Microscopic balance equations for the real-space transfer in multiple quantum wells

1998 J. Phys.: Condens. Matter 10 7391

(http://iopscience.iop.org/0953-8984/10/33/010)

J. Phys.: Condens. Matter **10** (1998) 7391–7406. Printed in the UK PII: S0953-8984(98)91545-4

Microscopic balance equations for the real-space transfer in multiple quantum wells

V Kubrak and P Kleinert

Paul-Drude-Institut für Festkörperelektronik, Hausvogteiplatz 5–7, D-10117 Berlin, Germany

Received 10 February 1998

Abstract. The parallel electronic transport in semiconductor multiple quantum wells and the associated real-space transfer of electrons from high-mobility quantum wells into low-mobility barriers is treated on the basis of the microscopic Lei–Ting balance-equation theory. Our model system consists of a quasi-two-dimensional subband that is bound to the wells and a quasi-three-dimensional band for the extended states above the barriers. The real-space transfer between the two subbands is microscopically described as intersubband scattering by phonons. In order to capture the appropriate selection rules for these transitions, it is necessary to choose a set of orthogonal subband wavefunctions. A solution of the balance equations is presented for a GaAs–$Al_xGa_{1-x}As$ system in which the real-space transfer causes negative differential resistance. Our approach shows how the transport properties of the system are interrelated with the different transfer scattering processes, and how system parameters influence the real-space transfer. Furthermore, it is shown that the incorporation of screening requires a careful selection of the screening model, as both intersubband contributions and dynamical effects are found to modify the results in the random-phase approximation.

1. Introduction

In modulation-doped quantum well structures electrons can be thermionically emitted from the wells and thus repopulate the adjacent barriers if an electric field is applied parallel to the layers. This real-space transfer (RST) may lead to N-shaped negative differential resistance (NDR) if the transport mobility in the barriers is much lower than in the quantum wells [1]. A large number of theoretical and experimental results on this effect are reviewed in reference [2]. Most previous theoretical papers that addressed the RST on a microscopic basis relied on Monte Carlo simulations of the Boltzmann equation, and only considered RST between wide quantum wells, in which quantization effects are not significant, and adjacent wide barriers.

Quantization effects, however, should be considered for most systems of practical interest because such systems usually involve narrow quantum wells or narrow potential wells formed at single heterojunctions. In this case the RST incorporates a change of the effective dimensionality of the electrons. In thermal equilibrium, the electrons are confined in the wells and have a quasi-two-dimensional (Q2D) character. The application of an electric field heats the electron gas, so some electrons are emitted into states which extend over the barriers and exhibit quasi-three-dimensional (Q3D) behaviour.

Describing this Q2D–Q3D RST is difficult because of the complicated energy dispersion relation of the Q3D states. In a periodic multiple-quantum-well system the extended states above the barriers belong to numerous minibands separated by narrow gaps, leading to a

0953-8984/98/337391+16$19.50

complicated multi-subband problem. The main physical aspects of the problem, however, can be more easily clarified if the complexity is reduced by assuming that all Q3D states belong to a single band (termed a Q3D subband for convenience). This approximation cannot take strong band bending in highly doped heterostructures into account; in such a case there would exist states almost localized in the potential wells of the doped barrier regions. In that case, one could treat the states in the barriers as Q2D states, too, albeit at the expense of approximating minibands by sharp energy levels.

To our knowledge, an adequate treatment of the Q2D–Q3D RST within transport theories has not yet been published even for the simplified system with a single Q3D subband. None of the previous microscopic transport calculations that included Q2D–Q3D RST provided a systematic comparison between the different transfer scattering mechanisms. Some approaches did not treat the RST microscopically—for example, they assumed that any electron with an energy above the bottom of the Q3D subband would always stay in this subband [3], or calculated the population of a Q3D subsystem that had the same chemical potential as the Q2D subsystem despite having a different electron temperature [4]. A number of papers included the RST microscopically but employed plane waves for the wavefunction of the Q3D subband [5–7], which are not orthogonal to the Q2D wavefunctions and therefore do not preserve the quantum mechanical selection rules of intersubband transitions. This serious deficiency was addressed in reference [5], but only in a semi-classical approximation. Finally, in reference [8] it was assumed that the dominant transfer scattering mechanism was elastic, thus *a priori* ruling out processes that involve LO phonons. Within this approximation, the authors found an analytical solution. In contrast, the RST problem has at least been solved for a system with a single (compositional) quantum well with strong bend bending, in which minibands are not an issue [9]. In this case, electrons in the barriers are confined in a potential well in the doped region, and the RST can be treated as Q2D-to-Q2D intersubband transfer. Considerable progress has also been made in describing the dependence of the electron emission from and capture by quantum wells [10–16] (Q2D-to-Q3D and Q3D-to-Q2D scattering respectively) on the electron energy and the well width, but these results were not obtained in the context of transport calculations.

In this paper we treat the Q2D–Q3D RST in a multiple-quantum-well system with narrow quantum wells and wide barriers. A two-subband model is considered with one parabolic Q2D subband for the states in the quantum wells and one parabolic Q3D subband for the states extending over the barriers. In order to preserve the selection rules of the intersubband transfer, it is essential that the model wavefunctions of the two subbands are orthogonal to each other. In our model the orthogonality is reintroduced explicitly by orthogonalizing plane waves to the wavefunction of the Q2D subband.

The transport characteristics of the system are determined in the framework of the Lei–Ting balance-equation theory [17]. The balance-equation method has proven to be a versatile approach for various phenomena of hot-carrier transport in semiconductors [18], including quasi-two-dimensional multi-subband systems [19] and bulk multi-valley systems [20]. The method, which describes the non-linear transport via a parametrized Fermi-type distribution function, provides a reliable, computationally much more efficient alternative to solving the Boltzmann equation by means of Monte Carlo simulation. It is applicable if the carrier density is sufficiently high to rapidly thermalize the carriers within each subband through carrier–carrier interactions. This prerequisite is not completely fulfilled in the system that we consider. However, it has been demonstrated in the literature that the balance-equation approach can be applied successfully even in such cases. We derive scattering rates for the Q2D–Q3D transfer of electrons by emission and absorption of phonons, which are used

to formulate a particle-transfer balance. This equation, the momentum and energy balance equations for each subband, and particle conservation form a closed set of equations for the parameters of the non-linear transport, i.e. the drift velocity, electron temperature, chemical potential, and relative occupation of each subband.

Another feature of the Lei–Ting balance-equation approach is that dynamical screening can be easily incorporated within the random-phase approximation (RPA). In the case of RST, the screening needs special attention because each of the two subsystems with rather different properties (Q2D and Q3D) may also significantly contribute to the screening of the intrasubband scattering in the other subsystem. Such intersubband contributions to the screening are known to be important in quasi-two-dimensional electron gases if the transport mobility is limited by ionized impurity scattering [21]. In addition, the RST may be affected by dynamical screening effects through their enhancement of LO-phonon emission rates [22], since the transfer is induced by heating the Q2D electrons.

