Coherent charge transport in semiconductor quantum cascade structures

2004 J. Phys.: Condens. Matter 16 R25

(http://iopscience.iop.org/0953-8984/16/4/R01)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS: CONDENSED MATTER

J. Phys.: Condens. Matter **16** (2004) R25–R48 PII: S0953-8984(04)59609-1

TOPICAL REVIEW

Coherent charge transport in semiconductor quantum cascade structures

Michael Woerner, Klaus Reimann and Thomas Elsaesser[1]

Max Born Institut für Nichtlineare Optik und Kurzzeitspektroskopie, Max-Born-Strasse 2A, 12489 Berlin, Germany

E-mail: elsasser@mbi-berlin.de

Received 16 December 2003
Published 16 January 2004
Online at stacks.iop.org/JPhysCM/16/R25 (DOI: 10.1088/0953-8984/16/4/R01)

Abstract

Quantum cascade structures have found extensive application in electrically driven semiconductor lasers working in the mid- to far-infrared spectral range. Optical amplification in such unipolar devices is based on a population inversion between quasi-two-dimensional conduction subbands in coupled quantum wells. The population inversion in the active region is generated by electrons tunnelling from an injector region through a barrier into the upper laser subband and by ultrafast extraction of these electrons out of the lower laser subband through a barrier into the next injector region. Such transport processes on ultrafast timescales have been the subject of extensive experimental and theoretical work without, however, reaching a clear physical picture of the microscopic electron dynamics. In this review, we report a comprehensive experimental study of electron transport in electrically driven quantum cascade structures. Ultrafast quantum transport from the injector into the upper laser subband is investigated by mid-infrared pump–probe experiments directly monitoring the femtosecond saturation and subsequent recovery of electrically induced optical gain. For low current densities, low lattice temperatures and low pump pulse intensities, the charge transport is dominantly coherent, leading to pronounced gain oscillations due to the coherent motion of electron wavepackets. For higher current densities, lattice temperatures, or pump intensities, the gain recovery shows an additional incoherent component, which essentially follows the pump-induced heating and subsequent cooling of the carrier gas in the injector.

(Some figures in this article are in colour only in the electronic version)

[1] Author to whom any correspondence should be addressed.

Contents

1. Introduction

Carrier transport in external electric fields has been a topic of research since the early days of semiconductor physics. Different regimes of electron and hole transport, e.g., drift-diffusive motion, hopping processes and ballistic motion, have been identified for a large variety of systems and an analysis of their basic microscopic interactions has been given. Quantum coherent transport processes, which are governed by the wave nature of electrons, represent a particularly interesting class of phenomena, which have been studied mainly in low-dimensional semiconductors such as quantum wells and superlattices [1–9]. Quantum coherent transport is connected with a nonstationary superposition of the wavefunctions of different electronic states, i.e., with the formation of electron wavepackets. Wavepacket motions underlie coherent tunnelling through barriers and—in a periodic form—charge oscillations in quantum wells [10] and Bloch oscillations in superlattices [11, 12].

Coherent carrier motions are destroyed by scattering processes within the carrier gas and by scattering with acoustic and optical phonons, resulting in decoherence, i.e., in the destruction of the quantum mechanical phase. For very low electron densities and extremely low temperatures, decoherence can be suppressed to a large extent. In this regime, basic processes of quantum transport have been discovered and analysed in great detail. For elevated carrier densities or for higher temperatures—regimes typical for most semiconductor devices—quantum coherence is limited to femtosecond or picosecond periods of time. Under such conditions, ultrafast optical techniques allow for the real-time observation of carrier dynamics, which are monitored through nonlinear changes of the steady-state optical properties of the system under investigation. Optical excitation of low-dimensional semiconductors by femtosecond optical pulses allows for the preparation of electron wavepackets with well-defined properties. Both unidirectional and periodic wavepacket motions persisting for up to several picoseconds have been temporally resolved [10–13].

The decoherence processes limiting quantum coherent transport at elevated carrier densities are far from being understood. Most experiments were performed with interband excitation of excitons or free carriers at a fixed density. Apart from the difficulty of separating electron and hole motions, carrier–carrier scattering mediated through the long-range Coulomb interaction has remained a major issue [14, 15]. A quantitative theoretical description of decoherence by carrier–carrier scattering is still lacking. Qualitative information has been derived from experimental studies of Bloch oscillations, where oscillatory motions were found

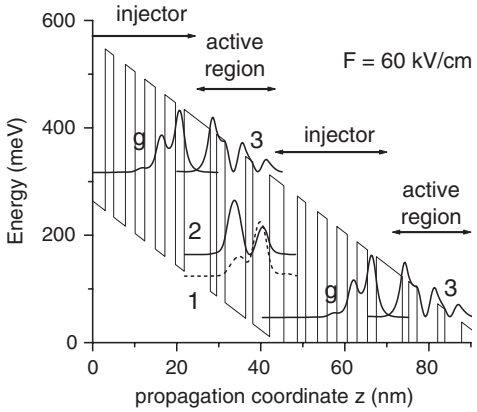

Figure 1. A conduction band diagram of the GaAs/$Al_{0.33}Ga_{0.67}As$ quantum cascade laser structure (sample A). Probability densities $|\Psi(z)|^2$ are shown for the wavefunctions relevant for the QCL dynamics: $|g\rangle$ (ground state in the injector), $|3\rangle$ (upper laser state), $|2\rangle$ (lower laser state) and $|1\rangle$ are eigenstates of the electronic Hamiltonian *without* the tunnel coupling through the injection barrier (the wide barrier to the left of the active region).

for electron concentrations of up to several 10^9 cm^{-2}. In contrast, a rapid damping of Bloch oscillations occurs at higher carrier densities [16].

Low-dimensional semiconductors, in which unipolar transport of electrons or holes can be induced by ultrafast optical excitation and followed by optical techniques, are particularly promising structures for investigating the issues outlined above. The quantum cascade laser (QCL) [17–21] represents such a structure in which electrons are transferred from an electron reservoir, the so-called injector region, through the injection barrier into the optically active region (figure 1). Under steady-state electric bias in forward direction, a population inversion builds up between subbands 3 and 2, resulting in optical gain and laser action on the $3 \leftrightarrow 2$ intersubband (IS) transition. Subband 2 is depopulated very efficiently by emission of longitudinal optical phonons, transferring carriers into energetically lower subbands, e.g., into subband 1. Electrons leave the active region by tunnelling through a thin exit barrier into the next injector region. Interaction with ultrashort optical pulses resonant to the $3 \leftrightarrow 2$ transition allows one to change the population inversion transiently and to follow the time evolution of optical gain. Since the populations of the optically coupled levels are inherently linked to the transport through the structure, such measurements directly reflect the transport dynamics.

In this article, we review recent experiments [22–25] in which coherent electron transport in quantum cascade structures has been studied by ultrafast spectroscopy. Such studies have provided clear evidence for the quantum coherent character of electron motion from the injector into the active region. They demonstrate how the microscopic injection process depends on the design parameters of the quantum cascade structure and on external parameters such as the injection current or the lattice temperature. This review is organized as follows. In section 2, we summarize previous work on electron transport in quantum cascade structures; this is followed by a brief discussion of the basic properties of optical intersubband transitions. The experimental techniques used in our study are discussed in section 3. The experimental results and their analysis are presented in section 4. We start with a phenomenological description (section 4.1), which is followed by an in-depth study of electron transport as a function of different structural parameters of the quantum cascade samples and of external parameters in the experiments (section 4.2). A summary is given in section 5.

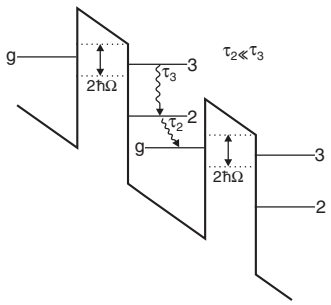

Figure 2. Schematics of the structure considered in the work of Kazarinov and Suris [26, 27]. $2\hbar\Omega$ is the coupling between subsequent quantum wells. It is assumed that τ_2, the electron lifetime in subband 2, is much less than τ_3, the electron lifetime in subband 3.

2. Quantum cascade structures: electron transport and optical properties

2.1. Electron transport in quantum cascade structures

Theoretical work on electron transport in electrically driven superlattices began with the seminal work of Kazarinov and Suris in 1971 [26, 27]. They considered vertical transport through such structures, i.e., the electron motion perpendicular to the quasi-two-dimensional layers, which is equivalent to tunnelling through a sequence of barriers. Using a density matrix approach in the tight binding approximation, they calculated current–voltage characteristics and predicted the occurrence of a negative differential resistance under appropriate bias.

In the nomenclature used nowadays for quantum cascade lasers, Kazarinov and Suris [26, 27] obtained the following result for the current density j (see figure 2 in the limit of $\tau_2 \ll \tau_3$):

$$j = eN_s \frac{2|\Omega|^2 T_2}{1 + (\frac{E_g - E_3}{\hbar})^2 T_2^2 + 4|\Omega|^2 T_2 \tau_3}, \tag{1}$$

where $E_g - E_3$ is the energy detuning from resonance between the subbands g and 3, which depends on the applied bias, N_s is the sheet electron density in the injector, $2\hbar\Omega$ is the tunnel coupling, T_2 is the irreversible dephasing time of the g $\leftrightarrow$ 3 intersubband polarization and τ_3 is the population lifetime of subband 3. At an appropriate bias V_{res}, subband g is in resonance with subband 3, i.e., $E_g - E_3 = 0$, and the maximum current density

$$j_{max} = eN_s \frac{2|\Omega|^2 T_2}{1 + 4|\Omega|^2 T_2 \tau_3} \tag{2}$$

is obtained. Increasing the bias further, equation (1) predicts a decrease of the current. Thus, there should exist a region with negative differential resistance above the resonance [9].

Sirtori *et al* [28] studied the resonant alignment of subbands g and 3, which is a prerequisite for resonant tunnelling through the injection barrier. They measured the voltage (V–I curve) and light output power (L–I curve) as a function of injection current for various GaInAs/AlInAs quantum cascade devices (figure 3 in [28]). The authors interpreted specific features in the V–I curve as an indication of resonant tunnelling through the injection barrier, arising from the alignment of subbands 3 and g. The theoretically predicted negative differential resistance was not observed. This was attributed to

(i) the comparably high impedance of the electric circuitry used in the experiments, which was much bigger than $|dV/dI|$, and to

(ii) structural inhomogeneities washing out any pronounced structures in the V–I characteristics.

More recent results, presumably obtained with better-quality samples, indeed show a region with the theoretically predicted negative differential resistance [29].

