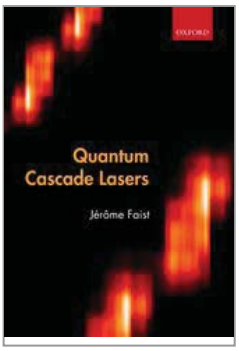


Quantum Cascade Lasers
Jérôme Faist

Print publication date: 2013
Print ISBN-13: 9780198528241
Published to Oxford Scholarship Online: May 2013
DOI: 10.1093/acprof:oso/9780198528241.001.0001

Quantum devices

Jérôme Faist

DOI:10.1093/acprof:oso/9780198528241.003.0001

Abstract and Keywords

This chapter discusses the development of quantum devices. It describes interband and intersubband transitions, which are transitions between electrons and holes, and between the electron (or holes) confined states, respectively. It then reviews previous studies on intersubband transitions, and this is followed by a discussion of mid-infrared sources.

Keywords: quantum wells, interband transitions, intersubb, mid-infrared sources

1.1 Quantum devices

Quantum mechanics was first developed and used to understand the behavior of matter at the microscopic scale. Although its birth was marked by controversies and skepticism, it was a hugely successful enterprise that managed to provide a firm theoretical background to chemistry and solid-state physics, not to speak about atomic, nuclear, and particle physics. Like virtually all successful theories, while they are used to explain ever more sophisticated phenomena, they are also turned around, becoming also a design tool used to conceive and predict properties of man-made objects.

The invention of the transistor, followed by the development of solid-state electronics has continuously pushed the fabrication of semiconductor materials with ever increasing purity as well as fabrication technologies able to pattern structures at microscopic scales. It is therefore not a surprise that, leaving aside chemistry, the first man-made quantum structures were fabricated using semiconductor technology. While being initially a topic of

research, they eventually became ubiquitous in our technology environment for both electronics and photonics. The most famous examples are the semiconductor quantum well laser, used in telecommunications, data storage, display; and the high electron mobility transistor (HEMT), used in the very-high-frequency applications.

The realization of quantum devices required a fabrication technique with nanometer accuracy. A major step was realized by the invention and development, in the end of the sixties and early seventies, of the molecular beam epitaxy (MBE) by A. Y. Cho and J. R. Arthur. MBE opened up for the first time the possibility of growing epitaxial structures with well-defined interfaces down to the atomic level; it should legitimately be considered as one of the key scientific step towards the development of nanotechnology [1]. As these interfaces exhibited a sharpness and a flatness well below the de Broglie wavelength of the charge carriers in semiconductors, they enabled the creation of quantum states controlled by confinement.

A schematic drawing illustrating the growth of a quantum well by MBE is shown in Fig. 1.1. In the situation illustrated in Fig. 1.1, the growth takes advantage of the existence of two materials, in this case GaAs and AlAs, with the same lattice spacing and crystallographic arrangement. In an ultrahigh vacuum environment, the co-evaporation of the elements of group III (Ga,Al) with the group V (As) is performed on a single crystal substrate of GaAs. Because the epitaxial growth occurs at a rate of about a monolayer per second, shutters in front of the cells enable the control of the evaporation at the atomic level. As shown in Fig. 1.1, a quantum well structure **(p.2)**

is grown when an ultrathin layer of a material with a narrower gap, say GaAs, is surrounded by layers with a material with a wider bandgap, in our case AlAs. As a result, the potential seen by the electrons and holes varies spatially and abruptly at the interface between the two materials.

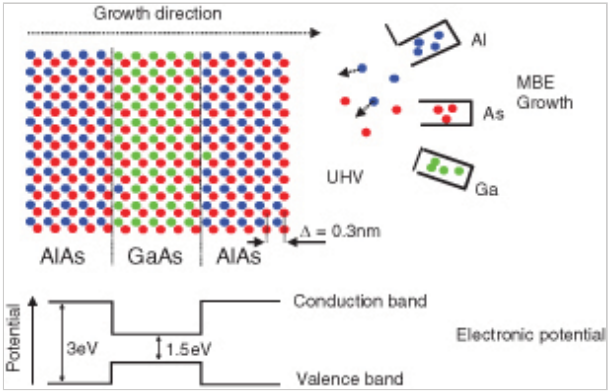


Fig. 1.1. Basic schematic showing the growth by molecular beam epitaxy of a GaAs/AlAs structure.

Because the two materials (such as GaAs and AlAs) have very similar electronic core structures, electrons and holes may cross the interface while experiencing very little disturbance apart from a change in effective mass and potential. As shown later in the text, mathematically one may describe the carrier by the same Bloch function and a slowly varying envelope function. Because the interfaces are atomically flat, the motion perpendicular and parallel to the interface remain essentially decoupled. This situation enables a very powerful semiclassical description to be applied in which the interaction between carriers and crystal can be

described by carriers with an effective mass interacting with a one-dimensional quantum well potential.

1.2 Interband and intersubband

In a celebrated paper [2], Esaki and Tsu pointed out that the additional potential introduced by a layered sequence of semiconductor layers could mimic a crystal with a periodicity much larger than the atomic one. They introduced the concept of miniband and minigap to describe the electronic state created by this artificial one-dimensional crystal and predicted the transport properties of this new structure. In particular, they realized that such an artificial structure could potentially lead to the direct observation of "Bloch oscillations" of electrons—a key concept in solid-state physics.

Another very important result was the first observation by Dingle and co-workers [3] of the characteristic blueshift and stepwise shape of the interband absorption in a series of quantum wells. This characteristic shape of the absorption edge is an image of the joint density of state of two-dimensional systems. It was soon realized by C. Henry [4] and co-workers that this property of two-dimensional systems could lead to semiconductor lasers with a much reduced threshold current and in general much higher performances. Indeed, after a few years quantum well semiconductor lasers had replaced its bulk counterpart completely.

(p.3)

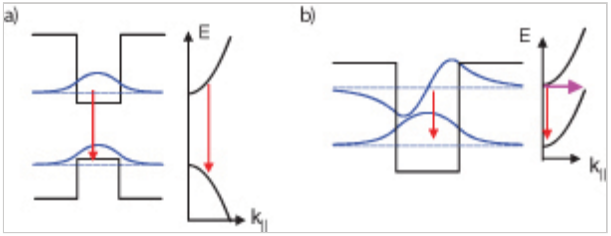


Fig. 1.2. a) An **interband** transition occurs between states (electrons and holes) that exhibit an opposite curvature of their dispersion curve. The two states are separated by a true bandgap. In clean samples, a radiative transition is the most likely recombination process. b) Intersubband transitions occur between states that belong to the same band and therefore have the same in-plane dispersion. Highly e?cient elastic or quasi-elastic non-radiative processes (interface roughness or optical phonon emission) limit the upper-state lifetime to very short values.

However, the fabrication of quantum wells had enabled consideration of not only **interband** transitions between electrons and holes, but also transitions between the electron (or holes) confined states, that were referred to as intersubband. As shown in Fig. 1.2, these transitions have a very different character. For materials with the so-called type **I** alignment, such as the one depicted in Fig. 1.2 a) where the minimum of the conduction band and the maximum of the valence band occur in the same material, the transition energy can be written as the sum of the energy gap of the well material and the confinement energies of the electrons and holes. As a result, the transition energy is limited towards lower value by the gap of the quantum well material. In contrast, in intersubband transitions, the transition energy is just the difference between the confinement energies of the individual electronic states. As a result, the

latter will tend towards zero as the well width is increased. This property is fundamental to the intersubband transitions because it enables the fabrication of devices over a very wide frequency range based on the same heterostructure materials.

As shown in Fig. 1.2, intersubband transitions occur between subband states with the same in-plane dispersion. As a result, the states are not separated by an energy gap; any elastic or inelastic transition that provides the necessary momentum exchange allows the scattering of an electron from the upper to the lower state. The electron lifetime is therefore very short—typically of the order of the picosecond—and is dominated by non-radiative processes even in the cleanest samples. This contrasts with the interband case where radiative transitions dominate the recombination in clean samples.

As mentioned already, the joint density of states for optical transitions is very different for the interband and intersubband transitions. This difference is illustrated in Fig. 1.3, showing the gain profile for the two systems. The joint density of states in intersubband transitions is delta-like, hence the system behaves optically very much like an atomic system. Gain and absorption lines are mostly broadened by scattering and exhibiting essentially symmetric lines centered around the transition energy. Like its atomic counterpart, the intersubband system is essentially transparent on both (high and low energy) sides of the transition. In contrast, interband systems are **(p.4)**

characterized by a gain spectrum that is broadened by the respective electron and hole distribution within the bands, and always displays an absorption for energies on the high energy side of the gain spectrum.

1.3 Intersubband transitions: historical aspects

Historically, intersubband transitions were first observed in Si/SiO_2 inversion layers. A number of important physical features on both experimental and theoretical aspects of intersubband transitions were already discussed in the context of this first system; these results are summarized in the landmark review of Ando, Fowler, and Stern [5]. However, these transitions occurred in the terahertz

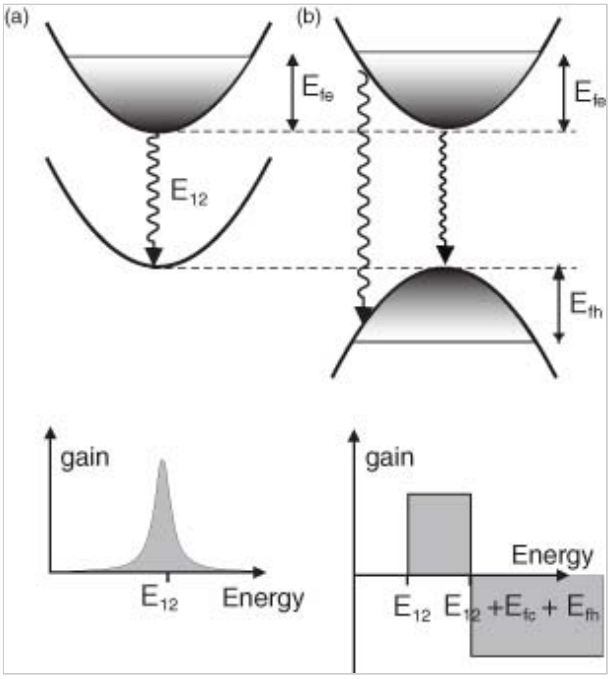


Fig. 1.3. Comparison between gain in an interband and an intersubband transitions system, shown for simplicity at zero temperature. (a) The intersubband transition joint density of states is essentially delta-like as for atomic transitions. The broadening originates from non-parabolicity as well as the broadening of the levels (b). In contrast, the joint

density of states of an interband transition has the characteristic step-like feature.

region of the spectrum, making their observation difficult experimentally; in addition the transition energy could not be tuned over a wide frequency range. As a result, the measurement of these transitions was meant as a pure spectroscopy tool enabling a better measurement of the subband spacing in these inversion layers. A significant progress was the observation of intersubband absorption in a GaAs/AlGaAs multi-quantum well by West and Eglash [6] in 1985. This work demonstrated two important features of the intersubband transitions **(p.5)** in this material system: the large value of the dipole matrix element and the fact that the transition energy could easily be tuned by changing the quantum well width.

The theoretical work of Kazarinov and Suris in 1972 must be credited with the first proposal for light amplification between discrete subband states [7]. In their work, the authors investigated the behaviour of a periodic superlattice consisting of alternating quantum well and quantum barrier layers under application of a strong applied electric field. In contrast to the work of Esaki and Tsu, they did not consider the regime of miniband transport but the high field one in which the states are mostly localized in each well. Their work did yield a number of important results. First of all, it demonstrated how intersubband transitions could be harnessed to create gain and a laser. A second important result was their analysis of the intersubband processes in the framework of a density matrix formalism [8] and the clear distinction it introduced between momentum and energy relaxation times. This work spurred a large activity, mostly theoretical, proposing variants of this structure that used resonant tunneling between quantum wells for the injection and extraction of the electrons [9–13].

Another related idea was the original proposal, in 1960, from B. Lax [14] of a laser based on an inversion between Landau levels created by a magnetic field in a solid. The “cyclotron laser” was demonstrated experimentally about twenty years later in the far infared using very lightly p-doped (about 10^{14} cm^{-3}) germanium (for a review, see [15]). In this device, holes perform a streaming motion under application of crossed electric and magnetic fields at liquid helium temperature. Depending on the magnetic field strength, the population inversion is either obtained between the light- and heavy-hole band (for the weaker fields ($B < 2T$) or between Landau levels at higher magnetic fields. Depending on the magnetic and electric field, gain is achieved in a wide wavelength range of the far-infrared, from 50 μ m to about 200 μ m. Large peak powers, in the watt range, are also obtained. However, being a bulk device, the electrical power requirements remain fairly high and prevent continuous wave operation of the device.

The first observation of emission from intersubband transitions was realized by Tsui and Gornick in an electron gas heated by a parallel current in a Si/SiO_2 inversion layer [16]. An important step was the observation of intersubband luminescence in a periodic superlattice by M. Helm in 1989 [17]. In this work, population of the excited states was achieved using resonant tunneling process; emission at the successive transition pairs in the quantum well

were correlated with the successive steps in the current–voltage characteristic of the superlattice.

The development of intersubband detectors (dubbed QWIP—Quantum Well Infrared Photocondutors) is discussed in the book by Schneider and Liu [18]. After a number of proposals and early experimental attempts, the first efficient detector using intersubband transitions was achieved at Bell Laboratories by the group led by B. Levine [19]. This work was rapidly emulated in other research labaratories, notably Thomson CSF (today Thales), and universities. Progress was rapid, as the maturity of the GaAs technology enabled shortly the fabrication of large arrays for thermal imaging. However, to reach a competitive detectivity, these detectors had to be cooled close to liquid nitrogen temperature. This can be readily achieved using portable, lightweight close-circuit coolers; nevertheless the price of the final system restricts its **(p.6)** use to high-end applications. Commercial camera systems based on QWIP detectors are now commercially sold by a number of manufacturers, mainly for military applications.

The first intersubband laser, called the quantum cascade laser (QCL) was demonstrated in Bell Laboratories in 1994 [20]. The first QCLs operated with limited optical power at cryogenic temperatures. Progress was, however, rapid. Due to their significance for applications, two important milestones were the first room-temperature operation of a single frequency, distributed-feedback QCL in 1997 [21], and the first continuous-wave operation of a quantum cascade laser at room temperature in 2002 [22]. These results were especially significant, since no other semiconductor laser source was operating at room temperature in the mid-infrared wavelength range. These achievements enabled the development of applications based on these sources; at the time of writing, the vast majority of the commercial applications are either for gas sensing or infrared countermeasures. An important aspect of quantum cascade laser technology is its versatility, as the same heterostructure material can be used to produce devices operating in a very wide wavelength range. A very important milestone in this respect was the first demonstration, in 2001, of a quantum cascade laser operating in the Terahertz region of the spectrum [23]. Operation of quantum cascade lasers has been demonstrated for frequencies spanning between 1.2 and 4.9 THz (corresponding to wavelengths λ between 60 to 250 μ m), but is up to now still limited to

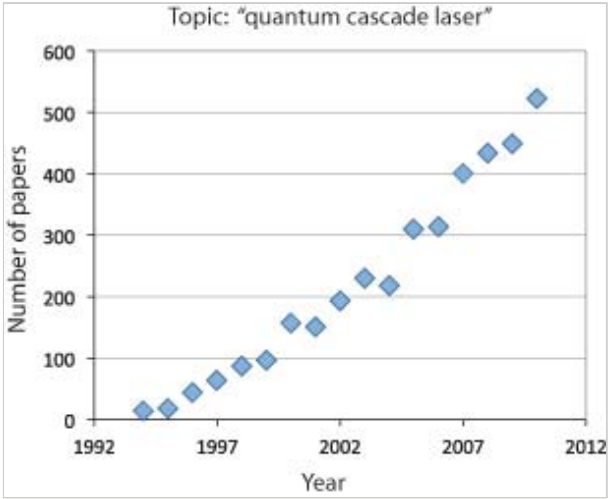


Fig. 1.4. Number of papers with "quantum cascade laser" as a topic, published every year. Source: ISI Web of Science.

cryogenic temperatures (T 〈
180 K).
Research in quantum cascade lasers remains very active, driven both by the large field of development still untapped and by the applications. A good indication of the vitality of the topic is the number of papers which, according to ISI Web of Science, **(p.7)** mention "quantum cascade laser". As shown by the graph presented in Fig. 1.4, this number, equal to 500 in 2010, keeps increasing year after year.

1.4 Mid-infrared sources

Most of the lasers based on intersubband transitions operate in the mid-infrared portion of the spectrum. The reason is twofold: on one hand, materials with a large enough conduction band discontinuity to allow emission in the near-infrared or visible, such as InGaN/AlN heterostructures, are not mature enough to allow the fabrication of quantum wells with a high enough quality to enable the realization of intersubband devices. On the other hand, interband devices are not performing well in the mid-infrared, further stimulating the development of cascade lasers as a substitution.

A schematic comparison of various mid-infrared sources is done in Fig. 1.5. CO and CO_2 gas lasers are limited to fixed emission lines. Large coverage is achieved using sources based on non-linear downconversion from near-infrared fiber lasers, but at the cost of relatively complex and expensive systems. High performance operation of (AlGaIn)(AsSb)-based quantum well (QW) interband diode lasers has been achieved, but room temperature has a limitation towards long wavelength cuto? at about λ 〈 3 μm. Interband cascade lasers (ICLs) are semiconductor lasers exploiting a cascade scheme similar to that of quantum cascade lasers, but in which the radiative transition occurs between the conduction and valence bands. The ICL concept has been first proposed by Yang [24] in 1995, one year after the first realization of a QCL. ICLs **(p.8)** have been operated at or close to room temperature in the wavelength range λ =3 − 4 μm [25]. The IV-VI lead salt semiconductor lasers can be fabricated to cover a wide frequency range [26], but have always been limited in their applications because of their limited reliability, low power and need for cryogenic operation. Quantum cascade lasers, in contrast, have demonstrated room temperature operation between 3 μm and 16 μm wavelength with large peak powers, and are therefore attractive for a number of applications.

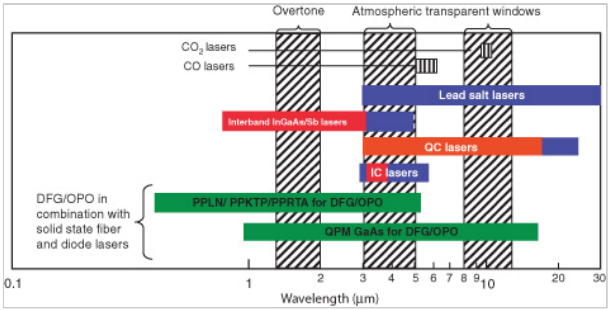
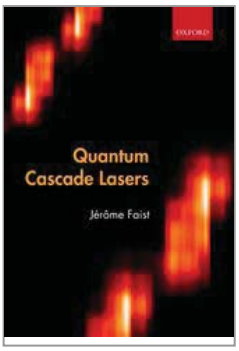


Fig. 1.5. Comparison of the wavelength coverage of various coherent sources in the near and mid-infrared. Large coverage is achieved using sources based on non-linear downconversion from near-infrared fiber lasers (two lower bars). In contrast, CO and CO_2 gas lasers are limited to fixed emission lines. Semiconductor lasers are represented by light or dark rectangles, depending on whether they operate at room temperature or cryogenic temperatures, respectively.



Quantum Cascade Lasers
Jérôme Faist

Print publication date: 2013
Print ISBN-13: 9780198528241
Published to Oxford Scholarship Online: May 2013
DOI: 10.1093/acprof:oso/9780198528241.001.0001

Technology

Jérôme Faist

DOI:10.1093/acprof:oso/9780198528241.003.0002

Abstract and Keywords

The development of quantum cascade lasers requires a series of sophisticated process steps starting with the epitaxy of the semiconductor structure, followed by the fabrication of the waveguide and contacting structures, and finishing with the mounting and encapsulation of the device. This chapter reviews the various steps, with an emphasis on the aspects relevant to the device performances.

Keywords: quantum cascade lasers, epitaxy, semiconductors, mounting, fabrication

The realization of quantum cascade lasers requires a series of sophisticated process steps, starting with the epitaxy of the semiconductor structure, followed by the fabrication of the waveguide and contacting structures, and finishing with the mounting and encapsulation of the device. We will briefly review the various steps, with an emphasis on the aspects that are relevant to the device performances.

2.1 Epitaxial layers

Well-defined and long-lived electronic states are obtained only in epitaxial heterostructures, where the spatial coherence of the crystal is achieved over large distances. In addition, to provide the necessary well-defined electronic states, the high crystalline quality of the epitaxial material is essential for obtaining the necessary mechanical stability under the huge thermal and electrical stresses under which the device has to operate. Obtaining devices with an active area in the order of 0.01 mm^2 with essentially *no* crystalline defects,

necessary for the achievement of devices operating reliably, is a feat that could be achieved only because of the large investment in the base technologies of optical telecommunications, which could be recycled for quantum cascade structures.

The first and essential step in the fabrication of quantum cascade lasers is the growth of the active region, consisting of thousands of layers with a high crystalline and chemical purity, along with a very tight control of the layer thickness down to the submonolayer level. Two key techniques are used to grow quantum cascade lasers: molecular beam epitaxy (MBE) and metalorganic vapor phase epitaxy (MOCVD).

2.1.1 Molecular beam epitaxy

The molecular beam epitaxy (MBE) is a thin-film deposition technique which allows fabrication of films of exceptional crystalline quality. Initially developed by Cho in Bell Laboratories, it quickly became a source of research interest and a tool for the fabrication of many novel semiconductor structures for both research and production.

An MBE system is basically a thermal evaporator with an exceptionally low base pressure (around 10^{-11} mbar) and very clean environment. As shown schematically in Fig. 2.1, the principal components of the MBE are thermal sources, also called Knudsen cells or simply cells, which hold the source materials, the substrate manipulator carrying a bare substrate on which the fabricated layer is grown and the growth chamber with ultra-high vacuum equipped with high pumping capacity. Other MBE components are represented by the analytical tools which allow in situ characterization **(p.10)**

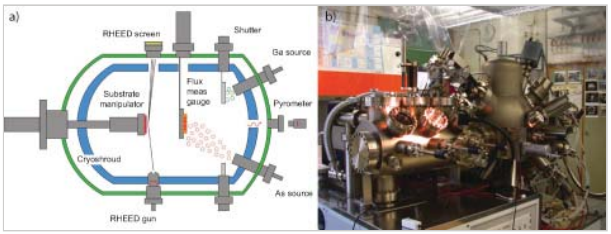


Fig. 2.1. a) Deposition chamber of an MBE system. Two sources and common in situ analytical tools are shown. b) Picture of a small research-scale system V80H by the company VG Semicon.

of the growth: the substrate temperature measurements achieved by pyrometers and thermocouples, the mass spectrometer for the determination of the background species present in the chamber, and reflection high-energy electron diffraction (RHEED) system used for surface sensitive characterization of the grown layer. Pumping is achieved by connecting the chamber to a high-capacity vaccum pump, typically ion and cryopumps, and by surrounding the growth chamber by a jacket cooled constantly at 77 K by liquid nitrogen.