We apply our approach to a GaAs–$Al_xGa_{1-x}As$ multiple-quantum-well system with narrow, undoped wells and highly doped barriers, and compare the different transfer scattering mechanisms. The qualitative dependences of the current–voltage characteristics, which exhibit NDR, on the energetic separation of the two subbands and on the electron density are investigated. We also show how the results are affected if the dynamical or the intersubband contributions to the screening are neglected.

2. The model

Consider a periodic multiple-quantum-well system with wells of width a and a period $d \gg a$. For sufficiently small a, the wells will accommodate just one subband. Due to the wide barriers, the tunnelling between the wells and the energetic width of this subband are negligibly small. This means that the electrons in different wells decouple and can be treated as being quasi-two-dimensional. For energies higher than the barriers, there is a sequence of minibands and minigaps, the gaps being smaller than 1 meV for the parameters considered below. The wavefunctions belonging to the states of these minibands extend over the wide barriers, so these states have a quasi-three-dimensional character. The detailed miniband structure of the Q3D states in the growth direction (the z-direction) does not significantly influence the parallel-transport properties (in the x–y plane). In the direction of the applied electric field, all minibands of the Q3D states have dispersion relations that are nearly the same. If the quantum wells are undoped and the barriers are highly doped, the mobility in the Q2D subband is high, while the mobilities in the Q3D minibands are all low and are assumed to be of similar value. Thus, the overall current is mainly determined by the total fraction of electrons occupying the Q3D minibands, but not by the exact distribution among these minibands. As the minigaps are narrow compared to the width of the minibands and the thermal energy, they also do not significantly influence the Q2D–Q3D transfer scattering. It is therefore concluded that an exact treatment of this miniband structure is not necessary for obtaining a qualitative picture of the relationship between the Q2D–Q3D RST and the transport properties of the system.

Disregarding the minigaps, we combine the Q3D minibands into a single parabolic, isotropic subband. Hence our model consists of two subbands with the energy dispersion relations

$$\varepsilon_{1\boldsymbol{k}_\parallel} = \frac{\hbar^2 \boldsymbol{k}_\parallel^2}{2m_1^*} \tag{1}$$

$$\varepsilon_{2\boldsymbol{k}} = \frac{\hbar^2 \boldsymbol{k}^2}{2m_2^*} + E_2 \tag{2}$$

where the Q2D (Q3D) subband is labelled by $j = 1$ ($j = 2$), $\|$ denotes vectors in the x–y plane, and E_2 is the energetic separation between the lower edges of the two subbands. In this paper we only consider the case where the effective-mass discontinuity between the well and barrier materials is small, and assume the same effective mass $m^* = m_1^* = m_2^*$ for the two subbands.

Neglecting the exponential tails in the barriers, the wavefunction of the Q2D subband is given by

$$\Phi_1(\boldsymbol{r}, \boldsymbol{k}_\|) = \sqrt{1/A}\,\mathrm{e}^{\mathrm{i}\boldsymbol{k}_\| \cdot \boldsymbol{r}_\|} \sum_\nu \phi_1(z - \nu d) \tag{3}$$

where the sum is taken over the well index ν, and

$$\phi_1(z) = \begin{cases} \sqrt{2/a}\cos(\pi z/a) & \text{for } -a/2 \leqslant z \leqslant a/2 \\ 0 & \text{otherwise.} \end{cases} \tag{4}$$

For the Q3D subband, it is tempting to choose plane waves as model wavefunctions [5–7]. This choice, however, does not account for effects that the quantum wells have on the Q3D eigenstates. The wavefunctions of these states must oscillate more rapidly in the wells than in the barriers because the difference between their energy and the bottom of the conduction band is higher in the wells. From a semi-classical point of view, this means that the electrons in Q3D states have a non-zero minimum quasi-momentum k_z in the growth direction when crossing the wells. As was pointed out in reference [12], this requires a minimum wavevector change $|q_z| > 0$ when a Q2D electron is scattered into a Q3D state, for example by absorption or emission of a phonon (q_z being the z-component of the phonon wavevector). This was incorporated in reference [5] by setting the matrix element to zero for $|q_z| < \sqrt{2m^*\,\Delta E_\mathrm{C}/\hbar^2}$, where ΔE_C is the conduction band discontinuity. Implicitly, this approximation requires one to assume that the Q2D electrons have a zero wavevector, $k_z = 0$, in the growth direction.

From the quantum mechanical point of view, however, such an approximate criterion is not required. We note that quantum mechanics prohibits any intersubband transitions by phonons with $q_z = 0$ via the orthogonality of the subband wavefunctions. This can be easily seen from the form factors of the intersubband scattering (cf. equation (8) below). The use of non-orthogonal wavefunctions would also wrongly allow direct intersubband transitions. In order to preserve the selection rule $|q_z| > 0$ for phonon-mediated intersubband scattering processes, we construct a Q3D wavefunction that is orthogonal to the Q2D wavefunction. We employ the simplest choice, which is that of orthogonalized plane waves:

$$\Phi_2(\boldsymbol{r}, \boldsymbol{k}) = \sqrt{1/A}\,\mathrm{e}^{\mathrm{i}\boldsymbol{k}_\| \cdot \boldsymbol{r}_\|} \phi_2(z, k_z) \tag{5}$$

$$\phi_2(z, k_z) = \Gamma(k_z)\left[\mathrm{e}^{\mathrm{i}k_z z} - \sum_\nu \mathrm{e}^{\mathrm{i}k_z \nu d} I^{(12)}(k_z)\phi_1(z - \nu d)\right]. \tag{6}$$

The orthogonalization coefficient $I^{(12)}$ and the normalization factor Γ are given in the appendix. Note that the probability of finding a Q3D electron in the barriers, where it still has plane-wave character, is close to one, because we assume that $d \gg a$. The Q3D electrons can thus be approximately described by plane waves when calculating the matrix elements of intrasubband scattering processes.

As intrasubband scattering processes are expected to occur much more frequently than intersubband processes, it is possible to treat the subbands as two separate subsystems, each

of which has its own drift velocity v_j, electron temperature T_j, and chemical potential μ_j. For either subband the centre-of-mass motion is separated from the relative motion and is treated semi-classically. Balance equations for the momentum and the energy are derived for each subband separately as in reference [7] using the Lei–Ting approach. We take into account intrasubband scattering on LO (Fröhlich coupling) and LA (deformation potential coupling) phonons and on ionized impurities. The determination of the vibrational properties of the heterostructure is a non-trivial problem itself, which requires detailed calculations. Regarding our qualitative investigation of the electronic high-field transport, however, we believe that details of the electron–phonon interaction are not of great relevance, and therefore describe lattice vibrations in terms of bulk phonons that are in thermal equilibrium at the lattice temperature. We do not include the effects of non-equilibrium phonons, the number of which may be large if the electron volume concentration is high. Then, hot phonons can reduce the energy relaxation rate [23, 24] and the mobility [25] of the Q2D electrons.