Further support for the picture of a resonant alignment of subbands 3 and g for an appropriate bias comes from measurements of Barbieri *et al* [30]. They studied electroluminescence (EL) spectra of GaAs/AlGaAs quantum cascade laser structures without a resonator. The lineshape of this luminescence, which originates from transitions between subbands 3 and 2, is sensitive to the tunnel splitting caused by a resonant coupling of subbands 3 and g. Such a coupling should lead to two distinct emission peaks, separated by the tunnel splitting $2\hbar\Omega$. Although it was not possible to observe two distinct peaks because of the strong homogeneous and inhomogeneous broadening of the band, the width of the emission band displays an increase with decreasing barrier width, in good agreement with the calculated tunnel splitting $2\hbar\Omega$. An indication of resonant alignment was also obtained from photocurrent and electroluminescence measurements by Wilson *et al* [31, 32].

The role of the dephasing time T_2 for transport through the injection barrier was investigated by Sirtori *et al* [28]. Using the calculations of Kazarinov and Suris (equation (2)) they identified two different transport regimes depending on the magnitude of the quantity $Q = 4|\Omega|^2 T_2 \tau_3$. For $Q \ll 1$ we have weak coupling between injector and active region, for $Q \gg 1$ strong coupling. The weak-coupling regime implies that the dephasing time T_2, which is assumed to be similar to τ_3, is much shorter than the inverse tunnelling rate $(2\Omega)^{-1}$. In this limit, the current density is strongly reduced by the dephasing of the g $\leftrightarrow$ 3 intersubband coherence, leading to a maximum current density of $j_{\max} = eN_s 2|\Omega|^2 T_2$. In the strong-coupling regime ($Q \gg 1$), the transport through the injection barrier is dominated by resonant tunnelling. The maximum current density in this case is given by $j_{\max} = eN_s/(2\tau_3)$, i.e., the current density is controlled by the lifetime τ_3 of subband 3. This is considered 'the configuration in which we want to operate the lasers in order to always ensure very fast electron injection into the upper laser state ($n = 3$), without being limited by the tunnelling rate' [28].

To calculate the quantity Q, the dephasing time T_2 is needed. Since no experimental values are available, T_2 was estimated from the electroluminescence width, giving a value of $T_2 \approx 50$–100 fs [28]. However, this only gives a lower limit for T_2, since the electroluminescence width is not only determined by the dephasing time T_2, which is the dephasing time of coherent superpositions of states in subband g and subband 3, but also by the energy splitting $2\hbar\Omega$, by the 3–2 intersubband dephasing and by structural inhomogeneities. As a consequence, T_2 may be considerably longer than the value obtained from the electroluminescence width. Assuming a dephasing time of $T_2 = 100$ fs results in values of Q between 0.7 and 11.6 for different laser devices. This suggests that most of the QCLs investigated work in between the strong- and the weak-coupling regimes and, hence, that the transport from the injector into the active region cannot be described in a strictly coherent picture. If, however, T_2 is considerably longer, most QCLs would operate in the strong-coupling regime.

To clarify the role of coherence in charge transport, Iotti and Rossi [7, 33, 34] carried out semiclassical Monte Carlo simulations taking into account carrier–carrier and carrier–phonon scattering. From their calculations they derived voltage–current characteristics in good agreement with experimental data. Moreover, voltage–current characteristics were calculated with a quantum mechanical density matrix approach neglecting carrier–carrier scattering. This model gives similar results to the semiclassical calculation. This agreement led to the conclusion that the semiclassical picture, in which quantum coherences are neglected, is sufficient to describe transport processes in quantum cascade laser. The authors claimed that

'energy-relaxation and dephasing processes are strong enough to destroy any phase-coherence effect on a sub-picosecond timescale' [7].

It is important to note that the voltage–current characteristic is a macroscopic phenomenon, reflecting the global properties of transport through the entire structure. The properties of microscopic transport, e.g., transport through the injection barrier, are very difficult to derive from such measurements. In contrast, a full quantum calculation including the microscopic dephasing by scattering processes is needed to account properly for the tunnelling processes. From this point of view, it is not surprising that the analysis of current–voltage characteristics has led to somewhat conflicting results. In addition to an appropriate theoretical treatment, experimental techniques probing the degree of quantum coherence in electron injection are required. In section 4, we demonstrate that ultrafast nonlinear spectroscopy gives direct and detailed insight into this issue.

2.2. Optical intersubband transitions and carrier relaxation after intersubband excitation

Optical gain in quantum cascade lasers relies on optical intersubband transitions of electrons. Such transitions have been studied extensively in heterostructures and quantum wells. Both the strength and shape of the linear absorption spectra and the underlying carrier dynamics have been addressed. Intersubband dipole transitions occur between subbands with different envelope wavefunctions for the quantized carrier motion. The intersubband dipole moment has a direction perpendicular to the quantum well layers. Since the envelope wavefunctions extend over the width of the quantum well (typical dimensions are 10 nm), the dipole moments can reach very high values. The in-plane dispersion of conduction subbands in quantum wells is nearly parallel, resulting in an intersubband transition energy practically independent of the in-plane wavevector of the electrons[2]. This fact results in a δ-like joint density of states for the optical transition, if one neglects broadening mechanisms and many-body effects. Moreover, intrasubband redistribution of electrons plays a minor role for gain dynamics, since all electrons contribute to emission at a single spectral position. Narrow absorption lines of approximately 3 meV width have been observed for electron concentrations lower than 5×10^{10} cm^{-2} in GaAs/AlGaAs quantum wells of high structural quality [35][3]. For higher electron concentrations, many-body effects mediated by the long-range Coulomb interaction come into play and alter both the position and the lineshape of the intersubband transition. A brief review of such many-body effects has been given in [36].

There are several mechanisms contributing to the width of intersubband absorption lines. Femtosecond photon echo experiments [35], in which the dynamics of coherent intersubband polarizations was followed in real time, have demonstrated that the intersubband absorption lines in high-quality GaAs/AlGaAs quantum wells are predominantly homogeneously broadened. The dephasing times, which are determined mainly by carrier–carrier scattering, are of the order of several hundred femtoseconds. Such dephasing times translate into homogeneous linewidths of several millielectronvolts. In samples containing structural disorder, e.g., quantum well thickness or alloy fluctuations, a distribution of transition energies results in additional inhomogeneous broadening.

The population lifetime of electrons in higher ($n > 1$) subbands has values between several hundreds of femtoseconds and several tens of picoseconds. For an energy separation of the

[2] The subbands would be exactly parallel if the band dispersion was parabolic. As the conduction band nonparabolicity in GaAs is quite small, it is of minor importance for GaAs/AlGaAs quantum wells. For other materials, e.g., InSb, nonparabolicity plays an important role.

[3] Even narrower lines are found if the subband separation is below the optical phonon energy, preventing optical phonon scattering.

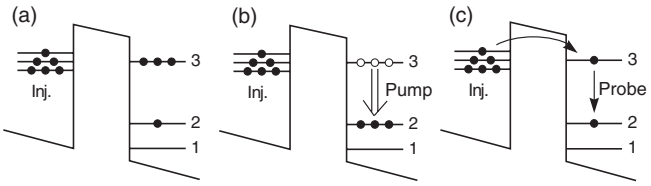

Figure 3. The experimental concept for the ultrafast electron transport measurement. (a) Under forward bias the electronic system is in a quasi-stationary nonequilibrium state. (b) A strong pump pulse depletes the electrically induced optical gain and (c) a weak probe pulse monitors the subsequent gain recovery.

optically coupled subbands larger than the energy of optical phonons, intersubband phonon scattering via the polar–optical interaction—in nonpolar systems such as Si or Ge via the optical deformation potential—represents the predominant relaxation mechanism, resulting in lifetimes less than a few picoseconds. For small subband separations, carrier–carrier scattering and—to a lesser extent—acoustic phonon scattering govern the population relaxation on a timescale of up to tens of picoseconds.

Intersubband scattering is followed by carrier thermalization, i.e., by the formation of a carrier distribution that can be described by a common carrier temperature, which is higher than the lattice temperature. The time needed to obtain such a thermalized carrier distribution depends on the electron concentration, the fraction of electrons excited and on the excess energy of the back-scattered carriers, which is given by the intersubband transition energy. In most cases, thermalization occurs on a timescale between 100 fs and 2 ps. Eventually, the hot carrier gas is cooled to lattice temperature, which means that the excess energy contained in the carrier gas is transferred to the lattice by emission of optical and acoustic phonons. In most cases, the cooling of the carrier gas occurs on a timescale of several ten picoseconds. The exact value depends on the amount of excess energy and on the lattice temperature. A review of carrier relaxation after intersubband excitation has been presented in [36].

The relaxation scenario described in this section has been derived from measurements with quantum wells, into which electrons were introduced by (modulation) doping. In such cases, a stationary equilibrium distribution of electrons in the lowest subband defines the system prior to intersubband excitation. In electrically biased quantum cascade structures, a much more complex situation exists: carrier injection on a timescale long compared to the microscopic transport and scattering processes generates a quasi-stationary *nonequilibrium* distribution [9], in which transport through barriers, i.e., real-space transfer, occurs on similar ultrafast timescales to carrier relaxation. So far, the intrinsic nonequilibrium dynamics under such conditions has remained mainly unexplored.

3. Experimental techniques

A real-time study of electron transport in quantum cascade structures requires femtosecond time resolution, i.e., it has to rely on optical techniques. The experiments reviewed here are based on a pump–probe scheme illustrated schematically in figure 3. The quantum cascade structure is held under forward bias, generating a quasi-stationary electron distribution in the sample (figure 3(a)). A strong pump pulse, resonant to the laser transition between subbands 3 and 2, depletes the electrically induced optical gain (figure 3(b)) via stimulated emission, promoting carriers from subband 3 into subband 2. A weak probe pulse at the same spectral position monitors the subsequent gain recovery (figure 3(c)) as a function of the pump–probe

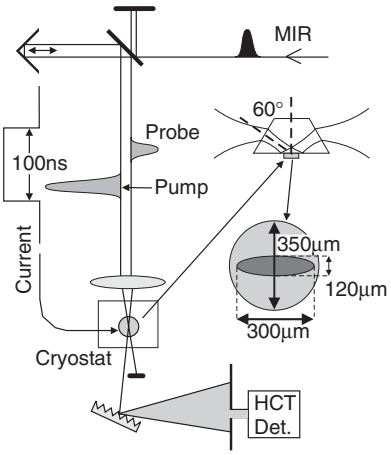

Figure 4. The experimental set-up for femtosecond pump–probe measurements on electrically driven quantum cascade laser structures. Mid-infrared pulses ($\lambda = 10\ \mu$m, pulse duration: 130 fs) are split into intense pump and weak probe pulses (intensity ratio 30:1). Subsequently they are focused onto the quantum cascade structure in the cryostat. The probe pulses are spectrally resolved (resolution 0.15 μm) after interaction with the sample. The pump–probe experiment is carried out within the time frame of 100 ns long current pulses of variable amplitude. Further shown is the geometry of beam propagation through the prism-shaped sample and the overlap between the mid-infrared laser spot and the disc-shaped quantum cascade structure (shown is the situation for sample A; in samples B and C the quantum cascade structure has elliptical shape, leading to an even better overlap with the laser spot).

delay t_D. The gain is proportional to the population difference between subbands 3 and 2 and gives direct information on the transport of electrons into subband 3 and out of subband 2.