A typical MBE system also has several separated vacuum chambers which allow introduction of the substrate without the need to break the vacuum. In an MBE chamber, epitaxial growth proceed at a rate of 0.5–1 μm per hour, corresponding to 0.5–1 ML per second. Changing material is achieved by the sequential opening and closing of mechanical shutters.

For the growth of quantum cascade lasers, semiconductor layers with a total thickness between 3 μm and 10 μm must be grown with a layer accuracy of about 1 Å for individual quantum wells and barriers. These requirements place strong constraints on the transition time and temporal accuracy of the mechanical shutters of the cells to characteristic times τ below 100 ms. In addition, the overall accuracy for the layer thicknesses as well as for the alloy composition requires flux accuracy of about 1% or better, during the whole duration of the growth, depending on the specific laser structure. This is especially true for thick structures grown from lattice-matched $Ga_{0.47}In_{0.53}As$/ $Al_{0.48}In_{0.52}As$ materials, as a deviation from the nominal indium content by as little as 1% yields to a measurable strain of $\epsilon = 7 \times 10^{-4}$. As a result, the cells must also have an excellent short and long term flux and temperature stabilites 〈 1°C. The calibration techniques needed to achieve such an accuracy rely either on optical reflectrometry or X-ray diffraction.

The development of MBE growth was instrumental in achieving the first intersubband devices, quantum well infrared photoconductors (QWIP), and QCL. In addition, the outstanding control of the doping profile in MBE makes it the preferred growth technique for QWIP.

(p.11) 2.1.2 Metalorganic chemical vapor deposition

In contrast to the MBE system in which the growth proceeds in a UHV environment, in the metalorganic chemical vapor deposition (MOCVD) reactor the epitaxial growth is achieved by the thermal decomposition, on the single crystalline substrate, of precursors that will free the elements III and V as well as the dopants species. As shown schematically in Fig. 2.2, group V hydrides and group III organometallics are flown using a carrier gas (usually H_2 or N_2) in the reactor. In the latter, the substrate, positioned on a susceptor, is heated to the growth temperature by halogen lamps. The decomposition of the hydrides and organometallics on the substrate free the group III and group V elements that will grow epitaxially. The remaining gases are pumped out from the chamber and are neutralized in a scrubber. Composition and layer sequences are controlled by the gas fluxes and gas source switching.

Because it does not operate under UHV, a MOCVD system has less in situ monitor capabilites than an MBE. For the same reasons, however, the downtime after a service is significantly reduced. Compared to the MBE, the MOCVD has the advantage that the fluxes, adjusted by mass flow controllers, are not susceptible to source depletion. For the $Ga_{0.47}In_{0.53}As$-based materials, the growth proceeds at a substrate temperature of

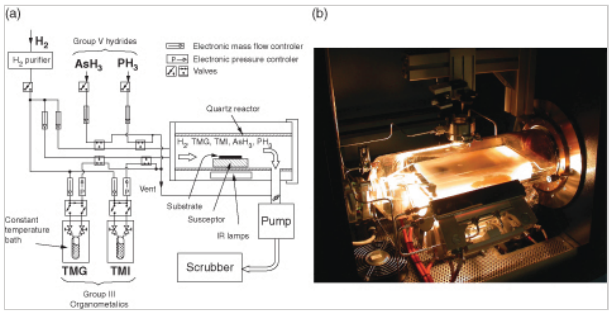


Fig. 2.2. a) Schematic diagram of a MOCVD system. b) Picture of the quartz reaction chamber of a horizontal flow MOCVD system during growth. Picture courtesy of Dr. E. Gini.

approximately 600°C, about 100°C higher than for MBE layers. Probably as a result of this, otherwise identical QCLs grown using MOCVD operate at a slightly smaller photon energy (30–50 cm^{-1}, or about 2–3% of the emission frequency) than their MBE grown counterparts. The higher growth temperature leads probably to interfaces that exhibit more grading in the composition.

A unique capability of MOCVD is the possibility of performing selective growth of InP on a SiO_2 or Si_3N_4 masked surface. In addition, in general the growth of phosphide materials is easier using MOCVD than MBE. For this reason, a combination of MBE **(p.12)** growth for the active region followed by MOCVD for the top InP cladding and lateral regrowth is often used.

The first QCL active regions grown by MOCVD were reported almost ten years after the QCL's first demonstration by the group at the University of Shefleld [27]. Shortly after this first demonstration, a collaboration between the group led by F. Capasso at Harvard and Agilent Technologies reported buried heterostructure quantum cascade lasers grown by MOCVD with very high performance levels [28]. Experience of nominally identical structures using lattice-matched $Ga_{0.47}In_{0.53}As$/ $Al_{0.48}In_{0.52}As$ material grown either by Andreas Bachle at AL Technologies using MOCVD or by Marcella Giovannini at the University of Neuchatel by MBE did not show any significant differences in performance levels.

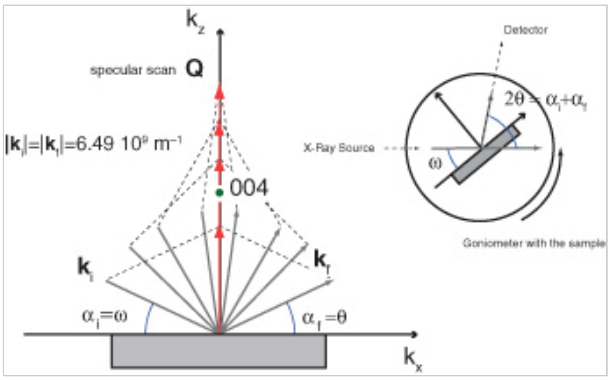


Fig. 2.3. Geometry of a typical HRXRD experiment in a specular scan. Incident (ω) and detection 2Θ angles are continuously varied. Maxima of the detected X-ray signal occur whenever the scattering vector $\mathbf{Q} = \mathbf{k_f} - \mathbf{k_i} = \mathbf{G}$, where **G** is a reciprocal lattice vector of the structure, is met. Inset: Typical arrangement of a HRXRD system: the X-ray source is fixed, while the sample and detectors are mounted on rotary stages.

2.1.3 Layer characterization

High-resolution X-ray diffraction. Because it yields highly accurate measurements of thicknesses and lattice parameters while remaining non-destructive, high-resolution X-ray diffraction (HRXRD) is one of the key analytical tools for quantum cascade lasers. In fact, X-ray spectra with broadened peaks are usually well correlated with lower QCL performances. For a deeper discussion of HRXRD, the reader is referred to the books by Bowen [29] and Ulrich [30], in which the fundamentals of HRXRD are described in detail. As shown in the inset of Fig. 2.3, in a high-resolution X-ray diffraction (HRXRD) spectrometer, a collimated source beam is incident on the crystal surface, diffracted by the

parallel atomic planes of the epitaxial layers and substrates and its intensity measured by a detector.

As the X-ray reflection is an elastic process the magnitude of both incident $\mathbf{k_i}$ and diffracted $\mathbf{k_f}$ vectors is identical and equals $\frac{2\pi}{\lambda}$, with λ = 0.154 nm in case of a Copper **(p.13)** Kα spectral line that is a standard radiation source. In a typical experiment the X-ray source is fixed while the sample and detector are mounted on two independent rotation stages, allowing the free selection of the incident and exit angles α_i and α_f that will set the directions of $\mathbf{k_i}$ and $\mathbf{k_f}$. This leads to a well-defined scan of the scattering vector

$$\mathbf{Q} = \mathbf{k_f} - \mathbf{k_i}$$

(2.1.1)

through the reciprocal space. In typical notation used on the goniometer, the rotation of the sample is denoted by ω and that of the detector by 2θ. The relationships between the goniometer angles and α_i and α_f are therefore $\omega = \alpha_i$ and $2\theta = \alpha_i + \alpha_f$.

An example of an X-ray experiment which probes the reflectivity in the vicinity of the reciprocal vector **G** corresponding to the (400) direction is shown in Fig. 2.3. This particular arrangement, called specular scan, is obtained if the direction of scattering vector **Q** is kept constant and parallel with respect to surface normal while its magnitude is gradually increased. A diffraction maximum is then expected when the Bragg condition

$$\mathbf{Q} = \mathbf{G}$$

(2.1.2)

is met, i.e when the scattering vector is a reciprocal vector of the lattice. Typical reflections from the (100) face of a zincblende, named also face centered cubic (fcc), lattice correspond to reciprocal vectors (400), (200), (311), (511). While the reflections (400), (200) are from the atomic planes parallel with the sample surface (symmetric reflection), the (311) and (511) planes are tilted (asymmetric reflections).

The maximum of diffracted intensity described by eqn. 2.1.2 corresponds to the one generated by an homogeneous crystal. Because its unit cell in the growth direction will be slightly expanded or contracted, a layer under light strain will have its diffraction maximum at a slightly different angle as compared to the substrate peak. The measurement of the difference in angle $\delta\theta$ between the two peaks enables very accurate and sensitive measurement of lattice strain and therefore of the alloy composition. The strain, or relative mismatch between the epilayer and substrate $\delta a/a$, can be evaluated from the splitting between layer and substrate peaks with the equation

$$\frac{\delta a}{a} = -\frac{\delta\theta}{\tan\theta}$$

(2.1.3)

where $\delta\theta$ corresponds to peak separation and is small with respect to θ. The separation, for a strain of 7×10^{-4}, corresponds to an angular separation of 76 arcsec or 0.37 mrad for the (400) reflection.

A generalization of eqn. 2.1.2 shows that the plot of diffraction intensity as a function of angle is approximately the Fourier transform of the structure, with the peak amplitude modulated by the form factor of the diffraction. As a result, a periodic arrangement of layers, such as the repetition of the active region in the gain region of a QCL, will lead to a spectrogram consisting of a periodic repetition of satellite peaks. The length of the superlattice period can be extracted from the spacing of the two satellites observed in the X-ray reflection using the formula

$$a_{SL} = \frac{\lambda}{2}\frac{(i-j)}{\sin\theta_i - \sin\theta_j}$$

(2.1.4)

(p.14) where i and j denote satellite order, counted from the layer peak, θ_i and θ_j is the angular position of the maxima on the reflectivity curve, and λ is the wavelength of the X-ray radiation.

The complete analysis of the rocking X-ray curves are performed by comparing the experimental result to the predictions produced by simulation software. In particular, the number of sharpness of the superlattice peaks will yield important information on the abruptness of the interfaces and possible flux drifts during the growth, while the position of the center peak measures the possible departure from the lattice matching to the substrate.

As an example, the experimental and simulated rocking curves of a $Ga_{0.28}In_{0.72}As$/ $Al_{0.48}In_{0.52}As$/ AlAs strain-compensated QCL designed for the operation at 3.3 μm, and measured in the (004)*InP* reflection, are shown in Fig. 2.4. The high degree of crystal quality and thickness homogeneity enable the observation of more than seventy well-defined satellite peaks with full width at half-maximums (FWHM) of 18– 30 arcsecs (FWHM of the *InP* substrate is 19 arcsecs [31]). From the spacing between the satellite peaks an average period thickness of 44.8 nm is extracted with a deviation smaller than 0.5% from the designed value of 44.6 nm. The 0th-order satellite peak is slightly separated from the substrate peak by 165 arcec, indicating that the strain is not perfectly balanced between the 3.5% compressive strained AlAs and the 1.3% tensile strained $Ga_{0.28}In_{0.72}As$ layers (in respect to InP), though the overall strain is still less than 0.1%.

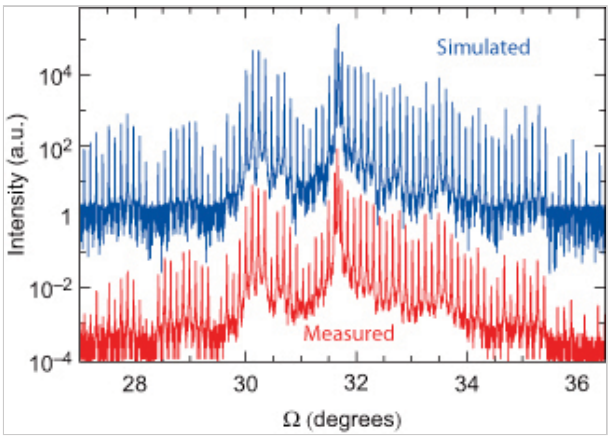


Fig. 2.4. Measured and simulated XRD curves. The simulated curve is offset for clarity. Measured data fit well with the expected compositions InGaAs/InAlAs-AlAs. An average period thickness of 44.2 nm can be derived in good agreement with

(p.15) *Transmission electron microscopy*.In a transmission

the expected one (44.6 nm). Reprinted with permission from [31]. Copyright 2011, American Institute of Physics (AIP).

electron microscope (TEM), a high-energy (200kV typical) electron beam is focused on a thin sample (20 nm) and the transmitted beam contains the information on the image with a resolution down to atomic dimensions. Using a TEM, individual quantum wells and barriers can be imaged, giving direct information on the quality of the interface and the thickness of the layers. As such, a TEM cross-section is a very complementary technique to the X-ray that convolute the overall information of the active region. However, the slightly lower accuracy as well as the time-consuming and destructive nature of the sample preparation technique makes it a less practical sample evaluation technique for routine observation.

Fig. 2.5 shows the transmission electron micrograph image of the active region of a quantum cascade laser grown by molecular beam epitaxy [32]. In this structure, a three monolayer thick (9 Å) $Ga_{0.47}In_{0.53}As$ quantum well is sandwiched between two $Al_{0.48}In_{0.52}As$ barriers and nevertheless appears with no visible fluctuation.

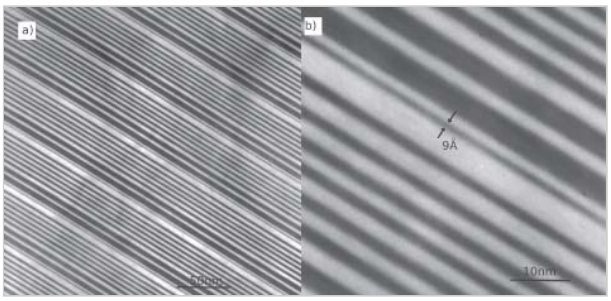


Fig. 2.5. a) TEM micrograph of a cleaved cross-section of a vertical transition quantum cascade laser active region [32]. Light regions are $Al_{0.48}In_{0.52}As$ barriers, while dark regions are $Ga_{0.47}In_{0.53}As$ quantum wells. b) One period of the active region shown at a larger magnification. The thinnest layer is a 9 Å $Ga_{0.47}In_{0.53}As$ quantum well. Micrograph courtesy of S. N. G. Chu; growth, D. L. Sivco and A. Y. Cho.

The same measurements, performed on two Si/SiGe quantum cascade electroluminescence structures [33, 34], show the importance of the growth conditions on the flatness of the interfaces, as shown in Fig. 2.6. In that case, the TEM micrograph enabled an optimization of the growth conditions.
One should note, however, that the TEM technique requires the measurement of an approximately 20 nm thick sample; as such the interface seen corresponds to the *average* over this thickness. Incidentally, this distance is of the same order of magnitude as the in-plane de Broglie wavelength ($\lambda = 36nm$) of an electron with a typical thermal energy at room temperature. A TEM micrograph may therefore **(p.16)**

measure long-distance thickness variations that would localize electrons, but does not provide useful information on the short-range disorder

responsible for the interface roughness scattering.

Scanning tunneling microscopy. In a scanning tunneling microscope [35], a very sharp tip is brought in close proximity to a surface and the tunnel current monitored. In this way, individual atoms may be resolved. Application of this technique to III–V semiconductors required the development of an instrument able to in situ cleave and measure a sample under UHV. The cleavage must proceed on the (110) face as the latter, in the absence of surface contaminants, does not exhibit band bending.

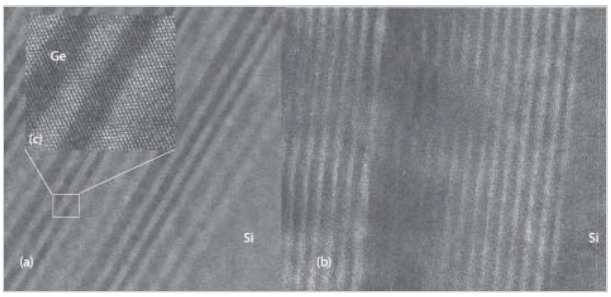


Fig. 2.6. a) TEM micrograph of a of a cross-section of aSi/SiGe quantum cascade structure, sample E027, grown on a Si wafer [33]. The Ge-rich regions are darker, the Si-rich lighter. The interfaces remain flat. b) Si/SiGe quantum cascade emitter, (L045), grown on a SiGe buffer layer [34]. The less optimal growth conditions and the larger strain induced a waviness of the interfaces, nucleating at the germanium layers. c) Close-up of sample E027, showing the individual atomic planes. Micrograph courtesy of Elisabeth Mu¨ller, LMN, Paul Scherrer Institute.

Using such an approach, scanning tunneling microscopy was performed on the cleaved edge of two $Ga_{0.47}In_{0.53}As$/$Al_{0.48}In_{0.52}As$ quantum cascade laser structures designed to operate at a wavelength of 9 μm [36]. As shown in Fig. 2.7, two samples were compared: one where the interfaces were abrupt, and one where the interfaces of the active region were intentionally digitally graded.

In agreement with the measurements that showed both samples to exhibit an identical linewidth of the electroluminescence peak, the two cross-sections showed very similar interfaces between the two samples. As shown in Fig. 2.7 (d) and (e), the aluminum profiles of two 22 Å thick $Al_{0.48}In_{0.52}As$ barriers could be fitted successfully by the same error function

$$c(x) = \frac{1}{2}\left[\mathrm{erf}\left(\frac{L/2+x}{2\sigma_x}\right) + \mathrm{erf}\left(\frac{L/2-x}{2\sigma_y}\right)\right]$$

(2.1.5)

(p.17)

with $\sigma_x = \sigma_y = 1\mathrm{Ml} = 2.8$ Å while L = 22 Å remained at the nominal value. The interpretation was that this interface grading arose as a result of the interdiffusion of the gallium and indium atoms

during growth. Note, however, that in those experiments the aluminum and gallium atoms themselves are not directly probed, but rather the electronic effects these two atomic species have on the surface arsenic atoms. The apparent height of an arsenic site should be dependent on the number of aluminum atoms in the nearest-neighbor positions at the cleaved (110) surface.

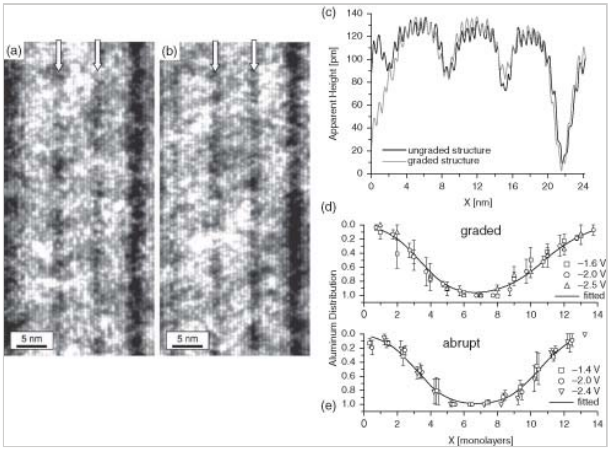


Fig. 2.7. STM image of the active region of (a) the structure with graded interfaces and (b) the structure with abrupt interfaces. (c) shows the averaged line profiles of the active region. Aluminum concentration of a 2.5 nm barrier of the structure with (d) graded interfaces and (e) abrupt interfaces. The aluminum concentration was derived from averaged line proles taken at three sample voltages. Adapted with permission from [36], courtesy of P. Offermans and P. M. Koenraad. Copyright 2003, AIP.

Atom probe.The atom probe is another technique that can resolve chemical structure at the atomic level. In this instrument, atoms are removed from a tip by a pulsed electric field. The magnification is achieved by the crowding of the field lines at the tip extremity and their spreading at long distance on a position-sensitive detector matrix. A time-of-flight mass spectrometer enables additional chemical sensitivity. The principle of operation was proposed already in the 1960s [37]. Recent progress in the instruments and in the reconstruction algorithm has now enabled three-dimensional **(p.18)** reconstruction of multi-layer semiconductor superlattices, yielding critical information on the shape of interfaces and interdiffusion effects [38]. This technique is potentially able to yield very critical information on the interface structure of QCLs.

Doping and background impurity level. Although systematic studies are still lacking, it is generally agreed that a low background impurity level, achieved by a low base vacuum pressure, is needed to achieve high-performance lasers. An inverse correlation between threshold current and HEMT mobility has been reported in a systematic growth study of terahertz QCLs [39]. As the average doping level in mid-infrared quantum cascade lasers is in the low 10^{16} cm^{-3}, the material should have a background impurity in the low 10^{15} cm^{-3}. Such a purity level is challenging to achieve in the high aluminum containing layers because of the high reactivity to the residual O_2 of the growth chamber. In [40] the performance of quantum cascade lasers were systematically measured as a function of intentional doping levels. In particular, two otherwise identical devices were compared—one (N68) grown at the beginning and another one (N120) at the end of a growth run. The growth chamber is typically cleaner, with a lower base pressure, at the end of a growth run because impurities are trapped against the walls or pumped away.

As shown in Fig. 2.8, the light- and voltage-versus-current of these two otherwise identical devices are compared. The shift in the threshold current density corresponded to a background impurity level dropping from 1.7×10^{16} cm^{-3} at the beginning of the growth run to 3×10^{15} cm^{-3} at the end. The sharper IV characteristics, as well as the larger slope efficiency of N120 compared to N68, is indeed compatible with a lower background impurity level

(p.19) .