The RST is accounted for by deriving a particle balance for the intersubband transitions due to absorption and emission of LO and LA phonons. Intersubband transitions due to scattering on ionized impurities can be included in analogy to reference [26], but they are found to be negligible here because the Q2D electron gas and the donors are spatially separated from each other, meaning that the form factor of the matrix element assumes very small values [27]. The part of the Hamiltonian that describes the Q3D-to-Q2D transition effected by absorption or emission of a phonon is given by

$$H_{\mathrm{ep}}^{12} = \sum_{\boldsymbol{q}\lambda} \sum_{\boldsymbol{k}\nu\sigma} M(\boldsymbol{q},\lambda) J^{(12)}(k_z, q_z) \mathrm{e}^{\mathrm{i}\boldsymbol{k}\cdot(\boldsymbol{R}_{1\nu}-\boldsymbol{R}_2)+\mathrm{i}\boldsymbol{q}\cdot\boldsymbol{R}_{1\nu}} (b_{\boldsymbol{q}\lambda} + b^{\dagger}_{-\boldsymbol{q}\lambda}) c^{\dagger}_{1\boldsymbol{k}_\parallel+\boldsymbol{q}_\parallel\nu\sigma} c_{2\boldsymbol{k}\sigma}. \tag{7}$$

Here $b_{\boldsymbol{q}\lambda}$ ($b^{\dagger}_{\boldsymbol{q}\lambda}$) is the annihilation (creation) operator for a phonon of wavevector $\boldsymbol{q}$ in branch λ, $c^{\dagger}_{1\boldsymbol{k}_\parallel\nu\sigma}$ is the creation operator of a Q2D electron with wavevector $\boldsymbol{k}_\parallel$ and spin σ in quantum well ν, and $c_{2\boldsymbol{k}\sigma}$ is the annihilation operator of a Q3D electron with wavevector $\boldsymbol{k}$ and spin σ. The centre-of-mass coordinates are defined as $\boldsymbol{R}_{1\nu} = (v_1 t, 0, \nu d)$ and $\boldsymbol{R}_2 = (v_2 t, 0, 0)$. The matrix elements $M(\boldsymbol{q},\lambda)$ are given by the usual bulk expressions [19], modified by the form factor

$$J^{(12)}(k_z, q_z) = \int \mathrm{d}z \; \mathrm{e}^{\mathrm{i}q_z z} \phi_1(z) \phi_2(z, k_z). \tag{8}$$

The explicit expression for $J^{(12)}$ in our model is given in the appendix. The Hamiltonian for the reverse transfer is the Hermitian conjugate of H_{ep}^{12}.

3. Balance equations

The derivation of the balance equations is similar to that in reference [20] for the Gunn effect in GaAs. The equations for the change of the centre-of-mass momenta $\boldsymbol{P}_j$, the kinetic energies $H_{j\mathrm{e}}$ of the relative electron motions, and the particle numbers N_j in both subbands result from a perturbative calculation of the expectation values of the respective time derivatives. The statistical averages are obtained using the density matrix of two decoupled electron systems in thermal equilibrium at electron temperatures T_j with chemical potentials μ_j and the free phonon system. Disregarding the contributions of intersubband scattering to frictional forces and energy relaxation rates, which are small compared to intrasubband contributions, we obtain for the steady state

$$\langle \dot{P}_{jx} \rangle = N_j e E + F_{j\mathrm{ei}} + F_{j\mathrm{ep}} = 0 \tag{9}$$

$$\langle \dot{H}_{je} \rangle = -v_j(F_{jei} + F_{jep}) - W_{jep} = 0 \tag{10}$$

$$\langle \dot{N}_1 \rangle = X_{12} = 0 \tag{11}$$

where E is the electric field, F_{jei} (F_{jep}) is the frictional force in subband j due to scattering on ionized impurities (phonons), W_{jep} is the energy relaxation rate in subband j due to scattering on phonons, and X_{12} is the difference between the Q2D-to-Q3D and Q3D-to-Q2D particle transition rates.

In our approximation the frictional forces and energy relaxation rates of the Q3D subband are given by the bulk expressions [17]. For the Q2D subband, the frictional forces and energy relaxation rates have been derived in reference [28] (cf. equations (32), (33) and (47) therein). The frictional force on the Q2D electrons due to impurity scattering for instance is calculated from

$$F_{1ei} = \frac{1}{d}\left(\frac{e^2}{2\epsilon_0\epsilon_s}\right)^2 \sum_{\boldsymbol{q}_\parallel} \frac{q_x}{q_\parallel^2} \tilde{N}_r(q_\parallel)\, \mathrm{Im}\, \hat{\Pi}_1(\boldsymbol{q}_\parallel, q_z = 0, q_x v_1) \tag{12}$$

where ϵ_s is the static dielectric constant. The effective impurity density $\tilde{N}_r$ is the product of the remote-impurity sheet concentration and a form factor (see the appendix). $\mathrm{Im}\,\hat{\Pi}_1$ is the imaginary part of the screened density–density correlation function of subband 1, which is modified compared to that of reference [28] in order to account for intersubband contributions to the screening (see the next section).

Our main theoretical result is a microscopic expression for the rate of the phonon-assisted particle exchange between the two subbands, and is given by

$$\begin{aligned} X_{12} = \frac{4\pi}{d\hbar} \sum_{\boldsymbol{q}\lambda} \sum_{\boldsymbol{k}} |M(\boldsymbol{q},\lambda)|^2 |J^{(12)}(k_z, q_z)|^2 \\ \times \Bigg\{ -\left[1 + n\left(\frac{\hbar\Omega_{\boldsymbol{q}\lambda}}{k_B T}\right)\right] f\left(\frac{\xi_{1\boldsymbol{k}_\parallel+\boldsymbol{q}_\parallel}}{k_B T_1}\right)\left[1 - f\left(\frac{\xi_{2\boldsymbol{k}}}{k_B T_2}\right)\right] \\ \times \delta(E_{1\boldsymbol{k}_\parallel+\boldsymbol{q}_\parallel} - E_{2\boldsymbol{k}} - \hbar\Omega_{\boldsymbol{q}\lambda}) \qquad (13a) \\ - n\left(\frac{\hbar\Omega_{\boldsymbol{q}\lambda}}{k_B T}\right) f\left(\frac{\xi_{1\boldsymbol{k}_\parallel+\boldsymbol{q}_\parallel}}{k_B T_1}\right)\left[1 - f\left(\frac{\xi_{2\boldsymbol{k}}}{k_B T_2}\right)\right] \delta(E_{1\boldsymbol{k}_\parallel+\boldsymbol{q}_\parallel} - E_{2\boldsymbol{k}} + \hbar\Omega_{\boldsymbol{q}\lambda}) \qquad (13b) \\ + n\left(\frac{\hbar\Omega_{\boldsymbol{q}\lambda}}{k_B T}\right) f\left(\frac{\xi_{2\boldsymbol{k}}}{k_B T_2}\right)\left[1 - f\left(\frac{\xi_{1\boldsymbol{k}_\parallel+\boldsymbol{q}_\parallel}}{k_B T_1}\right)\right] \delta(E_{1\boldsymbol{k}_\parallel+\boldsymbol{q}_\parallel} - E_{2\boldsymbol{k}} - \hbar\Omega_{\boldsymbol{q}\lambda}) \qquad (13c) \\ + \left[1 + n\left(\frac{\hbar\Omega_{\boldsymbol{q}\lambda}}{k_B T}\right)\right] f\left(\frac{\xi_{2\boldsymbol{k}}}{k_B T_2}\right)\left[1 - f\left(\frac{\xi_{1\boldsymbol{k}_\parallel+\boldsymbol{q}_\parallel}}{k_B T_1}\right)\right] \\ \times \delta(E_{1\boldsymbol{k}_\parallel+\boldsymbol{q}_\parallel} - E_{2\boldsymbol{k}} + \hbar\Omega_{\boldsymbol{q}\lambda}) \Bigg\} \qquad (13d) \end{aligned}$$

