D} = \sqrt{\epsilon_0\epsilon_\mathrm{s}kT/q^2N_\mathrm{D}}$. If eqn (3) is inserted into the Schrödinger equation for the envelope wave function, one obtains:

$$\left[-\frac{\mathrm{d}^2}{\mathrm{d}\xi^2} + \tfrac{1}{4}\xi^2 - \kappa_\mathrm{S}^2\right]\psi(\xi) = 0, \quad (4)$$

with the eigenvalue:

$$\kappa_\mathrm{S}^2 = \epsilon - \epsilon_\mathrm{c} - \frac{m_\mathrm{M}}{m_\mathrm{c}}\epsilon\sin^2\Theta. \quad (5)$$

In eqn (4) normalized quantities were introduced and labeled by small greek letters: $\xi = (x_\mathrm{B} - x)/\lambda$, $\kappa_\mathrm{S} = k_{\mathrm{S},x}\lambda\sqrt{m_\mathrm{eff}/m_\mathrm{c}}$, $\kappa_\mathrm{M} = \kappa_{\mathrm{M},x}\lambda m_\mathrm{eff}/m_\mathrm{M}$, $\epsilon = E/E_\lambda$, and $\epsilon_\mathrm{c} = E_\mathrm{c}/E_\lambda$. The angle Θ is measured between $\mathbf{k}_\mathrm{M}$ and the current (x-) axis. Normalization length and energy are defined by:

$$\lambda = \left(\frac{\hbar^2\lambda_\mathrm{D}^2}{4m_\mathrm{eff}kT}\right)^{1/4}, \quad E_\lambda = \frac{\hbar^2}{2m_\mathrm{eff}\lambda^2}. \quad (6)$$

E_λ corresponds to $2E_{00}$ in Refs [18–20] and to $2W_{00}$ in Ref. [16]. The effective mass m_eff has to be distinguished from m_c, since tunneling electrons can have energies reaching from E_c to $E_{\mathrm{F,M}} + \Phi_\mathrm{B}$. Therefore, a model of the tunneling mass $m_\mathrm{eff}(\epsilon)$ would be necessary that describes mixing of valence band states (and eventually metal states).

The differential equation (4) is solved by parabolic cylinder functions U and V ([21], p. 686):

$$\psi(\xi) = c_1 U(-\kappa_\mathrm{S}^2, \xi) + c_2 V(-\kappa_\mathrm{S}^2, \xi). \quad (7)$$

The matching conditions at $x = 0$ and $x = x_\mathrm{B}$, i.e. continuity of ψ and $(1/m)\,\mathrm{d}\psi/\mathrm{d}x$, yield for the transmission probability[17] (see Appendix A):

$$\mathscr{T}(\mathbf{k}_\mathrm{M}) = \frac{8}{\pi}\frac{\kappa_\mathrm{S}'}{\kappa_\mathrm{M}}|(V_0' + i\kappa_\mathrm{S}'V_0)(U_{\xi_\mathrm{B}} + i\kappa_\mathrm{M}^{-1}U_{\xi_\mathrm{B}}') - (U_0' + i\kappa_\mathrm{S}'U_0)(V_{\xi_\mathrm{B}} + i\kappa_\mathrm{M}^{-1}V_{\xi_\mathrm{B}}')|^{-2}, \quad (8)$$

with the abbreviations $U_0 = U(-\kappa_\mathrm{S}^2, 0)$, $U_{\xi_\mathrm{B}} = U(-\kappa_\mathrm{S}^2, \xi_\mathrm{B})$ and:

$$\kappa_\mathrm{S}' = \sqrt{\frac{m_\mathrm{c}}{m_\mathrm{eff}}}\kappa_\mathrm{S}, \quad \kappa_\mathrm{M} = \sqrt{\frac{m_\mathrm{eff}}{m_\mathrm{M}}}\epsilon\cos\Theta. \quad (9)$$

After some algebra the contact current density can be written as:

$$j = -2A_\mathrm{M}T^2\int_{\eta_\mathrm{c}}^{\infty}\mathrm{d}\eta\eta\left[\frac{1}{\exp(\eta - \eta_{\mathrm{F,M}}) + 1} - \frac{1}{\exp(\eta - \eta_{\mathrm{F,M}} - u) + 1}\right]\int_0^{v_\mathrm{max}(\eta)} v\,\mathrm{d}v\,\mathscr{T}(\eta, v), \quad (10)$$

with:

$$\mathscr{T}(\eta, v) = \frac{2}{1 + g(\eta, v)} \quad (11)$$

$$g(\eta, v) = \frac{\pi}{4}\left[\left(\frac{\kappa_\mathrm{M}}{\kappa_\mathrm{S}'}\right)(V_0'U_{\xi_\mathrm{B}} - U_0'V_{\xi_\mathrm{B}})^2 + \left(\frac{\kappa_\mathrm{S}'}{\kappa_\mathrm{M}}\right)(U_0V_{\xi_\mathrm{B}}' - V_0U_{\xi_\mathrm{B}}')^2 + \left(\frac{1}{\kappa_\mathrm{M}\kappa_\mathrm{S}'}\right)(V_0'U_{\xi_\mathrm{B}}' - U_0'V_{\xi_\mathrm{B}}')^2 + \kappa_\mathrm{M}\kappa_\mathrm{S}'(V_0U_{\xi_\mathrm{B}} - U_0V_{\xi_\mathrm{B}})^2\right] \quad (12)$$

$$\kappa_M = \sqrt{\frac{m_{eff}}{m_M}}\sqrt{\eta(1-v^2)/\eta_\lambda} \quad (13)$$

$$\kappa_S' = \sqrt{\frac{m_c}{m_{eff}}}\sqrt{\left[\eta\left(1-\frac{m_M}{m_c}v^2\right)-\eta_c\right]\Big/\eta_\lambda} \quad (14)$$

$$A_* = qm_*k^2/(2\pi^2\hbar^3) \quad \text{effective Richardson constant} \quad (15)$$

($\eta = E/kT$, $\eta_c = E_c/kT$, $\eta_\lambda = E_\lambda/kT$, $v = \sin\Theta$, $u = qU_{appl}/kT$). The derivation of eqn (11) with (12) from eqn (8) can be found in Appendix A.

The lower limit of the energy integral follows from the condition that the square root (14) must remain real with the lowest value of v ($v = 0$). It reflects the inability of electrons to transit from or into forbidden states of the semiconductor, i.e. the gap in the density of states. The condition of a real κ_S' also restricts the angular integration to:

$$v_{max}(\eta) = \sqrt{\frac{m_c}{m_M}}\sqrt{1-\eta_c/\eta}. \quad (16)$$

This limit, which is generally small since $(E - E_c) \ll E_{F,M}$, results from conservation of transverse momentum. Whereas in the initial integral (1) integration is symbolically over the entire k-space, which can be interpreted as taking the second moment of the boundary condition for the Boltzmann equation, conservation laws impose restrictions to the velocity vector, when the electron changes from one band structure to the other.

The magnitude of the parameter $\eta_\lambda = E_\lambda/kT$ determines the predominant current mechanism—thermionic emission (TE) if $\eta_\lambda \ll 1$ (low doping, Schottky contact), thermionic field emission (TFE) if $\eta_\lambda \approx 1$ (medium doping), and field emission (FE) if $\eta_\lambda \gg 1$ (heavy doping, Ohmic contact).

Often a simplified ansatz with a transmission probability depending only on E_x[22] or $k_{S,x}$[23] is used, and the k-space integration is performed over the Brillioun zone of the semiconductor instead of the Fermi sphere of the metal:

$$j = -\frac{q\hbar}{4\pi^3 m_c}\int_0^\infty dk_{S,x}k_{S,x}\mathcal{T}(k_{S,x}) \times \int d^2\mathbf{k}_\perp[f_M(\mathbf{k}_M) - f_S(\mathbf{k}_S)]. \quad (17)$$

In this case one obtains after changing to energy variables:

$$j = -A_cT^2\int_{\eta_c}^\infty d\eta_x\mathcal{T}(\eta_x) \times \int_0^\infty d\eta_\perp\left[\frac{1}{\exp(\eta_x+\eta_\perp-\eta_{F,M})+1} - \frac{1}{\exp(\eta_x+\eta_\perp-\eta_{F,M}-u)+1}\right]. \quad (18)$$

Introducing the variables E and v as before, this transforms to:

$$j = -2A_cT^2\int_{\eta_c}^\infty d\eta\eta\left[\frac{1}{\exp(\eta-\eta_{F,M})+1} - \frac{1}{\exp(\eta-\eta_{F,M}-u)+1}\right] \times \int_0^{v_{max}(\eta)} v\,dv\mathcal{T}[\eta(1-v^2)], \quad (19)$$

with $v_{max}(\eta) = \sqrt{1-\eta_c/\eta}$. Here the dependence of $\mathcal{T}$ on η and v is of the special form $\eta(1-v^2)$, which agrees with the general expression of the transmission probability equations (11)–(15) only if: (a) $m_c = m_M$; or (b) $\mathcal{T}(\eta, v) \to \mathcal{T}(\eta, 0)$.

Case (a) follows immediately from (14) and (15). In case (b) the v-integration becomes trivial and yields identical current densities (10) and (19). The *ad hoc* ansatz $\mathcal{T} = \mathcal{T}(E_x)$ has the advantage of avoiding a numerical double integral, because the $\eta_\perp$-integration in (18) is straightforward, but it requires modifications in the general expression of $\mathcal{T}$ (11).

Since the aim of the paper is twofold: replacing the WKB approximation by a better approach and deriving an analytical expression of the contact current suitable for device simulation, it seems to be most feasible to set m_M/m_c to unity in (14), or, which has the same effect, to set $v = 0$ ($\eta = \eta_x$) in order to enable an approximate analytical integration. This is justified by the above mentioned fact that conservation of transverse momentum only permits small angles between momentum vector $\mathbf{k}_M$ and x-axis. In particular, the approximation $\mathcal{T}(\eta, v) \to \mathcal{T}(\eta, 0)$ does not affect the results of the next section, where the WKB approximation is replaced by an interpolation scheme.