The experimental set-up is shown schematically in figure 4. Mid-infrared pulses (centred at 10 μm, 130 fs duration, bandwidth 1.5 μm, repetition rate 1 kHz) are generated by difference frequency mixing in a GaSe crystal [37]. They are split into intense pump and weak probe pulses. A translation stage allows one to adjust the temporal delay between pump and probe pulses. Both pulses are subsequently focused onto the quantum cascade structure. After interaction with the sample the probe pulses are spectrally dispersed (resolution: 0.15 μm) and detected with a cooled HgCdTe detector. For an optimal coupling of the incident light pulses with the p-polarized intersubband gain, we use a 60° prism geometry. The samples were processed into mesas either of circular shape with a diameter of 350 μm (sample A) or of elliptical shape with dimensions of 200 μm $\times$ 400 μm (samples B and C). These sizes allow for a sufficient overlap with the elliptically shaped beam profile in the plane of the quantum cascade structure and keep at the same time the total current at reasonably small values (<10 A) (see the discussion in [38]). The electric current is applied in the form of rectangular 100 ns long pulses of variable amplitude synchronized with the optical pulses.

Even using rather large mesa dimensions, a non-negligible amount of pump light is scattered from the mesa onto the probe detector. Scanning the delay between pump and probe, this scattered light leads to large interference fringes in the pump–probe signal. Suppressing these interference fringes is possible by modulation of the current pulses: we apply current pulses only for every second optical pulse and subtract from the pump–probe signal with current that obtained without current. Hence, the resulting transients represent the current-induced transmission change for the probe as a function of the time delay between pump and probe pulses.

Table 1. Parameters for the GaAs/$Al_xGa_{1-x}As$ samples A, B and C. The data for ΔE_c (conduction band offset), τ_2 and τ_3 are taken from [39, 40].

	A	B	C
Nominal Al fraction (%)	33	45	
Injection barrier width (nm)	6.2	4.0	4.6
Exit barrier width (nm)	3.4	2.8	
Number of periods	10	10	
Sheet density per period (cm^{-2})	4×10^{11}	4×10^{11}	
ΔE_c (meV)	295	390	
τ_2 (ps)	0.3	0.3	
τ_3 (ps)	1.5	1.4	
Mesa shape	Circular	Elliptical	
Mesa dimensions	Diameter 350 μm	200 μm × 400 μm	

Time-resolved experiments were carried out on three different samples. The parameters of the samples are summarized in table 1. Sample A is a GaAs/$Al_xGa_{1-x}As$ quantum cascade structure (aluminium content: $x = 33\%$) grown at Thales-CSF (previously Thomson) in Paris. It contains ten periods of the active layer structure, similar to the one in [39], and represents the layer structure of the first quantum cascade laser based on the GaAs/$Al_xGa_{1-x}As$ material system.

The conduction band diagram of the active region and the relevant wavefunctions of sample A are shown in figure 1. The active region consists of three coupled GaAs quantum wells. The injector is a graded superlattice containing five coupled GaAs quantum wells. The injector region and the active region are separated by a 6.2 nm thick AlGaAs injection barrier and a 3.4 nm thick AlGaAs exit barrier. To reduce the lifetime of subband 2 (lower laser states), the 1–2 subband spacing corresponds to the energy of an LO phonon ($\hbar\omega_{LO} = 36$ meV). The calculated lifetimes of subband 3 and subband 2 are $\tau_3 = 1.5$ ps and $\tau_2 = 0.3$ ps [39].

Samples B and C were grown at the Technical University of Vienna. The layer structures of these samples are similar to that of the GaAs/$Al_xGa_{1-x}As$ quantum cascade laser presented in [40], the first GaAs-based quantum cascade laser operating at room temperature. The aluminium content of this structure is $x = 45\%$. The conduction band offset $\Delta E_c \approx 390$ meV for Γ-valley electrons at this aluminium content is 95 meV higher than the offset at 33% aluminium content [40]. The active region consists of three GaAs quantum wells and the injector contains five GaAs quantum wells, i.e., the layer sequence is the same as for sample A (figure 1). The calculated lifetimes of subband 3 and subband 2 are $\tau_3 = 1.4$ ps and $\tau_2 = 0.3$ ps [40]. The only difference between samples B and C is the thickness d_b of the injection barrier, which is $d_b = 4.0$ nm for sample B and $d_b = 4.6$ nm for sample C.

4. Time-resolved studies of electron transport

4.1. Oscillatory electron transport

Our experiments reveal a prototypical oscillatory behaviour of electron transport after gain saturation. The basic features observed and a qualitative explanation are presented in this section.

Sample A, which has the conduction band diagram shown in figure 1, displays electroluminescence at wavelengths around $\lambda = 10\,\mu$m for a current density of $j = 7$ kA cm^{-2} (figure 5 (f)). This spontaneous emission is due to optical transitions between subbands 3 and 2 in the active region (figure 1).

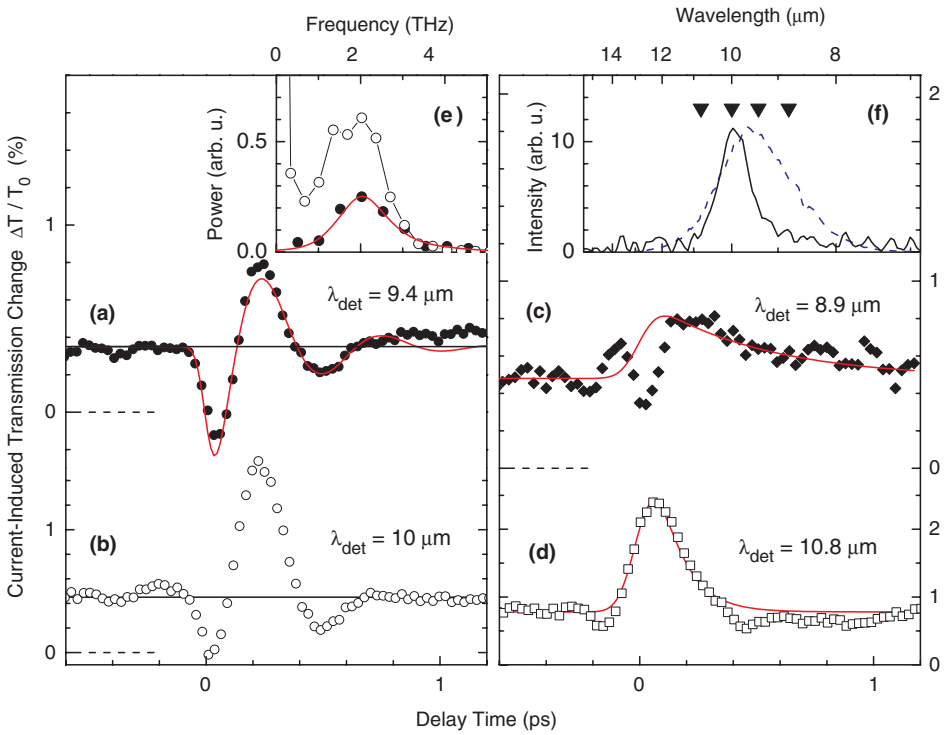

Figure 5. (a)–(d) Symbols: pump–probe transients obtained from sample A at various detection wavelengths λ_{det} (see the arrows in (f)) for a current density of 7 kA cm^{-2} and a lattice temperature of 10 K. The pump pulse energy is 2.5 nJ and the probe pulse energy is 75 pJ. The current-induced optical transmission change $\Delta T/T_0$ is plotted as a function of the time delay t_D between pump and probe pulses. Solid lines: exponentially decaying cosine function (a) or decaying exponential functions ((c) and (d)) convoluted with the cross-correlation between pump and probe. (e) Fourier transforms of transients (a) and (b). (f) Electroluminescence spectrum for a current density of 7 kA cm^{-2} (solid curve) and spectrum of pump and probe pulses (dashed curve).

In figure 5, femtosecond transients are shown for various detection wavelengths λ_{det} (see the arrows in figure 5(f)). The current-induced optical transmission change $\Delta T/T_0(t_D) = [T(j, t_D) - T(j = 0, t_D)]/T(j = 0, t_D)$ is plotted as a function of the time delay between pump and probe pulses. $T(j, t_D)$ is the transmission of the sample for a current density j and a pump–probe delay t_D. The sample temperature is held at 10 K, and the pulse energies of the pump and probe are 2.5 nJ and 75 pJ, respectively. The pump has sufficient energy to completely deplete the current-induced gain, whereas the probe is in the regime of linear response. In figures 5(a) and (b), we show experimental results for detection wavelengths resonant to the $3 \leftrightarrow 2$ transition, where the unbiased quantum cascade structure is transparent. At negative delay times, all transients show a finite increase of transmission, which is due to stimulated emission on the $3 \leftrightarrow 2$ transition, i.e., to current-induced gain [38]. In figure 5(a), an ultrafast saturation of the current-induced gain occurs at time delay zero, followed by an oscillation with a period of ≈500 fs and a decay rate of 3–4 ps^{-1}. The Fourier transform of this transient (figure 5(e)) exhibits a peak centred at 2 THz. Tuning the detection wavelength to the electroluminescence maximum $\lambda_{det} = 10\ \mu$m (figure 5(b)) we find an oscillation with larger amplitude and identical frequency (compare the Fourier transforms in figure 5(e)). In contrast, for detection wavelengths outside the electroluminescence band,

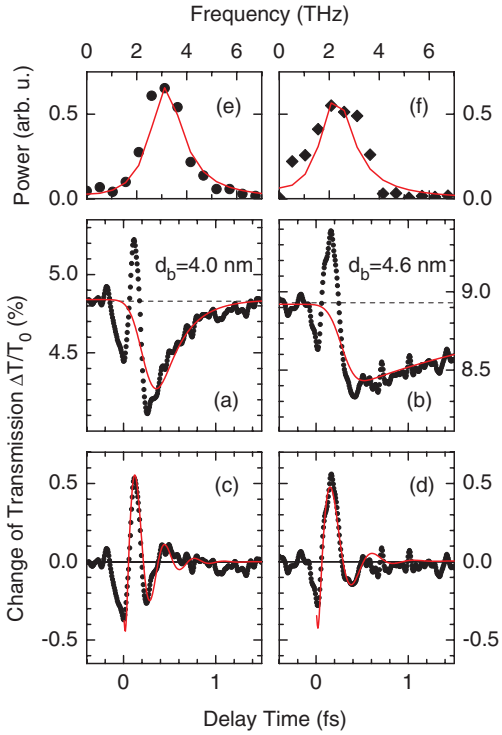

Figure 6. (a), (b) Symbols: pump–probe transients of the gain saturation and recovery for samples B (injection barrier $d_b = 4.0$ nm) and C ($d_b = 4.6$ nm) measured at a detection wavelength of $\lambda_{det} = 9.0$ μm for a current density of $j = 4.8$ kA cm^{-2} and a lattice temperature of $T_L = 150$ K. The pump pulse energy is 0.75 nJ and the probe pulse energy is 75 pJ. The pump pulse is centred at 9.0 μm. Solid curves: gradually rising and decaying component of the signal obtained by a low-pass Fourier filter. (c), (d) Symbols: oscillatory component of the pump–probe signal extracted by subtracting the slowly varying non-oscillating component (curves in (a) and (b)). Solid lines: exponentially decaying cosine fit functions. (e), (f) Fourier transforms of the oscillatory components shown in (c) and (d).

the transients follow a different dynamics (figures 5(c) and (d)): upon femtosecond excitation, the transmission increases within the time resolution and decays subsequently with a decay time strongly depending on λ_{det}. In particular, there are no oscillatory features for $t_D > 0$. (The modulation on the transient in figure 5(c) is caused by the coherent coupling of pump and probe pulses [41][4].) Those pump–probe signals outside the electroluminescence band stem from various intersubband transitions within the injector. The absorption coefficients of such intersubband transitions within the injector are typically much larger and have a stronger electric field dependence than the laser transition [42]. In the following we will concentrate exclusively on spectral positions dominated by the electrically induced gain.