with $f(x) = [\mathrm{e}^x + 1]^{-1}$ denoting the Fermi function, $n(x) = [\mathrm{e}^x - 1]^{-1}$ the Bose function, and using the definitions

$$\xi_{j\boldsymbol{k}_{(\parallel)}} = \varepsilon_{j\boldsymbol{k}_{(\parallel)}} - \mu_j \qquad \text{and} \qquad E_{j\boldsymbol{k}_{(\parallel)}} = \varepsilon_{j\boldsymbol{k}_{(\parallel)}} + \hbar k_x v_j + m^* v_j^2/2.$$

Terms (13*a*) and (13*b*) describe the scattering of electrons out of the Q2D subband into the Q3D subband via emission and absorption of a phonon respectively, while terms (13*c*) and (13*d*) describe the opposite processes.

The particle numbers in the two subbands are related with each other and with the electron temperatures T_j and chemical potentials μ_j via

$$N_1 = \sum_{\boldsymbol{k}_\parallel \nu\sigma} f(\xi_{1\boldsymbol{k}_\parallel}/k_B T_1) \qquad \text{and} \qquad N_2 = N - N_1 = \sum_{\boldsymbol{k}\sigma} f(\xi_{2\boldsymbol{k}}/k_B T_2).$$

4. Screening

In the Lei–Ting approach, screening is incorporated in the calculation of intrasubband scattering rates by using the screened density–density correlation functions $\hat{\Pi}_j$ and the bare matrix elements (cf. equation (12)). The real and imaginary parts of the bare (unscreened) density–density correlation functions of the two subbands are given by

$$\mathrm{Re}\,\Pi_j(\boldsymbol{q}_{(\|)},\omega) = 2\sum_{\boldsymbol{k}_{(\|)}} f\left(\frac{\xi_{j\boldsymbol{k}_{(\|)}}}{k_\mathrm{B}T_j}\right)\left[\frac{1}{\hbar\omega - \varepsilon_{j\boldsymbol{k}_{(\|)}+\boldsymbol{q}_{(\|)}} + \varepsilon_{j\boldsymbol{k}_{(\|)}}} - \frac{1}{\hbar\omega + \varepsilon_{j\boldsymbol{k}_{(\|)}+\boldsymbol{q}_{(\|)}} - \varepsilon_{j\boldsymbol{k}_{(\|)}}}\right] \tag{14}$$

$$\mathrm{Im}\,\Pi_j(\boldsymbol{q}_{(\|)},\omega) = 2\pi\sum_{\boldsymbol{k}_{(\|)}}\left[f\left(\frac{\xi_{j\boldsymbol{k}_{(\|)}}}{k_\mathrm{B}T_j}\right) - f\left(\frac{\xi_{j\boldsymbol{k}_{(\|)}+\boldsymbol{q}_{(\|)}}}{k_\mathrm{B}T_j}\right)\right]\delta(\hbar\omega + \varepsilon_{j\boldsymbol{k}_{(\|)}+\boldsymbol{q}_{(\|)}} - \varepsilon_{j\boldsymbol{k}_{(\|)}}) \tag{15}$$

where the restriction to the x–y plane is appropriate for the Q2D subband ($j = 1$). The screened correlation functions are calculated from the bare ones within the random-phase approximation (RPA).

If the subband j was the only polarizable subband in the system, the required imaginary part of the screened correlation function would be given by

$$\mathrm{Im}\,\hat{\Pi}_j(\boldsymbol{q},\omega) = \frac{\mathrm{Im}\,\Pi_j(\boldsymbol{q}_{(\|)},\omega)}{|\epsilon_j(\boldsymbol{q},\omega)|^2} \tag{16}$$

with the dielectric function

$$\epsilon_j(\boldsymbol{q},\omega) = 1 - V_j(\boldsymbol{q})\Pi_j(\boldsymbol{q}_{(\|)},\omega) \tag{17}$$

where $V_j(\boldsymbol{q})$ is the matrix element of the electron–electron interaction in subband j. The matrix element V_1 of the Q2D subsystem also has a q_z-dependence because of the Coulomb interaction between electrons in different wells. It is given in reference [28]. V_2 is the standard 3D Coulomb matrix element. Using equations (16) and (17) in a multi-subband system is known as diagonal RPA (DRPA). This approximation, however, neglects the fact that the polarizabilities of the individual subbands are coupled. These intersubband contributions to the screening are incorporated in an approximate manner by adding the polarizabilities of the two subbands in the dielectric function ϵ, which gives the result

$$\mathrm{Im}\,\hat{\Pi}_j(\boldsymbol{q},\omega) = \frac{\mathrm{Im}\,\Pi_j(\boldsymbol{q}_{(\|)},\omega)}{|\epsilon(\boldsymbol{q},\omega)|^2} \tag{18}$$

with

$$\epsilon(\boldsymbol{q},\omega) = 1 - \sum_{j=1}^{2} V_j(\boldsymbol{q})\Pi_j(\boldsymbol{q}_{(\|)},\omega). \tag{19}$$

We call this approximation additive RPA (ARPA) in the following section. Note that the conventional multi-subband RPA formalism, which would normally give additional Π^2-terms in equation (19), is obstructed by the difference in dimensionality of the two model subbands.

The screening of intersubband scattering processes cannot be easily accounted for in the balance-equation approach. However, screening is expected to have little influence on the intersubband scattering, because these processes always involve a relatively large wavevector change q whereas screening is only significant for small q.

5. Numerical results and discussion

The balance equations are solved numerically for a periodic, modulation-doped GaAs–$Al_xGa_{1-x}As$ multiple-quantum-well system with wells of width $a = 4.5$ nm and a period $d = 50$ nm. We focus on the NDR that is caused by the RST at $T = 77$ K. Narrow quantum wells are required to investigate the Q2D–Q3D RST in the quantum limit where there is just one Q2D subband. This ensures that the behaviour of the Q2D–Q3D RST is not obscured by additional intersubband transitions between several bound subbands. If there was a second subband bound in the wells, it would also be increasingly populated with growing electric field. The device characteristics would then sensitively depend on how the electrons distribute between the upper bound subband and the Q3D subband, and on the electron mobility in the upper bound subband. We also require the quantum wells to be so narrow that there are no resonances between the Q3D subband and quantum well states just above the bottom of the Q3D subband, as this would make our wavefunction model invalid and would increase the intersubband scattering.