With this modification of the transmission probability (11) we can use eqn (8) as the starting point of the further calculation. Analytical integration in Section 4 demands also an approximation for the Fermi functions. We choose:

$$\frac{1}{\exp(x)+1} \to e^{-x}\Theta(x) + \Theta(-x). \quad (20)$$

With this simplification the $\eta_\perp$-integral can be solved and eqn (18) turns into:

$$\begin{aligned} j = -A_cT^2\int_{\eta_c-\eta_{F,M}}^\infty & d\eta\mathcal{T}(\eta+\eta_{F,M}, 0) \\ & \times\{\Theta(\eta_c-\eta_{F,M})e^{-\eta} + \Theta(\eta_{F,M}-\eta_c)[e^{-\eta}\Theta(\eta) \\ & + (1-\eta)\Theta(-\eta)] - \Theta(\Phi_N)e^{-(\eta-u)} \\ & - \Theta(-\Phi_N)[e^{-(\eta-u)}\Theta(\eta-u) \\ & + (1-\eta+u)\Theta(u-\eta)]\}, \end{aligned} \quad (21)$$

where $\mathcal{T}$ is given by eqns (11)–(15). We have changed the meaning of the integration variable η for convenience: $\eta = \eta_x - \eta_{F,M}$, i.e. the energy zero for the variable η now is given by the metal Fermi level. Equation (21) will be considered as "exact" reference model in this paper. The term "exact" only applies

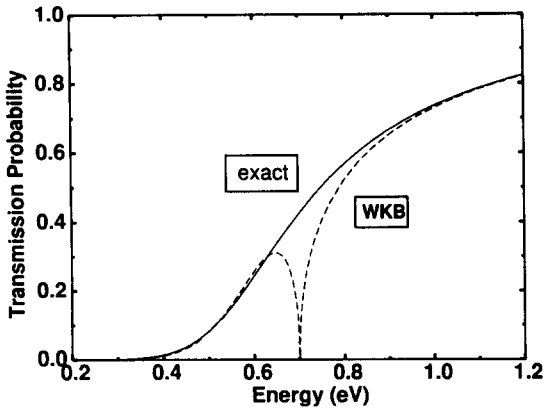

Fig. 2. Comparison of the transmission probabilities using parabolic cylinder functions (solid curve) and WKB wave functions (dashed curve). Parameters: $E_{F,M} = 11.7$ eV, $m_M = m_0$, $m_c = 0.258 m_0$ (⟨111⟩-silicon), $\Phi_B = 0.7$ eV, $N_D = 10^{19}\,\mathrm{cm}^{-3}$, $U_{appl} = 0$ V.

to the usage of the transmission probability in terms of the eigenfunctions of the potential barrier.

The well-known TE current formula follows immediately, if $\mathcal{T}$ is replaced by the step function $\Theta(\eta - \Phi_B/kT)$:

$$j_{TE} = -A_c T^2 \exp\left(-\frac{\Phi_B}{kT}\right)[1 - \exp(u)]. \quad (22)$$

3. INTERPOLATION SCHEME FOR THE TRANSMISSION PROBABILITY

In this section we derive a simplified expression of the transmission probability $\mathcal{T}(\eta, 0)$ suitable for analytical integration. It is the aim to substitute the WKB approximation by a better approach which essentially preserves the accuracy of eqn (11) (restriction to the case $v = 0$ is only for convenience in this section). The proposed interpolation scheme uses the asymptotic forms of the parabolic cylinder functions which read, e.g. in the case of U_{ξ_B} ([21], p. 690):

$$U_{\xi_B}^{appr} = \frac{\sqrt{\Gamma(\frac{1}{2} + \kappa_S^2)}}{(2\pi)^{1/4}|\xi_B^2 - 4\kappa_S^2|^{1/4}} \times \begin{cases} \frac{1}{2}e^{-S} & \text{below top} \\ \sin\left(\frac{\pi}{4} + |S|\right) & \text{above top} \end{cases}, \quad (23)$$

where S denotes the action integral:

$$S = \frac{1}{2}\int_{2\kappa_S}^{\xi_B} d\xi \sqrt{\xi^2 - 4\kappa_S^2}. \quad (24)$$

The corresponding expressions for $V_{\xi_B}^{appr}$, $U_{\xi_B}'^{appr}$, and $V_{\xi_B}'^{appr}$ are listed in Appendix B. These asymptotic forms agree with the WKB wave functions up to a prefactor. They approach U_{ξ_B}, V_{ξ_B}, etc. for energies much smaller or much larger than the barrier height, respectively, but they diverge at the top of the barrier. Figure 2 compares $\mathcal{T}^{appr}(\eta, 0)$ (with $U_{\xi_B}^{appr}$, $V_{\xi_B}^{appr}$, etc.) against $\mathcal{T}(\eta, 0)$ (using the exact solutions). Due to the breakdown of the WKB approximation the transmission probability drops off to zero at the top of the barrier ($\kappa_S = \xi_B/2$).

The correct shape of $\mathcal{T}(\eta, 0)$ can be reproduced, if a function is used which yields the same asymptotic limits, but interpolates over the divergency. Such a function is represented in the case of U_{ξ_B} by:

$$U_{\xi_B}^{Ai} = \frac{(2\pi)^{1/4}\sqrt{\Gamma(\frac{1}{2} + \kappa_S^2)}}{\left|\frac{\xi_B^2}{4} - \kappa_S^2\right|^{1/4}} (\tfrac{3}{2}|S|)^{1/6}\mathrm{Ai}([\tfrac{3}{2}S]^{2/3}). \quad (25)$$

If S is imaginary (above top), the real root of $(-1)^{1/3}$ has to be used, i.e. -1. For a triangular barrier a similar function can be obtained in a straightforward way, if a transformation of the Schrödinger equation from ξ to the new variable $S(\xi)$ is performed (as first done by Fowler and Nordheim[24]; see also the book of Moll[25]). It is easy to prove that $U_{\xi_B}^{Ai}$ has exactly the limits of eqn (23) and that it turns into the finite value:

$$\lim_{\kappa_S \to \xi_B/2} U_{\xi_B}^{Ai} = (2\pi)^{1/4}\sqrt{\left(\frac{2}{\xi_B}\right)^{1/3}\Gamma\left(\frac{1}{2} + \frac{\xi_B^2}{4}\right)}\mathrm{Ai}(0), \quad (26)$$

where $U_{\xi_B}^{appr}$ diverges. The maximum deviation of $U_{\xi_B}^{Ai}$ from U_{ξ_B} at the top of the barrier is of the order of 0.2% only. The same holds for the other functions in $\mathcal{T}(\eta, 0)$, which are substituted according to $V \to \mathrm{Bi}$, $U' \to \mathrm{Ai}'$, and $V' \to \mathrm{Bi}'$. Full expressions are given in Appendix B. After these substitutions, the function $g(\eta, 0)$ in the denominator of the transmittance can be written as:

$$\begin{aligned} g(\eta + \eta_{F,M}, 0) = {} & 2\pi\sqrt{\frac{(\eta + \eta_{F,M})|Y|}{|\varphi_B - \eta|}}\sqrt{\frac{m_c}{m_M}} \\ & \times \left\{\mathrm{Re}^2[e^{i(\pi/3)}\mathrm{Ai}(Y e^{-i(2\pi/3)})] \right. \\ & \left. + \frac{m_{eff}}{m_c}\mathrm{Im}^2[e^{i(\pi/3)}\mathrm{Ai}(Y e^{-i(2\pi/3)})]\right\} \\ & + 2\pi\sqrt{\frac{|\varphi_B - \eta|}{(\eta + \eta_{F,M})|Y|}}\sqrt{\frac{m_M}{m_c}} \\ & \times \left\{\frac{m_c}{m_{eff}}\mathrm{Re}^2[e^{-i(\pi/3)}\mathrm{Ai}'(Y e^{-i(2\pi/3)})] \right. \\ & \left. + \mathrm{Im}^2[e^{-i(\pi/3)}\mathrm{Ai}'(Y e^{-i(2\pi/3)})]\right\} \end{aligned} \quad (27)$$

with

$$Y = (\tfrac{3}{2}S)^{2/3}. \quad (28)$$

In deriving eqn (27) the ratio $\Gamma(\frac{3}{4} + \kappa_S^2/2)/\Gamma(\frac{1}{4} + \kappa_S^2/2)$ was approximated with the help of Stirling's formula by $|\kappa_S|$ without any visible influence on the transmission probability. The origin of η is again $\eta_{F,M}$. Furthermore, with eqns (3), (5) and (6) we substituted:

$$\frac{\xi_B^2}{4} - \kappa_S^2 = \frac{\varphi_B - (\eta - \eta_{F,M})}{\eta_\lambda}, \quad (29)$$

with $\varphi_B = \Phi_B/kT$.

Equation (27) also shows that the assumption of a tunneling mass m_{eff} different from the bulk effective mass m_{c} introduces an essential complication. In order to keep the expressions short, we now restrict ourselves to the simplest model for m_{eff}:

$$m_{\mathrm{eff}}(\epsilon) \equiv m_{\mathrm{c}}. \tag{30}$$

Equation (27) then can be written in the compact form:

$$\begin{aligned} g(\eta + \eta_{\mathrm{F,M}}, 0) = {} & 2\pi \sqrt{\frac{(\eta + \eta_{\mathrm{F,M}})|Y|}{|\varphi_{\mathrm{B}} - \eta|}} \\ & \times \sqrt{\frac{m_{\mathrm{c}}}{m_{\mathrm{M}}}} |\mathrm{Ai}(Y\,\mathrm{e}^{-i(2\pi/3)})|^2 \\ & + 2\pi \sqrt{\frac{|\varphi_{\mathrm{B}} - \eta|}{(\eta + \eta_{\mathrm{F,M}})|Y|}} \sqrt{\frac{m_{\mathrm{M}}}{m_{\mathrm{c}}}} \\ & \times |\mathrm{Ai}'(Y\,\mathrm{e}^{-i(2\pi/3)})|^2. \end{aligned} \tag{31}$$

We note that Y is positive below the top of the barrier (S real) and negative above (S imaginary).