The gain oscillation for forward bias shows a pronounced dependence on the width and height of the injection barrier. In figures 6(a) and (b), pump–probe transients are shown for

[4] Components from different interaction sequences of pump and probe fields can contribute to pseudo-two-colour pump–probe signals (see, e.g. [41]): (i) perturbed free induction decay, (ii) coherent pump–probe coupling and (iii) strictly sequential pump probe interaction. Whereas (i) and (ii) contribute only at negative delay times and during pump–probe overlap ($t_D < 70$ fs), the dynamics of both population differences and quantum coherences can be unambiguously observed in the period (iii) ($t_D > 70$ fs).

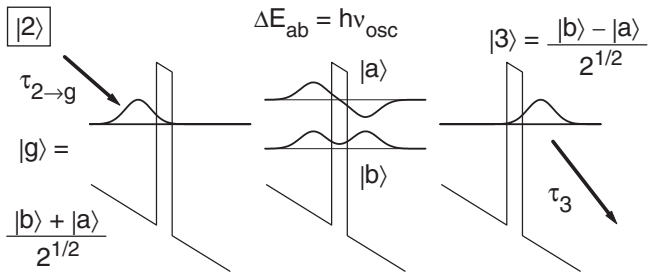

Figure 7. A schematic diagram of the coherent transport of electrons from the ground state in the injector |g⟩ into the upper laser state |3⟩ in the dynamic picture of resonant tunnelling. |b⟩ and |a⟩ are binding and anti-binding energy eigenstates of the electronic Hamiltonian *including* the tunnel coupling; |g⟩ and |3⟩ are coherent superpositions of |b⟩ and |a⟩.

sample B with a thin injection barrier ($d_b = 4.0$ nm) and for sample C with a thick injection barrier ($d_b = 4.6$ nm). For better comparison we have subtracted in figures 6(c) and (d) the slowly varying contribution of the signal. The two transients display distinctly different oscillation periods, as can be seen directly from the transients in figures 6 (c) and (d) and also from the Fourier transforms in figures 6(e) and (f).

The time evolution of gain saturation and recovery directly reflects the underlying transport dynamics of electrons in the quantum cascade structure. Transport in a quantum cascade structure represents a quasi-stationary *nonequilibrium* state of the electron gas [9], which is governed both by tunnelling and by energy and phase relaxation of electrons via scattering processes (disorder, electron–phonon and electron–electron scattering). In contrast to an equilibrium state, this quasi-stationary nonequilibrium state contains nonstationary, propagating wavefunctions made up of coherent superpositions of different electronic eigenstates.

The transients measured at wavelengths where stimulated emission dominates the nonlinear transmission changes (figures 5 and 6) show three dominant features for positive time delays, i.e., during gain recovery: (i) a pronounced oscillation, (ii) a strong damping of the oscillation and (iii) at half the oscillation period a prominent gain increase well above the quasi-stationary level. These three features will be discussed in the following.

(i) *Transport oscillation through the injection barrier.* Gain saturation by the intense mid-infrared pump pulse leads to a strong reduction of the electron population of subband 3, which is subsequently compensated for by transferring electrons from the injector region through the injection barrier into subband 3. Under quasi-stationary conditions, subband 2 is essentially unpopulated, since electrons leave this subband very quickly by resonant LO phonon emission into subband 2 and transfer to the strongly coupled miniband of the next injector region. Therefore, the contribution from subband 2 to the nonlinear pump–probe signal is negligible for $t_D > 100$ fs.[5] Consequently, the gain saturation and recovery dynamics is dominated by the population dynamics in the upper laser subband 3. Our data clearly demonstrate the oscillatory character of electron motion after gain saturation. This behaviour gives evidence for coherent resonant tunnelling of electrons through the injection barrier, as suggested in the theoretical work of [26, 27]. It is depicted schematically in figure 7: under bias, the subbands |g⟩ and |3⟩ are in resonance and the

[5] Due to the huge bandwidth (≈60 meV) of the injector miniband and to the resonant optical phonon coupling between state |2⟩ and other injector states the electron wavepacket leaves its initial position in the widest quantum well extremely fast (<50 fs) and loses phase coherence within the time resolution of our experiment.

tunnel coupling forms binding $|b\rangle$ and anti-binding $|a\rangle$ energy eigenstates, which are energetically split[6] by $\Delta E_{ab} = 2\hbar\Omega$. The size of this splitting is determined by the width and height of the injection barrier and by the applied voltage. Since the voltage across the quantum cascade structure is nearly constant for high enough current densities [39], ΔE_{ab} is independent of current density for $j \geqslant 1$ kA cm^{-2}. Coherent superpositions of states $|a\rangle$ and $|b\rangle$ represent the basis for wavepacket propagation from the injector ground state $|g\rangle$ into the upper laser state $|3\rangle$. In the experiment, the femtosecond pump pulse depletes the subband $|3\rangle = \sqrt{1/2}\,(|b\rangle - |a\rangle)$, and initiates a coherent wavepacket motion, by which electrons initially residing in the injector tunnel through the injection barrier into the active region. According to the energy splitting ΔE_{ab} between the binding and anti-binding energy eigenstates $|b\rangle$ and $|a\rangle$, the tunnelling carriers arrive after half the oscillation period $T_{osc} = h/\Delta E_{ab}$ in the state $|3\rangle$ of the active region. As the electron lifetime in state $|3\rangle$ ($\tau_3 \approx 1$ ps) is substantially longer than the oscillation period T_{osc} (see figures 5 and 6), most of the electrons will move back into the injector resulting in a gain depletion around $t_D = T_{osc}$. This picture of coherent transport is strongly supported by comparing the results for the three different samples. The data for sample B and C (figure 6) demonstrate that for identical barrier height the oscillation period increases with increasing barrier width, i.e., with decreasing tunnel splitting. Inspection of the data in figure 6 gives a ratio of oscillation frequencies $\nu_B/\nu_C \simeq 1.6$, in good agreement with results from eight-band $\mathbf{k}\cdot\mathbf{p}$ band structure calculations. Sample A contains a lower injection barrier (reduced Al content), which has, however, a larger width of $d_b = 6.2$ nm, resulting in an oscillation period similar to sample C.

(ii) *Damping of the oscillatory electron motion.* The fact that one observes one full oscillation period sets a lower limit for the homogeneous dephasing of the wavepacket motion close to the oscillation period: $T_2 \gtrsim T_{osc}$. This definitely rules out sub-100 fs dephasing times, which were derived indirectly from current–voltage measurements [28], as already mentioned in section 2.1. There are two mechanisms damping the oscillatory electron motion. First, there is inhomogeneous broadening of the tunnel couplings due to injection barrier width and height fluctuations. This results in a destructive interference of oscillations of different frequencies. Second, there is homogeneous dephasing, mainly by Coulomb scattering in the dense electron gas in the injector, which has a density of 4×10^{11} cm^{-2}. Recent Monte Carlo simulations of quantum cascade lasers give rather high electron–electron scattering rates $\gtrsim$10 ps^{-1} [33], much larger than $1/T_{osc}$. Obviously, as we find that T_2 is much longer than the time between two scattering events, only a fraction of scattering events can lead to a dephasing of the coherent superposition of states $|a\rangle$ and $|b\rangle$. This is in contrast to Fermi's golden rule, which assumes that every scattering event leads to a complete loss of all phase coherences. This is equivalent to assuming that particles are scattered only into eigenstates of the unperturbed Hamiltonian. On short timescales, however, quantum mechanics allows also coherent superpositions of such final states, a concept known as scattering-induced coherence [43, 44]: at scattering rates higher than $1/T_{osc}$, electrons are mainly redistributed between states localized in the injector. The coherent superposition $|g\rangle$ of the electronic eigenstates $|a\rangle$ and $|b\rangle$ is among such localized states. As a result, scattering events populating $|g\rangle$ on a timescale much faster than T_{osc} generate a coherent excitation in phase with the oscillation induced by the pump pulse. Such an intraband coherence after energy relaxation was previously observed for Bloch oscillations at much lower electron densities [45]. After the decay of

[6] In principle, this picture implies a doublet structure of the electroluminescence (EL) spectrum. However, the strong broadening of the EL spectra (see the inset of figure 5) smears out such a line structure [30].

the coherence between states |a⟩ and |b⟩ (for $t_D > 500$ fs), the optical gain returns to its quasi-stationary value by incoherent processes.

(iii) *Transient gain overshoot and scattering-induced coherence.* The gain overshoot observed at half the oscillation period (figures 5 and 6) gives additional evidence for this picture. Right after the gain depletion, electrons in state |g⟩ tunnel from the injector into the active region. This depletion in the |g⟩ subband recovers on a timescale faster than the oscillation period, allowing for additional electrons to contribute in phase to the oscillatory gain, which is temporarily much higher than the quasi-stationary value.

4.2. Quantitative analysis of coherent electron motion

After explaining the basic features of the coherent electron transport observed, we now present a systematic study of the amplitude, the frequency and the damping of the oscillatory electron motion as a function of three different parameters: the current density j, the lattice temperature T_L and the excitation density generated by the mid-infrared pump pulse.

For a systematic analysis, the oscillatory components of the pump–probe signals must be separated from other incoherent signal contributions. The Fourier transforms of the complete transients (see figure 5(e)) typically show two features: (i) a pronounced peak at the transport oscillation frequency and (ii) a narrow peak centred around $\nu = 0$ which is caused both by the stationary electrically induced gain and by an additional, slowly varying component of gain saturation and recovery typically dominating at late time delays (figures 6(a) and (b)). To isolate the incoherent signal contributions, the time-resolved data are Fourier transformed; the low-frequency part of the spectrum is selected by an appropriate low-pass filter and transformed back into the time domain by an inverse Fourier transform. We use a Gaussian filter centred at $\nu = 0$ with an appropriately chosen bandwidth given by the spectral position of the dip between the two maxima (see, e.g., the open circles in figure 5(e)). The result of such a procedure is shown in figure 6. The solid curves in figures 6(a) and (b) show the measured pump–probe transients of samples B and C. The dashed lines represent the gradually rising and decaying component of the respective transient after applying the low-pass Fourier filter. The solid curves in figures 6(c) and (d) show the oscillatory component of the pump–probe signal, which is obtained by subtracting the slowly varying non-oscillating component. In figures 6(e) and (f) we show the Fourier transforms (symbols) of the oscillatory components.