2.2 Quantum cascade laser processing

A laser consists of a gain medium inserted in an optical resonator cavity. In a quantum cascade laser, optical gain is achieved at the cost of a thermal dissipation of 20–100 kW/cm^2. This number is a factor of 10^2–10^3 larger than for a solid-state laser and 10–10^2 for a semiconductor interband laser. This power constraint strongly restricts the kind of cavity geometries that can be used for such lasers. As a result, as in in-plane semiconductor lasers, the optical cavity is formed by the epitaxial layers themselves, as discussed more in detail in Chapter 6. In the growth direction, optical confinement is achieved by total internal reflection between the high refractive index gain region and the lower refractive index substrate and cladding layers. Laterally, confinement is achieved by defining a stripe. The typical one-dimensional waveguide geometry allows a long interaction length while minimizing the total volume, and is also favorable for thermal extraction.

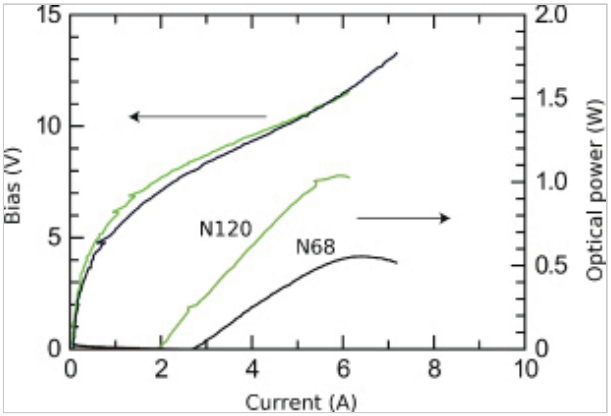


Fig. 2.8. Comparison between the light-versus-current characteristics of a laser grown at the beginning (N68, black line) and at the end (N120, grey line) of a growth run. The shift in the threshold current is attributed to the larger background impurity level at the beginning of the run [40].

2.2.1 Ridge process

In conventional semiconductor lasers, a stripe is usually achieved by etching only a fraction of the top cladding, slightly reducing the effective index of the mode on the side of the ridge. The low effective index step in such rib waveguide structures means that it supports a single transverse mode even for relatively wide stripes (about 10 λ). In addition, the low index step is also favorable for reducing the optical scattering on the sidewalls.

The same approach has been realized in QCLs, and has indeed yielded high-peak power devices [41]. The problem that arises with shallow etched QCLs is their very large operation voltage, and the anisotropic conduction properties of the active region result in a

very large current spreading. In these initial experiments an effective QCL width of 60 μm was estimated for a physical ridge width of 10 μm, greatly increasing the threshold current.

For these reasons, a ridge process where the active region is etched completely is usually preferred. In contrast to interband devices, such etching does not create non-radiative recombination sites, but a depletion layer. An example of a ridge process flow follows the steps outlined in Fig. 2.9. After MBE growth of the active region structure (a), the samples are transferred to the MOCVD and the cladding planar regrowth is performed (b). A layer of SiO_2 is then grown by plasma enhanced vapor deposition (PECVD) as a hard masking layer (c). The patterning of the laser waveguides is defined (d) and the pattern is then transferred to the SiO_2 layer by reactive ion etching (RIE), using an Ar/CHF_3 chemistry (e). The use of a hard masking layer allows the use of an HCl-based etching solution ($HCl:CH_3COOH$ (1:3)) to etch the InP cladding. This allows us to obtain vertical sidewalls and to reduce the undercutting (g). The active region is then etched using an isotropic $HBr:HNO_3:H_2O$ (1:1:10)(**h**) solution. Exhibiting a nearly uniform etch rate for all the materials (InP, AlInAs, InGaAs) a smooth surface is obtained after etching. Unfortunately, due to the isotropic behavior of the used solution, the etched structures show an undercut larger than the vertical etching depth (undercut ≈ 1.2 times the etch depth).

Once the etching of the active stack is performed, the hard mask layer is removed by HF etching and an insulating layer is deposited by PECVD on the etched structure**(i)**.

(p.20)

The overlap of the optical mode with this dielectric layer is one of the most important sources of optical losses in this waveguide configuration. In order to minimize the losses, silicon dioxide or silicon nitride are used as insulation materials, depending on the laser spectral range. Using optical lithography and RIE etching the insulating layer is opened on the ridge head. An ohmic contact is then deposited (Ti/Pt/Au, 5 nm/40 nm/100 nm) by e-beam evaporation on the opening using standard lift-off **(k)**. Finally electroplated gold pads are deposited **(m)** and the substrates are thinned to ~ 150 μm in order to reduce thermal resistance. Back ohmic contact is finally deposited on the wafer

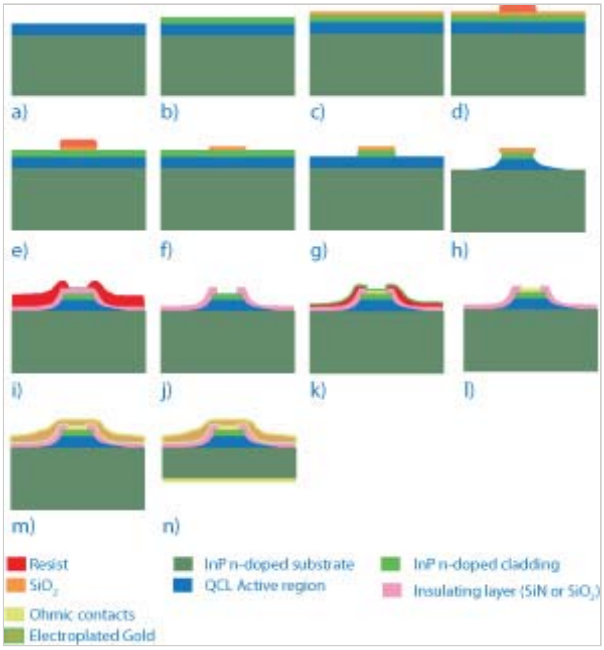


Fig. 2.9. Processing steps for the fabrication of a Fabry–Perot cavity quantum cascade laser. (a)–(b) MBE or MOCVD growth of the active region and cladding. (c)–(h) Defining and etching the ridge. (i)–(l) PECVD deposition of an insulating oxide or

bottom side (Ge/Au/Ni/Au, 15 nm/50 nm/10 nm/150 nm) **(n)**. An example of a device fabricated using this process flow is shown in Fig. 2.10.

nitride and opening it on top of the ridges. (k)–(n) Top and bottom contact metalization.

2.2.2 Buried heterostructure process

The ridge laser process has many advantages. The process is relatively simple, uses standard clean-room processes, and does not require additional regrowth. The devices have a good electrical and mechanical stability. However, large additional waveguide losses are observed for devices operating at long wavelength (λ 〉 8 μm) using narrow ridges, as both Si_3N_4 and SiO_2 have large absorption coefficients at these long wave-lengths. As a result, there was interest in substituting these dielectrics by infrared transparent materials such as ZnSe [42] or chalcogenides [43]. Nevertheless, the best **(p.21)**

results were achieved by regrowing laterally semi-insulating InP [22]—a process known in interband lasers as "buried heterostructure." This material is highly transparent, has a lower index step with the active region, therefore reducing the effect of lateral roughness to the optical loss, and has a very large thermal conductivity, enabling thermal extraction also through the sidewalls.

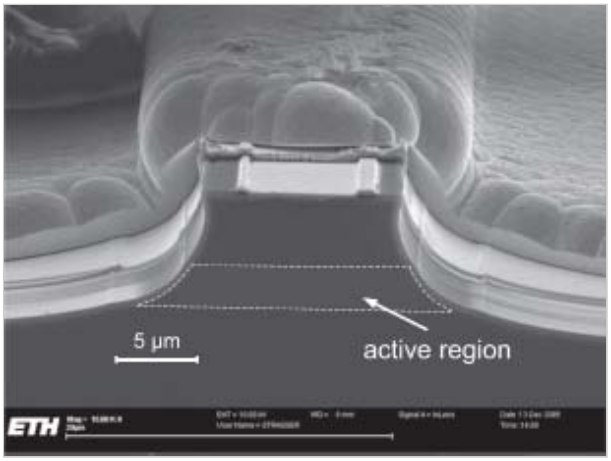


Fig. 2.10. Scanning electron micrograph of a ridge waveguide device.

The buried heterostructure process flow is shown in Fig. 2.11 and starts with processing steps which are identical to the ridge process up to the active region wet etching (a)–(h). Using the SiO_2 etching mask, an insulating InP layer (Fe doped) is regrown using MOCVD selectively on the sides of the ridge. After stripping the masking layer, by HF etching, an Si_3N_4 insulating layer is then deposited using PECVD (j) to prevent parasitic injection through defects. Since the optical mode is far from the insulating layer, its influence on the waveguide losses is irrelevant, and consequently only Si_3N_4 is used due to its slightly higher thermal conductance compared to SiO_2. The devices are then finished and metalized in a similar manner as in the ridge process (j)–(p).

Recently, a variation of this process has been developed where the order of the regrowth has been inverted [44]. The ridges are etched and regrown laterally *before* the top cladding is grown. This allows an easier fabrication of very narrow devices with vertical sidewalls, as shown in Fig. 2.12. The highly planar surface achieved after the final regrowth allows a good junction-down mounting.

Leakage currents. As InP is not an insulator, leakage through the lateral confining layers is a concern for buried heterostructure devices. As shown in Fig. 2.13a, such a current leakage path would start from the n-doped InP inside the semi-insulating InP and end up in the substrate. The leakage current is strongly dependent on the nature of the blocking layer and on the characteristics of the junction between the latter and the top and bottom claddings.

(p.22)

The characteristics of MOCVD grown n-i-n mesa structures with different blocking layers were investigated. More specifically, three 2 μm thick barrier structures are compared. In the reference sample the blocking barrier consists of an undoped InP layer. In the second sample the undoped InP layer is interrupted by 8 $Al_{0.48}In_{0.52}As$ barriers. In the final sample the blocking barrier consists of an Fe homogeneously doped InP layer. As shown in Fig. 2.13a where the current–voltage characteristics of these structures are compared, the i-InP layer does not efficiently block the current flow, and current densities above $100 A/cm^2$ are measured at 1V already. The current density through **(p.23)** the multi-barrier $Al_{0.48}In_{0.52}As$/ InP structure is already strongly reduced compared to the homogeneous layer. However, leakage currents of 100–$300 A/cm^2$ are still flowing at the QCL operating voltage of 12–15 V. In contrast, in the sample with the Fe-doped InP barrier the leakage current remains below $10^{-2} A/cm^2$ up to a maximum bias of 15 V. The blocking behavior of the Fe:InP layer may be understood by considering the band profile along the leakage path as shown schematically in Fig. 2.14a for zero applied bias.

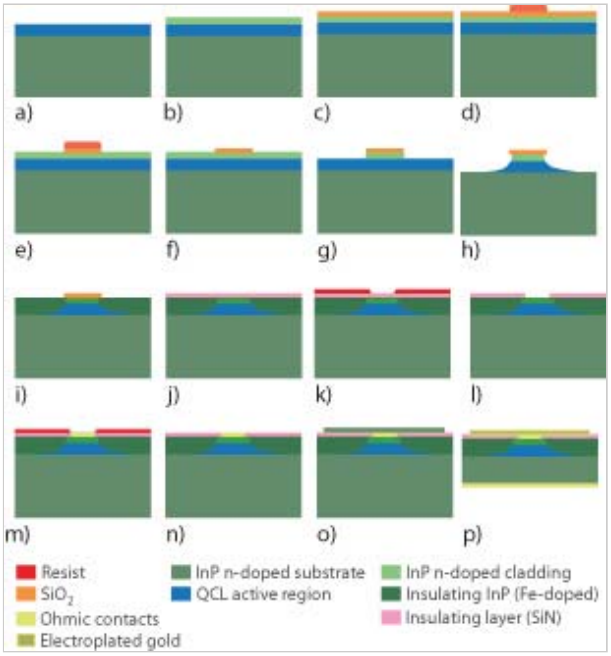


Fig. 2.11. Process flow for the standard buried heterostructure fabrication in mid-infrared

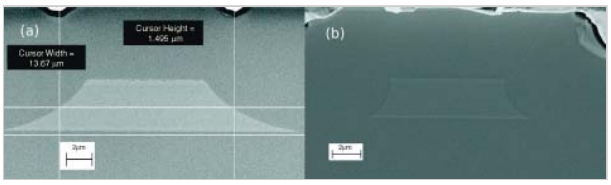


Fig. 2.12. Facet of a buried heterostructure device. a) normal process. b) Process where the order of the regrowth (lateral and vertical) have been inverted [44].

The iron doping creates deep donor states in the InP, pinning the Fermi level at an energy of $E_{Fe} = -0.63$ eV from the conduction band [46]. Assuming room temperature and a relatively low doping concentration N_d for the upper cladding, as required for a low-loss waveguide, the position of the chemical potential, in respect to the conduction band edge, is given by

(p.24)

$$\mu = kT \ln\left(\frac{N_d}{N_c}\right) \quad (2.2.6)$$

where

$$N_c = 2\left(\frac{m^* kT}{2\pi\hbar^2}\right)^{3/2} \quad (2.2.7)$$

and is $N_c = 5.7 \times 10^{17}$ cm^{-3} for InP. The build-in potential of the junction is then

$$V_{bi} = -\mu - E_{Fe} \quad (2.2.8)$$

and we find $V_{bi} = 0.57$eV for a typical cladding doping of $N_d = 8 \times 10^{16}$ cm^{-3}. V_{bi} is the effective barrier that will block the thermoinic current flowing through the barrier.

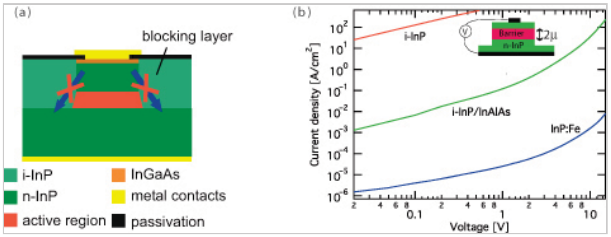


Fig. 2.13. a) Leakage paths (arrows) in a buried heterostructure device. b) Current–voltage characteristics of n-i-n structures with different barrier materials, as indicated.

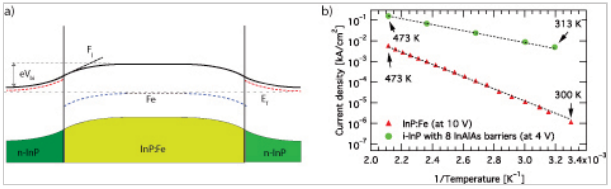


Fig. 2.14. a) Schematic band structure of the n-i-n profile, at zero applied bias, along the leakage path. b) Arrhenius plot of the current through the blocking layer at fixed applied bias. Activation energies of $E_a = 588$ meV and $E_a = 289$ meV are found for the structure with the InP:Fe and the eight multibarrier InP/ $Al_{0.48}In_{0.52}As$ material sequence, respectively [45].

To assess the temperature stability of these blocking layers, the current was measured at fixed bias as a function of the temperature for both the Fe:InP and $Al_{0.48}In_{0.52}As$/ InP samples. The result is plotted in Fig. 2.14 as an Arrhenius plot. The current is found to be activated thermally with an activation energy of 588 meV for the Fe:InP sample, in good agreement with our previous estimate, and 289 meV for the $Al_{0.48}In_{0.52}As$/ InP sample, in good agreement with the expected band discontinuity.

As shown in Fig. 2.13, the current starts to increase super-exponentially at a bias of 2–5 V, suggesting a barrier lowering of the junctions. A rough estimate of the order of magnitude of the field at which this happens can be obtained by computing the maximum field in the junction. For an ideal abrupt junction between the n-doped and the Fe-doped InP, the maximum field in the junction F_i is given by

$$F_i = \sqrt{\frac{2eV_{bi}}{\epsilon_s}\left(\frac{N_d N_{Fe}}{N_d + N_{Fe}}\right)}$$

(2.2.9)

We obtain a field $F_i = 80kV/cm$ for a typical Fe doping of $N_{Fe} = 8 \times 10^{16}$ cm^{-3}. This field F_i will be decreased, on the negative side of the junction, by an applied bias to the junction. Assuming an operation bias of the QCL of 10 V, the average applied field would be of 50 kV/cm and the field in the junction to be lowered to 30 kV/cm. We do, however, expect a field redistribution to occur in the junction, and the field lowering to be weaker than this rough estimate.

2.3 Mounting techniques

The initial mounting techniques of quantum cascade lasers, junction up mounting using indium solder on copper submounts were developed with cryogenic operation in mind. Oxygen-free copper has one of the best low-temperature conductivities of all materials (10, 800 Wm^{-1} K^{-1} at 20 K, and still 400 Wm^{-1} K^{-1} at 300 K). The indium soldering will accommodate the large thermal expansion mismatch between the InP substrate $\alpha = 4.6 \times 10^{-6}$ K−1 and Cu $\alpha = 15 \times 10^{-6}$ K^{-1} submount. In a junction-up mounting, any residual dislocation would have to cross the whole substrate before reaching the active region.

(p.25)

To achieve the best long-term reliability with continuous wave devices operating near room temperature, however, the techniques originally developed for telecom devices and high-power lasers are now implemented in QCL. The latter are mounted episide-down on AlN submounts using a high temperature, fluxless AuSn solder in an hermetically sealed, organic-free package. AlN exhibits very good thermal conductivity (200 Wm−1 K^{-1}) along with a room-temperature expansion coefficient 4.6×10^{-6} K^{-1}, virtually equal to that of InP. The absence of volatile organic compounds as well as the use of a solder with a relatively high melting temperature (about 280° C) minimizes long-term chemical reactions in the device. An example of such a mounted device is shown in Fig. 2.15.

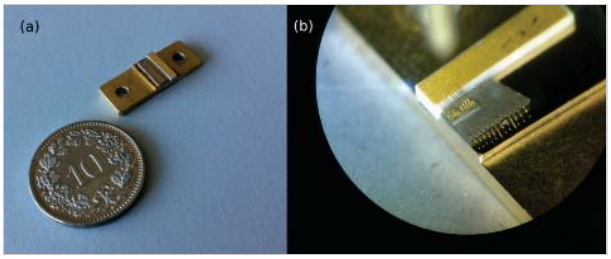
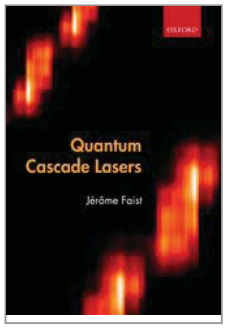


Fig. 2.15. a) General view and b) close-up of a device mounted episide-down on an AlN submount [47].



Quantum Cascade Lasers
Jérôme Faist

Print publication date: 2013
Print ISBN-13: 9780198528241
Published to Oxford Scholarship Online: May 2013
DOI: 10.1093/acprof:oso/9780198528241.001.0001

Electronic states in semiconductor quantum wells

Jérôme Faist

DOI:10.1093/acprof:oso/9780198528241.003.0003

Abstract and Keywords

This chapter discusses the computation of electronic states in semiconductor quantum wells. It covers the k·p approximation, envelope function approximation, Hartree potential, active region building blocks, in-plane dispersion, and the valence band.

Keywords: k·p approximation, envelope function approximation, hartree potential, active region building blocks, in-plane dispersion, valence band

3.1 Band structure of semiconductors in the k·p approximation: origin of the effective mass

3.1.1 The k·p approximation

An accurate computation of the energy states in complicated structures formed by quantum wells and barriers is a requirement for the design of working devices. Although in principle a number of approaches can be used to compute the energy bands in a solid, in practice the vast majority of the computations are performed using the $\mathbf{k} \cdot \mathbf{p}$ approximation. The latter is a powerful approach to the computation of the band structure that relies on the knowledge of the band structure at $k = 0$ and expands the wavefunctions on this basis.

Shown in Fig. 3.1 is the band structure of the GaAs and InAs; the $\mathbf{k} \cdot \mathbf{p}$ is able to predict accurately the band structure close to the Γ edge of the bandgap near the Fermi energy. For most devices, they are the relevant states for the optical and transport properties.

3.1.2 Basic model

The Schrodinger equation for a crystal is:

$$(\frac{p^2}{2m_0} + V\left(\mathbf{r}\right) + \frac{\hbar}{4m_0^2c^2}(\sigma \times \nabla V) \cdot \mathbf{p})\psi(\mathbf{r}) = E\psi(\mathbf{r})$$

(3.1.1)

For simplicity, we will drop the spin–orbit coupling term from the following derivation. The latter arises as a relativistic term. The motion of the electon in the electric field of the ion creates an equivalent magnetic field that interacts with the spin of the electron. Let us first compute the action of p on ψ, written in terms of Bloch wavefunctions $u_{n,\mathbf{k}}$ of band n and wavevector $\mathbf{k}$ so that

$$\psi_{n\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\mathbf{r}}u_{n,\mathbf{k}}(\mathbf{r})$$

(3.1.2)

We then obtain

$$\mathbf{p} \cdot (e^{i\mathbf{k}\cdot\mathbf{r}}u_{n,\mathbf{k}}(\mathbf{r})) = -i\hbar\nabla \cdot (e^{i\mathbf{k}\cdot\mathbf{r}}u_{n,\mathbf{k}}(\mathbf{r}))$$

(3.1.3)

$$= \hbar\mathbf{k}e^{i\mathbf{k}\cdot\mathbf{r}}u_{n,\mathbf{k}}(\mathbf{r}) + e^{i\mathbf{k}\cdot\mathbf{r}}\mathbf{p} \cdot u_{n,\mathbf{k}}(\mathbf{r})$$

(3.1.4)

$$= e^{i\mathbf{k}\cdot\mathbf{r}}(\mathbf{p} + \hbar\mathbf{k})u_{n,\mathbf{k}}(\mathbf{r})$$

(3.1.5)

(p.27)

Using the above relation, the Schrödinger equation is obtained for $u_{n,\mathbf{k}}(\mathbf{r})$:

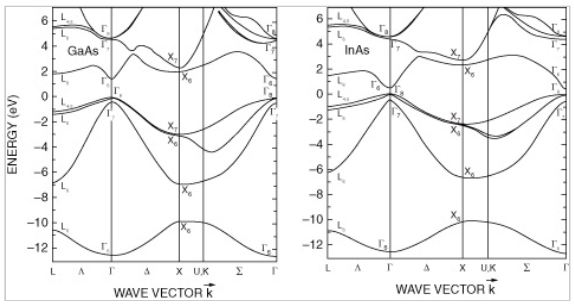


Fig. 3.1. Computed band structure of GaAs and InAs, as indicated. The relevant states considered in this work are those around the fundamental gap at the Γ point. The lateral valleys at X and L play only a parasitic role. Reprinted with permission from [48]. Copyright 1976 by the American Physical Society (APS).

$$\left(\frac{p^2}{2m_0} + \frac{\hbar}{m_0}\mathbf{k} \cdot \mathbf{p} + \frac{\hbar^2k^2}{2m_0} + V(\mathbf{r})\right)u_{n,\mathbf{k}}(\mathbf{r}) = E_{nk}u_{n,\mathbf{k}}(\mathbf{r})$$

(3.1.6)

The Hamiltonian $H = H_0 + W(\mathbf{k})$ may be split into a k-independent

$$H_0 = \frac{p^2}{2m_0} + V(\mathbf{r})$$

(3.1.7)
and *k*-dependent part

$$W(\mathbf{k}) = \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar}{m_0}\mathbf{k}\cdot\mathbf{p}.$$

(3.1.8)
The solution of the equation

$$H_0 u_{n,0}(\mathbf{r}) = E_{n0} u_{n,0}(\mathbf{r})$$

(3.1.9)
shows the energies of the band structure at the Γ point $k = 0$. The fundamental idea of the $\mathbf{k}\cdot\mathbf{p}$ approximation is to use the $u_{n,0}(\mathbf{r})$ as a basis for the expansion of the wavefunction and energies at finite **k** value.