To derive an analytical formula of the contact current, the mathematical form of the transmittance has to be further simplified. Therefore, it is desirable to substitute the Airy functions. The first maximum of the Airy function $\mathrm{Ai}(t)$ with a real argument t is located at $t_0 = -1.0187929(7)$ which corresponds to a certain energy level above the maximum of the barrier. From the bottom of the barrier up to that point it is possible to replace the Airy function by a Gaussian. Figure 3(a) shows how this is done: the center of the Gaussian is placed at t_0 and two points are used to fix the shape of the Gaussian. The first point is $t = 0$ which corresponds to the top of the barrier and ensures a good agreement in the TE regime, the second point is fixed by fitting the attenuation constant such that best agreement is achieved in the TFE and FE regimes. Therefore, we replace [see Fig. 3(a)]:

$$\mathrm{Ai}(t) \to \mathrm{Ai}(0)\,\mathrm{e}^{-bt(t-2t_0)}. \tag{32}$$

For the transmittance we need the absolute value of Ai with the complex argument $Y \exp(i(2\pi/3))$, i.e.

$$|\mathrm{Ai}(Y\,\mathrm{e}^{-i(2\pi/3)}| = \tfrac{1}{2}\sqrt{\mathrm{Ai}^2(Y) + \mathrm{Bi}^2(Y)}. \tag{33}$$

Applying (32) to that case yields:

$$|\mathrm{Ai}(Y\,\mathrm{e}^{-i(2\pi/3)})| \to \mathrm{Ai}(0)\,\mathrm{e}^{1/2bY(Y-2t_0)}, \tag{34}$$

i.e. (34) includes a corresponding approximation to Bi as well. Figure 3(b) shows the accuracy of (34). The fit of the attenuation parameter b was found to be unique for all doping concentrations. Best agreement with the exact transmission probability was obtained for $b = 0.38$. The effect of a variation of b will be discussed in Section 4.

The Gaussian approximation results in a further simplification of the function g in the denominator of the transmittance:

$$\begin{aligned} g(\eta + \eta_{\mathrm{F,M}}, 0) = {} & 2\pi\,\mathrm{Ai}^2(0)\,\mathrm{e}^{bY(Y-2t_0)} \\ & \times \left[\sqrt{\frac{m_{\mathrm{c}}}{m_{\mathrm{M}}}} \sqrt{\frac{(\eta + \eta_{\mathrm{F,M}})|Y|}{|\varphi_{\mathrm{B}} - \eta|}} \right. \\ & + \sqrt{\frac{m_{\mathrm{M}}}{m_{\mathrm{c}}}} \sqrt{\frac{|\varphi_{\mathrm{B}} - \eta|}{(\eta + \eta_{\mathrm{F,M}})|Y|}} \\ & \left. \times\, 4b^2(Y^2 + Yt_0 + t_0^2) \right]. \end{aligned} \tag{35}$$

This approximation may be used up to an energy $\eta_{\max}$ given by the condition $Y(\eta_{\max}) = t_0$. The energy $\eta_{\max}$ can be determined analyticaly if $Y(\eta)$ is developed up to first order in η around the top of the barrier, resulting in:

$$\eta_{\max} = \varphi_{\mathrm{B}} + \left(\frac{\xi_{\mathrm{B}}}{2}\right)^{2/3} \eta_\lambda |t_0|. \tag{36}$$

Note again, that η is measured from $\eta_{\mathrm{F,M}}$. The derivation of (36) is given in Appendix C. For energies larger than $\eta_{\max}$ the WKB form of g is sufficient:

$$\begin{aligned} g_{\mathrm{WKB}}(\eta + \eta_{\mathrm{F,M}}, 0) = {} & \frac{1}{2}\sqrt{\frac{m_{\mathrm{c}}}{m_{\mathrm{M}}}} \sqrt{\frac{\eta + \eta_{\mathrm{F,M}}}{\eta - \varphi_{\mathrm{B}}}} \\ & + \frac{1}{2}\sqrt{\frac{m_{\mathrm{M}}}{m_{\mathrm{c}}}} \sqrt{\frac{\eta - \varphi_{\mathrm{B}}}{\eta + \eta_{\mathrm{F,M}}}}, \quad \eta > \eta_{\max}. \end{aligned} \tag{37}$$

Appendix D contains the derivation of the last

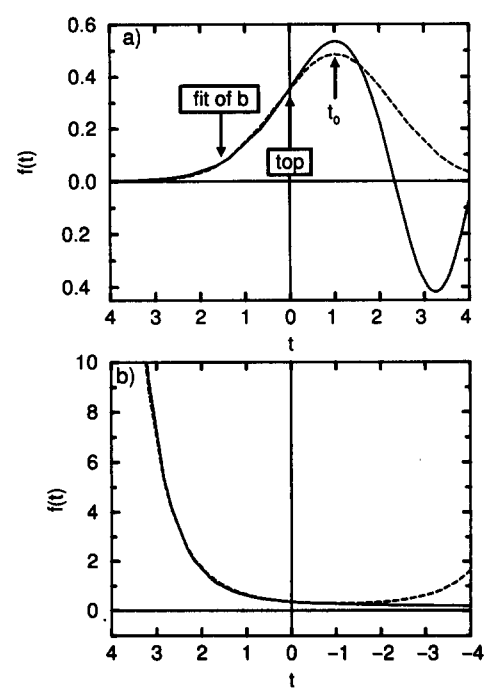

Fig. 3. (a) Approximation of the Airy function $\mathrm{Ai}(t)$ (solid curve) by a Gaussian (dashed curve). t_0 denotes the position of the first maximum of the Airy function. The point $t = 0$ corresponds to the barrier maximum, where both curves coincide. The fit is tuned by the attenuation parameter b. (b) The same fit for the function $[\mathrm{Ai}^2(t) + \mathrm{Bi}^2(t)]^{1/2}/2$.

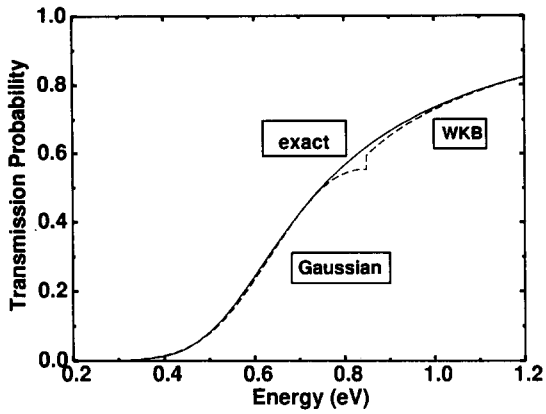

Fig. 4. Comparison of the transmission probabilities using parabolic cylinder functions (solid curve) and the Gaussian approximation (dashed curve), which is used up to an energy E_{max}. For higher energies the WKB form is used. Parameters: $b = 0.38$, for other parameters see Fig. 2.

equation. In Fig. 4 we show $\mathscr{T}^{\text{Gauss}}$ [eqns (35) and (37)] in comparison with $\mathscr{T}(\eta, 0)$ (using the exact solutions). Now, the dip at φ_B is removed and only a small error occurs well above the top of the barrier. This error has practically no effect on the current in the TFE and FE regimes and only litte effect in the TE regime. A further advantage of $\mathscr{T}^{\text{Gauss}}$ is that in this form an approximate analytical integration becomes possible.

4. ANALYTICAL MODEL OF THE CONTACT CURRENT

To enable analytical integration with the distribution function model (20) the transmission probability $\mathscr{T}(\eta + \eta_{F,M}, 0)$ has to be transformed into a Gaussian with respect to the energy η. Therefore, $Y(\eta)$ must be linearized in η at the energy level of maximum current density. The major failure in the analytical model of the contact current results from the inability for finding a satisfactory analytical expression of that energy level. Furthermore, such an expression depends on the particular case, i.e. on the sign of the applied voltage and if the semiconductor is degenerate or not (see Ref. [18]). The various parts of the integrand in the current integral yield different approximations for the level of maximum of these parts [compare eqn (21)]. In any case, they have to be borrowed from the WKB approach. Thus, instead of $\mathscr{T}(\eta + \eta_{F,M}, 0)$ we must use $\exp[-2S(\eta)]$ to find a solution of:

$$\frac{\mathrm{d}}{\mathrm{d}\eta}\,\mathrm{e}^{-2S} f(\eta) = 0, \tag{38}$$

where $f(\eta)$ is one of the various terms in the braces of eqn (21). Inserting the action S of the parabolic barrier, this gives the implicit relation:

$$\eta - \varphi_B = (\eta_c - \eta_{F,M} - \varphi_B)\tanh^2\left[\frac{\eta_\lambda}{2}\frac{\mathrm{d}}{\mathrm{d}\eta}\ln f(\eta)\right]. \tag{39}$$

From the last equation one can derive the following energy levels that approximately describe the positions of the current density maxima of the different terms in eqn (21):

$$\eta^+ = \eta_c - \eta_{F,M} + \frac{\varphi_B - (\eta_c - \eta_{F,M})}{\cosh^2\left(\frac{\eta_\lambda}{2}\right)}$$

$$\text{term} \sim \Theta(\eta_c - \eta_{F,M}) \quad \text{and} \quad \text{term} \sim \Theta(\Phi_N)$$

$$\eta_1^- = \max(0, \eta^+) \quad \text{term} \sim \Theta(\eta_{F,M} - \eta_c)\Theta(\eta)$$

$$\eta_2^- = -\tfrac{1}{2}\ln^{-1}\left(\frac{\sqrt{\varphi_B + \eta_{F,M} - \eta_c} + \sqrt{\varphi_B}}{\sqrt{\eta_{F,M} - \eta_c}}\right)$$

$$\text{term} \sim \Theta(\eta_{F,M} - \eta_c)\Theta(-\eta)$$

$$\eta_1^d = \max(u, \eta^+) \text{ term} \sim \Theta(-\Phi_N)\Theta(\eta - \mu)$$

$$\eta_2^d = u + \eta_2^- \text{ term} \sim \Theta(-\Phi_N)\Theta(u - \eta). \tag{40}$$