Results are presented for five different samples, the differences among which are summarized in table 1.

Table 1. Differences among samples A to E.

Sample	E_2 (meV)	N (10^{12} cm^{-2})	Screening model
A	165	1.25	ARPA
B	110	1.25	ARPA
C	165	0.75	ARPA
D	165	1.25	DRPA
E	165	1.25	Static

Our reference sample (sample A) has an energetic separation between the Q2D and the Q3D subband of $E_2 = 165$ meV. For the above geometrical parameters, this is appropriate for a conduction band discontinuity of 250 meV, corresponding to an aluminium mole fraction of $x \approx 0.28$, if band bending is neglected. Band bending is estimated to lower the centre of the barriers by 45 meV in sample A at thermal equilibrium. The main influence that band bending has (provided it is not so strong that the potential wells in the barriers give rise to strongly confined states) on the stationary transport comes from its reduction of E_2, which is one of the model parameters that is varied below. Note, however, that band bending may cause oscillatory current instabilities in the NDR regime [29, 30] and that it does affect the dynamics of the RST following optical excitation into Q3D states [31]. The instability in the NDR regime means that the static current–voltage characteristic cannot be measured directly, but it may be derived from measurements at microwave frequencies [32]. The electron concentration N of sample A is chosen to be 1.25×10^{12} cm^{-2} per well, and the impurity sheet concentration per barrier for the remote-impurity scattering of the Q2D electrons is set equal to this value. The effective impurity concentration for the intrasubband scattering of Q3D electrons is assumed to have a higher value of 10^{18} cm^{-3} in order to adjust their mobility to a value that is appropriate for $Al_xGa_{1-x}As$. This is required because we do not include alloy scattering or scattering from DX centres. Using the same effective mass $m^* = 0.067m_e$ for all electrons restricts our approach to aluminium mole fractions where the Γ valley is lowest in $Al_xGa_{1-x}As$ and to electric fields for which the Γ–L transfer is unimportant in GaAs. We employ a GaAs-like phonon spectrum with bulk material parameters for the coupling constants (cf. reference [7]). Dynamical screening is

incorporated in the ARPA.

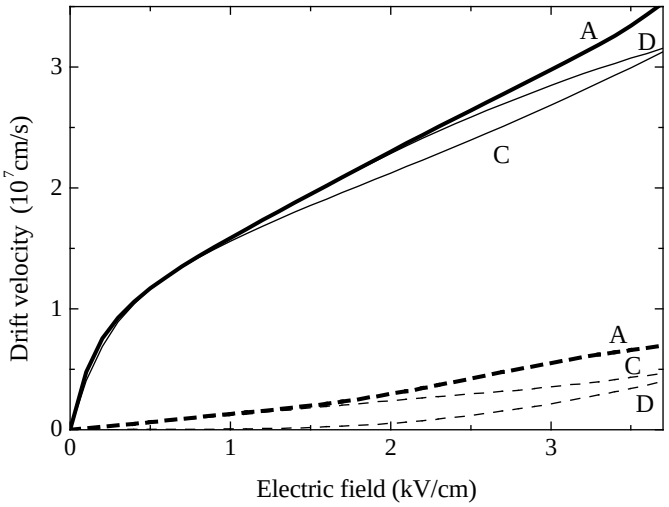

Figure 1. The dependences of the drift velocities of the Q2D (full lines) and the Q3D (dashed lines) subbands on the applied field. The labels refer to the different samples; the values for sample A are emphasized by thicker lines.

The resulting drift velocities are shown in figure 1. The Q2D electrons have a low-field mobility of approximately 50 000 $cm^2\ V^{-1}\ s^{-1}$, which is mainly limited by remote-impurity scattering. Above moderate fields of approximately 100 V cm^{-1}, the scattering by emission of optical phonons becomes more important, and this mechanism dominates at high fields. Due to the energy-independent density of states of the Q2D subband, the emission of optical phonons sets in more abruptly than in bulk systems [33]. This results in a reduction of the differential mobility at the LO-phonon emission threshold, which can be seen in the v_1–E characteristic, although this reduction does not occur abruptly because of the relatively high lattice temperature. The drift velocity of the Q3D electrons is limited by impurity scattering for all electric fields considered. For the chosen effective impurity concentration, we find a low-field mobility of approximately 1300 $cm^2\ V^{-1}\ s^{-1}$. The small increase in the mobility at fields higher than 2 kV cm^{-1} is correlated with the onset of the RST, which increases the occupation of the Q3D subband (see below).

The high drift velocity of the Q2D electrons is accompanied by a large increase of their temperature with the electric field shown in figure 2. The difference between the Q2D-electron temperature T_1 and the lattice temperature varies roughly as E^2; this variation would be exact if the mobility and energy relaxation time were independent of the electric field. The temperature T_2 of the Q3D electrons remains rather low due to the smaller power input from the drift motion. The change in its slope near $E = 2$ kV cm^{-1} is again correlated with the onset of the RST. For both subbands, the energy relaxation is dominated by emission of LO phonons even at low electric fields.

In thermal equilibrium, all electrons occupy the Q2D subband, and the chemical potentials of the two subbands are equal and approximately 45 meV (measured from the bottom of the Q2D subband). As this chemical potential is 120 meV or some $18k_BT$ below the bottom of the Q3D subband, the carriers have to be heated considerably before a significant fraction may be scattered into the Q3D subband. Thus, the Q3D subband is only populated for fields higher than 1.5 kV cm^{-1} (cf. figure 3). With increasing field,

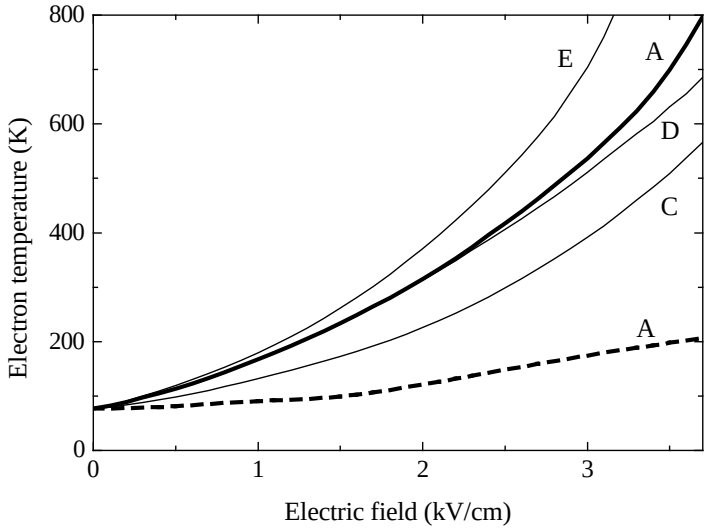

Figure 2. The dependences of the electron temperatures in the Q2D (full lines) and the Q3D (dashed lines) subbands on the applied field for different samples.