In the simplest cases, taking an interband transition across the gap and looking at the conduction band, taking the second-order perturbation expansion and first neglecting the spin–orbit term:

$$E_c(\mathbf{k}) = E_c(0) + \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar^2}{m_0}\sum_{m\neq c}\frac{|\langle u_{c,0}|\mathbf{k}\cdot\mathbf{p}|u_{m,0}\rangle|^2}{E_c - E_{m,0}}$$

(3.1.10)
(p.28) In the lowest-order approximation, all other bands except the valence band may be neglected. Furthermore, computing the dispersion of the conduction band we use the fact that the interband matrix element $\mathbf{p}_{cv}$ is isotropic, and the dependence of the energy is only in the magnitude of **k** and not in its direction. As a result the dispersion can be written as (taking the zero energy at the top of the valence band):

$$E_c(k) = E_g + \frac{\hbar^2 k^2}{2m_0} + \frac{\hbar^2 k^2}{m_0}\frac{|p_{cv}|^2}{E_g}$$

(3.1.11)
Defining the Kane energy E_P such that

$$E_P = \frac{2}{m_0}|\langle u_{c,0}|p|u_{v,0}\rangle|^2$$

(3.1.12)
the dispersion of the conduction band can be written as

$$E_c(k) = E_c + \frac{\hbar^2 k^2}{2m_0}\left(1 + \frac{E_P}{E_g}\right)$$

(3.1.13)
We then obtain the effective mass as:

$$(m^*)^{-1} = (m_0)^{-1}\left(1 + \frac{E_P}{E_g}\right)$$

(3.1.14)
The Kane energy is much larger than the gap $E_P \rangle\rangle E_g$ and is rather constant across the III–V semiconductors. As a result, the effective mass is roughly proportional to the bandgap. One should be careful that some authors use the definition of an interband operator matrix element P with the units of a velocity, yielding

$$E_P = 2m_0P^2$$

(3.1.15)

(Bastard [49], for example), while others use P with units of a square root of energy

$$E_P = P^2$$

(3.1.16)

The Kane energy E_P is, however, always defined in the same manner.

3.1 Fundamental parameters for various III–V semiconductors. E_g and Δ_0 are the fundamental gap and the spin–splitting, respectively, m_c^*/m_0 the conduction band effective mass, E_P the Kane energy, and r_{vc} the interband matrix element.

	GaN	GaAs	InP	GaSb	InAs	InSb
′)	3.4	1.519	1.424	0.811	0.418	0.235
′)	0.017	0.341	0.11	0.75	0.38	0.81
$_0$	0.17	0.0665	0.079	0.0405	0.023	0.0139
′)	20.2	22.71	17	22.88	21.11	22.49
)	2.55	6.14	5.67	11.5	21.5	39.5

In the above expressions, the spin–orbit term can be included by replacing the operator p by

$$\pi = \mathbf{p} + \frac{\hbar}{4m_0c^2}(\sigma \times \nabla V)$$

(3.1.17)

(p.29) 3.1.3 Beyond the perturbation expansion

A very powerful procedure, first introduced by E. O. Kane in a celebrated paper [50], is to expand formally the solutions of eqn. 3.1.6 for the *n*th band in the solutions at **k** = 0, writing formally:

$$u_{n,k}(\mathbf{r}) = \sum_m c_m^{(n)}(\mathbf{k}) u_{m,0}(\mathbf{r})$$

(3.1.18)

and restricting the sum to a limited relevant subset of *m* bands. Improved accuracy can be achieved by adding more bands. In this basis, and projecting the equation onto the state $u_{M,0}$, the Hamilton equation is written as

$$\left(E_{M0} + \frac{\hbar^2 k^2}{2m_0} - E_n(\mathbf{k})\right) c_M^{(n)}(\mathbf{k}) + \sum_{m \neq M} H_{Mm}^{kp} c_m^{(n)}(\mathbf{k}) = 0$$

(3.1.19)

where the **k** · **p** Hamiltonian is

$$H_{Mm}^{kp} = \frac{\hbar}{m_0}\mathbf{k} \cdot \langle u_{M,0}|\mathbf{p}|u_{m,0}\rangle$$

(3.1.20)

To the extent that these matrix elements are known, eqn. 3.1.19 can be solved.

3.1.4 Example: a two-band Kane model

Let us write explicitely a two-band Kane model. The latter can be a fairly realistic model if one is interested only in the effect of the valence band on the conduction band, replacing the three spin-degenerate valence bands (heavy hole, light hole, split-off) by an effective valence band. Using the subscripts *c* and *v* for conduction and valence bands, respectively, the $u_{n,\,\mathrm{k}}$ can then be expressed as:

$$u_{n,\mathbf{k}} = a_c^{(n)}(\mathbf{k}) u_{c,0} + a_v^{(n)}(\mathbf{k}) u_{v,0}$$

(3.1.21)

Replacing this expansion into eqn. 3.1.19, we obtain the following matrix equation:

$$\begin{pmatrix} E_c + \frac{\hbar^2 k^2}{2m_0} & \frac{\hbar}{m_0}\mathbf{k} \cdot \mathbf{p}_{cv} \\ \frac{\hbar}{m_0}\mathbf{k} \cdot \mathbf{p}_{cv}^* & E_v + \frac{\hbar^2 k^2}{2m_0} \end{pmatrix} \begin{pmatrix} a_c^{(n)}(\mathbf{k}) \\ a_v^{(n)}(\mathbf{k}) \end{pmatrix} = E_n(\mathbf{k}) \cdot \begin{pmatrix} a_c(\mathbf{k}) \\ a_v(\mathbf{k}) \end{pmatrix}$$

(3.1.22)

Assuming $E_c = 0$, $E_v = -E_g$, the solution of the matrix equation satisfies:

$$\left(\frac{\hbar^2 k^2}{2m_0} - E_n(\mathbf{k})\right)\left(-E_g + \frac{\hbar^2 k^2}{2m_0} - E_n(\mathbf{k})\right) - \frac{\hbar^2}{m_0^2}|\mathbf{k} \cdot \mathbf{p}_{cv}|^2 = 0$$

(3.1.23)

This second-order equation in $E_n(\mathbf{k})$ can, of course, be solved directly; it is simplified by the assumption that the interband matrix element $\mathbf{p}_{cv}$ is isotropic and the dependence is only in the magnitude of $\mathbf{k}$ and not in its direction. It is, however, more instructive to write it in the form of a pseudo-effective mass equation. We take as an example **(p.30)** the conduction band, remembering $E_P = \frac{2}{m_0}|p_{cv}|^2$ and neglecting the square of the kinetic energy of the free electron:

$$E_c(\mathbf{k}) = \frac{\hbar^2 k^2}{2m_0} \frac{E_P + E_g + 2E_c(\mathbf{k})}{E_c(\mathbf{k}) + E_g}$$

(3.1.24)

For $k \to 0$ the above expression reduces itself to the result of the perturbation expansion:

$$(m^*)^{-1} = (m_0)^{-1}(1 + \frac{E_P}{E_g})$$

(3.1.25)

Eqn. 3.1.24 can be expressed (by neglecting the term $2E/(E_P + E_g)$), equivalent to neglecting the diagonal free-electron kinetic energy in the Hamiltonian 3.1.22, in the somewhat simplified form

$$m^*(E) = m^*(0)\left(1 + \frac{E}{E_g}\right)$$

(3.1.26)

that is commonly used in the literature. The above expression can be rewritten, using a Taylor expansion of the square root, to express the dependence of the energy in k [4]

$$E(k) = \frac{\hbar^2 k^2}{2m^*(0)}(1 - \gamma k^2)$$

(3.1.27)

where

$$\gamma = \frac{\hbar^2}{2m^*(0)E_g}$$

(3.1.28)

is the non-parabolicity coefficient. Setting γ and $m^*(0)$ uniquely defines the Kane energy E_P and the gap E_g through eqn. 3.1.28 and 3.1.25.

3.2 Envelope function approximation

The problem we will try to solve now is that of a heterostructure, in which two materials A and B are sandwiched together. Of course, such a material could be seen as a new material by itself, and its band solved by ab initio techniques, but such a computation is very time-consuming and, moreover, does not present much physical insight into the result. The envelope function approximation solves this problem in a very efficient and elegant manner. It is widely used to predict the optical and electrical properties of semiconductor nanostructures.

3.2.1 Multi-band case

At the core of the envelope function approximation is a generalization of the $\mathbf{k} \cdot \mathbf{p}$ approximation. It is postulated that the wavefunction can be written as a sum of slowly varying envelope functions $f_l^{A,B}(\mathbf{r})$ that will modulate the Bloch function of the material: namely,

$$\Psi(\mathbf{r}) = \sum_l f_l^{A,B} u_{l,\mathbf{k}_0}^{A,B}(\mathbf{r})$$

(3.2.29)

Behind eqn. 3.2.29 is the idea that at each point, the wavefunction is described by a $\mathbf{k} \cdot \mathbf{p}$ decomposition around a point $\mathbf{k}_0$ in the Brillouin zone (taken in the following **(p.31)** at the Γ point, i.e. $\mathbf{k}_0 = 0$) and that this decomposition depends on the position. Furthermore, it is assumed that

1. the envelop function $f_l^{A,B}(\mathbf{r})$ is slowly varying compared to the Bloch wavefunction, if $f_l^{A,B}(\mathbf{r})$ is written in a Fourier decomposition, the wavevectors are close to the center of the Brillouin zone.
2. the Bloch functions are identical in both materials, i.e. $u_{n,0}^{A}(\mathbf{r}) = u_{n,0}^{B}(\mathbf{r})$. This also implies that the interband matrix element p_{cv} is identical in both materials.

It allows us to write the wavefunction as

$$\Psi(\mathbf{r}) = \sum_l f_l^{A,B} u_{l,0}(\mathbf{r})$$

(3.2.30)

Let us assume first a quantum well, in which a layer of material A is clad on both sides by a barrier material B. Because of the in-plane translational invariance, the wavefunction may be written as plane waves in the x–y direction:

$$f_l(r_{\|}, z) = \frac{1}{\sqrt{S}} e^{i k_{\|} r_{\|}} \chi_l(z)$$

(3.2.31)

where z is chosen as the growth direction and $k_{\|} = (k_x, k_y)$ is the in-plane wavevector. Note that there is confusion in the literature, since the in-plane direction is referred either with the sign ∥, as is done here, or with the sign ⊥, meant as "perpendicular to the growth axis." We finally used the first notation as it was the one enabling the best consistency in the various topics (strain, scattering, lifetimes) where such a convention was needed. The Hamiltonian is then

$$H = \frac{p^2}{2m_0} + V_A(\mathbf{r}) Y_A(z) + V_B(\mathbf{r}) Y_B(z)$$

(3.2.32)

where the functions $Y_A(z)$ and $Y_B(z)$ "turn on" the potential in the respective layers. We will develop our system close to $k = 0$. To solve the system we must

1. Let H act upon $\Psi(\mathbf{r})$.
2. Multiply on the left by $u_{m0}^{*}(\mathbf{r}) e^{-i k_{\|} r_{\|}} \chi_m^{*}(z)$.
3. Integrate over space.

We have to use the following relations. As the envelop function is slowly varying over the lattice cell, we may write

$$\int_{cell} d^3 r f_l^{*} f_m u_l^{*} u_m = f_l^{*} f_m \int_{cell} d^3 r u_l^{*} u_m = f_l^{*} f_m \delta_{lm}$$

(3.2.33)

and take advantage of the fact that the band edges are eigenfunctions of the Hamiltonian at ($\mathbf{k} = 0$):

$$\left(\frac{p^2}{2m_0} + V^{A,B}(\mathbf{r})\right) u_{m,0}(\mathbf{r}) = E_{m,0}^{A,B} u_{m,0}(\mathbf{r})$$

(3.2.34)

The derivation is rather tedious, but one should note the similarity with the normal $\mathbf{k} \cdot \mathbf{p}$ technique by considering the action of the operator $\mathbf{p}$ on the wavefunction:

(p.32)

$$\mathbf{p}(e^{k_{\|}r_{\|}}\chi_l(z)u_{l,0}(\mathbf{r})) = (\hbar k_{\|} - i\hbar\frac{\partial}{\partial z} + \mathbf{p})e^{ik_{\|}r_{\|}}\chi_l(z)u_{l,0}(\mathbf{r})$$

(3.2.35)

and we then may consider the substitution

$$\mathbf{p} \rightarrow (\hbar k_{\|} - i\hbar\frac{\partial}{\partial z} + \mathbf{p})$$

(3.2.36)

where it is understood that $\mathbf{p}$ then acts only on the Bloch part of the wavefunction. Using the above substitution in the Hamiltonian, one finally obtains the following set of differential equations written in a matrix form:

$$\underline{D}(z, -i\hbar\frac{\partial}{\partial z})\chi = E\chi$$

(3.2.37)

where the elements of the matrix $\underline{D}$ are given by

$$D_{lm} = (E_l^A Y_A + E_l^B Y_B + \frac{\hbar^2 k_{\|}^2}{2m_0} - \frac{\hbar^2}{2m_0}\frac{\partial^2}{\partial z^2})\delta_{l,m} + \frac{\hbar k_{\|}}{m_0}\langle l|p_{\|}|m\rangle - \frac{i\hbar}{m_0}\langle l|p_z|m\rangle\frac{\partial}{\partial z}$$

(3.2.38)

In the above equation, $E_l^{A,B}$ is the energy of the band edge *i* (at $\mathbf{k} = 0$) in materials A and B respectively, the *z* -derivatives apply to χ, and the matrix elements $\langle l|p|m\rangle = \int_{cell} u_l^* p u_m d^3\mathbf{r}$.

The part of the band structure show in Fig. 3.1 used for the expansion of eqn. 3.2.38 are the bands around the fundamental gap at the Γ point, in the center of the Brillouin zone. They are shown schematically in Fig. 3.2 and consist of one conduction band, originating from an s-like orbital, and three valence bands originating from p-orbitals from the atoms. Because of the spin–orbit coupling, the sixly-degenerate valence band Γ edge splits into a doubly degenerate split-off band and a quadruplet heavy-hole light-hole edge separated by the gap Δ_0. To keep the system 3.2.38 as simple as possible, it is most helpful to use the quantization of the angular momentum axis in the growth direction with the basis wavefunction indicated in Fig. 3.2. Since each band is doubly degenerate, the choice of this basis yields an 8x8 system of equations.

One shows, however, that an accurate description of the valence band (that in particular yields the correct sign of the mass for the heavy-hole band) requires the introduction of additional bands, most conveniently achieved by using perturbative terms additional to eqn. 3.2.38. The final result, given here only for indicative purpose, is (see Bastard [49]):

$$\sum_{m=1}^{8}([E_m^A + V_m(z) + \frac{\hbar^2 k_\parallel^2}{2m_0} - \frac{\hbar^2}{2m_0}\frac{\partial^2}{\partial z^2}]\delta_{l,m} + \frac{\hbar k_\parallel}{m_0}\langle l|p_\parallel|m\rangle$$
$$-\frac{i\hbar}{m_0}\langle l|p_z|m\rangle\frac{\partial}{\partial z} - \frac{\hbar^2}{2}\frac{\partial}{\partial z}\frac{1}{M_{lm}^{zz}}\frac{\partial}{\partial z}$$
$$-\frac{i\hbar^2}{2}\sum_{\alpha=x,y}[k_\alpha\frac{1}{M_{lm}^{\alpha z}}\frac{\partial}{\partial z} + \frac{\partial}{\partial z}\frac{1}{M_{lm}^{z\alpha}}k_\alpha] + \frac{\hbar^2}{2}\sum_{\alpha,\beta=x,y}k_\alpha\frac{1}{M_{lm}^{\alpha\beta}}k_\beta)\chi_m = E\chi_l$$

(3.2.39)

(p.33)

where the indices α and β run over the three axis x, y and z; E_m^A is the energy of the m band edge and $V_m(z)$ its position dependence between materials A and B. The effective mass parameters $M_{lm}^{\alpha\beta}$ are defined as

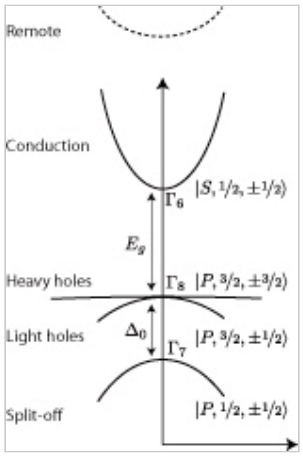


Fig. 3.2. Schematic dispersion of the bands taken into account for the **k** · **p** approximation: the various band edges, their group symmetry notations, and the basis states expressed by their angular momentum and its projection.

$$\frac{m_0}{M_{lm}^{\alpha\beta}} = \frac{2}{m_0}\sum_\nu \langle l|p_\alpha|\nu\rangle\frac{1}{E - E_{\nu 0} - V_\nu(z)}\langle m|p_\beta|\nu\rangle$$

(3.2.40)

where the sum ν runs over the remote bands.

3.2.2 One-band model

As an example, let us consider first a pure one-band model in the effective mass approximation. Setting

$$E_l^A Y_A + E_l^B Y_B = V(z)$$

(3.2.41)

we obtain the Schrödinger equation:

$$(-\frac{\hbar^2}{2m^*}\frac{\partial^2}{\partial z^2} + V(z))\chi(z) = E\chi(z)$$

(3.2.42)

In this equation the band is assumed to be parabolic with a curvature given by an effective mass m^*. For an isolated band such as the heavy-hole band, it is a rather good approximation. It can also

be used successfully for thick quantum well structures in the conduction band, when the confinement energies are much smaller than the material energy gap.

One first formal difficulty arises because the mass of the barrier material is in general different than the one of the quantum well. For this reason the effective mass should be considered as a position-dependent quantity $m^* = m^*(z)$. In that case, one **(p.34)** shows that the proper boundary conditions, derived by integrating the Schrödinger equation across the interface, implies the continuity of

$$\chi(z)$$

(3.2.43)
and of the quantity

$$\frac{1}{m^*}\frac{\partial \chi(z)}{\partial z}$$

(3.2.44)
across the interface.

One point, often overlooked, concerns the choice of the mass for the barrier material. Very often the value barrier mass is taken at the band edge. This is usually a poor approximation, since often the conduction band discontinuity is a sizable fraction of the bandgap. It is much better to assume a *constant* value corrected for non-parabolicity using eqn. 3.1.26.

3.2.3 Two-band model

For the conduction band a very nice model is that in which one valence band is kept, creating a two-band model. For simplicity, let us look at the states at $k_{\parallel} = 0$. Starting from eqn. 3.2.38, we neglect the diagonal free electron term and retain only the firstorder derivative of the envelope functions, as $|\partial u_{n,0}/\partial z| \rangle \rangle |\partial \chi/\partial z|$. We then obtain the system of equations given by

$$\begin{pmatrix} V_c(z) & -\frac{i\hbar}{m_0}p_{cv}\frac{\partial}{\partial z} \\ \frac{-i\hbar}{m_0}p_{cv}^*\frac{\partial}{\partial z} & V_v(z) \end{pmatrix}\begin{pmatrix}\chi_c \\ \chi_v\end{pmatrix} = E\begin{pmatrix}\chi_c \\ \chi_v\end{pmatrix}$$

(3.2.45)
Extracting χ_v from the second equation yields

$$\chi_v = \frac{1}{V_v(z) - E}\frac{i\hbar}{m_0}p_{cv}^*\frac{\partial \chi_c}{\partial z}$$

(3.2.46)
Replacing in the first equation, after substitution, the following result is obtained:

$$-\frac{\hbar^2|p_{cv}|^2}{m_0^2}\frac{\partial}{\partial z}\frac{1}{E - V_v(z)}\frac{\partial}{\partial z}\chi_c + V_c(z)\chi_c = E\chi_c$$

(3.2.47)
Recalling the definition of the Kane energy $E_P = \frac{2}{m_0}|p_{cv}|^2$ and defining an energy-dependent effective mass,

$$\frac{1}{m(E,z)} = \frac{1}{m_0}\frac{E_P}{E - V_v(z)}$$

(3.2.48)
we obtain finally a Schrodinger-like equation:

$$-\frac{\hbar^2}{2}\frac{\partial}{\partial z}\frac{1}{m(E,z)}\frac{\partial}{\partial z}\chi_c + V_c(z)\chi_c = E\chi_c$$

(3.2.49)
This model is very useful for modeling the electronic states in the conduction band with the inclusion of the non-parabolicity. It is very widely used in the study of intersubband transitions **(p.35)** .