After linearizing $Y(\eta)$ at these levels:

$$Y(\eta) = Y(\tilde{\eta}) + \frac{\eta - \tilde{\eta}}{\gamma(\tilde{\eta})}, \tag{41}$$

with

$$\gamma(\tilde{\eta}) = Y'^{-1}(\tilde{\eta}) = -\sqrt{Y(\tilde{\eta})} \times \ln^{-1}\left(\frac{\sqrt{\varphi_B + \eta_{F,M} - \eta_c} + \sqrt{\varphi_B - \tilde{\eta}}}{\sqrt{\tilde{\eta} + \eta_{F,M} - \eta_c}}\right), \tag{42}$$

where $\tilde{\eta}$ denotes one of the energy levels (40) (with the origin $\eta_{F,M}$), the transmittance can be written in Gaussian form for all energies in the range $[\eta_c - \eta_{F,M}, \eta_{max}]$:

$$\mathscr{T}(\eta + \eta_{F,M}, 0) \approx \frac{2\,\mathrm{e}^{b[Y(\tilde{\eta}) - t_0]^2}}{1 + g(\tilde{\eta} + \eta_{F,M}, 0)} \times \exp\left[-b\left\{\frac{\eta - \tilde{\eta}}{\eta_\lambda \gamma(\tilde{\eta})} + Y(\tilde{\eta}) - t_0\right\}^2\right], \quad \eta \leqslant \eta_{max}. \tag{43}$$

To allow analytical integration for energies larger than η_{max} too, the transmittance is approximated there by (see eqn (37) and Appendix D):

$$\mathscr{T}(\eta + \eta_{F,M}, 0) \approx \frac{4\sqrt{\frac{m_M}{m_c \eta_{F,M}}}\sqrt{\eta - \varphi_B}}{\left(1 + \sqrt{\frac{m_M}{m_c \eta_{F,M}}}\sqrt{\eta_{max} - \varphi_B}\right)^2}, \quad \eta > \eta_{max}. \tag{44}$$

With eqns (43) and (44) the contact current can be calculated without any further approximations. The final result reads (omitting details of the lengthy but straightforward calculation):

$$j = j^+ + j_1^- + j_2^- + j_1^d + j_2^d + j_3, \tag{45}$$

with

$$j^+ = -j_0(\eta^+)[\Theta(\eta_c - \eta_{F,M})$$

$$- \Theta(\Phi_N)\, e^{u}]e^{-\eta^+ + \vartheta^2(\eta^+, \eta^+)}$$
$$\times \operatorname{Erf}[\vartheta(\eta_c - \eta_{F,M}, \eta^+), \vartheta(\eta_{max}, \eta^+)] \tag{46}$$

$$j_1^- = -j_0(\eta_1^-)\Theta(\eta_{F,M} - \eta_c)\, e^{-\eta_1^- + \vartheta^2(\eta_1^-, \eta_1^-)}$$
$$\times \operatorname{Erf}[\vartheta(0, \eta_1^-), \vartheta(\eta_{max}, \eta_1^-)] \tag{47}$$

$$j_1^d = +j_0(\eta_1^d)\Theta(-\Phi_N)\, e^{-\eta_1^d + u + \vartheta^2(\eta_1^d, \eta_1^d)}$$
$$\times \operatorname{Erf}[\vartheta(u, \eta_1^d), \vartheta(\eta_{max}, \eta_1^d)] \tag{48}$$

$$j_2^- = -j_0(\eta_2^-)\Theta(\eta_{F,M} - \eta_c)\, e^{\vartheta_1^2(\eta_2^-, \eta_2^-)}$$
$$\times \left\{\left[1 + \frac{|\gamma(\eta_2^-)|\eta_\lambda}{\sqrt{b}}\vartheta_1(0, \eta_2^-)\right]\right.$$
$$\times \operatorname{Erf}[\vartheta_1(\eta_c - \eta_{F,M}, \eta_2^-), \vartheta_1(0, \eta_2^-)]$$
$$\left. + \frac{|\gamma(\eta_2^-)|\eta_\lambda}{\sqrt{\pi b}}\left(e^{-\vartheta_1^2(0, \eta_2^-)} - e^{\vartheta_1^2(\eta_c - \eta_{F,M}, \eta_2^-)}\right)\right\} \tag{49}$$

$$j_2^d = +j_0(\eta_2^d)\Theta(-\Phi_N)\, e^{\vartheta_1^2(\eta_2^d, \eta_2^d)}$$
$$\times \left\{\left[1 + u + \frac{|\gamma(\eta_2^d)|\eta_\lambda}{\sqrt{b}}\vartheta(0, \eta_2^d)\right]\right.$$
$$\times \operatorname{Erf}[\vartheta_1(\eta_c - \eta_{F,M}, \eta_2^d), \vartheta_1(u, \eta_2^d)]$$
$$\left. + \frac{|\gamma(\eta_2^d)|\eta_\lambda}{\sqrt{\pi b}}\left(e^{-\vartheta_1^2(u, \eta_2^d)} - e^{\vartheta_1^2(\eta_c - \eta_{F,M}, \eta_2^d)}\right)\right\} \tag{50}$$

$$j_3 = -\frac{2A_c T^2\sqrt{\pi}\sqrt{\dfrac{m_M}{m_c \eta_{F,M}}}}{\left(1 + \sqrt{\dfrac{m_M}{m_c}}\sqrt{\dfrac{\eta_{max} - \varphi_B}{\eta_{F,M}}}\right)^2}(1 - e^{u})$$
$$\times \left[e^{-\varphi_B}\operatorname{erfc}(\sqrt{\eta_{max} - \varphi_B})\right.$$
$$\left. + \frac{2}{\sqrt{\pi}}\sqrt{\eta_{max} - \varphi_B}\, e^{-\eta_{max}}\right], \tag{51}$$

where

$$j_0(\eta) = \frac{A_c T^2\sqrt{\pi}|\gamma(\eta)|\eta_\lambda}{\sqrt{b}[1 + g(\eta + \eta_{F,M}, 0)]}. \tag{52}$$

The energies η^+, η_1^-, η_1^d, η_2^- and η_2^d are the quantities defined in (40), the function $g(\eta + \eta_{F,M}, 0)$ is given by eqn (35) and $\gamma(\eta)$ by eqn (42). $\operatorname{Erf}(x, y)$ denotes the generalized error function $\operatorname{Erf}(x, y) = \operatorname{erf}(y) - \operatorname{erf}(x)$, and $\operatorname{erfc}(x)$ the complementary error function $\operatorname{erfc}(x) = \operatorname{Erf}(x, \infty)$ ([21], p. 297). The action S reads explicitly:

$$S(\eta) = \sqrt{\varphi_B + \eta_{F,M} - \eta_c}\sqrt{\varphi_B - \eta} - (\eta + \eta_{F,M} - \eta_c)$$
$$\times \ln\left(\frac{\sqrt{\varphi_B + \eta_{F,M} - \eta_c} + \sqrt{\varphi_B - \eta}}{\sqrt{\eta + \eta_{F,M} - \eta_c}}\right). \tag{53}$$

Two new functions were introduced for convenience:

$$\vartheta(\eta_1, \eta_2) = \frac{\sqrt{b}}{|\gamma(\eta_2)|\eta_\lambda}$$
$$\times \left\{\eta_1 - \eta_2 + \eta_\lambda\gamma(\eta_2)[Y(\eta_2) - t_0] + \frac{\gamma^2(\eta_2)\eta_\lambda^2}{2b}\right\} \tag{54}$$

$$\vartheta_1(\eta_1, \eta_2) = \vartheta(\eta_1, \eta_2) - \frac{|\gamma(\eta_2)|\eta_\lambda}{2\sqrt{b}}. \tag{55}$$

To complete the list of quantities we repeat that:

$$\eta_c = \eta_{F,M} + \Phi_N/kT + u$$
$$\eta_{max} = \varphi_B - t_0[\eta_\lambda^2(\varphi_B + \eta_{F,M} - \eta_c)]^{1/3}$$
$$\eta_\lambda = \frac{\sqrt{\hbar q}}{kT}\sqrt{\frac{N_D}{m_c \epsilon_0 \epsilon_s}}$$
$$t_0 = \mathrm{Ai}'^{-1}(0) = -1.0187929(7)$$
$$\mathrm{Ai}(0) = 0.3550280(5)$$
$$A_c = 120.155\left(\frac{m_c}{m_0}\right)\mathrm{AK}^{-2}\,\mathrm{cm}^{-2}$$
Richardson constant
$$b = 0.38.$$

We should note that the analytical model (45) describes the contact current through a parabolic barrier for the whole range of doping concentrations (from Schottky to Ohmic) and for the whole range of external voltages (from strong reverse to strong forward bias). Therefore, the appearance of different terms is natural. All terms have a similar structure and each term contains an error function as the most complicated ingredient. For the case of a Schottky contact (low doping, TE regime) only the terms j_3 and $j^+ + j_1^-$ are essential, in the TFE regime the important terms are j^+ and j_1^-, whereas for an Ohmic contact (FE regime) only j_2^- (reverse bias) and $j_1^d + j_2^d$ (forward bias) contribute considerably to the total current density.

In Fig. 5 the analytical model (45) is compared with the "exact" reference model (21), where the transmittance is expressed by parabolic cylinder functions and the energy integral is performed numerically. We used typical parameters for an Al-contact on n-type silicon varying the donor concentration from $10^{14}\,\mathrm{cm}^{-3}$ to $10^{20}\,\mathrm{cm}^{-3}$. Φ_N was calculated with Fermi statistics. The agreement is excellent over the whole range, i.e. over 14 orders of magnitude in the current density. One observes the typical change from a rectifying behaviour to an almost linear behaviour as the doping level is increased.

Both current densities are compared in more detail in Fig. 6, where their ratio is plotted over the whole doping range for -2 and $+0.2$ V bias, respectively. The maximum relative error is about 40% and sharply peaked at a particular doping concentration. Here the energy level of maximum current density crosses the metal Fermi level, and the statistics changes from Boltzmann to total degeneracy according to the simplified model (20). Note that the disagreement is mainly caused by the analytical

integration and not by the approximation for the transmittance. Leaving one numerical integration would not only allow for Fermi statistics, but also practically reproduce the curves of the "exact" reference model.