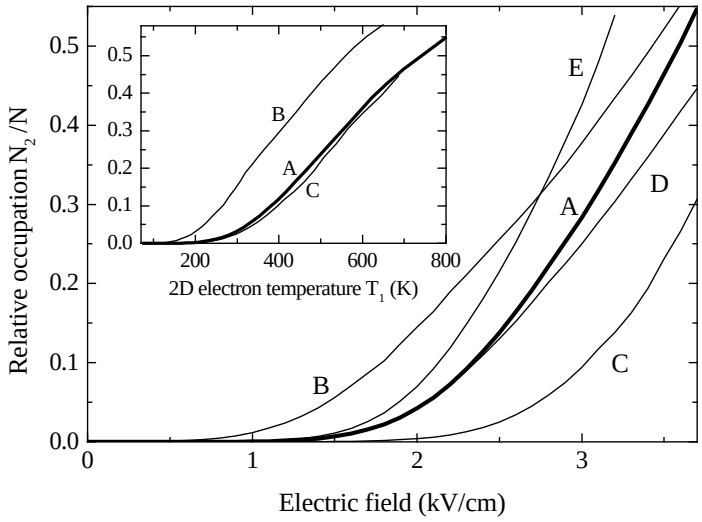

Figure 3. The dependences of the relative numbers of electrons occupying the Q3D subband on the applied field for all samples and on the electron temperature in the Q2D subband (inset).

more and more carriers are transferred into the Q3D subband and inversion is reached at 3.6 kV cm^{-1}. At this field value, the chemical potential of the Q2D subband has fallen from its equilibrium value to -57 meV whilst the chemical potential of the Q3D subband has risen to 156 meV (i.e. 9 meV below the bottom of this subband). The inset of figure 3 shows the dependence of the transfer into the Q3D subband on the electron temperature of the Q2D subband. It can be seen that the RST sets in when the Q2D electrons are heated to about 250 K and that inversion is reached at $T_1 \approx 750$ K, the latter corresponding to a thermal energy of nearly 40% of the intersubband separation E_2.

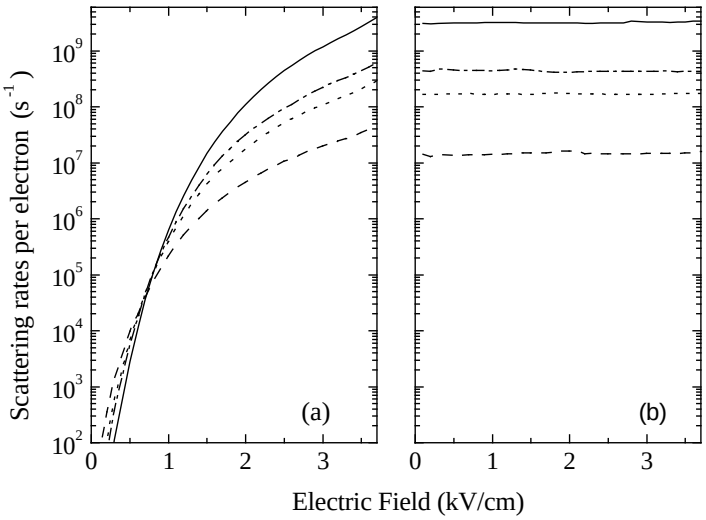

Figure 4. The dependences of the Q2D-to-Q3D (a) and Q3D-to-Q2D (b) scattering rates per electron on the applied field for sample A. Full (dashed) lines represent transfer by emission (absorption) of an LO phonon, and chain (dotted) lines the transfer by emission (absorption) of an LA phonon.

The various contributions to the RST scattering rate per electron are shown in figure 4, where the Q2D-to-Q3D rates refer to the number of Q2D electrons and the Q3D-to-Q2D rates to the number of Q3D electrons. For low fields, the dominant mechanism for the transfer of Q2D electrons into the Q3D subband (figure 4(a)) is by absorption of an LO phonon, reflecting the fact that this process is possible for a lower initial energy of the transferred Q2D electron than the other processes. Because of the small number of available LO phonons at 77 K, the transfer with LO-phonon absorption is exceeded by transfer with absorption or emission of LA phonons and with emission of LO phonons as soon as an appreciable number of Q2D electrons have sufficient kinetic energy to participate in these processes. The emission of an LO phonon is the dominant mechanism of Q2D-to-Q3D transfer for fields larger than 1 kV cm^{-1}. This is due to the stronger coupling of electrons to LO phonons than to LA phonons in GaAs, although the effect is reduced because fewer Q2D electrons have the kinetic energy required for transfer by LO-phonon emission ($\approx 165 + 35$ meV) than have that required for transfer by LA-phonon emission ($\approx 165 + 3$ meV). For the backward transfer (figure 4(b)) the scattering rates per electron are nearly independent of the electric field. This reflects the fact that an electron in the Q3D subband always has sufficient energy to transfer into the Q2D subband, and that there are always a large number of unoccupied final states in the Q2D subband. Again, the Q3D-to-Q2D scattering by LO-phonon emission dominates over that by LA-phonon emission, and the transfer by phonon absorption is less likely owing to the low number of available phonons.

The average drift velocity is given by

$$v_{\mathrm{d}} = \frac{N_1 v_1 + N_2 v_2}{N} \tag{20}$$

and is shown in figure 5. For fields up to approximately 2 kV cm^{-1} it is nearly equal to the drift velocity v_1 of the Q2D subband. When the field is further increased, the number of carriers in the Q2D subband steadily reduces, and the Q2D part $N_1 v_1 / N$ of v_{d} falls off

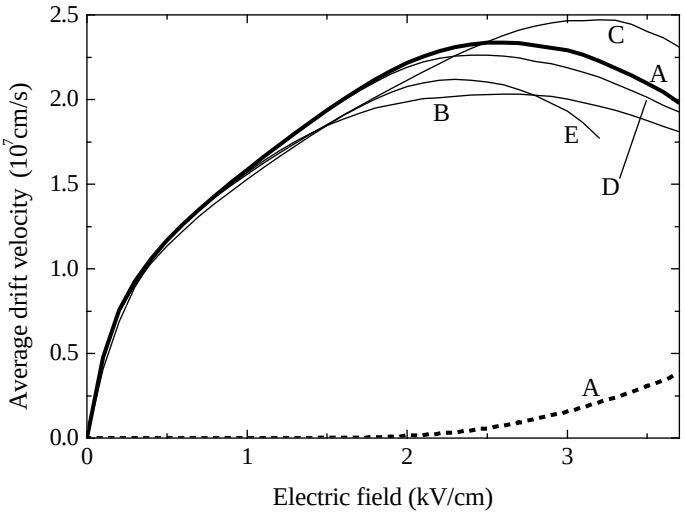

Figure 5. The dependences of the average drift velocity v_d on the applied field for all samples. The dotted line shows the contribution of the Q3D subband to v_d for sample A.

above 2.4 kV cm^{-1}. The Q3D part $N_2 v_2/N$ rises with increasing field, but this cannot compensate for the decrease of the Q2D contribution due to the lower drift velocity in the Q3D subband. This leads to NDR for fields higher than 2.6 kV cm^{-1}. However, a sufficiently large difference between the mobilities in the two subbands is necessary for a NDR to occur. For sample A the mobilities differ by a factor of 5 at 3.5 kV cm^{-1}. It can be estimated from figure 5 that the NDR would be quenched if the mobility in the Q3D subband was doubled because the increase in the Q3D contribution would then compensate for the reduction of the Q2D contribution.