After the oscillatory components have been extracted in this way, the data are subject to the following analysis: for positive delay times $t_D > 0$, an exponentially decaying cosine function (after convolution with the cross-correlation of pump and probe pulses) with amplitude A, angular frequency ω_{ab} and damping rate γ_{ab} is fitted to the data. The fit curves and their Fourier transforms for the data of figures 6(a) and (b) are shown as solid curves in figures 6(c)–(f). This procedure gives reliable estimates for the values of A, ω_{ab} and γ_{ab}. It should be noted, however, that the absolute values derived display some uncertainty originating from the limited accuracy of the subtraction procedure, in particular when comparing different samples. The ratios of such values measured for a particular quantum cascade structure display a much smaller uncertainty than the absolute values. Therefore, we will concentrate in the following on experimental results obtained from sample A.

4.2.1. Transport oscillations at different current densities. In a first series of experiments, the wavepacket motion of electrons was studied as a function of the current density j. In figure 8, we present data for current densities between 70 A cm^{-2} and 8 kA cm^{-2}. The injection current of $j = 70$ A cm^{-2} is the lowest current density for which a significant current-induced pump–probe signal is observed for $t_D > 100$ fs. For even lower j (not shown; see figure 3

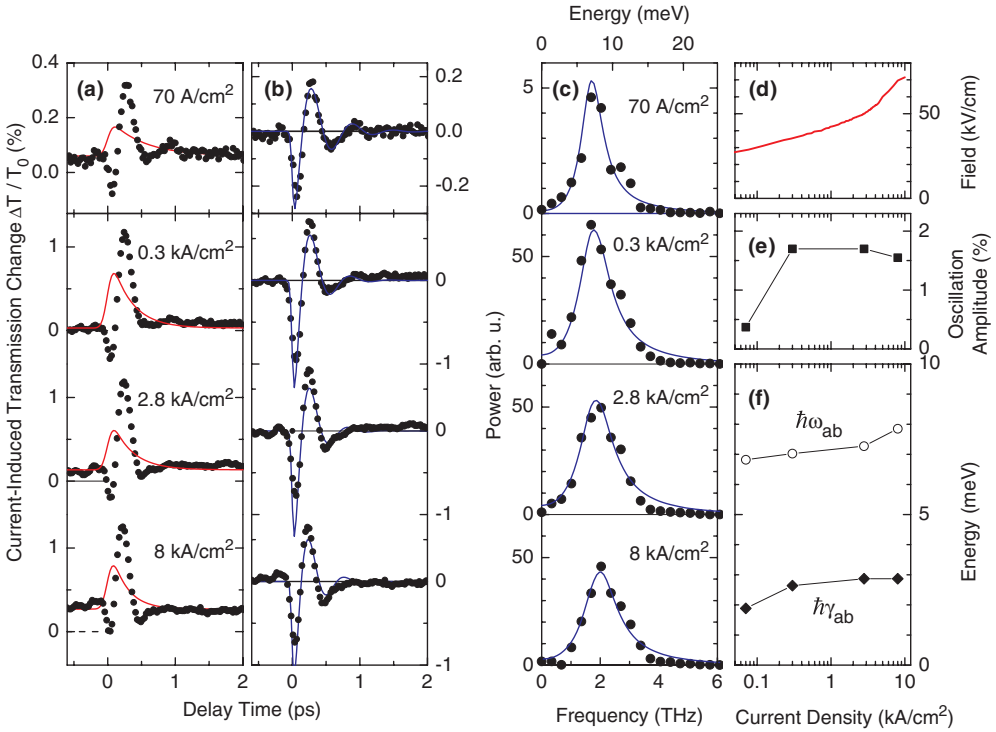

Figure 8. Amplitude, frequency and damping of the coherent transport oscillation as a function of the current density. (a) Dots: measured pump–probe transients from sample A at a detection wavelength of $\lambda_{\rm det} = 10\,\mu$m for different current densities (lattice temperature 10 K). Solid curves: the slowly varying component of the signal (i.e., the component with frequencies below 1 THz). (b) Dots: the oscillatory component of the signal, i.e., the slowly varying component has been subtracted. Solid lines: exponentially decaying cosine fit curves (after convolution with the cross-correlation of pump and probe) with amplitude A, angular frequency $\omega_{\rm ab}$ and damping rate $\gamma_{\rm ab}$. (c) Fourier transforms of the extracted oscillations (symbols) and fit curves (lines). (d) Current density versus applied electric field. (e), (f) Amplitude (squares), frequency (circles) and damping rate (diamonds) as a function of the current density.

in [22]) one finds only a signal due to coherent pump–probe coupling. This signal is centred at $t_{\rm D} = 0$ and has an amplitude that depends strongly on the intensities of pump and probe pulses. Although $j = 70$ A cm^{-2} is well below the range of current-induced quasi-stationary gain ($j > 1$ kA cm^{-2}), still a pronounced gain oscillation occurs. For this low j the oscillation has an extraordinarily long dephasing time of $T_2 = 350$ fs. For higher current densities the dephasing time shortens and the damping $\gamma_{\rm ab}$ increases, as shown in figure 8(f). In contrast to the damping, the oscillation frequency $\omega_{\rm ab}$ (circles in figure 8(f)) depends only weakly on the current density j. Since the current density is inherently connected to the applied electric field (figure 8(d)), the dependence of $\omega_{\rm ab}$ on the electric field is equally weak.

Such surprising findings can be understood from the electronic levels and wavefunctions for different applied electric fields. In figure 9, the conduction band diagrams of sample A are plotted together with probability densities $|\Psi(x)|^2$ for the relevant wavefunctions for different electric fields. The energies in figure 9(a) show clearly that in the range of electric fields used in the experiment (30–65 kV cm^{-1}) there are three distinct anticrossings between the upper laser state $|3\rangle$ and one of the levels in the preceding injector or in the active region with almost identical tunnel splittings of $\hbar\omega_{\rm ab} \approx 7$ meV.

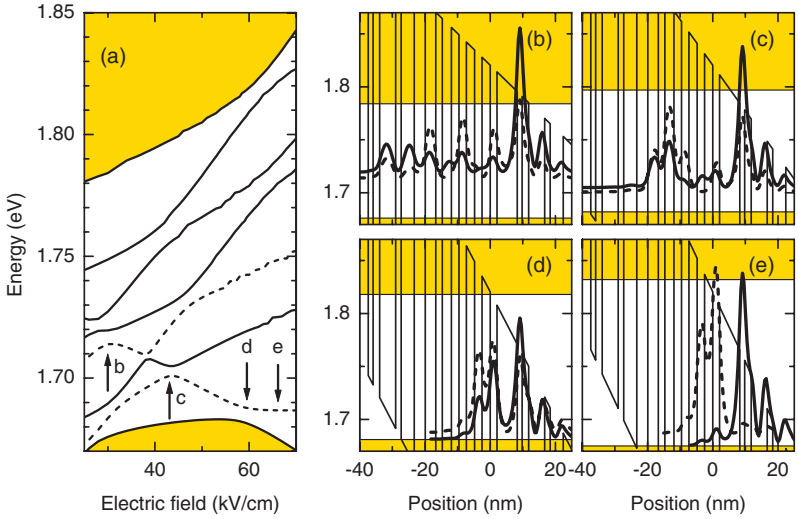

Figure 9. (a) Energies of the electronic levels in the conduction band of sample A as a function of the applied electric field (the zero of energy is the valence band maximum at $z = 0$). (b)–(e) Conduction band diagrams with probability densities $|\Psi(x)|^2$ for the relevant eigenstates of the electronic Hamiltonian for electric fields of (b) 30 kV cm^{-1}, (c) 43 kV cm^{-1}, (d) 59.5 kV cm^{-1} and (e) 66 kV cm^{-1} (see the arrows in (a)). The value in (d) is approximately equal to the laser threshold.

Such results demonstrate that the dominant channel of electron transport through the injection barrier is resonant tunnelling with relatively long decoherence times. In a real quantum cascade structure with disorder caused by well width and alloy fluctuations together with local fluctuations of the applied electric field, the major part of the current probably flows in small spatial filaments that are in exact resonance for tunnelling through the injection barrier. This picture also suggests oscillation frequencies of the electron wavepackets that are rather insensitive to the current density j, as found in the experiments (figure 8).

An interesting situation occurs at 30 kV cm^{-1} (figure 9(b)), where the lower laser state $|2\rangle$ in one period is in resonance with the upper laser state $|3\rangle$ in the next period. In this situation—which closely resembles the experimental situation of figure 8 for $j = 70$ A cm^{-2}—the femtosecond gain depletion starts the corresponding $|a\rangle$–$|b\rangle$ quantum coherence leading to a pronounced oscillation between gain and absorption on the $|3\rangle$–$|2\rangle$ transition without any current-driven stationary gain before the impulsive excitation.

In conclusion, resonant tunnelling dominates the electron transport through the injection barrier for applied fields between 30 and 60 kV cm^{-1}, i.e., even for electric fields too low to produce current-induced stationary gain. This behaviour is due to the comparably long dephasing times of the $|a\rangle$–$|b\rangle$ quantum coherence. The dephasing rates (diamonds in figure 8(f)) gradually increase with the current density j. This increase is caused by a higher carrier temperature in the injector at higher current densities j leading to higher rates of those scattering mechanisms responsible for the irreversible dephasing of the $|a\rangle$–$|b\rangle$ quantum coherence. A detailed theoretical description of this phenomenon is still to come.