As already mentioned, the agreement was achieved with an unique numerical fit parameter $b = 0.38$ in the approximate mapping of the Airy function to a Gaussian. The sensitivity of the analytical jV-characteristics to that parameter is demonstrated in Fig. 7 where we chose the intermediate case $N_D = 10^{18}\,\mathrm{cm}^{-3}$. Increasing the parameter b decreases the reverse current, but not very pronounced. The sensitivity of the results to b depends on the doping concentration—in the TE regime the influence is negligible, whereas it is largest in the FE regime.

In order to prove if the interpolation method may be extended to other barrier shapes, we acted as if we did not know the exact eigenfunctions of the parabolic potential but only the WKB solutions. Proceeding with them in the same manner as described above in detail leads to the same final expressions. This is because of the close relationship between asymptotic forms and WKB wave functions and due to the fact that the unknown normalization constants of the latter cancel out in the expression of the transmittance. Therefore, the interpolation scheme is not restricted to the parabolic barrier, although its accuracy can not be proved directly for more complicated problems.

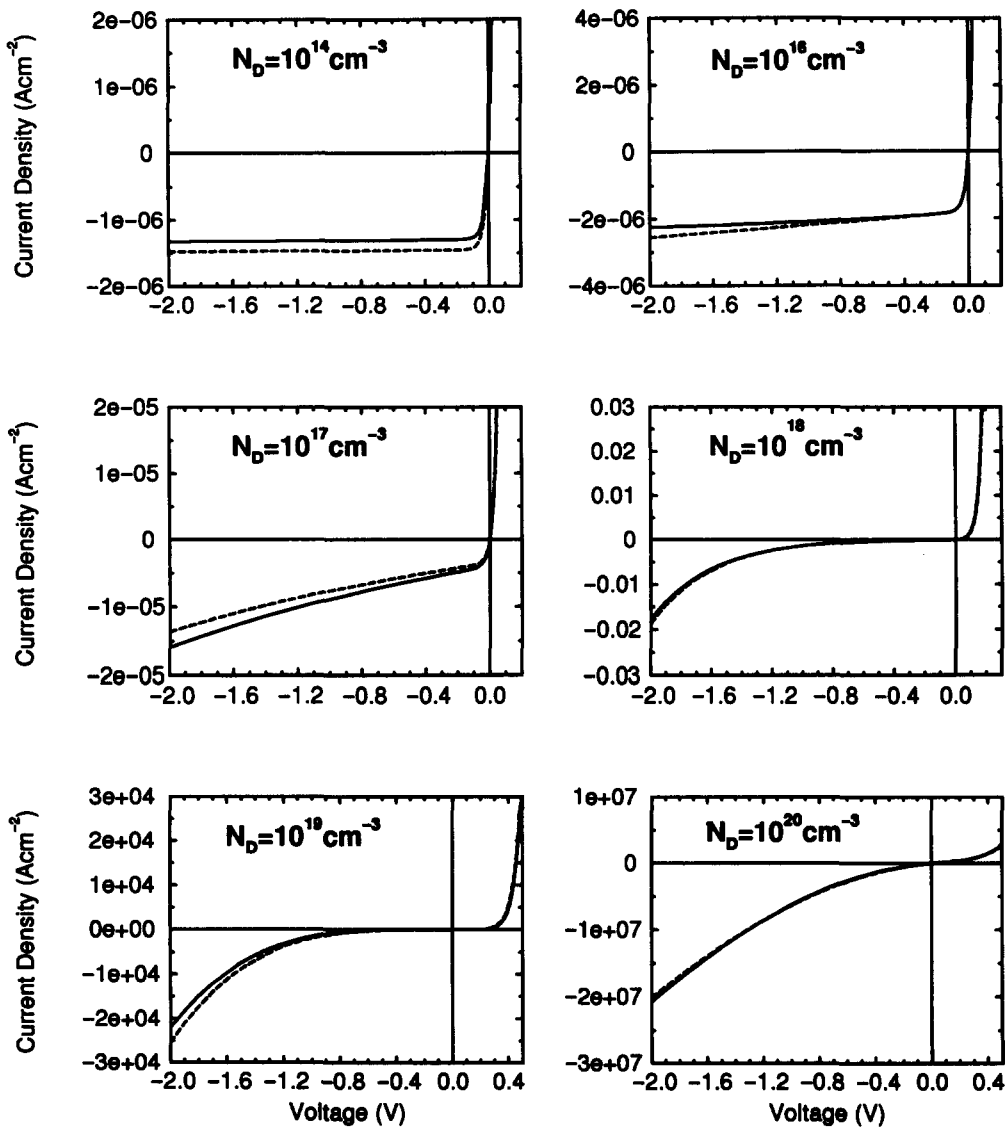

Fig. 5. Calculated jV-characteristics of an Al/n-Si contact for various donor concentrations. Solid curves: exact transmission probability and numerical integration, dashed curves: fully analytical model. Parameters: $E_{F,M} = 11.7$ eV, $m_M = m_0$, $m_c = 0.258 m_0$ (⟨111⟩-silicon), $\Phi_B = 0.7$ eV, $b = 0.38$.

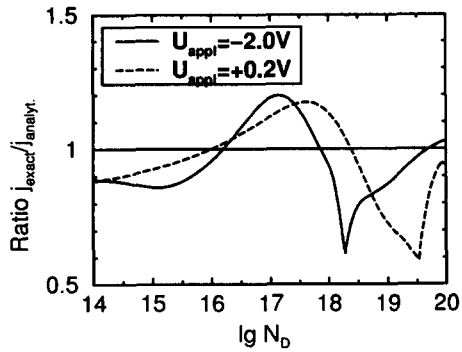

Fig. 6. Ratio of the calculated current densities $j_{exact}/j_{analyt.}$ at $U_{appl} = -2.0$ V (solid curve) and $U_{appl} = +0.2$ V (dashed curve) as a function of the donor concentration. For parameters see Fig. 5.

5. BOUNDARY CONDITIONS FOR DEVICE SIMULATION

The analytical model of the current through a parabolic barrier (45) can be used to set up boundary conditions for device simulation. We will consider here the drift-diffusion approach, where the unknowns are the electrostatic potential and the carrier densities (or the quasi Fermi levels, respectively). The region of the physical contact may be defined as that region, where transport is governed by emission and not by drift-diffusion. Here, a Fermi level is not defined. In the previous sections the whole space-charge layer was considered as such a region. In the case of low doping concentrations the barrier width becomes large and can cover a considerable fraction of the device. Therefore, we restrict the emission region now to a certain part of the space-charge layer. This part may be determined by the maximum of the spectral current density[26] or, following an idea of Schroeder[23], an energy Φ_T is defined where the spectral density of the emission current has dropped off to a small fraction of its maximum value. The corresponding coordinate x_T then may be considered as the "boundary" for the drift-diffusion simulator, as illustrated in Fig. 8.

Feasible boundary conditions can only be derived under simplifying assumptions. If we assume that a

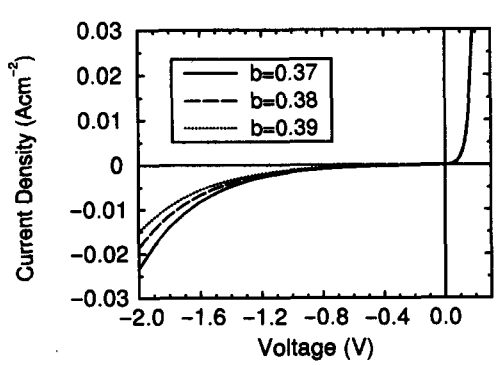

Fig. 7. Influence of the fit parameter b on the jV-characteristic for $N_D = 10^{18}\,cm^{-3}$. For other parameters see Fig. 5.

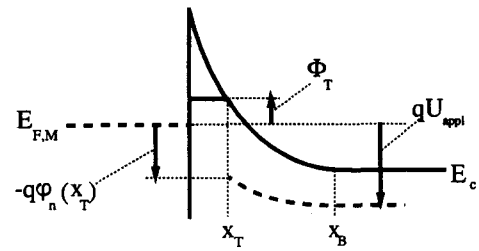

Fig. 8. Illustration of the energies Φ_T and $-q\varphi_n(x_T)$ at the reverse-biased contact. x_T denotes the extension of the contact region for device simulation.

degenerate semiconductor produces ideal Ohmic contacts, Φ_T is given by $E_c - E_{F,M}$ and it cannot become smaller. Then, we have only to care about the partial current densities j^+, j_1^-, and j_2^-. For j^+ a reasonable value of Φ_T is easily found from the condition:

$$\exp[-\vartheta^2(\varphi_T, \eta^+)] = e^{-\delta}, \tag{56}$$

where $\varphi_T = \Phi_T/kT$ and δ denotes a number. If, e.g $\delta = 3$, the spectral current density has dropped off to e^{-3} at the energy level Φ_T. The left-hand side of (56) is the Gaussian factor under the integral of the current density, and we used the function ϑ (54) for convenience. Equation (56) has the explicit solution:

$$\varphi_T(\eta^+) = \frac{|\gamma(\eta^+)|\eta_\lambda}{\sqrt{b}}[\vartheta(0, \eta^+) + \sqrt{\delta}]. \tag{57}$$

For j_1^- one obtains $\varphi_T = \varphi_T(\eta^-)$, whereas no explicit expression is available in the case of j_2^-. Since the contribution of j_2^- exceeds that of j_1^- as the doping level and the reverse bias are increased, a formula of φ_T for the total current, i.e. including all the different partial current densities and reflecting the falling-off of the leading term of the integrand to $e^{-\delta}$, would be desirable but cannot be found analytically. Instead, we use eqn (57), but limit φ_T by the value for an ideal Ohmic contact:

$$\varphi_T = \max[\eta_c - \eta_{F,M}, \varphi_T(\eta^+)]. \tag{58}$$

The effect of eqn (58) for the higher doping concentrations (above $10^{17}\,cm^{-3}$ in the case of silicon) is an overestimation of Φ_T and hence of x_T, which, on the other hand, only means that δ is effectively larger for these heavy doping cases. Note, that x_T is only a reasonable value for the transition from emission to the drift-diffusion mechanism of transport and depends on the choice of δ. The influence of that choice and, therefore, of the relation for Φ_T [eqn (58)] on numerical device simulation will be discussed elsewhere[27].