5.1. Variation of parameters

To study the influence of the energetic separation between the subbands we compare the results for sample A with those for a sample with $E_2 = 110$ meV that is otherwise unchanged (sample B). In the experiment, E_2 could be varied by using different aluminium mole fractions in the barriers or different well widths. We find that the drift velocities and electron temperatures of the two subbands do not significantly differ from those of sample A. The smaller subband separation means that the threshold kinetic energy of the Q2D electrons for experiencing RST is reduced. The Q2D electrons can therefore be transferred into the Q3D subband for lower electron temperatures and thus lower electric fields (cf. figure 3). A lower electron temperature T_1 at the field for which the RST sets in is connected with a smaller temperature increase $\mathrm{d}T_1/\mathrm{d}E$ at this field as T_1 roughly rises quadratically with the field. This causes the fraction of Q3D carriers to rise more slowly with increasing field. The effect on the average drift velocity is a reduction of the maximum velocity (figure 5) and a flattening of the characteristics on both sides of the maximum. A qualitatively similar reduction of the maximum velocity with decreasing energy separation was also found with the approximate analytical solution of [8] and in Monte Carlo simulations of the RST for wide quantum wells [2]. We conclude that a high subband separation is favourable both for a high saturation velocity and a steep decrease of the current in the NDR regime.

The other important parameter for the device characteristics is the electron sheet

concentration. For sample C we use a lower value of $N = 7.5 \times 10^{11}$ cm^{-2} (the other parameters are as for sample A). In this case the drift velocities of both subbands are lower than in sample A at high electric fields (figure 1). For the Q2D subband this is due to the reduced screening of the interaction between the Q2D electrons and LO phonons. The smaller screening efficiency at lower electron density leads to an increase of the energy relaxation rate and thus to a lower electron temperature (figure 2). For the Q3D subband the change in the drift velocity results from the higher field that is required for the RST (cf. figure 3) as its mobility is influenced by its occupation (as in sample A). The RST sets in at a higher field mainly because the Q2D-electron temperature rises more slowly with increasing field. It can be seen from the inset of figure 3 that the dependence of the RST on the Q2D-electron temperature is remarkably close to that of sample A. This means that the separation between the bottom of the Q3D subband and the chemical potential at thermal equilibrium, which is larger at lower electron density, has less influence on the RST than the change in the screening efficiency. The smaller gradient of the T_1–E curve for sample C also explains why the increase in the occupation of the Q3D subband with rising field after the onset of the RST is not steeper than for sample A. Thus, the maximum of the average drift velocity is shifted to a higher field, and its value is slightly higher as compared to that for sample A (figure 5). So, a lower electron concentration also favours a high saturation drift velocity.

If the well width is increased to 5.0 nm (with the other parameters as for sample A), the transport properties do not differ appreciably from the results for sample A. An increase of the period d to 60 nm also does not cause any significant deviation if the electron volume number density is reduced such that the sheet concentration remains unaltered. Although the average drift velocity then is nearly the same as in sample A, the total current is of course lowered accordingly. In experiment, an alteration of a or d would of course affect the results because it induces a change of E_2.

5.2. Comparison of screening models

In order to show the influence of intersubband contributions to the screening, we solve the balance equations for the same parameters as for sample A, but using the DRPA instead of the ARPA (sample D). It can be seen from figure 1 that this drastically changes the drift velocity in the Q3D subband. The low occupation of this subband at low fields means that its contribution to the screening is extremely small, and hence that the low-field mobility in the Q3D subband is close to zero in the DRPA (approximately 25 times smaller than in the ARPA). For fields higher than the RST threshold, the Q3D intrasubband screening contribution rises with increasing occupation, and the Q3D drift velocity varies superlinearly with the field. Similarly, the drift velocity in the Q2D subband in the DRPA is reduced compared to that in the ARPA case when the occupation of this subband decreases due to the RST. However, the difference is less pronounced than in the Q3D subband because the mobility in the Q2D subband is limited by LO-phonon emission at the relevant fields. This process is less sensitive to screening than the impurity scattering, which limits the mobility in the Q3D subband. The electron temperatures are affected in the same way as the drift velocities. The slightly smaller temperature of the Q2D electrons at high fields (cf. figure 2) results in a smaller occupation of the Q3D subband (figure 3). This and the superlinear dependence of v_2 on the field lead to a flatter v_d–E characteristics (figure 5).

Finally, we compare dynamical screening with static screening. Static screening is obtained in the ARPA by setting $\omega = 0$ in the denominator of equation (18). We find that in both subbands the drift velocities are not significantly altered if static screening is used

instead of the dynamical ARPA (sample E, with the same parameters as for sample A). This result, which is well known, is sometimes used to justify the use of static screening rather than dynamic screening. However, in a system in which different subbands have dramatically different mobilities, the total current also depends on the electron temperatures via the redistribution of carriers among the subbands. The energy relaxation rates are higher if dynamical screening is used, resulting in lower electron temperatures. This is a sign of enhanced LO-phonon emission due to their coupling to plasmons [22], which is neglected in static screening. Because of the higher temperature of the Q2D electrons for static screening (figure 2), the RST sets in at a lower electric field (figure 3), and the average drift velocity becomes maximal at a lower field, too (figure 5; values for fields higher than 3.2 kV cm^{-1} were not obtained because of numerical problems). We therefore conclude that using dynamical instead of static screening causes a perceptible change in the calculated current–voltage characteristics of a RST device. Note, however, that hot-phonon effects would give a modification in the opposite direction, and this correction is expected to be stronger for dynamical screening than for static screening.

6. Summary

The microscopic balance-equation approach to hot-electron transport developed by Lei and Ting has been applied to the thermionically induced real-space transfer of electrons between a subband that is bound to narrow quantum wells and states that extend above the barriers. We have treated the case where the bound subband exhibits quasi-two-dimensional behaviour while the extended states can be described as a quasi-three-dimensional subband. It is pointed out that the wavefunction of the quasi-three-dimensional subband must be orthogonal to that of the bound subband for an adequate description of the transfer. The energy and momentum relaxation of the two subbands are calculated with separate balance equations that include scattering on ionized impurities, and LA and LO phonons. Dynamical screening is accounted for in an approximate two-subband RPA treatment. An additional balance equation microscopically describes the particle exchange between the two subbands, for which we have derived the transfer rates due to phonon emission and absorption.