4.2.2. Influence of the lattice temperature on coherent electron transport. Resonant tunnelling through the injection barrier is quite sensitive to the lattice temperature of the quantum cascade structure. Using the same analysis as in figure 8, we show in figure 10 the amplitude, frequency and damping rate of the coherent transport oscillation as a function of the

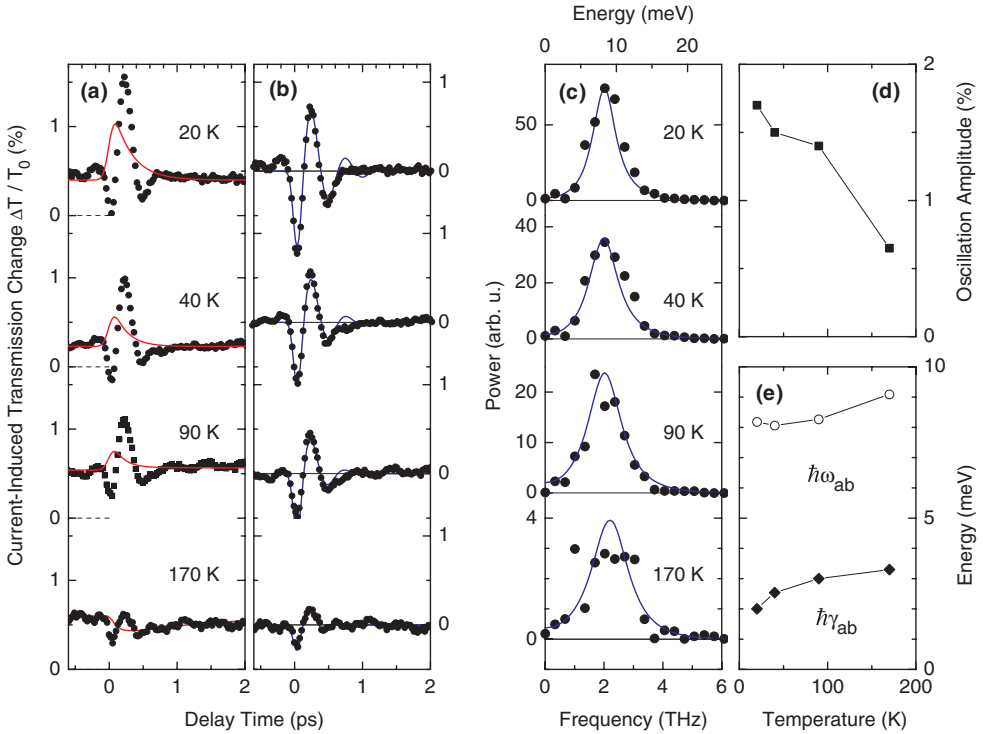

Figure 10. Like in figure 8 we show the amplitude, frequency and damping of the coherent transport oscillation as a function of the lattice temperature (current density 7 kA cm^{-2}). Increasing the lattice temperature leads to a significant reduction of the oscillation amplitude and to a pronounced enhancement of the damping rate γ_{ab}. For high lattice temperatures (e.g., at 170 K) one observes an additional slow component of gain recovery at later times.

lattice temperature T_L. In this series of measurements with sample A, the current density was set to 7 kA cm^{-2}. The energy of the pump pulse was 2.5 nJ; the energy of the probe pulse 75 pJ.

The oscillation amplitude decreases substantially when increasing the lattice temperature from 20 to 170 K (squares in figure 10(d)). In addition, the damping rate γ_{ab} (diamonds in figure 10(e)) is enhanced moderately. In contrast, the current-induced stationary gain (this manifests itself in $\Delta T/T_0$ at negative delay times in figure 10(a)) is almost independent of the lattice temperature up to $T_L = 170$ K. This shows that the population difference $n_3 - n_2 \simeq n_3$ remains nearly unchanged while raising T_L from 10 to 170 K.

The strong decrease of the oscillation amplitude (squares in figure 10(d)), which is a measure for the resonant tunnelling contribution to the gain recovery at an almost constant oscillation frequency (circles in figure 10(e)), shows that an additional incoherent carrier transfer from the injector into the upper laser subband $|3\rangle$ is activated by heating the crystal lattice. Our experimental data give no specific information on the nature of this incoherent transport. It could involve the following mechanisms:

(i) non-resonant tunnelling from thermally populated injector subbands at higher energies,
(ii) direct hopping over the injection barrier, i.e., thermal excitation of continuum states above the barrier,
(iii) thermal activation of new scattering processes directly transferring carriers across the injection barrier.

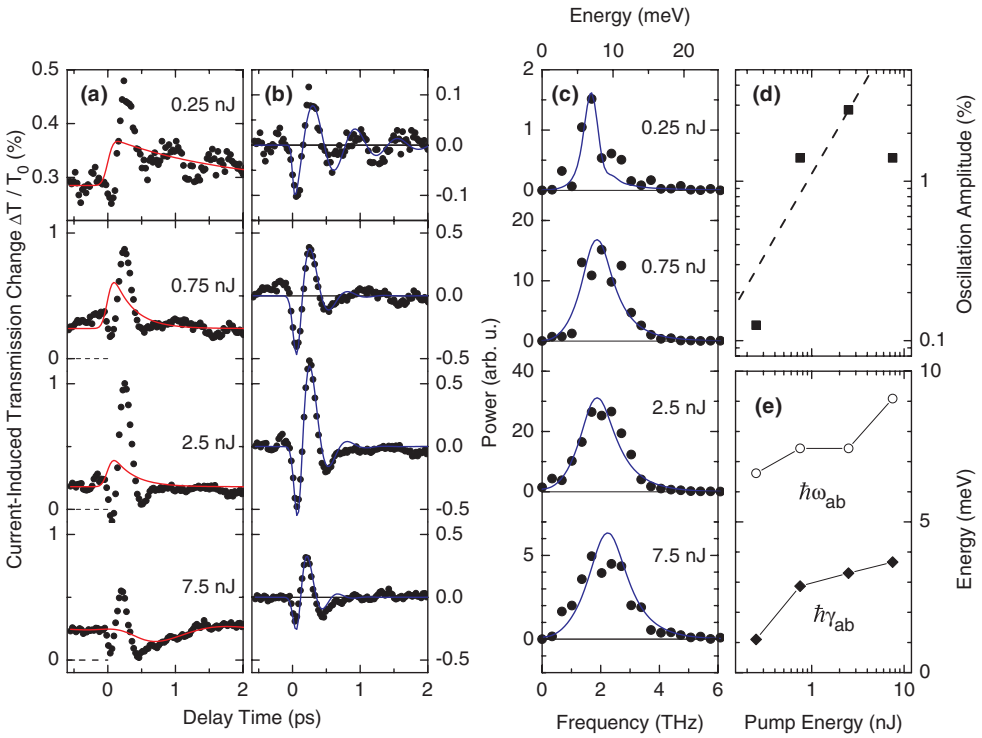

Figure 11. Excitation density dependence of the gain saturation and recovery measured in sample A at a detection wavelength of $\lambda_{\mathrm{det}} = 10\ \mu$m for a current density of 7 kA cm^{-2} and a lattice temperature of 10 K (labelling similar to figure 10). The pump intensity has a pronounced influence on both the coherent transport oscillation and the additional slow component of gain recovery at later times. For extremely low excitation densities (e.g., for 0.25 nJ) one observes quite small oscillation amplitudes but rather long decay times up to 700 fs. At higher excitation densities (e.g., at 7.5 nJ) the oscillatory motion of the electrons becomes less important and the slow component of gain recovery at later times dominates. (d) Dashed line: expected amplitude for a linear dependence on the pump pulse energy.

Possible new scattering processes that can be activated through lattice heating are intrasubband and intersubband scattering due to the absorption of longitudinal optical (LO) phonons. The rate of any LO phonon absorption process is directly proportional to the LO phonon distribution function. Using an equilibrium Bose–Einstein distribution function for the LO phonons predicts that LO absorption processes set in significantly around $T_{\mathrm{L}} = 150$ K, in close agreement with our experimental observations (figure 10).

For even higher lattice temperatures $T_{\mathrm{L}} > 170$ K, i.e., for temperatures at which quantum cascade lasers made from this structure do not work any longer [28], one observes an additional slow component of gain recovery at later times. A first indication of this component can be observed for the pump–probe transient shown in figure 10. This effect is much more pronounced for higher excitation densities as will be discussed in the following.

4.2.3. Excitation density dependence of gain saturation and recovery. We now discuss the excitation density dependence of gain saturation and recovery. Figure 11 shows the amplitude, frequency and damping of the coherent transport oscillation measured in sample A at a detection wavelength of $\lambda_{\mathrm{det}} = 10\ \mu$m for a current density of 7 kA cm^{-2} (threshold of the working

QCL [28]) and a lattice temperature of 10 K. The pump energy has a pronounced influence on both the coherent transport oscillation and the additional slow component of incoherent gain recovery at late delay times. For extremely weak excitation (e.g., for 0.25 nJ in figure 11) one observes a rather small oscillation amplitude and an extraordinarily long dephasing time of $T_2 = 700$ fs. Under such conditions of small signal excitation, the amplitude of the impulsively excited gain oscillation $A \approx 0.1$ is distinctly smaller than the electrically induced stationary gain observed at negative delay times (figure 11(a)). The amplitude of the impulsively excited gain oscillation depends almost linearly on the excitation density up to a pump pulse energy of 2.5 nJ (dashed line in figure 11(d)). Beyond this point, which corresponds to a full saturation of the optical gain, we observe a slight decrease of the oscillation amplitude A, as seen from the measurement with a pump pulse energy of 7.5 nJ shown in figure 11.

The saturation of the oscillation amplitude has a rather destructive influence on the coherence time of the $|a\rangle$–$|b\rangle$ quantum coherence as shown in figure 11(e). This phenomenon can be explained as follows: every stimulated emission process triggered by the ultrafast pump pulse will bring the energy $eEL - \hbar\omega_{23}$ into the carrier gas residing in the downstream injector (e: elementary charge, E: applied electric field, L: length of one period of the quantum cascade structure, $\hbar\omega_{23}$: photon energy of the laser transition). The total excess energy pumped into the electron gas of the injector is quite large in most of our experiments. A rough estimate of the excess energy for a pump pulse completely saturating the electrically induced gain is $j\tau_3(eEL - \hbar\omega_{23})/(eN_s)$ per carrier in the injector. For $j = 7$ kA cm^{-2} this estimate yields a value of 25 meV/carrier, thus allowing for a considerable heating of the injector plasma. The hotter the carrier gas the higher the scattering rates for irreversible dephasing of the $|a\rangle$–$|b\rangle$ quantum coherence. This leads to pronounced excitation-induced decoherence as observed in the pump–probe experiments shown in figure 11.

4.2.4. Incoherent gain recovery at high excitation densities. The physical processes discussed in the previous section also account for the incoherent gain recovery at later times observed for high pump pulse energies (see the 7.5 nJ transient in figure 11(a)). Here, the oscillatory motion of the electrons becomes less important and the slow component of gain recovery at later times dominates. Ultrafast thermalization due to the strong Coulomb interaction among the carriers forms a unified plasma of electrons in the upper laser subband and in the preceding injector at an elevated carrier temperature. The population inversion n_3–n_2 depends sensitively on the carrier temperature T_C through thermal back filling of lower laser subband $|2\rangle$ and depopulation of the subbands $|a\rangle$ and $|b\rangle$ by redistribution of carriers among several subbands within the injector. As a result, the slow dynamics of gain recovery at later times contains a significant contribution of heating and cooling dynamics of the electron plasma in the upper laser subband and the preceding injector.

In figure 12, we show pump–probe transients using a high excitation density (pump pulse energy: 7.5 nJ) for sample A for different lattice temperatures T_L (current density 7 kA cm^{-2}). At low temperatures, e.g., at 10 K, the coherent electron transport through the injection barrier still dominates. However, the time constant of the slow component of gain recovery at later times increases with lattice temperature and dominates for $T_L > 160$ K.