3.2.4 The three-band model

The relation between the two-band and the full 8x8 Hamiltonian can be better explored by considering a three-band model, again assuming that $k_1 \approx 0$. In this approximation the heavy hole is decoupled from the other valence bands and one is left with two equivalent 3x3 Hamiltonians (one for each spin direction)

$$H = \begin{pmatrix} E_c(z) & \sqrt{\frac{2}{3}}\frac{p_{cv}}{m_0}p_z & -\sqrt{\frac{1}{3}}\frac{p_{cv}}{m_0}p_z \\ -\sqrt{\frac{2}{3}}\frac{p^*_{cv}}{m_0}p_z & E_{lh}(z) & 0 \\ \sqrt{\frac{1}{3}}\frac{p^*_{cv}}{m_0}p_z & 0 & E_{so}(z) \end{pmatrix}$$

(3.2.50)
acting on the three-dimensional vector of the envelope function $\psi = (\chi_c, \chi_{lh}, \chi_{so})$, where *c, lh* and *so* label the conduction, light-hole, and split-off band edges, respectively. In the above expression, p_z acts on ψ as $-i\hbar\partial/\partial z$. Note also that in the Hamiltonian 3.2.50, as we did in the 2x2 model, we have also neglected the diagonal "free electron" terms p_z^2/m_0 that can be shown to contribute to terms of the order of $E_g/E_p \ll 1$. Moreover, as we are interested only in the energy levels located above the edge of the conduction band, the problem can be solved by using the second and third rows of the matrix 3.2.50 to express the equation in the first row as

$$p_z\frac{1}{2m(E,z)}p_z\chi_c + E_c(z)\chi_c = E\chi_c$$

(3.2.51)
but with now the energyand position-dependent effective mass given by

$$\frac{1}{m(E,z)} = \frac{1}{m_0}\left[\frac{2}{3}\frac{E_p}{E-E_{lh}} + \frac{1}{3}\frac{E_p}{E-E_{so}}\right]$$

(3.2.52)
The solution of the differential eqn. 3.2.51 gives the conduction component χ_c and the energy of the stationary states. Mathematically it is very similar to a one-band model, but with an energydependent effective mass. Moreover, the difference between the twoand three-band models lies entirely in the form of the energy dependence of the effective mass.

We must recall, however, that the total stationary wavefunctions are given by the three components χ_c, χ_{lh}, and χ_{so}, weighted by their corresponding Bloch functions, so that the only knowledge of χ_c is insufficient for the complete physical description of the stationary state. The 2x2 Hamiltonian can be obtained approximatively, defining an "effective" valence band by $E_v = (2E_{lh} + E_{so})/3$, and therefore the gap is now an effective gap

$$E_g^* = E_g + \Delta_0/3,$$

(3.2.53)

which has also a contribution coming from the split-off edge Γ_7. In doing so, we neglect contributions to the wavefunctions and energies of the order of

$$(\Delta/|E_c - E_v|)^2 = (\Delta/E_g)^2$$

(3.2.54)

with $\Delta = \frac{\sqrt{2}}{3}\Delta_0$. For typical III–V semiconductors used for QC lasers this factor is quite small: in GaAs $(\Delta/E_g)^2_{GaAs} \simeq 0.01$, and in InGaAs (lattice matched on InP) **(p.36)** $(\Delta/E_g)^2_{GaInAs} \simeq 0.04$. As a result, the corrections never exceed more than a few percent. For devices based on InAs, however, this approximation is likely to be poor, as $\Delta \approx E_g$ in this material.

A feeling for these approximations can be obtained by comparing the various expressions for the energy-dependent effective mass. This is done in Fig. 3.3, where the three-band expression 3.2.52 is compared to the two-band case, using either the parameters used in the literature for the InGaAs/AlInAs system lattice-matched to InP, or using fitting parameters. The value of the non-parabolicity coefficient γ reported in the literature [51] ($\gamma = 1.13 \times 10^{-18}\ m^2$), obtained by fitting experimental absorption data, is slightly higher than that derived from fitting the three-band model ($\gamma = 1.06 \times 10^{-18}\ m^2$). This discrepancy is attributed to the inherent inaccuracies of the model and of the experiments

.

3.3 Hartree potential

For the carrier densities used in QCLs (typically about 10^{16} cm^{-3}), the electron– electron interaction is not a dominant term but must be taken into account if the electronic states are to be computed with a good accuracy, especially in the terahertz. The effects are especially strong in devices where the electrons are Δte from the ionized donors. It is usually treated in a mean-field approximation by adding a Hartree potential $V_H(z)$ to the Hamiltonian:

$$H = -\frac{\hbar^2}{2m^\star}\nabla^2 + V_{QC}(z) + V_H(z)$$

(3.3.55)

This term is computed from the local charge density

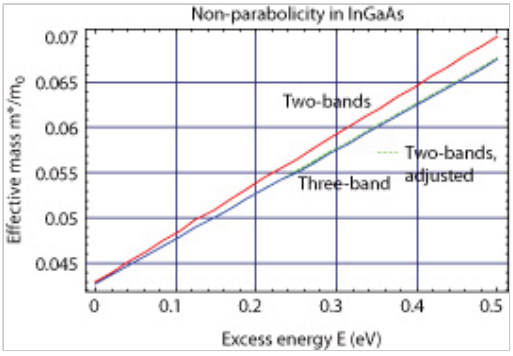


Fig. 3.3. Comparison between various models for the energy-dependent effective mass for $Ga_{0.47}$ $In_{0.53}$ As material. The lower curve is the full three band expression 3.2.52, using E_P = 20.4 eV, E_g = 0.78 eV and a spin–orbit splitting Δ = 0.36 eV; the upper curve is the two-band model approximation given by eqn. 3.1.26 with the parameters usually found in the literature E_P = 17.3 eV, E_g = 0.786 eV, and finally, in dashed, a two-band model with the parameters fitted to the values of the three-band model (E_P= 18.4, E_g = 0.84)

(p.37)

$$\rho(z) = eN_D(z) - e\sum_i n_i|\chi_i(z)|^2$$

(3.3.56)

where N_D is the doping profile of ionized dopants, $|\chi_i(z)|^2$ is the probability density, and n_i is the number of electrons per unit area in the ith subband. The potential $V_H(z)$ is computed from $\rho(z)$ using Poisson's equation:

$$\frac{\partial^2 V_H(z)}{\partial z^2} = \frac{\rho(z)}{EE_0}$$

(3.3.57)

The electronic densities n_i on the subbands are not known a priori, and in the general case depend on the transport in the device. However, a good starting approximation is to assume that the electron distribution is thermal, characterized by Fermi distribution with a common chemical potential μ in each period, and that charge neutrality is achieved in each period: i.e.,

$$\sum_i n_i = \sum_i \int \mathscr{D}_i(E) f(E) dE(\text{electrons})$$
$$= \int N_D(z) dz = n_s(\text{impurities})$$

(3.3.58)

where

$$\mathscr{D}_i(E) = \frac{m_i^*}{\pi\hbar^2}\theta(E - E_i)$$

is the density of states of the ith subband, $\theta(E - E_i)$ is the Heaviside function,

$$f(E) = \frac{1}{\exp\left(\frac{E-\mu}{kT}\right) + 1}$$

(3.3.59)

is the Fermi distribution function, and n_S is the total sheet electron concentration, which, because we assumed charge neutrality within each period, is equal to the impurities concentration. The Hartree potential depends on the solution of Schrödinger's equation; therefore Schrödinger's and Poisson's equations must be solved iteratively until convergence is achieved.

3.4 Active region building blocks

When designing intersubband structures, it is important to acquire an intuitive grasp of simple building blocks. This is done most easily in a simple one-band model. The numerical results are obtained using the following approach:

1. The wavefunctions are found for the regions in which the potential is constant (or assumed constant). They will be plane waves or decaying exponentials.
2. Those solutions are then matched at each interface, using the boundary conditions compatible with the Hamiltonian. For a one-band model, it means matching the value of the wavefunction and of the derivative, divided by the mass, at the interfaces.
3. Boundary conditions are then imposed on the edge of the sample: decaying exponential for bound states, or periodicity for a periodic structure.

We will review some of these results.

(p.38) 3.4.1 The single quantum well

The simplest case is, of course, the quantum well with infinite barrier height. The wavefunctions are simply sine and cosine functions (if the well is taken symmetric around zero) or sine only if the quantum well is lying in the interval [0,L]. In the latter case, the condition on the value of k_n

$$k_n = n\frac{\pi}{L}$$

(3.4.60)

with n integer enables us to immediately write the energy as

$$E_n = \frac{\hbar^2 k_n^2}{2m^*} = \frac{n^2\pi^2\hbar^2}{2m^*L^2}$$

(3.4.61)

Of course, in real structures one should take into account the finite barrier height. Following Bastard [49], we assume a potential with the form

$$V_b(z) = \begin{cases} -V_b & if \quad |z| \leqslant L/2 \\ 0 & if \quad |z| \rangle L/2 \end{cases}$$

The bound energy solutions (E 〈 0) are solutions of the transcendental equation:

$$k_w \tan\left(k_w \frac{L}{2}\right) = \kappa_b$$

(3.4.62)

for even states and

$$k_w \cot\left(k_w \frac{L}{2}\right) = -\kappa_b$$

(3.4.63)

for odd states. The wavector k_w and κ_b are defined by

$$k_w = \sqrt{\frac{2m^*}{\hbar^2}(E + V_b)}$$

(3.4.64)

and

$$\kappa_b = \sqrt{\frac{-2m^*E}{\hbar^2}}$$

(3.4.65)

For positive energies, one no longer obtains real bound states, but resonances for which the transmission function of the quantum well (corresponding to the transmission probability of an incident electron at energy ϵ) has a maximum. The transmission function can be computed, and is given by

$$T(\epsilon) = |t(\epsilon)|^2 = |[\cos(k_w L) - \frac{i}{2}(\xi + \frac{1}{\xi})\sin(k_w L)]^{-1}|^2$$

(3.4.66)

where ξ is given by the weighted ratio of the wavevectors

$$\xi = \frac{m_b k_w}{m_w k_b}$$

(3.4.67)

where m_b and m_w refer to the electron mass in the barrier and well, respectively, and $k_b = \sqrt{2m^*\epsilon}/\hbar$.

(p.39)

In Fig. 3.4 the energy level of an electron in a square GaAs/$Al_{0.3}$ $Ga_{0.67}$ As quantum well is a function of the well thickness L.

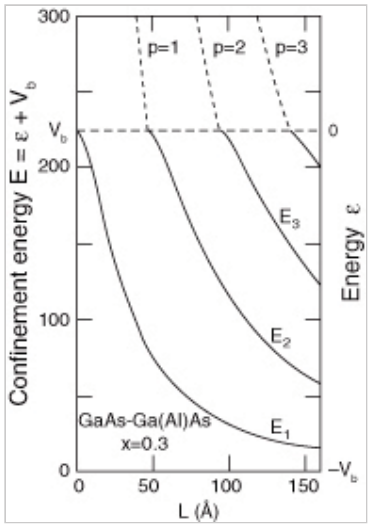


Fig. 3.4. Computed confinement energy of the electrons in the conduction band as a function of well width. Bound states are shown with full lines, and resonances with dotted lines.(Adapted from Bastard [49])

A very attractive feature of the two-band model is that some of the numerical implementations can be done the same way as the one-band one by simply introducing the energy-dependent effective mass as a small change in the code. This is of course also valid in the formula given above. The importance of introducing the non-parabolicity via the interaction of the conduction and valence band is shown clearly in Fig. 3.5, where the energy states of a single quantum well are compared in the case where the non-parabolicity is neglected (dotted line) and where it is taken into account (solid lines). The discrepancy for the position of the highest level is 90 meV, much larger than any experimental uncertainty.

3.4.2 The coupled well system

The system to be considered here is two-quantum-well coupled through a tunnel barrier. Of course this system can be modeled directly by solving the Hamiltonian equation of the whole structure. We, however, wish to study this system in a tight-binding model in which we use a base formed by the solution of the individual wells. Again, following Bastard [49], we consider the Hamiltonian

$$H = \frac{p_z^2}{2m^*} + V_b(z - z_1) + V_b(z - z_2)$$

(3.4.68)

where $V_b(z)$ has the same form as in eqn. 3.4.1.

(p.40)

We solve the problem using the basis wavefunction of the isolated wells,

satisfying

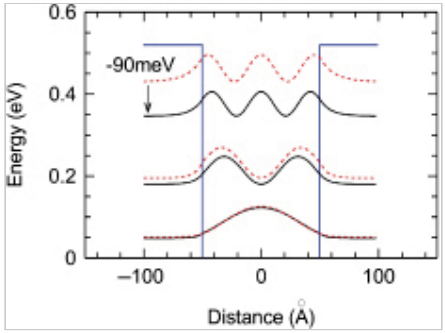


Fig. 3.5. Energy states of a 100 Å thick quantum well computed with a two-band model, and compared with a one-band model (dashed lines). The growing importance of non-parabolicity as one moves away from the gap is clearly apparent.

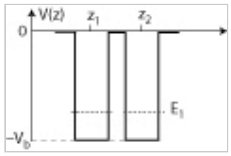


Fig. 3.6. A symmetric coupled quantum well system.

$$\frac{p_z^2}{2m^*} + V_b(z - z_1)\chi(z - z_1) = E_1\chi(z - z_1)$$

(3.4.69)

Considering only the ground states of the isolated wells, the complete wavefunction can be expanded in terms of the basis function using

$$\psi(z) = \alpha\chi(z - z_1) + \beta\chi(z - z_2).$$

(3.4.70)

Introducing the expansion in the Hamiltonian yields the following matrix equation:

$$\begin{pmatrix} E_1 + s - \epsilon & (E_1 - \epsilon)r + t \\ (E_1 - \epsilon)r + t & E_1 + s - \epsilon \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = 0$$

(3.4.71)

Setting the determinant of the matrix to zero yields the solution for ϵ:

$$\epsilon = E_1 \mp \frac{t}{1 \mp r} + \frac{s}{1 \mp r}$$

(3.4.72)

where r is the overlap,

$$r = \langle \chi_1(z - z_1) | \chi_1(z - z_2) \rangle$$

s is the shift

$$s = \langle \chi_1(z - z_1) | V_b(z - z_2) | \chi_1(z - z_1) \rangle$$

(p.41) and *t* the transfer integral

$$t = \langle \chi_1(z - z_1) | V_b(z - z_1) | \chi_1(z - z_2) \rangle.$$

This result can be expressed graphically, as shown in Fig. 3.7. The most important term is then the transfer integral because it is responsible for the splitting between the two states. The larger the transfer integral, the wider the spacing between the two states.

3.4.3 The superlattice

In a superlattice the potential $V_b(z)$ is defined as a sum of potential of the individual wells:

$$V_b(z) = \sum_{n=-\infty}^{\infty} V_b(z - nd)$$

with each potential well being

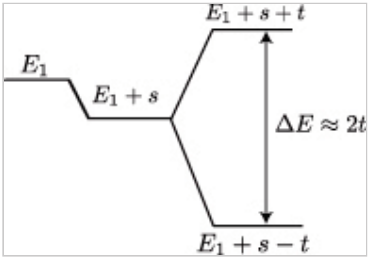


Fig. 3.7. Energy levels in a coupled quantum well system. The term (1 ± *r*) in eqn. 3.4.72 has been neglected.

$$V_b(z - nd) = \begin{cases} -V_b & if \quad |z - nd| \leqslant L/2 \\ 0 & if \quad |z - nd| \rangle L/2 \end{cases}$$

The wavefunctions for the quantum wells are then given by

$$\chi(z) = \alpha \exp(ik_w(z - nd)) + \beta \exp(-ik_w(z - nd)); |z - nd| \leqslant L/2$$

For the barriers, for positive energies ($\epsilon \rangle$ 0), the wavefunction is then given by

$$\chi(z) = \gamma \exp(ik_b(z - nd - \frac{d}{2})) + \delta \exp(-ik_b(z - nd - \frac{d}{2})); |z - nd - \frac{d}{2}| \leqslant h/2$$

In addition to the boundary conditions at both interfaces of the quantum well, we should also impose the periodicity of the structure. As is well-known, it will force the wavefunction to follow the Bloch theorem: that is,

$$\chi_q(z + nd) = \exp(iqnd)\chi_q(z)$$

(3.4.73)
for a periodicity *d.*

(p.42)

The resultant transcendental equation
is given by

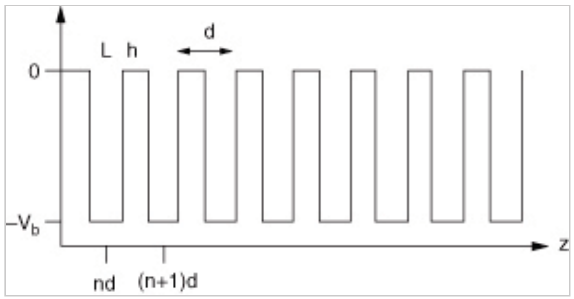


Fig. 3.8. Schematic band structure of a superlattice.

$$\cos(qd) = \cos(k_w L)\cos(k_b h) - \frac{1}{2}(\xi - \frac{1}{\xi})\sin(k_w L)\sin(k_b h)$$

(3.4.74)

where $\xi = k_b/k_w$ has the same meaning as for the single-quantum-wellcase. For negative energies, $\xi \rightarrow i\bar{\xi}$ and the transcendental equation becomes:

$$\cos(qd) = \cos(k_w L)\cosh(\kappa_b h) - \frac{1}{2}(\bar{\xi} - \frac{1}{\bar{\xi}})\sin(k_w L)\sinh(\kappa_b h)$$

(3.4.75)

The above equation defines a dispersion relation between the energy ϵ (appearing in k_w and k_b) and q. An example of such a dispersion, showing clearly the minibands and minigaps, is shown in Fig. 3.9. In particular, the increase in the miniband width with energy is clearly apparent. The superlattice dispersion of the first miniband is compared to a sinusoidal fit, showing that the latter is a very good approximation for narrow minibands. The equivalent real space picture is shown in Fig. 3.11 with the energies and wavefunctions of a finite eight-quantum-well long superlattice

3.5 In-plane dispersion

Up to now, the energy states have been computed for vanishing in-plane momentum, i.e. at $k_{\parallel} = 0$. This assumption was critical to allow the derivation of the pseudoSchrödinger equation 3.2.51 with the energy-dependent effective mass 3.2.52. In fact, keeping the terms dependent on $k_{\parallel}$ in the 2x2 matrix yields a differential equation that contains both first and second derivatives of the conduction band wavefunction.

The system can be solved easily if non-parabolicity can be neglected, i.e. if all the other bands are treated in the perturbation approximation. In that case, the $k_{\parallel}$ dependence appears in the Hamiltonian as an additive term to the potential. As a result the in-plane dispersion can be expressed as an effective mass that is a weighted average of the masses of the quantum well and barriers by the probability density of the state.

As the next order of approximation one may keep a parabolic dispersion, but using the energy-dependent effective mass computed at $k_{\parallel} = 0$. As shown in Fig 3.12, where **(p.43)**

this approximation is compared to the solution of the full 8x8 Hamiltonian for a single 48 Å InGaAs/AlInAs quantum well, the two models yields comparable values of energy for small values of $k_{\parallel}$.

A better approximation is obtained by treating the dispersion of the subband as a bulk material, assuming that the confinement energy E_i computed at $k_{\parallel} = 0$ yields a value of $k_{z,i}$, i.e.

$$k_{z,i} = \frac{\sqrt{2m^*(E_i)}}{\hbar}$$

(3.5.76)

(p.44)

and writing then for the energy of the ith subband at a arbitrary value of $k_{\parallel}$:

$$E_i(k_{\parallel}) = \frac{\hbar^2}{2m^*(E_i(k_{\parallel}))}(k_{z,i}^2 + k_{\parallel}^2).$$

(3.5.77)

Assuming the simple expression 3.1.26 for the energy-dependent effective mass, one obtains for the dispersion

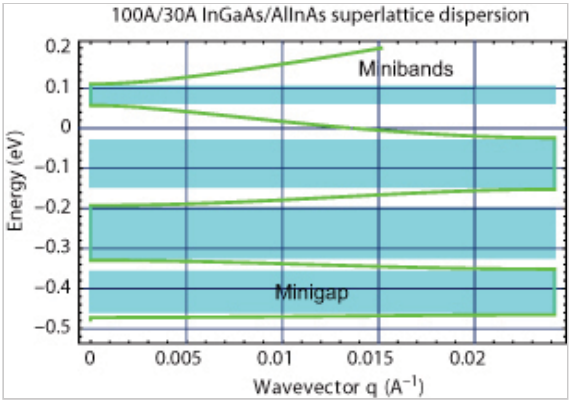


Fig. 3.9. Computed dispersion of a superlattice composed from 100 Å $Ga_{0.47}$ $In_{0.53}$ As quantum wells followed by 30 Å $Al_{0.48}$ $In_{0.52}$ As barriers. A conduction-band discontinuity of $\Delta E_c = 0.52$ eV as well as a Kane energy of 18.3 eV was assumed in an effective two-band model. The minigaps are indicated by shaded regions.

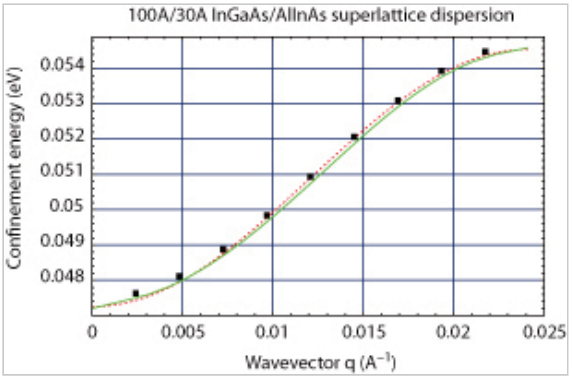


Fig. 3.10. Computed confinement energy $\epsilon + V_b$ dispersion of the first miniband, compared to a sinusoidal fit (dotted line). The results of the finite superlattice are also plotted as full squares.

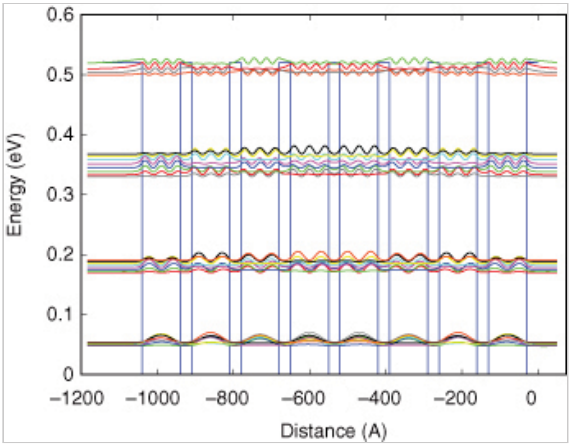


Fig. 3.11. Square of the wavefunctions for a finite superlattice formed by eight 100 Å InGaAs wells separated by 30 Å AlInAs barriers.

$$E_i(k_{\|}) = \frac{E_g}{2}\left\{-1+\sqrt{1+4\left(\frac{E_i}{E_g}\left(1+\frac{E_i}{E_g}\right)+\frac{\hbar^2 k_{\|}^2}{2m^*(0)E_g}\right)}\right\}$$

(3.5.78)

where E_i is the confinement energy at $k_{\|} = 0$. This equation can be inverted to obtain $k_{\|}(E)$. As shown in Fig 3.12, this approximation matches very well the result of the eight-band model. We attribute this good agreement to the fact that, unlike the case of valence band states, all the other bands are far away energetically **(p.45)**

.