We have numerically solved the equations for a GaAs–$Al_xGa_{1-x}As$ multiple-quantum-well system with undoped wells and highly doped barriers. In this structure, the heating of the electrons in the quasi-two-dimensional subband by an applied DC electric field results in the emission of electrons from the high-mobility quantum wells into the low-mobility barriers. This leads to a saturation of the average drift velocity at a field of about 2.6 kV cm^{-1} and negative differential resistance at higher fields. The maximum drift velocity increases with increasing energetic separation of the two subbands and with decreasing electron density. It is also shown that the results depend on which screening model is used, which indicates that intersubband contributions and dynamical effects should not be neglected.

As far as we are aware, real-space transfer has not yet been investigated experimentally for multiple quantum wells with such narrow wells. We believe that the negative differential resistance caused by the transfer can in principle be observed. As the current–voltage characteristic is unstable to oscillations in this regime, either a high-frequency resolution is required, as demonstrated in reference [30], or the sample must be excited at microwave frequencies [32]. These measurements were made for multiple quantum wells with wider wells and a single heterojunction respectively. In both cases several bound subbands were present in the quantum well(s), which complicates the interpretation of how the negative differential resistance is caused by the real-space transfer. For the narrower quantum wells

considered here, it will be rather difficult to produce samples with a sufficiently high mobility of the Q2D electrons owing to limitations caused by interface roughness.

A possible extension of our approach is the inclusion of an additional subband that is bound to the quantum wells. One could then investigate how the average drift velocity depends on the energetic position of the bottom of this subband relative to the energetic separation between the ground state and the lowest extended state.

Acknowledgments

The authors gratefully acknowledge discussions with M Asche, E Schöll, and N J M Horing.

Appendix

The orthogonalization coefficient in equation (6) is given by

$$I^{(12)}(k_z) = \frac{2\pi\sqrt{2a}}{\pi^2 - k_z^2 a^2}\cos(k_z a/2) \tag{A1}$$

and the normalization factor of the Q3D wavefunction is

$$\Gamma(k_z) = \sqrt{\frac{d}{d - |I^{(12)}(k_z)|^2}}. \tag{A2}$$

The form factor of the intersubband scattering on phonons (cf. equation (8)) is given by

$$J^{(12)}(k_z, q_z) = \Gamma(k_z)[I^{(12)}(k_z + q_z) - I^{(12)}(k_z)I^{(1)}(-\mathrm{i}q_z)] \tag{A3}$$

with

$$I^{(1)}(Q) = \frac{4\pi^2}{Qa(4\pi^2 + Q^2a^2)}(\mathrm{e}^{Qa/2} - \mathrm{e}^{-Qa/2}). \tag{A4}$$

The square modulus of $J^{(12)}$ approximately vanishes as q_z^2 for small phonon wavevectors q_z. For large values of $|q_z|$ it is maximal for $k_z = -q_z$.

The effective impurity density for the scattering of Q2D electrons on remote impurities (cf. equation (12)) is calculated in the same way as in reference [28]. For a homogeneous distribution of donors in the barriers and no impurities in the quantum wells, we obtain

$$\tilde{N}_\mathrm{r}(q_\parallel) = n_\mathrm{ir}\left[\frac{2}{q_\parallel(d-a)}I^{(1)}(q_\parallel)\frac{\sinh(q_\parallel(d-a)/2)}{\sinh(q_\parallel d/2)}\right]^2 \tag{A5}$$

where n_ir is the impurity sheet concentration per barrier.

References

[1] Gribnikov Z S 1972 *Fiz. Tekh. Poluprov.* **6** 1380 (Engl. Transl 1973 *Sov. Phys.–Semicond.* **6** 1204)
[2] Gribnikov Z S, Hess K and Kosinovsky G A 1995 *J. Appl. Phys.* **77** 1337
[3] Brennan K F and Park D H 1989 *J. Appl. Phys.* **65** 1156
[4] Kostial H, Ihn Th, Kleinert P, Hey R, Asche M and Koch F 1993 *Phys. Rev.* B **47** 4485
[5] Ravaioli U and Ferry D K 1986 *Superlatt. Microstruct.* **2** 377
[6] Nederveen K and Zimmermann J 1989 *Superlatt. Microstruct.* **6** 133
[7] Kleinert P and Asche M 1994 *Phys. Rev.* B **50** 11 022
[8] Bryksin V V and Kleinert P 1997 *J. Phys.: Condens. Matter* **9** 7391
Bryksin V V and Kleinert P 1997 private communication
[9] Thobel J L, Baudry L, Bourel P, Dessenne F and Charef M 1993 *J. Appl. Phys.* **74** 6274

[10] Brum J A and Bastard G 1986 *Phys. Rev.* B **33** 1420
[11] Babiker M and Ridley B K 1986 *Superlatt. Microstruct.* **2** 287
[12] Lent C S, Liang L and Porod W 1989 *Appl. Phys. Lett.* **54** 2315
[13] Babiker M, Ghosal A and Ridley B K 1989 *Superlatt. Microstruct.* **5** 133
[14] Brunetti R, Jacoboni C and Price P J 1994 *Phys. Rev.* B **50** 11 872
[15] Yassievich I N, Schmalz K and Beer M 1994 *Semicond. Sci. Technol.* **9** 1763
[16] Abou-Khalil M, Goano M, Reid B, Champagne A and Maciejko R 1997 *J. Appl. Phys.* **81** 6438
[17] Lei X L and Ting C S 1985 *Phys. Rev.* B **32** 1112
[18] Lei X L and Horing N J M 1992 *Int. J. Mod. Phys.* B **6** 805
[19] Lei X L, Birman J L and Ting C S 1985 *J. Appl. Phys.* **58** 2270
[20] Lei X L, Xing D Y, Liu M, Ting C S and Birman J L 1987 *Phys. Rev.* B **36** 9134
[21] Hai G-Q, Studart N, Peeters F M, Koenraad P M and Wolter J H 1996 *J. Appl. Phys.* **80** 5809
[22] Collet J H 1989 *Phys. Rev.* B **39** 7659
[23] Kocevar P 1986 *Physica* B **134** 155
[24] Lei X L and Horing N J M 1987 *Phys. Rev.* B **35** 6281
[25] Gupta R, Balkan N and Ridley B K 1992 *Phys. Rev.* B **46** 7745
[26] Kubrak V, Kleinert P and Asche M 1998 *Semicond. Sci. Technol.* **13** 277
[27] Sawaki N and Akasaki I 1985 *Physica* B **134** 494
[28] Lei X L, Horing N J M and Zhang J Q 1986 *Phys. Rev.* B **34** 1139
[29] Schöll E and Aoki K 1991 *Appl. Phys. Lett.* **58** 1277
[30] Döttling R, Rudzick O, Schöll E, Straw A, Vickers A J, Balkan N and Da Cunha A 1994 *Semicond. Sci. Technol.* **9** 611
[31] Goodnick S M and Lary J E 1992 *Semicond. Sci. Technol.* B **7** 109
[32] Masselink W T 1989 *Semicond. Sci. Technol.* **4** 503
[33] Ridley B K 1982 *J. Phys. C: Solid State Phys.* **15** 5899