In Fig. 9 we plotted Φ_T against the voltage for $\delta = 3$ and the parameters of an Al-contact on n-type silicon varying the donor concentration from $10^{14}\,cm^{-3}$ to $10^{20}\,cm^{-3}$. In the TE regime Φ_T is only slightly smaller than the barrier height for all voltages. That limits the contact region to a small fraction of the barrier width.

With rising doping level Φ_T increases, and the contact comprises a growing part of the barrier. At a certain voltage Φ_T reaches the limit $E_c - E_{F,M}$, then the whole barrier defines the physical contact. As discussed above the slope of the curves for $N_D = (10^{18}\text{–}10^{19})\,\text{cm}^{-3}$ is actually smaller (more voltage drop over the barrier) than predicted by the approximation (58). However, it should be mentioned again that the overestimation of the matching point x_T is equivalent to the choice of a larger δ, which itself is somewhat arbitrary.

Now, we derive a Dirichlet condition for the quasi Fermi level φ_n at the boundary x_T by matching the emission current to the drift-diffusion current there:

$$j = -q\mu_n n \frac{d\varphi_n}{dx}\bigg|_{x=x_T}. \tag{59}$$

This is analogous to the treatment of Crowell and Sze[7] who matched the drift-diffusion current to a surface recombination current at the top of the barrier, and of Chang and Sze[12] who considered this matching at the maximum of the spectral current density. We again assume that degenerate semiconductors produce ideal Ohmic contacts and concentrate on the nondegenerate case. Then, the electron density can be expressed by Boltzmann statistics and eqn (59) can be integrated between x_T and x_B, if we further assume that the current density remains constant in that region (generation–recombination neglected):

$$\frac{\lambda_D j\, e^{\eta_c - \eta_{F,M}}}{kT\mu_n N_c}\sqrt{\frac{\pi}{2}}\,\text{Im erf}(i\sqrt{\varphi_T - (\eta_c - \eta_{F,M})}) = \exp(u) - \exp\left(\frac{-q\varphi_n(x_T)}{kT}\right), \tag{60}$$

where j is the emission current given by eqn (45) and $\text{erf}(x)$ the error function ([21], p. 297). We are prepared now to write down the boundary conditions for the electrostatic potential Ψ and the electron quasi Fermi potential φ_n:

$$\Psi(x_T) = -\frac{1}{q}\Phi_T, \tag{61}$$

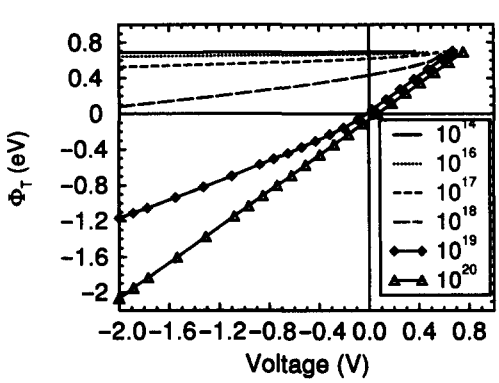

Fig. 9. Energy Φ_T, defining the contact region, as function of applied voltage for different doping levels. Parameter: $\delta = 3$, for other parameters see Fig. 5.

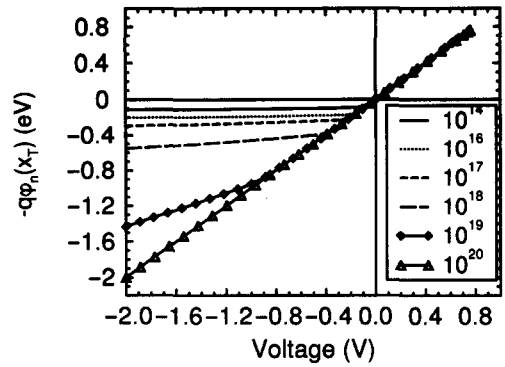

Fig. 10. Electron Fermi energy $-q\varphi_n$ at the boundary x_T as function of applied voltage for different doping levels. For parameters see Fig. 5.

$$\varphi_n(x_T) = -U_{appl} - \frac{kT}{q}\ln[1 - \beta(U_{appl})\exp(-u)], \tag{62}$$

with

$$\beta(U_{appl}) = \frac{\lambda_D j\, e^{\eta_c - \eta_{F,M}}}{kT\mu_n N_c}\sqrt{\frac{\pi}{2}} \times \text{Im erf}(i\sqrt{\varphi_T - (\eta_c - \eta_{F,M})}) \tag{63}$$

$$\approx \frac{\lambda_D j\, e^{\varphi_T}}{kT\sqrt{2[\varphi_T - (\eta_c - \eta_{F,M})]}\mu_n N_c}. \tag{64}$$

The last approximation involves a negligible error only. It is based on the asymptotic behavior ([21], p. 298):

$$\text{Im erf}(ix) \to \frac{e^{x^2}}{\sqrt{\pi}x}, \tag{65}$$

for large x. Only for large arguments of erf in (63), i.e. for reverse biases ($\eta_c < \eta_{F,M}$) deviations from a straight line occur in Fig. 10. In all other cases $\beta(U_{appl})$ remains negligibly small.

Figure 10 shows the slope of the quasi Fermi level under the same conditions as for Φ_T. The Caughey–Thomas model[28] was used for the mobility. Under forward bias there is almost no deviation from qU_{appl} at the boundary, i.e. only a negligible voltage drop between x_T and x_B in agreement with previous studies (e.g. [26]). Under reverse bias this voltage drop is largest for the lightly doped material and goes to zero as x_T approaches x_B for the degenerately doped silicon.

Both boundary conditions depend on the parameter "U_{appl}" which equals the "applied voltage" for the considered example of a simple Schottky diode with negligible series resistance. Devices which are subject to numerical simulation give rise to the complicated problem that under operating conditions the Fermi level it tilted at the boundary of the space-charge layer and the amount of dropped

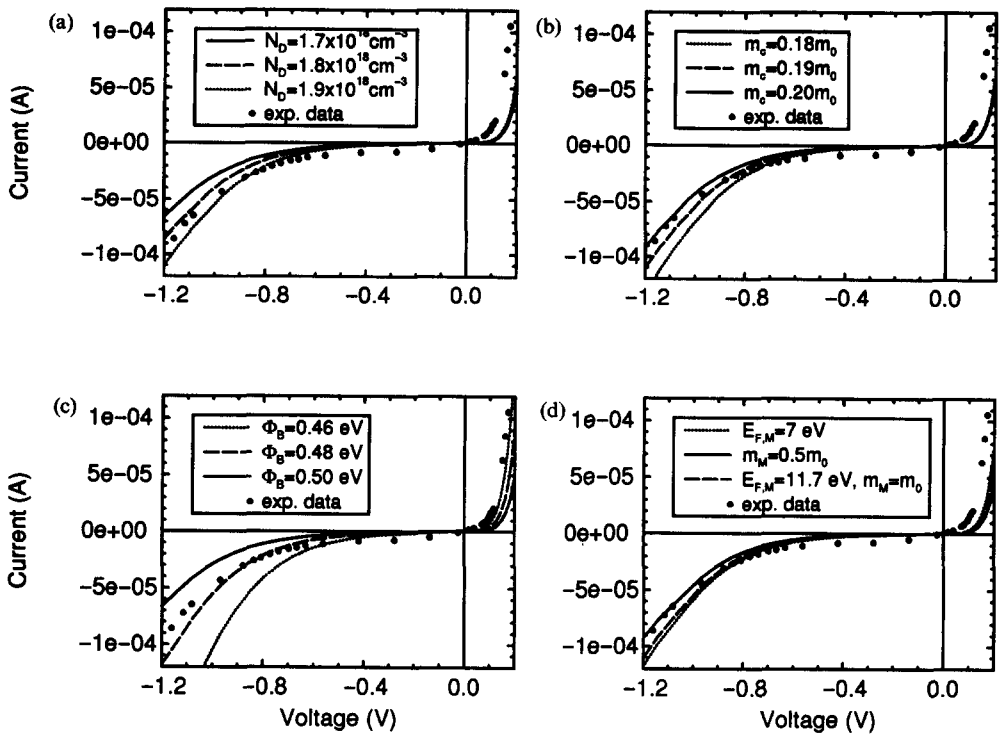

Fig. 11. Comparison of the analytical model with a measured $I(V)$-characteristic of Ti/n-Si with $N_D = (1.8\text{–}2.2) \times 10^{18}\,\text{cm}^{-3}$ and Area $= 3.4 \times 10^{-7}\,\text{cm}^2$ (dots). Theoretical curves are based on the parameter set: $m_c = 0.19m_0$, $\Phi_B = 0.50$ eV, $m_M = m_0$, and $E_{F,M} = 11.7$ eV. (a) Variation of the doping level, (b) variation of the tunneling mass with $N_D = 1.9 \times 10^{18}\,\text{cm}^{-3}$, (c) variation of the barrier height with $N_D = 1.7 \times 10^{18}\,\text{cm}^{-3}$, (d) effect of a change in the metal Fermi energy and effective mass with $N_D = 1.9 \times 10^{18}\,\text{cm}^{-3}$.

voltage, applied at some contacts of the device, is on principle unknown there. This problem will be considered elsewhere[27].

The corresponding boundary condition for holes can be obtained if for the hole injection current the rate of diffusion into the semiconductor bulk as in a pn-junction is used[5]. This will not be discussed in detail here.

6. COMPARISON WITH EXPERIMENT

The $I(V)$-characteristics of a titanium contact on intermediately doped silicon is used to test the model and to discuss the effect of the most important physical parameters on the fit†. A flat doping profile in ⟨100⟩-silicon was obtained by implanting phosphorus with a dose of $1.83 \times 10^{14}\,\text{cm}^{-2}$ and a subsequent high temperature annealing at 1100°C for 160 min in pure nitrogen. The surface donor concentration was determined by SIMS measurements and process simulation‡ to be $(1.8\text{–}2.2) \times 10^{18}\,\text{cm}^{-3}$. Contact windows were opened by reactive ion etching. Finally, a Kelvin structure was formed to avoid the voltage drop over a series resistance.