3.6 Full model: the valence band

These approximations completely fail in the valence band because the in-plane wavevector couple states belonging to different bands. For this reason, the analysis of intersubband devices operating in the valence band requires the solution of at least a 6x6 **k** · **p** Hamiltonian coupling the valence bands. Accurate predictions are obtained by implementing the multi-band Hamiltonian 3.2.39 that includes Δte bands effects.

(p.46)

It is observed that:

- It is usually convenient to use the growth direction as the quantization

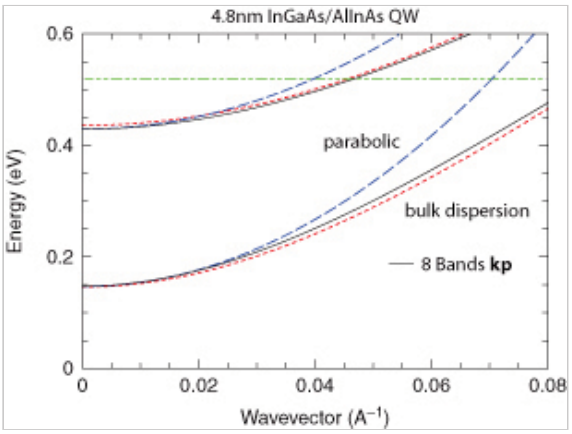


Fig. 3.12. In-plane dispersion for a 48 Å quantum well computed using an eight-band **k** · **p** (full lines), a parabolic approximation with the energy-dependent effective mass computed at $k = 0$ (dashed) and assuming the bulk-like dispersion given by eqn. 3.5.78 (dotted). For InGaAs, values $E_g = 0.78eV$ and $m^*(0)/m_0 = 0.043$ for the gap and the mass were assumed.

direction for the angular momentum.

• The confinement potential lifts the degeneracy between the heavyand light-hole bands, because their different mass induce a different confinement energy.

(p.47)

• The in-plane dispersion is highly non-parabolic because of the coupling between the bands induced by the in-plane momentum. In particular, this coupling prevents any crossing between the light-hole and the heavy-hole band (see dashed lines).

• In some cases the mass is inverted: the bottom of the LH1 band has an electronlike character over some portion of reciprocal space because of the repulsion and its proximity to the HH2 state.

• The presence of biaxial strain will split the heavy-hole and light-hole bands even in the absence of quantum confinement.

• Because of the **k** · **p** interaction, the bands rapidly lose their character (light hole, heavy hole, etc.) when one departs from $k = 0$, and are therefore mixed; each state is represented by a superposition of Bloch functions originating from many bands.

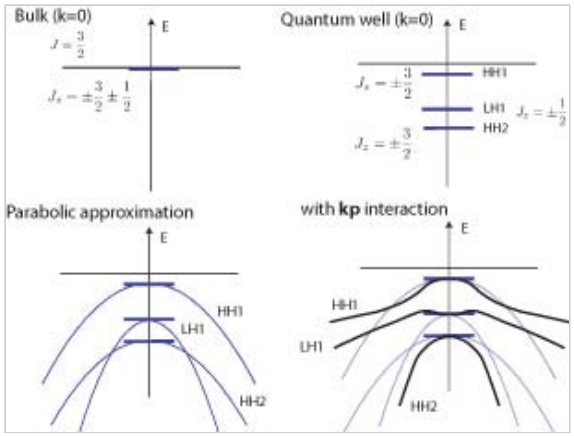


Fig. 3.13. Schematic description of the origin of the valence band dispersion in the quantum well showing schematically the effects of confinement and interactions.

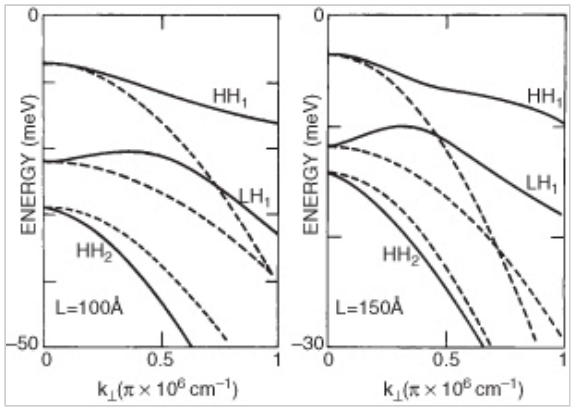
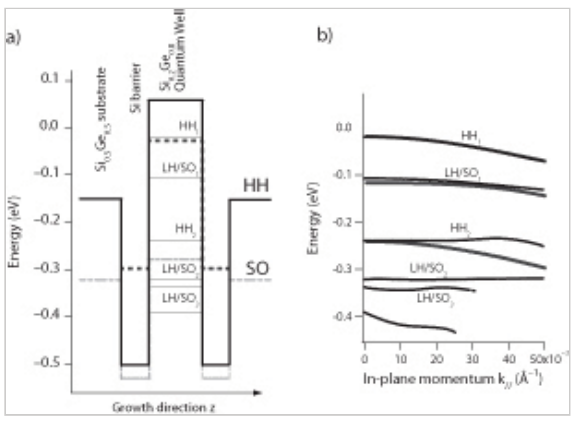
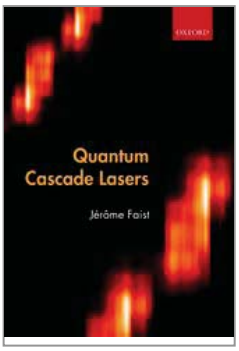


Fig. 3.14. Dispersion of the valence band states of a 100 Å (left) and 150 Å (right) thick GaAs/$Al_{0.3}Ga_{0.7}As$ quantum well. Note the reversal of curvature of the first light-hole band that becomes electron-like near $k_{\parallel} = 0$. Dashed lines: parabolic approximation. From Ref [49].



This effect is shown schematically in Fig. 3.13. As a result, the computed band structure is usually fairly complex, and yields results such as those shown in Fig. 3.14 for an unstrained GaAs/AlGaAs quantum wells, and in Fig. 3.15 for a Si/SiGe quantum well where strain has to be taken into account.

Fig. 3.15. a) Band-structure of a 35 Å $Si/Si_{0.2}Ge_{0.8}$ quantum well grown on $Si_{0.5}Ge_{0.5}$ substrate as calculated in the six-band $\mathbf{k} \cdot \mathbf{p}$ transfer-matrix approach. The energy at $k = 0$ of each bound state is shown by a line. The band onsets are indicated with different styles, solid black, dashed gray and dashed-dot grey for respectively the HH, the LH, and the SO bands. b) Dispersion relations calculated by either a four-bands (gray curve) or a six-bands (black curve) bands $\mathbf{k} \cdot \mathbf{p}$ models. Adapted from [52].



Quantum Cascade Lasers
Jérôme Faist

Print publication date: 2013
Print ISBN-13: 9780198528241
Published to Oxford Scholarship Online: May 2013
DOI: 10.1093/acprof:oso/9780198528241.001.0001

Optical transitions

Jérôme Faist

DOI:10.1093/acprof:oso/9780198528241.003.0004

Abstract and Keywords

Optical absorption and gain arise at the microscopic level resulting from the interaction between the electronic systems and the optical modes propagating in the waveguide of the device. In most cases, the optical fields are strong enough to modify the population of the states, by introducing transitions between them, but not strong enough to dress the states themselves. For this reason, the light-matter interaction can be introduced at the level of a perturbation. As a result, the Hamiltonian describing the electronic states H_0, discussed in Chapter 3, is complemented by an interaction term $H_{int}(t)$, such that $H = H_0 + H_{int}(t)$. Using perturbative methods, this chapter presents the computation of the transition probabilities between the stationary states of $H_0 0$ as a result of the interaction term.

Keywords: optical absorption, hamiltonian, intersubband transition, interband transition, optical fields

At the microscopic level, optical absorption and gain arise as a result of the interaction between the electronic systems and the optical modes propagating in the waveguide of the device. In most cases the optical fields are strong enough to modify the population of the states, by introducing transitions between them, but not strong enough to dress the states themselves. For this reason, the light–matter interaction can be introduced at the level of a perturbation. As a result, the Hamiltonian describing the electronic states H_0, discussed in the preceding chapter, is complemented by an interaction term $H_{int}(t)$, such that

$$H = H_0 + H_{int}(t)$$

(4.0.1)
Using perturbative methods, the transition probabilities between the stationary states of H_0 as a result of the interaction term will be computed.

4.1 Interaction Hamiltonian

Typically two different Hamiltonian can be used to compute the absorption. The first one is the dipole Hamiltonian given by

$$H_{int} = -e\mathbf{R} \cdot \mathscr{E} \sin(\omega t)$$

(4.1.2)
The other one is derived from the potential vector **A** by writing the kinetic energy term of a charge in an electromagnetic field as

$$H = \frac{(\mathbf{p} - e\mathbf{A})^2}{2m}$$

(4.1.3)
using the Coulomb gauge ($\nabla \mathbf{A} = 0$). In a single-band model the mass in eqn. 4.1.3 would be the effective mass $m = m^*$; in a multi-band case, as the band curvature arises from the $\mathbf{k} \cdot \mathbf{p}$ interaction, the mass would be the rest electron mass $m = m_0$. Such a case is treated later in this chapter.

For low intensity (neglecting the term in $|\mathbf{A}|^2$), we obtain as an interaction Hamiltonian

$$H_{int} = -\frac{e}{m}\mathbf{A} \cdot \mathbf{p}$$

(4.1.4)
Because of the large difference between the light wavelength and the atomic dimension, the spatial dependence of **A**(**r**) is neglected inside the matrix elements, and **A**(**r**) is **(p.49)** taken outside the integral. This is called the dipole approximation. The form commonly used is then

$$\langle \psi_i | H_{int} | \psi_j \rangle = \frac{-e\mathbf{A}(\mathbf{r}, t)}{m} \langle \psi_i(\mathbf{r}) | \mathbf{p} | \psi_j(\mathbf{r}) \rangle$$

(4.1.5)
For a plane wave, **A** is parallel to $\mathscr{E}$, and for a wave polarized along z and propagating in the y-direction:

$$\mathbf{A}(\mathbf{r}, t) = A_0 \mathbf{e}_z e^{i(ky - \omega t)} + A_0^* \mathbf{e}_z e^{-i(ky - \omega t)}$$

(4.1.6)
It is convenient to use a purely imaginary A_0, such that both electric $\mathscr{E}$ and magnetic fields B are real: that yields $\mathscr{E} = 2i\omega A_0$ and $B = 2ikA_0$.

The matrix elements of the position and momentum operators are related to each other. As the commutator between p_z and z is $[z, p_z] = i\hbar$, using this rule to compute the commutator between the displacement and Hamiltonian $[H_0, z]$ with $H_0 = \mathbf{p}^2/2m + V(r)$ one can derive the relation

$$\langle \chi_n | p_z | \chi_m \rangle = im\omega_{nm} \langle \chi_n | z | \chi_m \rangle$$

(4.1.7)

This relation between the momentum and position matrix elements enables the use of either interaction Hamiltonian equivalently, with the same remark concerning the mass being either the rest mass m_0 or the effective mass m^* of the electron, depending on the model used for the electron states.

4.2 Intersubband and interband transition

The intersubband and interband transitions have a very different character. The distinction can be easily understood when examining the action of the momentum operator between an initial state *i* in band *v* to the final state *j* in band *v*′:

$$\Psi^{(i,f)}(\mathbf{r}) = f^{(i,f)}(\mathbf{r}) u_{(\nu,\nu')}(\mathbf{r})$$

(4.2.8)

where $f^{(i,j)}$ is the envelope and $u_{v,v'}$ the Bloch part of the wavefunction. The matrix element can be written then as

$$\begin{aligned} \langle \Psi^{(i)} | p | \Psi^{(j)} \rangle &= \langle f^{(i)} u_\nu | p | f^{(j)} u_{\nu'} \rangle \\ &= \langle f^{(i)} | p | f^{(j)} \rangle \langle u_\nu | u_{\nu'} \rangle + \langle f^{(i)} | f^{(j)} \rangle \langle u_\nu | p | u_{\nu'} \rangle \\ &= \langle f^{(i)} | p | f^{(j)} \rangle \delta_{\nu\nu'} + \langle f^{(i)} | f^{(j)} \rangle \langle u_\nu | p | u_{\nu'} \rangle \end{aligned}$$

(4.2.9)

where we have taken advantage of the slow variation of the envelope function compared to that of the Bloch part. The first part of the last expression represents the intersubband transition, and will be non-zero when two envelope states are taken from the same band. Conversely, the second corresponds to an interband transition matrix element

(p.50) Note that this “clean” distinction between intersubband and interband formally disappears as soon as a multi-band model is used for the initial and final states: i.e., as soon as these states are written as

$$\Psi^{(i)}(\mathbf{r}) = \sum_\nu f_\nu^{(i)}(\mathbf{r}) u_\nu(\mathbf{r})$$

(4.2.10)

In these situations all transitions have various “intersubband” and “interband” components. This is especially true when examining transitions between valence band states, as they are strongly mixed as soon as $\mathbf{k} \neq 0$.

4.3 Selection rules and absorption geometries

For conduction band states in a one-band model, only the component of the electric field (and of the potential vector, as both are parallel) in the growth direction will induce a non-zero dipole matrix element. In fact, taking for the wavefunction the canonical form

$$f_i = \frac{1}{\sqrt{S}} e^{i\mathbf{k}_\parallel \mathbf{r}_\parallel} \chi_i(z)$$

(4.3.11)

then the matrix element corresponding to the x-component of the momentum p_x yields:

$$\langle f_i|\mathscr{E}_x p_x|f_j\rangle = \frac{1}{S}\langle\chi_i|\chi_j\rangle\hbar k_x\mathscr{E}_x\delta_{\mathbf{k}_\parallel,\mathbf{k}_\parallel'} = \frac{1}{S}\hbar k_x\mathscr{E}_x\delta_{\mathbf{k}_\parallel,\mathbf{k}_\parallel'}\delta_{i,j}$$

(4.3.12)

which, because of the term $\delta_{i,j}$ is zero for transitions between different states. Of course, a similar condition is valid in the y-direction. As a result, *only the z-component of the electric field couples to the intersubband transition*. This has profound experimental implications as it, for example, rules out absorption for light at normal incidence. As a result, a number of geometries have been developed to measure the absorption, as shown in Fig. 4.1. Because of the large refractive index (typically larger than $n_{refr} \rangle$ 3) of the semiconductor substrates, prism-like geometries are required to induce a large electric field component in the growth direction. The validity of the selection rule

(p.51) was checked in GaAs/AlGaAs samples by H. C. Liu and coworkers in a photocurrent experiment on specially designed samples [53]. As shown in Fig. 4.2, the absorption by the TE polarization was less than 0.2% of the TM absorption. The absorption in TE is non-zero when considering the wavefunctions in the multi-band case, and will be discussed shortly.

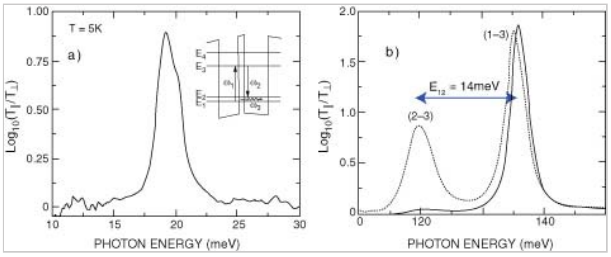


Fig. 4.1. Experimental geometries allowing the measurement of intersubband transitions. Because of the large refractive index of the semiconductor substrates, prism-like geometries are favorable, as they induce a large electric field component in the growth direction.

4.4 Absorption strength

4.4.1 Absorption for a two-dimensional system

When examining the transmission of an essentially two-dimensional system, one should realize that the correct number to represent the absorption is unitless, and represents the fraction of the light absorption by the system. The scattering rate, using Fermi's golden rule, is

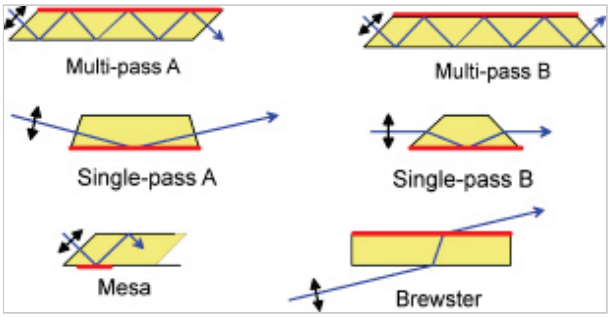


Fig. 4.2. Photocurrent in a quantum well photoconductor as a function of the polarization. The residual responsivity for the TE polarization is less than 0.2% than for the electric field normal to the layers. Reprinted with permission from [53]. Copyright 1998, AIP.

$$R = \frac{1}{\tau} = \frac{\pi}{2\hbar}\sum |\langle i|H_{int}|f\rangle|^2 \rho(E_f - E_f - \hbar\omega).$$

(4.4.13)

The absorption of the electromagnetic wave by the quantum system is responsible for a decay of the latter. Writing α_{2D} as the fraction of the wave absorbed by the system, we have that the power loss per unit area S is now

$$P_{diss}/S = I\alpha_{2D}$$

(4.4.14)

and the intensity is related to the electric field by

$$I = \frac{1}{2}\epsilon_0 n_{refr} c\mathscr{E}^2 \cos\theta$$

(4.4.15)

(p.52) where n_{refr} is the refractive index and c the velocity of light. Assuming that the beam arrives at an angle θ to the normal, only the component of $\mathscr{E}$ in the z-direction will couple to the intersubband system. Balancing the energy loss of the EM field with the energy gained by the quantum system yields

$$\alpha_{2D}(\omega) = \frac{2R\hbar\omega}{\epsilon_0 n_{refr} c\mathscr{E}^2}\frac{\sin^2\theta}{\cos\theta}$$

(4.4.16)

We defined the (unitless) two-dimensional absorption $\alpha_{2D}(\omega)$ as the fraction of the light absorbed by crossing the two-dimensional system, such that the light intensity is $I = I_0 \exp(-\alpha_{2D})$. Substituting the scattering rate by the one computed above, we obtain

$$\alpha_{2D}(\omega) = \frac{\pi\omega}{\epsilon_0 n_{refr} c}\sum_{i,f} |\langle i| - ez|f\rangle|^2 \rho(E_f - E_i - \hbar\omega)\delta(\mathbf{k}_f - \mathbf{k}_i)\frac{\sin^2\theta}{\cos\theta}$$

(4.4.17)

where ρ is now a density of state per unit area, the δ function forces the conservation of the momentum in the light absorption, and i and f are the initial and final states. In the one-band model the electron subbands are parallel, and the absorption is

$$\alpha_{2D}(\omega) = \frac{\pi e^2\omega}{\epsilon_0 n_{refr} c} z_{if}^2 n_i \delta(E_f - E_i - \hbar\omega)\frac{\sin^2\theta}{\cos\theta}$$

(4.4.18)

where n_i is the populations (per unit area) of the initial state (the final state is assumed to be empty) and z_{if} is the shorthand notation for the dipole matrix element.

Scattering of the electrons with lattice vibrations, as well as collisions with interface roughness, broadens the line in practice, and the δ function $\delta(E_f - E_i - \hbar\omega)$ can be replaced by a normalized Lorentzian line $\frac{\gamma/\pi}{(E_f - E_i - \hbar\omega)^2 + \gamma^2}$ of half width γ, and we have finally for the absorption:

$$\alpha_{2D}(\omega) = \frac{e^2 \omega z_{ij}^2 n_i}{\epsilon_0 n_{refr} c} \frac{\gamma}{(E_f - E_i - \hbar\omega)^2 + \gamma^2} \frac{\sin^2\theta}{\cos\theta}$$

(4.4.19)

For the geometries shown in Fig. 4.1 one should multiply the above result by the number of quantum wells in the stack and by the number of "passes" in the waveguide: i.e., the number of times the light crosses the multi-quantum well system.

The value of the dipole matrix element between state *i* and state *j* can be computed analytically for an infinite quantum well, and the value yields for the dipole in the growth direction, taken as *z*,

$$\langle \chi_i | z | \chi_j \rangle = z_{ij} = \frac{L}{\pi^2} \frac{8ij}{(i^2 - j^2)^2}$$

(4.4.20)

where *L* is the width of the quantum well. This relation stresses the linear relationship between the dipole matrix element and the size of the quantum well.

(p.53) 4.4.2 Waveguide

Another interesting case is that of the waveguide where the absorption occurs in a guided mode. In that case, the angle of incidence is $\theta = \pi/2$ and expression 4.4.19 diverges. In that case it is more convenient to think of the absorption as being from an "homogeneous material" of thickness *L*, defined as absorption coefficient $\alpha = \alpha_{2D} \cos\theta / L$ for an homogeneous material in which the light would be traveling. It is sufficient to reinterpret the above equation as

$$\alpha(\omega) = \frac{e^2 \omega z_{ij}^2 N_i}{\epsilon_0 n_{refr} c} \frac{\gamma}{(E_f - E_i - \hbar\omega)^2 + \gamma^2}$$

(4.4.21)

where $N_i = n_i / L$ is the volume electron density. The natural questions are of course, "What is the correct value of *L*? Is *L* the thickness of the well, the effective one of the wavefunction, the one of the quantum well and the barrier?" In most cases, happily, the particular choice of *L* has no practical influence on the relevant physical quantity to be measured. In fact, what is measured directly is *not* the material absorption but that of a guided mode. As shown in Chapter 6, the modal loss of the waveguide α_w is related to the material one by an overlap factor,

$$\alpha_w = \frac{\int_{-\infty}^{\infty} \mathcal{E}^2(z) \alpha(z) dz}{\int_{-\infty}^{\infty} \mathcal{E}^2(z) dz} = \Gamma \alpha$$

(4.4.22)

where the overlap factor is defined as a function of the electric field profile in the waveguide $\mathcal{E}(x)$ by

$$\Gamma = \frac{\int_{-L/2}^{L/2} \mathscr{E}^2(z)dz}{\int_{-\infty}^{\infty} \mathscr{E}^2(z)dx} \approx L\tilde{\mathscr{E}}^2$$

(4.4.23)

if $\tilde{E}^2$ is the normalized intensity in the guide at the location of the quantum well. The linear dependence of Γ in L just cancels the inverse dependence of α in the same quantity.