This behaviour is in agreement with the physical picture of plasma heating and cooling discussed above. The temperature dependent pump–probe signal (figure 12) shows a decreasing stationary gain for lattice temperatures above $T_L = 200$ K. The electrons excited by the pump pulse considerably heat up the entire electron gas as shown by the estimate presented in the previous section. Hence, for strong heating, we expect also a contribution to the pump–probe signal from the cooling of the heated electron gas. This is shown schematically in figure 13(a). In quasi-stationary equilibrium, the carrier temperature is T_C^{eq} and a certain gain

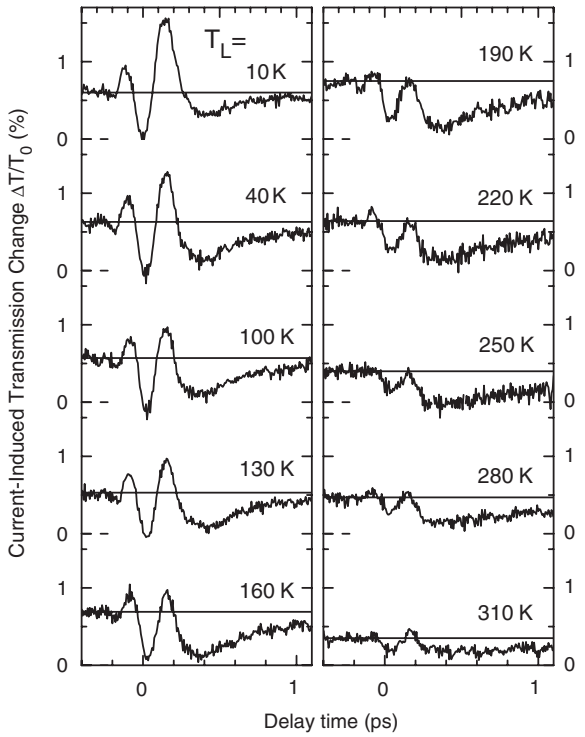

Figure 12. Pump–probe transients of the gain saturation and recovery using high excitation density (7.5 nJ) to sample A for different lattice temperatures T_{L} at a current density of 7 kA cm^{-2}. Whereas at low temperatures the coherent electron transport through the injection barrier (i.e., the oscillatory motion) dominates, the amplitude and the recovery time constant of the slow component of gain recovery increase with the lattice temperature and dominate for $T_{\mathrm{L}} > 160$ K.

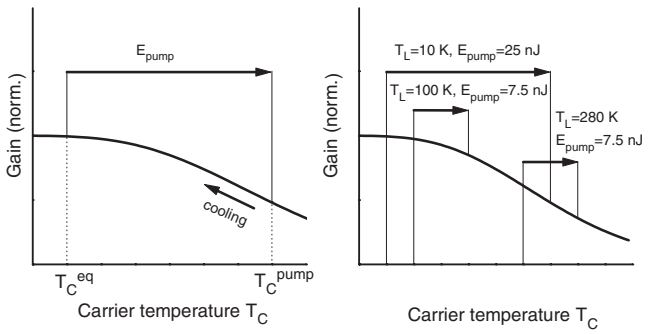

Figure 13. Schematics for the heating of the electron gas due to a strong pump pulse. (a) In quasi-stationary equilibrium the carrier temperature is $T_{\mathrm{C}}^{\mathrm{eq}}$. The pump-depleted electrons thermalize and heat up the carrier distribution ($T_{\mathrm{C}}^{\mathrm{pump}}$), which subsequently cools down again to $T_{\mathrm{C}}^{\mathrm{eq}}$. (b) Heating of the electron gas for several lattice temperatures and pump pulse energies.

is observed. At zero delay time the pump pulse promotes the carriers from subband 3 into subband 2. These 'hot' carriers thermalize on a sub-picosecond timescale [7] and heat up the entire electron gas ($T_{\mathrm{C}}^{\mathrm{pump}}$) leading to a decrease of the gain for the reasons mentioned above. Subsequently, the electron distribution cools down until it has reached $T_{\mathrm{C}}^{\mathrm{eq}}$ again.

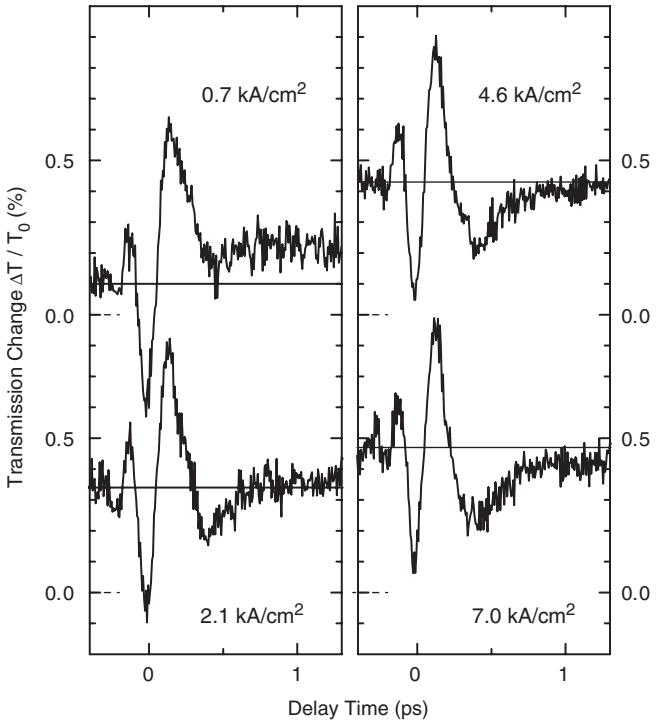

Figure 14. Pump–probe transients for sample A measured at a detection wavelength of $\lambda_{\text{det}} = 10.0\ \mu$m for various current densities j (lattice temperature: $T_{\text{L}} = 10$ K). The pump pulse energy is 7.5 nJ and the probe pulse energy is 250 pJ. The pump pulse is centred at 9.9 μm. In contrast to the case for measurements at low excitation densities (cf figure 8) an additional slow component of gain recovery at later times is observed. The amplitude of the slow gain recovery component increases monotonically with the current density and dominates for $j > 3$ kA cm^{-2}.

The heating and cooling dynamics is also observed in pump–probe transients measured with high pumping levels (pump energy 7.5 nJ) for various current densities j (figure 14, lattice temperature: $T_{\text{L}} = 10$ K). In contrast to the case for measurements at low excitation densities (figure 8), the additional slow component of gain recovery is observed at late delay times. The amplitude of the slow gain recovery component increases monotonically with the current density and dominates for $j > 3$ kA cm^{-2}. The high current density brings the system initially to an elevated carrier temperature T_{C}, from which an identical scenario starts as in the experiments shown in figure 12. Interestingly, for low current densities (e.g. 0.7 kA cm^{-2}) the pump initiated heating of the carrier gas leads after the decay of the oscillation to a higher value of gain than electrically induced before the optical excitation. In this case the upper laser subband is probably in resonance with an excited subband in the injector (e.g. figure 9(c)) which has to be thermally populated in order to maximize the optical gain at this low current density.

The heating and cooling dynamics is even more pronounced for sample B (4.0 nm injection barrier). Data for this sample measured at a lattice temperature of $T_{\text{L}} = 150$ K are shown in figure 15 for different current densities. Here the pump pulse energy is only 0.75 nJ. In contrast to the case for sample A one observes an additional slow component of gain recovery at later times also for these low excitation densities. The amplitude of the slow gain recovery component increases monotonically with the current density and dominates for $j > 6$ kA cm^{-2}.

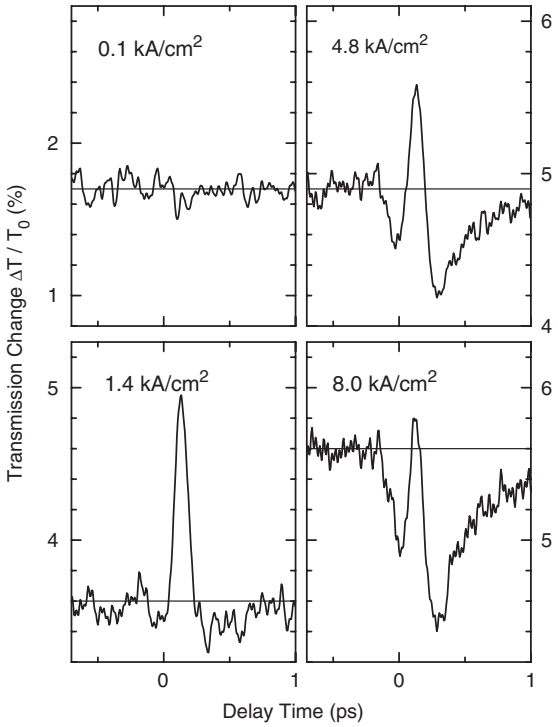

Figure 15. Pump–probe transients for sample B (4.0 nm injection barrier) measured at a detection wavelength of $\lambda_{\mathrm{det}} = 9.0\ \mu$m for various current densities (lattice temperature: $T_{\mathrm{L}} = 150$ K). The pump pulse energy is 0.75 nJ and the probe pulse energy is 75 pJ. The pump pulse is centred at 9.0 μm. In contrast to the case for sample A one observes an additional slow component of gain recovery at later times also for low excitation densities. The amplitude of the slow gain recovery component increases monotonically with the current density and dominates for $j > 6$ kA cm^{-2}.

5. Conclusions

In conclusion, we have presented a direct time-resolved study of coherent electron transport in an electrically driven quantum cascade laser structure. The ultrafast quantum transport of electrons from the injector into the upper laser state was investigated in femtosecond mid-infrared pump–probe experiments. Our data clearly show that even at the high electron densities present in a quantum cascade laser the coherence properties of the electron wavefunction play an important role for the microscopic injection process. This process is crucial for generating gain in quantum cascade lasers and represents a key step in the overall charge transport through the device. Our results strongly support the empirical finding that the design of the wavefunction overlap between the injector subbands and the upper laser subband is essential for the performance of a quantum cascade laser. Theoretical calculations that include both the quantum character of transport and the decoherence caused by electron–electron scattering are still lacking and pose a challenge for the future.

Acknowledgments

The measurements reported in this review were done by F Eickemeyer, now at Nanjing University, Nanjing, China. The samples used were provided by S Barbieri and C Sirtori from Thales-CSF, Orsay, France, and by G Strasser, T Müller, R Bratschitsch and K Unterrainer from

the Technical University of Vienna, Austria. We thank A Wacker, S-C Lee and A Knorr from the Technical University of Berlin for valuable discussions. We acknowledge financial support by the Deutsche Forschungsgemeinschaft through Forschergruppe 394, by the Fonds zur Förderung der wissenschaftlichen Forschung in Österreich (SFB-ADLIS) and by the European Union (IST SUPERSMILE).