Figure 11 shows a comparison of the data and eqn (45). Remarkable deviations occur both for low reverse bias and forward bias. The reasons may be manifold, but most likely a recombination current is superimposed resulting from deep levels due to the processing (reactive ion etching). Hence, the "breakdown" branch is more suitable for a comparison with the model.

There are five physical parameters in the analytical model that characterize the metal–semiconductor contact: the doping concentration N_D, effective (tunneling) mass m_c, barrier height Φ_B, effective mass on the metal side m_M, and the metal Fermi energy $E_{F,M}$. A reasonable fit is obtained using the set $N_D = 1.9 \times 10^{18}\,\text{cm}^{-3}$, $m_c = 0.19m_0$ (⟨100⟩-oriented material was used for the contacts), $\Phi_B = 0.50$ eV[29], $m_M = m_0$, and $E_{F,M} = 11.7$ eV. In Fig. 11(a) the doping concentration was varied from 1.7 to $1.9 \times 10^{18}\,\text{cm}^{-3}$ showing its strong influence on the tunneling current. The best fit is achieved with a value between $(1.8\text{–}1.9) \times 10^{18}\,\text{cm}^{-3}$ which agrees well with the measured range. A similar impact on the characteristics results from a change of the tunneling mass [Fig. 11(b)]. As the main tunneling path moves down

†The author is grateful to Dr W. Grabinski (ABB semiconductors, Baden-Dättwil) for providing the experimental data.

‡The author wishes to thank Dr Th. Feudel (ETH, Zürich) for discussions on the process and for providing the process simulation data.

with rising reverse bias, mixing-in of valence states (or even metal states) might reduce the effective electron mass. Lowering the barrier height [Fig. 11(c)] increases both the forward and reverse currents. Barrier height reduction is likely from several reasons, e.g. occupation of interface states or image forces. Although the observed discrepancy at forward bias could be removed with a reduction to $\Phi_B = 0.44\,\text{eV}$ which exactly corresponds to the image-force lowering at $U_{appl} = 0\,\text{V}$, the fit of the reverse branch then would require a far too low doping level or too large tunneling mass. Figure 11(d) shows the effect if the two metal parameters m_M and $E_{F,M}$ are changed drastically. Since the expression for the transmittance depends only weakly on both quantities, the $I(V)$-characteristics are not much affected.

7. CONCLUSIONS

Based on the concept of thermionic field emission a 1D analytical model of the metal–semiconductor contact was developed by-passing the WKB approximation. This was done by interpolating analytically between the asymptotic forms of the eigenfunctions of the parabolic potential and mapping the interpolation functions approximately to Gaussians. The resulting transmittance of the parabolic barrier was shown to be in remarkable agreement with the exact one. The method has the potential for an application to more complicated barrier shapes. Using a simplified carrier statistics model and expressing the transmittance by an integrable function (a Gaussian with respect to the energy) enabled the contact current to be calculated analytically. The final form contains error functions as most complicated ingredients and is therefore suitable for an implementation in device simulation programs to model boundary conditions for non-ideal contacts. It has been demonstrated that Dirichlet boundary conditions can be derived both for the electrostatic potential and the quasi Fermi level under certain assumptions. Therefore, a contact region was defined as that part of the barrier where tunneling gives a noticable contribution to the emission current. The local value of the quasi Fermi level at the inner boundary of that region was found by matching the drift-diffusion current to the TFE current there. In device simulation such a balancing is done numerically. A complicated problem arises from the fact that the input model for the barrier contains the potential value at the boundary of the space-charge zone, which is not accessible locally in numerical simulation because of the series resistance of the device. Simplifying assumptions can yield a way out, but further work is necessary to tackle this problem.

Acknowledgement—This work has been financially supported by the Swiss Research Program LESIT.

REFERENCES

1. J. L. Freeouf, *Appl. Phys. Lett.* **41**, 285 (1982).
2. X. Wu and E. S. Yang, *J. appl. Phys.* **65**, 3560 (1989).
3. C. R. Crowell, H. B. Shore and E. E. LaBate, *J. appl. Phys.* **36**, 3843 (1965).
4. J. M. Andrews and M. P. Lepselter, *Solid-St. Electron.* **13**, 1011 (1970).
5. H. C. Card and E. H. Rhoderick, *Solid-St. Electron.* **16**, 365 (1973).
6. S. J. Fonash, *J. appl. Phys.* **54**, 1966 (1983).
7. C. R. Crowell and S. M. Sze, *Solid-St. Electron.* **9**, 1035 (1966).
8. E. Y. Lee and L. J. Schowalter, *J. appl. Phys.* **70**, 2156 (1991).
9. A. Cola, M. G. Lupo, L. Vasanelli and A. Valentini, *Solid-St. Electron.* **36**, 785 (1993).
10. D. Bauza, *J. appl. Phys.* **73**, 1858 (1993).
11. A. G. Chynoweth, W. L. Feldmann and R. A. Logan, *Phys. Rev.* **121**, 684 (1961).
12. C. Y. Chang and S. M. Sze, *Solid-St. Electron.* **13**, 727 (1970).
13. J. H. Werner and H. H. Güttler, *J. appl. Phys.* **69**, 1522 (1991).
14. J. P. Sullivan, R. T. Tung, M. R. Pinto and W. R. Graham, *J. appl. Phys.* **70**, 7403 (1991).
15. J. W. Conley and G. D. Mahan, *Phys. Rev.* **161**, 681 (1967).
16. D. Schroeder, in *Proc. 4th Int. Conf. on Simulation of Semiconductor Devices and Processes*, Zürich, Switzerland (1991).
17. J. W. Conley, C. B. Duke, G. D. Mahan and J. J. Tiemann, *Phys. Rev.* **150**, 466 (1966).
18. F. A. Padovani and R. Stratton, *Solid-St. Electron.* **9**, 695 (1966).
19. C. R. Crowell and V. L. Rideout, *Solid-St. Electron.* **12**, 89 (1969).
20. A. Y. C. Yu, *Solid-St. Electron.* **13**, 239 (1970).
21. M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover Publications, New York (1972).
22. R. Stratton, *Phys. Rev.* **125**, 67 (1962).
23. D. Schroeder, in *Proc. of the NASECODE VIII Conf.*, Vienna, Austria (1992).
24. R. H. Fowler and L. Nordheim, *Proc. R. Soc. A* **119**, 173 (1928).
25. J. L. Moll, *Physics of Semiconductors*. McGraw-Hill, New York (1964).
26. C. R. Crowell and M. Beguwala, *Solid-St. Electron.* **14**, 1149 (1971).
27. A. Schenk and S. Müller, in *Simulation of Semiconductor Devices and Processes*, Vol 5, pp. 441–444, Vienna, Austria (1993).
28. D. M. Caughey and R. E. Thomas, *Proc. IEEE*, pp. 2192–93 (1967).
29. S. M. Sze, *Physics of Semicondcutor Devices*. Wiley, New York (1981).

APPENDIX A

Transmission Probability for a Parabolic Barrier

The matching conditions for the wave functions (wave supposed to be incoming from the metal side):

$$\psi_{\mathrm{I}}(x) = A\,\mathrm{e}^{ik_M x} + B\,\mathrm{e}^{-ik_M x} \quad \text{metal} \tag{A1}$$

$$\psi_{\mathrm{II}}(x) = \alpha U_\xi + \beta V_\xi \quad \text{barrier} \tag{A2}$$

$$\psi_{\mathrm{III}}(x) = C\,\mathrm{e}^{ik_S x} \quad \text{semiconductor bulk} \tag{A3}$$

at $\xi = \xi_B$ $(x = 0)$ and $\xi = 0$ $(x = x_B)$ read:

$$A + B = \alpha U_{\xi_B} + \beta V_{\xi_B} \tag{A4}$$

$$\frac{1}{m_M}(ik_M A - ik_M B) = -\frac{1}{\lambda m_{eff}}(\alpha U'_{\zeta_B} + \beta V'_{\zeta_B}) \quad (A5)$$

$$\alpha U_0 + \beta V_0 = C\, e^{ik_{S,x}x_B} \quad (A6)$$

$$-\frac{1}{\lambda}(\alpha U'_0 + \beta V'_0) = Cik_S\, e^{ik_{S,x}x_B} \quad (A7)$$

since $m_{eff}(x_B) = m_c$. The transmission probability is defined as:

$$\mathcal{T}(\kappa_S, \kappa_M) = \frac{|C|^2}{|A|^2}\cdot\frac{v_{S,x}}{v_{M,x}} = \frac{|C|^2}{|A|^2}\cdot\frac{\kappa'_S}{\kappa_M}. \quad (A8)$$

Resolving the system (A4)–(A7) and inserting the amplitudes gives:

$$\mathcal{T}(\kappa_S, \kappa_M) = \frac{8}{\pi}\frac{\kappa'_S}{\kappa_M}|(V'_0 + i\kappa'_S V_0)(U_{\zeta_B} + i\kappa_M^{-1}U'_{\zeta_B}) - (U'_0 + i\kappa'_S U_0)(V_{\zeta_B} + i\kappa_M^{-1}V'_{\zeta_B})|^{-2}, \quad (A9)$$

which is eqn (8). Here we have used the Wronskian $U_0V'_0 - V_0U'_0 = \sqrt{2/\pi}$ ([21], p. 687) and the definition of the normalized momenta $\kappa'_S = \kappa_{S,x}\lambda$, $\kappa_M = k_{M,x}\lambda m_{eff}/m_M$. If the absolute square in (A9) is evaluated, the mixed terms are reordered, and again the Wronskian of the parabolic cylinder functions is used, we end up with:

$$\mathcal{T}(\kappa_S, \kappa_M) = \frac{2}{1 + g(\kappa_S, \kappa_M)} \quad (A10)$$

$$g(\kappa_S, \kappa_M) = \frac{\pi}{4}\left[\left(\frac{\kappa_M}{\kappa'_S}\right)(V'_0U_{\zeta_B} - U'_0V_{\zeta_B})^2 + \left(\frac{\kappa'_S}{\kappa_M}\right)(U_0V'_{\zeta_B} - V_0U'_{\zeta_B})^2 + \left(\frac{1}{\kappa_M\kappa'_S}\right)(V'_0U'_{\zeta_B} - U'_0V'_{\zeta_B})^2 + \kappa_M\kappa'_S(V_0U_{\zeta_B} - U_0V_{\zeta_B})^2\right]. \quad (A11)$$