4.4.3 Gain and loss cross-sections

In an electronic system where the upper-state and not the lower state is populated, stimulated emission will occur with exactly the same rate as for the absorption. As a result, the same derivation as above can be performed, just changing the sign of P_{diss} and R, as the optical power is now generated instead of absorbed. For this reason, when both lower and upper-states are populated, the ground-state density in eqn. 4.4.21 should be substituted by the difference between ground- and upper-state densities.

Rewriting eqn. 4.4.21 using the wavelength $\frac{2\pi}{\lambda}$ instead of $\frac{\omega}{c}$, and writing the volume density $N_i = n_i/L_p$ as a function of the sheet density n_i, where L_p is (by convention) the period length of the active region of a quantum cascade laser, we obtain

$$\alpha(\omega) = \frac{2\pi e^2 z_{if}^2 (n_i - n_f)}{\epsilon_0 n_{refr} \lambda L_p} \frac{\gamma}{(E_f - E_i - \hbar\omega)^2 + \gamma^2}$$

(4.4.24)

(p.54) As α can now be either negative or positive, the gain is defined simply as a negative absorption:

$$g(\omega) = -\alpha(\omega)$$

(4.4.25)

An important quantity is the peak-gain cross-section, or the maximum gain at resonance ($\hbar\omega = E_f - E_i$) per unit of population of the upper-state:

$$g_c = g/(n_f - n_i) = \frac{2\pi e^2 z_{if}^2}{\epsilon_0 n_{refr} \lambda \gamma L_p}$$

(4.4.26)

4.5 Experimental results

The first measurement of the intersubband absorption in a heterojunction quantum well (as compared to that in an inversion layer) was performed in a fifty period GaAs/$Al_{0.3}Ga_{0.7}As$ multiquantum well by West and Eglash [6]. In these first measurements, shown in Fig. 4.3, the quantum wells were modulation doped at a level of 4×10^{11} cm^{-2} and measured at the Brewster angle. A much better signal over noise

can be achieved using the
waveguide configuration such
as the one shown in Fig. 4.1, as

the absorption strength is greatly enhanced. Indeed, the absorption 4.4.19 is given by the expression

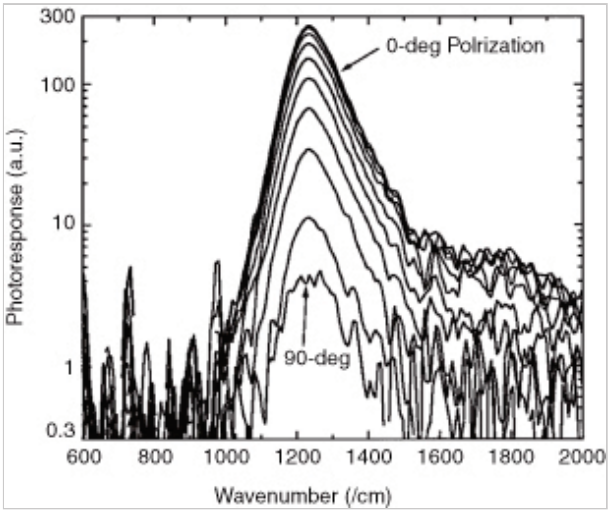


Fig. 4.3. Intersubband absorption in a series of 65 Å thick and 82 Å thick GaAs/$Al_{0.3}Ga_{0.7}As$ multi quantum wells. Reprinted with permission from [6]. Copyright 1985, AIP

$$\alpha_{2D}(\omega) = \frac{e^2 \omega z_{ij}^2 n_i}{\epsilon_0 n_{ref} c} \frac{\gamma}{(E_f - E_i - \hbar\omega)^2 + \gamma^2} \frac{\sin^2\theta}{\cos\theta} N_{QW} N_{pass}$$

(4.5.27)

where N_{pass} is the number of passes across the quantum well system, and N_{QW} isthe number of quantum wells in the stack. For a Brewster-angle experiment, the lastterm in eqn. 4.5.27 is 0.089 for $N_{QW} = 1$; for a waveguide geometry with a 45-degree **(p.55)** angle (such as "multi-pass A" in Fig. 4.1) with a typical length of $l = 4$ mm and a thickness of $t =$ 0.5 mm, the number of passes $N_{pass} = l/t = 8$, and the last term is equal to 5.65—an improvement of a factor of 63 in comparison with the Brewster-angle geometry. The latter is especially unfavorable for materials with a high refractive index, since the internal angle becomes very low, with a very small component of the field in the growth direction.

Fig. 4.4a shows the absorption spectrum of a series of twenty InGaAs/AlInAs quantum wells, 52 Å thick, doped at a sheet density of $n_s = 2.5 \times 10^{11}$ cm^{-2} in a waveguide geometry at 45-degree incident angle.

In contrast to interband transition, where the interband matrix element r_{cv} is the dominant term, the atomic-like nature of the joint density of state as well as the tailorability of the potential enables the fabrication of complex energy ladder structures. This capability was exploited to

create multi quantum wells with levels engineered to provide resonant non-linearities. The specific feature of these structures was to provide electronic states that were all coupled by strong dipole matrix elements. The example of a structure designed to provide a resonance for the third harmonic field [54], i.e. presenting a very large resonant $\chi_{3\omega}^{(3)}$, is shown in Fig. 4.4b. The structure consists of an 18 and 20 Å thick InGaAs quantum well pair, coupled to a thick 42 Å thick InGaAs well by 16 Å thick AlInAs barriers. The absorption of this structure is shown in Fig. 4.4b and clearly exhibits the absorption resonances related to the 1–2, 1–3, and 1–4 transitions.

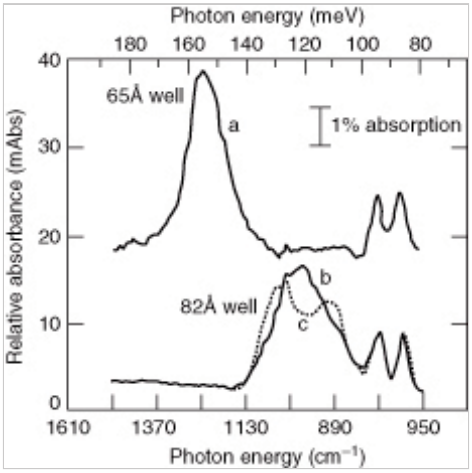
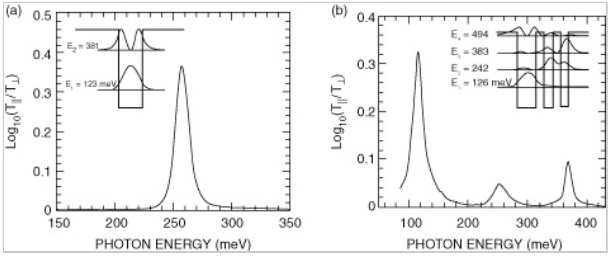


Fig. 4.4. a) Intersubband absorption between two bound states in a stack of twenty repetitions of a 52 Å InGaAs/AlInAs quantum well. The latter are doped with an electron sheet density of $n_s = 2.5 \times 10^{11}$ cm^{-2}. b) Intersubband absorption in a multi-quantum well designed for triply resonant non-linear susceptibility. Reprinted with permission from [51]. Copyright 1994, APS.

Another situation arises when a state (the upper one, for absorption measurements) is located in a continuum. In this case, the absorption is distributed among a continuum of transitions corresponding to the various allowed transitions. As both localized (for the ground states) and continuum states (for the excited states) have to be treated **(p.56)** together, normalization of the wavefunctions might be somewhat challenging. One common trick is to confine artificially the whole potential within a large "box", such that all states can be considered as bound states. The size of the box must be large enough such that the level spacing created by the additional confinement is smaller than the natural line width of the levels.

A comparison between a bound-to-continuum and bound-to-bound transition is shown in Fig. 4.5, where the absorption measured in a series of 32 Å thick quantum well is compared to the absorption in a system where the excited resonance of the system, localized some 60 meV above the onset of the continuum, is confined by electronic Bragg reflection [55]. The asymmetric line of the bound-to-continuum transition

displays a maximum close to the onset of continuum, and displays two features attributed to a superlattice effect due to the periodic arrangement of the quantum wells. Note that in this system the maximum of the absorption does *not* correspond to the position of the resonance



in transmission. In the second sample the main quantum wellis surrounded by a superlattice consisting of an L_w = 16 Å thick InGaAs quantum well and an L_b = 39 Å thick AlInAs barrier, built in order to fulfil the Bragg reflection **(p.57)** condition

$$k_w L_w = \frac{\pi}{2}$$

(4.5.28)
and

$$k_b L_b = \frac{\pi}{2}$$

(4.5.29)
with the electron wavevectors in the well k_w and in the barrier k_b evaluated at the energy of the resonance. These conditions are equivalent to those used to build optical quarter wave stack for high-reflectivity mirrors. As a result, the absorption shown in Fig. 4.5d is narrow and symmetric, as is the one for a bound-to-bound transition shown in Fig. 4.4.

Fig. 4.5. a) Energy levels in a sample consisting of a periodic stack of 32 Å thick quantum wells. The measured absorption (b) shows an asymmetric line with two features corresponding to superlattice resonances formed by hybridation of the resonances created by the individual wells. In the second sample (c), electron reflectors formed by 16 Å thick InGaAs quantum wells and 39 Å thick quantum barriers localize the state above the onset of the classical continuum. (d) The resultant absorption line is narrow and peaks at a higher energy than the onset of the continuum. (Adapted from [55]).

4.6 Sum rule in absorption

Let us consider first a one-band model. Combining the completeness of the (eigen)states

$$\sum_i |\chi_i\rangle\langle\chi_i| = 1$$

(4.6.30)
and the relationship between position and momentum matrix element 4.1.7, the sum rule for the oscillator strength can be derived:

$$\sum_j f_{ij} = m_0/m^*$$

(4.6.31)
where the oscillator strength is defined as

$$f_{ij} = \frac{2m_0}{\hbar^2}|z_{ij}|^2(E_j - E_i)$$

(4.6.32)
Writing the sum rule this way emphasizes the effect of the effective mass in the enhancement of the intersubband absorption that basically scales like m_0/m^*. In the $\mathbf{k}\cdot\mathbf{p}$ approach of Kane, the effective mass arises from the interaction between the conduction and valence bands, and is therefore proportional to the bandgap. As a consequence, narrow-gap semiconductors will exhibit a much larger intersubband transition strength than wide bandgap materials. This important fact is obscured when a “normalized” oscillator strength is used, $f' = fm^*$, in place of the usual definition 4.6.32.

The sum rule is illustrated graphically in Fig. 4.6, where the absorption spectra of samples with different quantum well arrangements are compared for similar electron densities and experimental conditions. The strength of a simple square well 1-2 intersubband absorption is “spread” further between more and more transitions when considering two or three quantum well systems.

Eqn. 4.6.31 also holds with the initial state being an excited state, in which case the “downward” transitions have a negative sign. Since the sum must remain constant, the upwards transitions have an oscillator strength that grows with the initial state index *i*. In the simplest case of an harmonic oscillator, the oscillator strength $f_{i,\,i+1}$ between adjacent states is simply proportional to the index *j* of the state:

$$f_{j,j+1} = jf_{1,2} = j(m^*/m_0)^{-1}$$

(4.6.33)

This very important fact shows that transitions between excited states, as they occur naturally in quantum cascade lasers, naturally yield larger intersubband absorption **(p.58)**

.

4.7 Absorption in a quantum well: a two-band model

The problem arising with the previous model is that it does not allow the introduction of non-parabolicity. It is valid, therefore, only for confinement energies much smaller than the bandgap (E_n, E_m 〈〈 E_g). One could be tempted to use the same approach when computing the absorption in the multi-band case, taking as the matrix element the envelope function of the conduction band state. This approach was used in some of the initial intersubband literature, but led to difficulties since, strictly speaking, the envelope functions of the conduction bands are no longer orthogonal to each other ($\langle \chi_m | \chi_n \rangle \neq \delta_{nm}$). A much better approach is to compute the matrix element directly in the multi-band model [51]. If one wants to treat a conduction-band state, an effective two-band model is a good approximation. In the spirit of eqn. 3.2.45, we assume a two-component wavefunction for the initial $\Psi^{(1)}$ and final $\Psi^{(2)}$ state, with

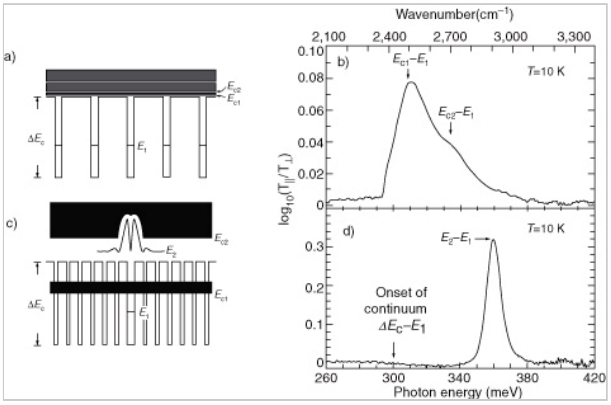


Fig. 4.6. Comparison of the intersubband absorption for various structures, with energy levels schematically drawn close to the curves, illustrating the spreading of the oscillator strength for samples with an increasing number of allowed transitions.

$$\Psi^{(1)} = \chi_c^{(1)} u_c + \chi_v^{(1)} u_v$$

(4.7.34)

$$\Psi^{(2)} = \chi_c^{(2)} u_c + \chi_v^{(2)} u_v$$

(4.7.35)

We then evaluate the matrix element p_z:

$$\begin{aligned}\langle \Psi^{(1)} | p_z | \Psi^{(2)} \rangle &= \langle \chi_c^{(1)} u_c + \chi_v^{(1)} u_v | p_z | \chi_c^{(2)} u_c + \chi_v^{(2)} u_v \rangle \\ &= \langle \chi_c^{(1)} u_c | p_z | \chi_c^{(2)} u_c \rangle + \langle \chi_c^{(1)} u_c | p_z | \chi_v^{(2)} u_v \rangle \\ &+ \langle \chi_v^{(1)} u_v | p_z | \chi_c^{(2)} u_c \rangle + \langle \chi_v^{(1)} u_v | p_z | \chi_v^{(2)} u_v \rangle\end{aligned}$$

(4.7.36)

Let us consider the first term, which yields

(p.59)

$$\begin{aligned}\langle \chi_c^{(1)} u_c | p_z | \chi_c^{(2)} u_c \rangle &= \langle \chi_c^{(1)} | p_z | \chi_c^{(2)} \rangle \langle u_c | u_c \rangle + \langle \chi_c^{(1)} | \chi_c^{(2)} \rangle \langle u_c | p_z | u_c \rangle \\ &= \langle \chi_c^{(1)} | p_z | \chi_c^{(2)} \rangle\end{aligned}$$

(4.7.37)

Similarly, the second term will yield

$$\langle \chi_c^{(1)} u_c | p_z | \chi_v^{(2)} u_v \rangle = \langle \chi_c^{(1)} | \chi_v^{(2)} \rangle p_{cv}$$

(4.7.38)

Using similar derivations for the third and fourth terms, the result can be summarized in a matrix form:

$$\mathbf{p}_z = \begin{pmatrix} p_z & p_{cv} \\ p_{cv}^* & p_z \end{pmatrix}$$

(4.7.39)

acting on the components (χ_c, χ_v) of the wavefunction. Dropping the diagonal terms as p_z 〈 〈 p_{cv}, we finally obtain

$$\langle \Psi^{(1)} | p_z | \Psi^{(2)} \rangle = \langle \chi_c^{(1)} | \chi_v^{(2)} \rangle p_{cv} + \langle \chi_v^{(1)} | \chi_c^{(2)} \rangle p_{cv}^*$$

(4.7.40)

Using the relationship between χ_c and χ_v given by eqn. 3.2.46, and using the definition of the energy-dependent effective mass 3.2.48, we finally obtain

$$\langle \Psi^{(1)} | p_z | \Psi^{(2)} \rangle = \langle \chi_c^{(1)} | (-i\hbar \frac{\partial}{\partial z}) \frac{m_0}{m(E,z)} + \frac{m_0}{m(E,z)} (-i\hbar \frac{\partial}{\partial z}) | \chi_c^{(2)} \rangle$$

(4.7.41)

In this picture, both intersubband and interband transitions are treated on the same footing.

Using a generalized version of this model, it is possible to show [56] that transitions with in-plane electric field couple states with opposite spins, and for a symmetric quantum well for a transition between state of opposite parity, the ratio between the two absorption is given approximately by

$$\frac{\alpha_x}{\alpha_z} \approx \frac{1}{3} \frac{E_{ij}\Delta_0}{(E_i + E_g + \Delta_0)(E_j + E_g)}$$

(4.7.42)

where E_{ij} is the intersubband transition energy, Δ_0 is the spin–orbit splitting, and E_g is the fundamental energy gap of the material. For a transition at $\lambda = 10\ \mu$m in a GaAs/AlGaAs quantum well, the predicted fraction of the absorption in TE is only 0.35% of the TM absorption, in rough agreement with the value observed experimentally.

4.8 Depolarization shift

In absorption experiments the electron density is easily of the order of $10^{17–18}$ cm^{-3}. As a result, electron–electron interactions are expected to add a correction to the transition energy. Besides the Hartree potential, treated in the preceding chapter, and whose influence will be felt mostly for potentials with a dipole charge, another important effect is the depolarization shift. This effect can be seen as a screening of the interaction of one electron by the rest of the electron bath. Following Ando, Fowler, and Stern [5], it can be treated in a relatively straightforward manner by considering **(p.60)** the quantum well system as a sheet of electronic plasma of sheet density n_s, and thickness d_{eff}, with a (three-dimensional) conductivity at frequency ω given by

$$\sigma_{zz} = \frac{n_s e^2 f_{12}}{\epsilon_0 m_0 d_{eff}} \frac{i\omega}{\omega_{12}^2 - \omega^2 - 2i\omega/\tau}$$

(4.8.43)

where τ characterizes the broadening of the transition, and f_{12} and ω_{12} the oscillator strength and bare electron resonance frequency respectively. The current per unit area is given as a function of the electric field $\mathcal{E}_z$ by

$$j_{zz} = \sigma_{zz}\mathcal{E}_z$$

(4.8.44)

Between the quantum well system and the semiconductor having itself a dielectric constant ϵ_{sc}^r, the boundary conditions of Maxwell's equation force the continuity of the z-component of the displacement vector $D_z = \epsilon_0 \epsilon \mathcal{E}_z$. As a result, following reference [5], we define the ratio between the the electric field outside $\mathcal{E}_{ext}$ the slab to that inside the slab by $\mathcal{E}_{zz}$:

$$\mathcal{E}_z = \frac{\mathcal{E}_{ext}}{\epsilon_{zz}}$$

(4.8.45)

This relative dielectric constant of the two-dimensional electron gas, in the z-direction, is given as a function of the conductivity by

$$\epsilon_{zz}(\omega) = 1 + \frac{i\sigma_{zz}(\omega)}{\epsilon_0 \epsilon_{sc}^r \omega}$$

(4.8.46)

The effective conductivity $\tilde{\sigma}_{zz}$, allowing us to write the current–field relation eqn. 4.8.44 in terms of the external field $\mathcal{E}_{ext}$, is then given by

$$\tilde{\sigma}_{zz} = \frac{\sigma_{zz}}{\epsilon_{zz}}$$

(4.8.47)

This quantity will control the dissipation of the two-dimensional system. The dissipation in the two-dimensional electron system is given by the effective conductivity

$$\frac{1}{2}\Re e(j_{zz} d_{eff} E^*_{ext}) = \frac{1}{2}\Re e(\tilde{\sigma}_{zz}) d_{eff} \mathscr{E}^2_{ext}$$

(4.8.48)

The expression for the effective conductivity is given as a function of frequency by substituting eqns. 4.8.43 and 4.8.46 into eqn. 4.8.47, and its expression (including the term d_{eff} for simplicity) may be written as

$$\tilde{\sigma}_{zz} d_{eff} = \frac{n_s e^2 f_{12}}{\epsilon_0 m_0} \frac{i\omega}{\tilde{\omega}^2_{12} - \omega^2 - 2i\omega/\tau}$$

(4.8.49)

where a resonance at frequency $\tilde{\omega}_{12}$

$$\tilde{\omega}^2_{12} = \omega^2_{12} + \tilde{\omega}^2_p$$

(4.8.50)

is blueshifted from the bare frequency by an effective plasma frequency given by

$$\tilde{\omega}^2_p = \frac{n_s e^2 f_{12}}{\epsilon_0 \epsilon^r_{sc} m_0 d_{eff}}$$

(4.8.51)

As a result, the absorption will be peaking at higher frequency than the bare transition. This difference is called the depolarization shift, and is especially strong for large electron densities and low frequencies. Its importance was appreciated very early, when **(p.61)** the subband structure of the Si/SiO_2 interfaces was measured. Fig. 4.7 shows the intersubband absorption in an asymmetric $GaAs/Al_xGa_{1-x}As$ coupled well system. This system exhibited a bare subband splitting of 14 meV for the E_{12} transition, and allowed transitions E_{13} and E_{23} in the mid-infrared range. As a result, the depolarization shift is negligible for these transitions, and the bare value of E_{12} can be directly extracted from the values of E_{13}, E_{23}:

$$E_{12} = E_{13} - E_{23} = 15 meV$$

(4.8.52)

The derived value compares well with the computed value (14 meV). In contrast, the measurement of the absorption in the terahertz of the absorption peaks at a photon energy $\tilde{E}_{12}$ = 19.3 meV that includes a significant contribution from the depolarization shift. Indeed, applying eqns. 4.8.50 and 4.8.51 with $n_s = 4 \times 10^{11}$ cm^{-2}, ϵ_{sc} = 13.1, d_{eff} = 15*nm*, f_{12} = 5.9, and yields $\tilde{E}_{12}$ = 19.1 meV, which is very close to the experimental value

.

4.9 Absorption linewidth

The lineshape function has a very important role in the modeling of quantum cascade lasers. For any given design, material choice, injection level, and the densities and oscillator strength will be fixed. The value of the gain will then be inversely proportional to the linewidth as the lineshape function is normalized to unity.