References

[1] Capasso F, Mohammed K and Cho A Y 1986 Resonant tunneling through double barriers, perpendicular quantum transport phenomena in superlattices, and their device applications *IEEE J. Quantum Electron.* **22** 1853–69
[2] Rammer J 1991 Quantum transport theory of electrons in solids: a single-particle approach *Rev. Mod. Phys.* **63** 781–818
[3] Haug H and Jauho A P 1996 *Quantum Kinetics in Transport and Optics of Semiconductors* (Berlin: Springer)
[4] Rauch C, Strasser G, Unterrainer K, Boxleitner W, Gornik E and Wacker A 1998 Transition between coherent and incoherent electron transport in GaAs/GaAlAs superlattices *Phys. Rev. Lett.* **81** 3495–8
[5] Schöll E (ed) 1998 *Theory of Transport Properties of Semiconductor Nanostructures* (London: Chapman and Hall)
[6] Haug H 2001 Quantum kinetics for femtosecond spectroscopy in semiconductors *Ultrafast Physical Processes in Semiconductors* ed K T Tsen (San Diego, CA: Academic) pp 205–29
[7] Iotti R C and Rossi F 2001 On the nature of charge transport in quantum-cascade lasers *Phys. Rev. Lett.* **87** 146603
[8] Kleinert P and Bryksin V V 2001 Theoretical study of carrier transport in two-band semiconductor superlattices *Physica* B **304** 60–6
[9] Lee S C and Wacker A 2002 Nonequilibrium Green's function theory for transport and gain properties of quantum cascade structures *Phys. Rev.* B **66** 245314
[10] Leo K, Shah J, Göbel E O, Damen T C, Schmitt-Rink S, Schäfer W and Köhler K 1991 Coherent oscillations of a wavepacket in a semiconductor double-quantum-well structure *Phys. Rev. Lett.* **66** 201–4
[11] Feldmann J, Leo K, Shah J, Miller D A B, Cunningham J E, Meier T, von Plessen G, Schulze A, Thomas P and Schmitt-Rink S 1992 Optical investigation of Bloch oscillations in a semiconductor superlattice *Phys. Rev.* B **46** 7252–5
[12] Waschke C, Roskos H G, Schwedler R, Leo K, Kurz H and Köhler K 1993 Coherent submillimeter-wave emission from Bloch oscillations in a semiconductor superlattice *Phys. Rev. Lett.* **70** 3319–22
[13] Shah J 1999 *Ultrafast Spectroscopy of Semiconductors and Semiconductor Nanostructures* 2nd edn (Berlin: Springer)
[14] Kaindl R A, Lutgen S, Woerner M, Elsaesser T, Nottelmann B, Axt V M, Kuhn T, Hase A and Künzel H 1998 Ultrafast dephasing of coherent intersubband polarizations in a quasi-two-dimensional electron plasma *Phys. Rev. Lett.* **80** 3575–8
[15] Lutgen S, Kaindl R A, Woerner M, Elsaesser T, Hase A, Künzel H, Gulia M, Meglio D and Lugli P 1996 Nonequilibrium dynamics in a quasi-two-dimensional electron plasma after ultrafast intersubband excitation *Phys. Rev. Lett.* **77** 3657–60
[16] Martini R, Klose G, Roskos H G, Kurz H, Grahn H T and Hey R 1996 Superradiant emission from Bloch oscillations in semiconductor superlattices *Phys. Rev.* B **54** 14325–8
[17] Faist J, Capasso F, Sivco D L, Hutchinson A L, Sirtori C and Cho A Y 1994 Quantum cascade laser *Science* **264** 553–6
[18] Gmachl C, Capasso F, Sivco D L and Cho A Y 2001 Recent progress in quantum cascade lasers and applications *Rep. Prog. Phys.* **64** 1533–601
[19] Faist J, Hofstetter D, Beck M, Aellen T, Rochat M and Blaser S 2002 High performance quantum cascade lasers: physics and applications *SPIE Proc.* **4651** 274–85
[20] Köhler R, Tredicucci A, Beltram F, Beere H E, Linfield E H, Davies A G, Ritchie D A, Iotti R C and Rossi F 2002 Terahertz semiconductor-heterostructure laser *Nature* **417** 156–9
[21] Scalari G, Ajili L, Faist J, Beere H, Linfield E, Ritchie D and Davies G 2003 Far-infrared ($\lambda \simeq 87\ \mu$m) bound-to-continuum quantum-cascade lasers operating up to 90 K *Appl. Phys. Lett.* **82** 3165–7
[22] Eickemeyer F, Reimann K, Woerner M, Elsaesser T, Lee S C, Wacker A, Barbieri S, Sirtori C and Nagle J 2002 Ultrafast coherent electron transport in GaAs/AlGaAs quantum cascade structures *Physica* B **314** 314–22
[23] Eickemeyer F, Reimann K, Woerner M, Elsaesser T, Barbieri S, Sirtori C, Strasser G, Müller T, Bratschitsch R and Unterrainer K 2002 Ultrafast coherent electron transport in semiconductor quantum cascade structures *Phys. Rev. Lett.* **89** 047402

[24] Eickemeyer F, Reimann K, Woerner M, Elsaesser T, Barbieri S, Sirtori C, Strasser G, Müller T, Bratschitsch R and Unterrainer K 2003 Ultrafast gain dynamics in quantum cascade laser structures *Proc. 26th Int. Conf. on the Physics of Semiconductors (Edinburgh, 2002)* ed A R Long and J H Davies (Bristol: Institute of Physics Publishing) p M1.3

[25] Eickemeyer F, Reimann K, Woerner M, Elsaesser T, Barbieri S, Sirtori C, Strasser G, Müller T, Bratschitsch R and Unterrainer K 2003 Ultrafast coherent electron transport in quantum cascade structures *Ultrafast Phenomena* vol 13, ed D R Miller, M M Murnane, N F Scherer and A M Weiner (Berlin: Springer) pp 356–8

[26] Kazarinov R F and Suris R A 1971 Possibility of the amplification of electromagnetic waves in a semiconductor with a superlattice *Fiz. Tekh. Poluprov.* **5** 797–800
Kazarinov R F and Suris R A 1971 *Sov. Phys.—Semicond.* **5** 707–9 (Engl. Transl.)

[27] Kazarinov R F and Suris R A 1972 Electric and electromagnetic properties of semiconductors with a superlattice *Fiz. Tekh. Poluprov.* **6** 148–62
Kazarinov R F and Suris R A 1972 *Sov. Phys.—Semicond.* **6** 120–31 (Engl. Transl.)

[28] Sirtori C, Capasso F, Faist J, Hutchinson A L, Sivco D L and Cho A Y 1998 Resonant tunneling in quantum cascade lasers *IEEE J. Quantum Electron.* **34** 1722–9

[29] Becker C, Sirtori C, Page H, Glastre G, Ortiz V, Marcadet X, Stellmacher M and Nagle J 2000 AlAs/GaAs quantum-cascade lasers based on large direct conduction band discontinuity *Appl. Phys. Lett.* **77** 463–5

[30] Barbieri S, Sirtori C, Page H, Stellmacher M and Nagle J 2001 Design strategies for GaAs-based unipolar lasers: optimum injector-active region coupling via resonant tunneling *Appl. Phys. Lett.* **78** 282–4

[31] Wilson L R, Keightley P T, Cockburn J W, Duck J P, Skolnick M S, Clark J C, Hill G, Moran M and Grey R 1999 Spectroscopic determination of the electron distribution in a quantum cascade structure *Appl. Phys. Lett.* **75** 2079–81

[32] Wilson L R, Keightley P T, Cockburn J W, Duck J P, Skolnick M S, Clark J C, Hill G, Moran M and Grey R 2000 Mid-infrared spectroscopic studies and lasing in GaAs–AlGaAs quantum cascade devices *Physica* E **7** 713–7

[33] Iotti R C and Rossi F 2000 Microscopic theory of hot-carrier relaxation in semiconductor-based quantum-cascade lasers *Appl. Phys. Lett.* **76** 2265–7

[34] Iotti R C and Rossi F 2001 Carrier thermalization versus phonon-assisted relaxation in quantum-cascade lasers: a Monte Carlo approach *Appl. Phys. Lett.* **78** 2902–4

[35] Kaindl R A, Reimann K, Woerner M, Elsaesser T, Hey R and Ploog K H 2001 Homogeneous broadening and excitation-induced dephasing of intersubband transitions in a quasi-two-dimensional electron gas *Phys. Rev.* B **63** 161308

[36] Woerner M and Elsaesser T 2001 Ultrafast nonequilibrium dynamics of intersubband excitations in quasi-two-dimensional semiconductors *Ultrafast Phenomena in Semiconductors* ed K T Tsen (New York: Springer) pp 93–160

[37] Kaindl R A, Wurm M, Reimann K, Hamm P, Weiner A M and Woerner M 2000 Generation, shaping, and characterization of intense femtosecond pulses tunable between 3 and 20 μm *J. Opt. Soc. Am.* B **17** 2086–94

[38] Eickemeyer F, Kaindl R A, Woerner M, Elsaesser T, Barbieri S, Kruck P, Sirtori C and Nagle J 2000 Large electrically induced transmission changes of GaAs/AlGaAs quantum cascade structures *Appl. Phys. Lett.* **76** 3254–6

[39] Sirtori C, Kruck P, Barbieri S, Collot P, Nagle J, Beck M, Faist J and Oesterle U 1998 GaAs/$Al_xGa_{1-x}As$ quantum cascade lasers *Appl. Phys. Lett.* **73** 3486–8

[40] Page H, Becker C, Robertson A, Glastre G, Ortiz V and Sirtori C 2001 300 K operation of a GaAs-based quantum-cascade laser at $\lambda \approx 9$ μm *Appl. Phys. Lett.* **78** 3529–31

[41] Chachisvilis M *et al* 1995 *Chem. Phys. Lett.* **234** 141

[42] Wacker A, Lee S C and Pereira M 2003 Simulation of transport and gain in quantum cascade structures *Advances in Solid State Physics* ed B Kramer (Berlin: Springer) at press
(Wacker A, Lee S C and Pereira M 2003 *Preprint* cond-mat0304562)

[43] Vaupel H, Thomas P, Kühn O, May V, Maschke K, Heberle A P, Rühle W W and Köhler K 1996 Dissipative tunneling in asymmetric double-quantum-well systems: a coherence phenomenon *Phys. Rev.* B **53** 16531–42

[44] Gurvitz S A, Bar-Joseph I and Deveaud B 1991 Quantum tunneling and relaxation in asymmetric coupled wells *Phys. Rev.* B **43** 14703–6

[45] Wolter F, Roskos H G, Haring Bolivar P, Bartels G, Kurz H, Köhler K, Grahn H T and Hey R 1997 Influence of LO-phonon emission on Bloch oscillations in semiconductor superlattices *Phys. Status Solidi* b **204** 83–6