APPENDIX B

Asymptotic Forms and Interpolation Functions

The asymptotic formulae of the parabolic cylinder functions $U_{\zeta_B} = U(-\kappa_S^2, \xi_B)$, $V_{\zeta_B} = V(-\kappa_S^2, \xi_B)$, $U'_{\zeta_B} = U'(-\kappa_S^2, \xi_B)$, and $V'_{\zeta_B} = V'(-\kappa_S^2, \xi_B)$ are given by ([21], p. 690):

$$U^{appr}_{\zeta_B} = \frac{\sqrt{\Gamma(\frac{1}{2} + \kappa_S^2)}}{(2\pi)^{1/4}|\xi_B^2 - 4\kappa_S^2|^{1/4}} \times \begin{cases} \frac{1}{2}e^{-S} & \text{for } \xi_B > 2\kappa_S \\ \sin\left(\frac{\pi}{4} + |S|\right) & \text{for } \xi_B < 2\kappa_S \end{cases}, \quad (B1)$$

$$V^{appr}_{\zeta_B} = \frac{2|\xi_B^2 - 4\kappa_S^2|^{-1/4}}{(2\pi)^{1/4}\sqrt{\Gamma(\frac{1}{2} + \kappa_S^2)}} \times \begin{cases} e^{S} & \text{for } \xi_B > 2\kappa_S \\ \cos\left(\frac{\pi}{4} + |S|\right) & \text{for } \xi_B < 2\kappa_S \end{cases}, \quad (B2)$$

$$U'^{appr}_{\zeta_B} = -\frac{\sqrt{\Gamma(\frac{1}{2} + \kappa_S^2)}}{2(2\pi)^{1/4}|\xi_B^2 - 4\kappa_S^2|^{-1/4}} \times \begin{cases} \frac{1}{2}e^{-S} & \text{for } \xi_B > 2\kappa_S \\ \cos\left(\frac{\pi}{4} + |S|\right) & \text{for } \xi_B < 2\kappa_S \end{cases}, \quad (B3)$$

$$V'^{appr}_{\zeta_B} = \frac{|\xi_B^2 - 4\kappa_S^2|^{1/4}}{(2\pi)^{1/4}\sqrt{\Gamma(\frac{1}{2} + \kappa_S^2)}} \times \begin{cases} e^{S} & \text{for } \xi_B > 2\kappa_S \\ \sin\left(\frac{\pi}{4} + |S|\right) & \text{for } \xi_B < 2\kappa_S \end{cases}, \quad (B4)$$

with

$$S = \frac{1}{2}\int_{2\kappa_S}^{\zeta_B} d\xi\sqrt{\xi^2 - 4\kappa_S^2}. \quad (B5)$$

The prime denotes the derivative with respect to the second argument. The interpolating functions with the same asymptotic behavior read:

$$U^{Ai}_{\zeta_B} = \frac{(2\pi)^{1/4}\sqrt{\Gamma(\frac{1}{2} + \kappa_S^2)}}{\left|\frac{\xi_B^2}{4} - \kappa_S^2\right|^{1/4}}(\tfrac{3}{2}|S|)^{1/6}\,\mathrm{Ai}([\tfrac{3}{2}S]^{2/3}), \quad (B6)$$

$$V^{Ai}_{\zeta_B} = \frac{(2\pi)^{1/4}\left|\frac{\xi_B^2}{4} - \kappa_S^2\right|^{-1/4}}{\sqrt{\Gamma(\frac{1}{2} + \kappa_S^2)}}(\tfrac{3}{2}|S|)^{1/6}\,\mathrm{Bi}([\tfrac{3}{2}S]^{2/3}), \quad (B7)$$

$$U'^{Ai}_{\zeta_B} = \frac{(2\pi)^{1/4}\sqrt{\Gamma(\frac{1}{2} + \kappa_S^2)}}{\left|\frac{\xi_B^2}{4} - \kappa_S^2\right|^{-1/4}}(\tfrac{3}{2}|S|)^{-1/6}\,\mathrm{Ai}'([\tfrac{3}{2}S]^{2/3}), \quad (B8)$$

$$V'^{Ai}_{\zeta_B} = \frac{(2\pi)^{1/4}\left|\frac{\xi_B^2}{4} - \kappa_S^2\right|^{1/4}}{\sqrt{\Gamma(\frac{1}{2} + \kappa_S^2)}}(\tfrac{3}{2}|S|)^{-1/6}\,\mathrm{Bi}'([\tfrac{3}{2}S]^{2/3}). \quad (B9)$$

APPENDIX C

Energy Limit for Gaussian Approximation

Taylor expansion of the action $S(\xi_B)$ in the vicinity of $\eta = \varphi_B$ in the range $2\kappa_S > \xi_B$ yields:

$$S(\xi_B) \approx \frac{4}{3\xi_B}\left(\frac{\xi_B^2}{4} - \kappa_S^2\right)^{3/2}, \quad (C1)$$

hence, the function Y becomes:

$$Y(S) \approx \left(\frac{2}{\xi_B}\right)^{2/3}\left(\frac{\xi_B^2}{4} - \kappa_S^2\right) \quad (C2)$$

there. The energy limit is given by the maximum of the Gaussian (32), i.e. by $Y(S) = t_0 = -|t_0|$. This leads to:

$$\epsilon_{max} - \epsilon_c = \frac{\xi_B^2}{4} + \left(\frac{\xi_B}{2}\right)^{2/3}|t_0|. \quad (C3)$$

Changing to the energy variable η (measured from $\eta_{F,M}$), we get:

$$\eta_{max} = \varphi_B + \left(\frac{\xi_B}{2}\right)^{2/3}\eta_\lambda|t_0| \quad (C4)$$

$$= \varphi_B + (\eta_{F,M} - \eta_c + \varphi_B)^{1/3}\eta_\lambda^{2/3}|t_0|. \quad (C5)$$

APPENDIX D

WKB Approximation for the Range $\eta > \eta_{max}$

The WKB form of the transmission probability, valid for energies much larger than the maximum of the barrier, is most easily obtained from eqn (31) inserting the asymptotic representations of the Airy functions for large negative arguments ([21], p. 448):

$$\mathrm{Ai}(-Y) \to \pi^{-1/2}Y^{-1/4}\sin\left(\tfrac{2}{3}Y^{3/2} + \frac{\pi}{4}\right), \quad (D1)$$

$$\mathrm{Ai}'(-Y) \to -\pi^{-1/2}Y^{1/4}\cos\left(\tfrac{2}{3}Y^{3/2} + \frac{\pi}{4}\right), \quad (D2)$$

$$\mathrm{Bi}(-Y) \to \pi^{-1/2}Y^{-1/4}\cos\left(\tfrac{2}{3}Y^{3/2} + \frac{\pi}{4}\right), \quad (D3)$$

$$\mathrm{Bi}'(-Y) \to \pi^{-1/2}Y^{1/4}\sin\left(\tfrac{2}{3}Y^{3/2} + \frac{\pi}{4}\right). \quad (D4)$$

Then we get:

$$|\mathrm{Ai}(Y\,\mathrm{e}^{-i(2\pi/3)})| = \tfrac{1}{2}\sqrt{\mathrm{Ai}^2(Y)+\mathrm{Bi}^2(Y)} \to \tfrac{1}{2}\pi^{-1/2}|Y|^{-1/4}, \tag{D5}$$

$$|\mathrm{Ai}'(Y\,\mathrm{e}^{-i(2\pi/3)}) = \tfrac{1}{2}\sqrt{\mathrm{Ai}'^2(Y)+\mathrm{Bi}'^2(Y)} \to \tfrac{1}{2}\pi^{-1/2}|Y|^{1/4}. \tag{D6}$$

Inserting into eqn (31) yields immediately (37). The limit $\eta \to \infty$ of $\mathscr{T}$ turns out to be:

$$\mathscr{T}(\eta+\eta_{\mathrm{F,M}},0) \xrightarrow{\eta\to\infty} \frac{2}{1+\frac{1}{2}\left(\sqrt{\frac{m_c}{m_M}}+\sqrt{\frac{m_M}{m_c}}\right)}, \tag{D7}$$

which actually tends to 1, since the effective masses have to approach the free electron mass for $\eta \to \infty$. The latter effect has not been modeled in the present paper, consequently the limit (D7) expresses quantum reflection at the boundary of two media with different effective masses.

For the purpose of analytical integration the WKB form $\mathscr{T}^{\mathrm{WKB}}$ has to be approximated in the vicinity of $\eta_{\max}$. Therefore, we write $\mathscr{T}^{\mathrm{WKB}}$ as:

$$\mathscr{T}^{\mathrm{WKB}}(\eta+\eta_{\mathrm{F,M}},0) = \frac{4\sqrt{\frac{m_M}{m_c(\eta+\eta_{\mathrm{F,M}})}}\sqrt{\eta-\varphi_B}}{\left(1+\sqrt{\frac{m_M}{m_c(\eta+\eta_{\mathrm{F,M}})}}\sqrt{\eta-\varphi_B}\right)^2}. \tag{D8}$$

An integrable approximation is obtained, if η is neglected compared to $\eta_{\mathrm{F,M}}$ (because $\eta_{\max} \ll \eta_{\mathrm{F,M}}$ can be assumed) and η is replaced by $\eta_{\max}$ in the denominator (because $\mathscr{T}^{\mathrm{WKB}}$ is only important for the lowly doped contacts, where contributions to the current originate from a range of a few kT above the top of the barrier only).