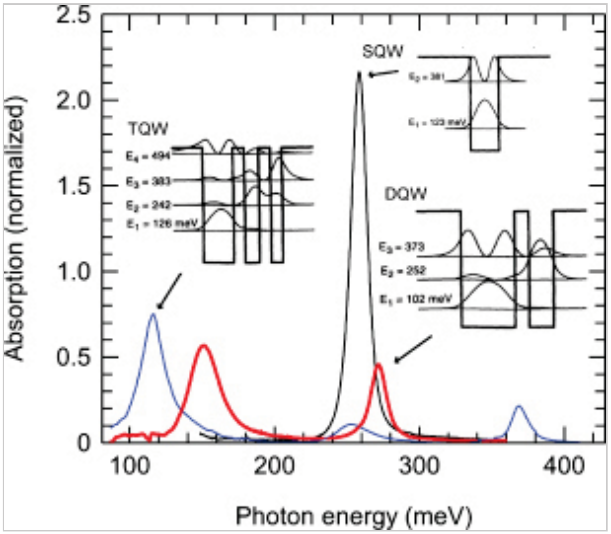


Fig. 4.7. Intersubband absorption in a modulation doped asymmetric GaAs/$Al_xGa_{1-x}As$ coupled well system, consisting of 61 Å and 70 Å GaAs well coupled through a 20 Å $Al_{0.33}Ga_{0.67}As$ barrier. The sheet carrier density is $n_s = 4 \times 10^{11}$ cm^{-2}. Inset: energy level scheme. a) Absorption in the terahertz, showing the (1–2) transition with the depolarization shift. b) Absorption in the mid-infrared, showing the (2–3) and (1–3) transitions where the depolarization shift is negligible.

4.9.1 Homogeneous and inhomogeneous broadening

In atomic systems, finite upper-state lifetime and inhomogeneities convert the lineshape of the optical transition from the δ function that arises from an energy conservation **(p.62)** requirement into a function with a finite energy width. One therefore distinguishes the *homogeneous* broadening component (lifetime broadening due to collisions and radiative transitions) from the *inhomogeneous* components (velocity distribution due to thermal motion of the individual atoms).

An homogeneous broadening will yield a Lorentzian lineshape such as the one assumed in the preceding section $g_{ij}(\omega)$:

$$g_{ij}(\omega) = \frac{1}{\pi} \frac{\gamma}{(\hbar\omega - \hbar\omega_{ij})^2 + \gamma^2}$$

(4.9.53)

where ω_{ij} is the resonant frequency and γ is the broadening parameter. In contrast to a simple atomic system, the scattering time defining the value of γ is not only given by the lifetime of the electron in the upper-state. In a quantum well system, upper and lower states are actually subbands, and a scattering process taking an electron from a state $\mathbf{k}_i$ to $\mathbf{k}_f$ within the same subband will not change the population of the lower or upper-state while still inducing loss of phase (dephasing) of the state, and therefore a lifetime broadening. Using the terminology of nuclear magnetic resonance, the time between such dephasing processes is called T_2, while the upper-state lifetime is referred to as T_1. As a result, the broadening parameter can be written as

$$\gamma = \hbar\left(\frac{1}{2T_1} + \frac{1}{T_2}\right)$$

(4.9.54)

A more complete model will take into account the correlation between the scattering mechanisms for both states[1] [5], and the computation of these two times is the topic of the next chapter.

In a quantum-well system, inhomogeneous broadening will arise from long-range quantum well interface fluctuations, from non-parabolicity, as well as in multi quantum well systems, from well-to-well width fluctuations. Interface roughness occurring on a spatial scale much shorter than the de Broglie wavelength of the electron, as we show in the next chapter, is best treated as an homogeneous broadening mechanism.

Limiting cases. According to these considerations, the lineshape should tend to a square, Lorenzian, or Gaussian lineshape in the following limiting cases, as shown schematically in Fig. 4.8a.

- Non-parabolicity. At low temperatures, for large electron densities and narrow bandgap materials, where non-parabolicity should be dominant, a square-shaped lineshape should be observed. This lineshape can be seen as an image of the step-like joint density of state of two subbands with different masses.
- Disorder. For narrow wells, large bandgap materials with a heavy mass, wheredisorder should play a dominant role, a Gaussian lineshape should be observed.
- Lifetime. In a clean system in which a short lifetime would be expected, a Lorenzian lineshape should be measured.

(p.63) The low-temperature measurement [57] of the 1–3 transition of the asymmetric quantum well system shown in Fig 4.7 is compared to a Gaussian and Lorenzian fit in Fig. 4.8b. The experimental measurement fits very well with a Lorenzian line shape with broadening parameter Γ = 1.33 meV. Note that the same fit on the 2–3 transition of the same system leads to a much more Gaussian lineshape. In the low-density carrier regime, a Gaussian line shape was also measured in a single quantum well in which the density could be modulated [58].

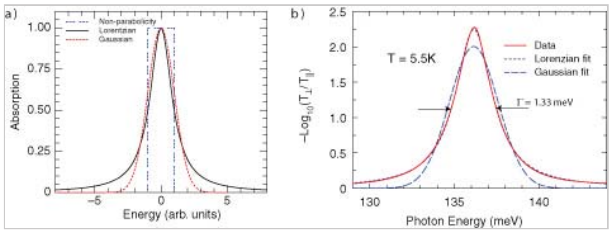


Fig. 4.8. a) Lineshape functions in various limiting cases, as discussed in the text. b) Comparison between the absorption measured in a coupled well system shown in Fig. 4.7 (solid line) with a Gaussian (dashed) and a Lorenzian (dotted line) fit. The

4.9.2 Interface disorder

In quantum-well samples the greatest source of disorder is the monolayer fluctuations of the quantum well interfaces. The first experimental

broadening corresponds to a full width at half maximum of 2 Γ = 2.66 meV.

indication that the linewidth was limited by interface roughness scattering came from the experimental work of Campman and coworkers [59], who measured a series of GaAs/AlGaAs quantum well samples of various width. They also measured samples in which the well was not from a binary but the ternary systems AlGaAs and InGaAs. As shown in Fig. 4.9, they found Lorenzian lineshapes whose widths *decreased* with increasing quantum well width, with little or no effect on the alloy scattering. This is in contrast to what would be expected if the linewidth were limited only by the optical intersubband phonon emission, as the strength of the optical phonon scattering is predicted to *increase* for wider quantum wells. A very successful model for interface scattering has been developed by Unuma *et al*. [60], based on earlier consideration by Ando on two-dimensional systems [5]. This model, described in more detail in the next chapter, is able to predict correctly the absolute value and the temperature dependence of the broadening in quantum well [60] and quantum cascade laser systems [61,62]. **(p.64)**

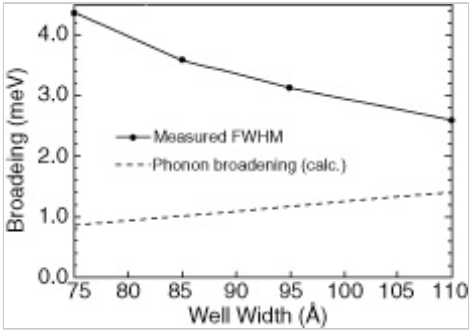


Fig. 4.9. Linewidth of the absorption of quantum wells of various widths. The dotted line shows the computed value of the broadening caused by the optical phonon scattering. The decrease in linewidth due to the phonon scattering as a function of quantum well width is a consequence of the increase in transferred momentum, and the overall increase is attributed to the increasing role of interface roughness scattering. Reprinted with permission from [59]. Copyright 1996, AIP.

4.9.3 Non-parabolicity and depolarization shift

Assuming that the non-parabolicity is treated in the two-band model, according to eqn. 3.1.26, the broadening ΔE is given by

$$\Delta E = E_F \frac{E_{21}}{E_G} \tag{4.9.55}$$

where E_F is the Fermi energy of the ground subband, E_{21} is the intersubband transition energy, and E_G is the bandgap. As an example, in a 15 nm thick InAs/AlSb quantum well ($E_G = 0.42eV$) with a 1–2 transition energy of 125 meV doped to a sheet density of $n_s = 1 \times 10^{12}$ cm^{-2} corresponding to a Fermi energy of about 60 meV, the predicted broadening from eqn. 4.9.55 is about 20 meV. The intersubband absorption was measured by Warburton *et al*. [63], and the result is shown in Fig. 4.10. Instead of the square-top lineshape that would be expected from the simple argument, a narrow line was observed. The authors explained this effect as the depolarization shift.

As discussed in a preceding paragraph, the depolarization shift arises from the screening of the incident field, preventing it from penetrating into the quantum-well system, and causing for a single transition a blueshift of the latter. Its effect is more subtle when a large collection of oscillators with different transition energies are interacting via the depolarization field. In essence, the oscillator strength of the transitions at lower energies is transferred to the one at the higher transition energy. As a result, the overall transition is not only blueshifted but is also narrower. This effect has been observed both when the broadening arises from non-parabolicity, as discussed here, as well as from inhomogeneous broadening [58].

This effect of the depolarization shift can be understood by considering eqn. 4.8.47. We could rewrite the conductivity, for an ensemble of oscillators with a distribution of transition energies ω_{12}, as

(p.65)

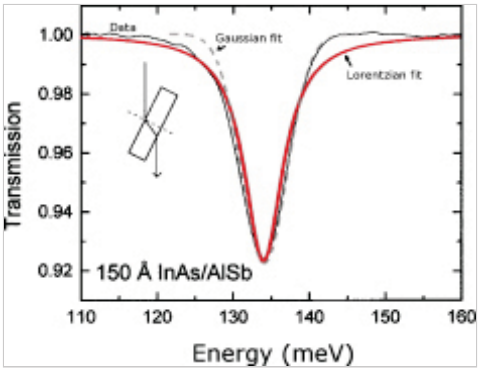


Fig. 4.10. Transmission through a 15 nm thick InAs/AlSb mult-quantum well system. Because of the large non-parabolicity of this system, a broad (about 20meV wide) peak would be expected, in contrast with the narrow experimental line observed. This narrowing was interpreted as a result of depolarization shift. (Adapted from [63].)

$$\tilde{\sigma}_{zz}(\omega) = \frac{\int d\omega_{12}\, \sigma_{zz}(\omega, \omega_{12})}{1 + \frac{4\pi i}{\epsilon^{r}_{ac} d_{eff} \omega} \int d\omega_{12}\, \sigma_{zz}(\omega, \omega_{12})}$$

(4.9.56)

For low densities, the denominator remains close to unity and the conductivity remains close to the average of an ensemble of oscillators. In the large density limit, however, the absorption peaks at the frequency ω where the denominator becomes close to zero. This effect is especially clear in a simple case where two absorption lines, assumed originating from the 1–2 and 2–3 transitions of a 22 nm thick GaAs square well, are interacting via the depolarization shift. A carrier density of $n_s = 3 \times 10^{11}$ cm^{-2} was assumed, leading to a

Fermi energy slightly above the $n = 2$ state, as shown in Fig. 4.11a, where the imaginary part of the dielectric constant and the real part of the effective conductivity $\tilde{\sigma}_{zz}$ are compared. The depolarization shift not only increases both transition energies by about 3 meV, but also enhances the strength of the high-energy transition at the expense of the lower-energy transition. This effect was observed experimentally in an InAs/AlSb highly doped quantum well [64].

In Fig 4.11b, the hypothetical case of a continuum of transition extending between a transition energy of 7 meV and 12 meV is assumed, as would occur if the broadening were originating from non-parabolicity. As observed experimentally in Fig. 4.10, the broadened line is blueshifted and narrowed by the depolarization shift.

4.10 Stark-tuning of intersubband absorption

The addition of a static electric field to the potential of the heterostructure will shift the levels. Using a perturbation expansion for the energy levels in a one-band model, the shift of the i th level ΔE_I can be written as

$$\Delta E_i(\mathscr{E}_z) = \langle \chi_i | - ez\mathscr{E}_z | \chi_i \rangle + \sum_{j \neq i} \frac{|\langle \chi_i | - ez\mathscr{E}_z | \chi_j \rangle|^2}{E_i - E_j}$$

(4.10.57)

(p.66)

as a function of the applied electric field $\mathscr{E}_z$ in the z-direction. Due to the lower effective mass of the electrons as compared to the bare mass, the wavefunctions exend over a larger distance, and the corrections to the energy position is usually much larger than in the case of atoms and molecules.

For a symmetric quantum well, the first-order correction in eqn. 4.10.57 vanishes by symmetry, and if one takes into account only the first two levels, the Stark shift for the 2–1 transition is written as

$$\Delta E_{21}(\mathscr{E}_z) = \frac{2e^2 z_{12}^2 \mathscr{E}_z^2}{E_{21}}$$

(4.10.58)

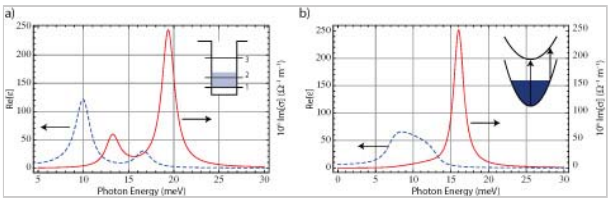


Fig. 4.11. Imaginary part of the dielectric constant (dashed line) and real part of the conductivity (solid line) showing the effect of the depolarization shift on an ensemble of oscillator. Case of a doped square well, with the Fermi energy above the first excited state. Note the transfer of oscillator strength to the 2–3 transition at the expense of the 1–2 transition. Case of a continuous distribution of transitions, as would arise in the case of a transition broadened by non-parabolicity at low temperature. Note the narrowing of the line, as was observed experimentally in Fig. 4.10.

where the shorthand notation $E_{21} = E_2 - E_1$ has been used. As shown in the above equation, the Stark shift can easily be as large as a few meV for a structure with a transition at 100 meV [65]. A very interesting feature is, however, that quantum-well structures need not to be symmetric, in which case the first-order term of the perturbation theory dominates and the Stark shift is now

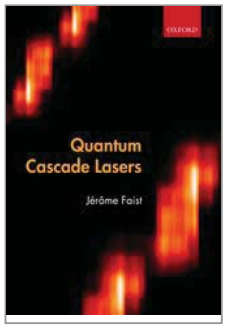
$$\Delta E_{21}(\mathscr{E}_z) = -e(z_{22} - z_{11})\mathscr{E}_z$$

(4.10.59)

where the term $z_{22} - z_{11}$ is the difference between the center of mass of the probability distribution between the upper and lower states. Using this effect, very large changes of the transition energy can be achieved [66]. Such tunability has been used in lasers [67, 68], modulators [69], and the tuning of non-linear optics structures [70].

Notes:

([1]) In particular, the broadening written in eqn. 4.9.54 assumes an infinite lifetime for the ground state as well as the same (uncorrelated) dephasing for both upper and lower states.



Quantum Cascade Lasers
Jérôme Faist

Print publication date: 2013
Print ISBN-13: 9780198528241
Published to Oxford Scholarship Online: May 2013
DOI: 10.1093/acprof:oso/9780198528241.001.0001

Intersubband scattering processes

Jérôme Faist

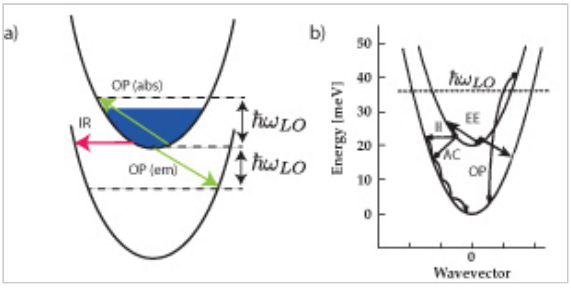
DOI:10.1093/acprof:oso/9780198528241.003.0005

Abstract and Keywords

This chapter analyzes intrasubband scattering processes, namely spontaneous emission, phonon scattering, and elastic scattering. It then discusses experimental measurements of intersubband lifetime using the following techniques: interband pump and probe, intersubband pump and probe, intersubband saturation experiments, and intersubband electroluminescence.

Keywords: spontaneous emission, phonon scattering, elastic scattering, intersubband lifetime

As shown in Fig. 5.1, an electron in an excited subband can recombine to the lower one through a variety of processes. It can change subband through spontaneous emission, inelastic scattering by a phonon, elastic scattering through an impurity or interface defect, or finally through electron–electron scattering. In stark contrast to the inter-band case, the radiative emission of photons is by far not the dominant intersubband mechanism even in a perfectly clean material. As shown schematically in Fig. 5.1, the



situation is very different if one considers two subbands that are spaced by an energylarger or smaller than the optical phonon energy. In the first case, for an electron sitting in the excited subband, optical phonon emission is always possible and leadsto lifetimes in the order of 1 ps. This process is usually the dominant scattering mechanism, as other processes that could compete with the

latter can be engineered away by correctly designing the structure, as discussed in Chapter 7. Another important process for short-wavelength lasers ($\lambda \leq 5\ \mu$m is the elastic scattering by interface roughness scattering.

In contrast, when the energy spacing between subbands is reduced below the optical phonon energy, scattering by the latter may be forbidden depending on the temperature and the electron densities, and the lifetime is now controlled by a competition **(p.68)** between optical phonons and a number of other processes, as shown in Fig. 5.1b. Typically, while optical phonon scattering by thermally excited electrons will always dominate at high temperatures ($T \rangle\ 60K$), electron–electron scattering or interface scattering will be the strongest scattering mechanisms at low temperature, depending on the structure and on the electron density. Only for extremely low densities, and for very clean systems, will the acoustic phonon scattering be of any significance.

Fig. 5.1. a) Schematic description of the relevant scattering mechanisms in the mid-infrared where the optical phonon scattering emission, absorption (abbreviated OP), and interface roughness (abbreviated (IR)) (for short-wavelength lasers) are dominant. b) Intersubband relaxations processes in a quantum well with an energy separation smaller than the optical phonon energy $E_{21} \langle \hbar\omega_{LO}$. In this configuration, various processes may play a role: optical phonon emission (OP), electron–electron scattering (EE), ionized-impurity scattering or interface roughness scattering (II, IR), acoustic-phonon emission (AC).

Intersubband and intrasubband scattering. The intersubband scattering rate will be evaluated to compute the population dynamics of the states and correspond to processes where an electron moves from one subband to the next. In general, for any given scattering Hamiltonian H_{scatt}, the scattering rate between the state with initial wavevector k_i in subband i to the available states in subband j will be evaluated using Fermi's golden rule:

$$\frac{1}{\tau_{ij}} = \frac{2\pi}{\hbar}\sum_{k_f}|\langle i,k_i|H_{scatt}|j,k_f\rangle|^2\delta(E(i,k_i)-E(j,k_f)-\delta E)$$

(5.0.1)

where δE is the energy exchanged during the scattering process, and as such is zero for elastic processes.

Intrasubband scattering, in contrast, corresponds to processes where the electron remains in its subband but changes its wavevector. The quantity of interest is the energy broadening Γ_{intra} it will induce to the corresponding intersubband transition. Following the approach of Ando [71], we write then

$$\Gamma_{intra}^{ij} = 2\pi\sum_{k_f}|(\langle i,k_i|H_{scatt}|i,k_f\rangle - \langle j,k_i|H_{scatt}|j,k_f\rangle|^2\delta(E_{i,j}(k_i)-E_{i,j}(k_f)-\delta E)$$

(5.0.2)

where the δ function specifies that the energy conservation must be held in both subbands. The evaluation of this δ function requires expansion of the square, and some extra care when taking non-parabolicity into account [72].

In an optical transition, the total broadening is then

$$\gamma = \frac{1}{2}\left(\Gamma_{intra}^{ij} + \Gamma_{inter}^{i} + \Gamma_{inter}^{j}\right)$$

(5.0.3)

Where

$$\Gamma_{inter}^{i,j} = \hbar \sum_{j'} \tau_{(i,j)j'}^{-1}$$

(5.0.4)

are the broadening arising from sum of the intersubband rates out of subbands *i* and *j*, respectively. In this approach, therefore, the dephasing time T_2 introduced in eqn. 4.9.54 can be computed, and is given by

$$T_2 = \frac{2\hbar}{\Gamma_{intra}}$$

(5.0.5)

(p.69) 5.1 Spontaneous emission

Between an initial state *i* and a final state *j* with a non-zero optical matrix element z_{ij}, spontaneous photon emission is possible and will occur with a rate given by

$$w_{sp} = \frac{e^2 n_{refr} z_{ij}^2 E_{ij}^3}{3\pi c^3 \epsilon_0 \hbar^4}$$

(5.1.6)

The energy dependence of the above equation can be better grasped when expressing it as a function of the oscillator strength between the two states:

$$w_{sp} = \frac{e^2}{6\pi m_0 c^3 \epsilon_0 \hbar^2} n_{refr} f_{ij} E_{ij}^2$$

(5.1.7)

as the oscillator strength will be of the order of $1/m^*$ for transition between the first two states in an infinite potential well $f_{12} = \frac{2^8}{\pi^2 3^3}/m^* \approx 0.96/m^*$ (using eqn. 4.4.20 for the dipole matrix element) and $f_{12} = 1/m^*$ for the parabolic well. As a result, the spontaneous emission rate is fundamentally proportional to the square of the photon energy. In Fig. 5.2 the spontaneous lifetime is plotted as function of transition energy

for a transition between the first two states in a $Ga_{0.47}In_{0.53}As/Al_{0.48}In_{0.52}As$ square well, assuming an oscillator strength $f_{12} = 23$. The lifetime varies between 60 ns at $\lambda = 10$ μm to 6 μs at $\lambda = 100$ μm. These long lifetimes are to be compared with the very short non-radiative lifetime observed in the same system, and explain the very low radiative efficiency observed in the spontaneous emission.

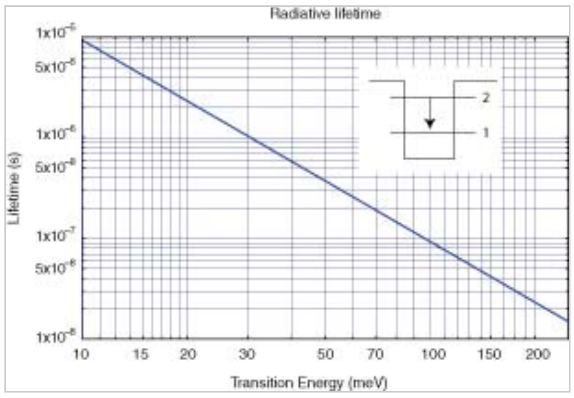


Fig. 5.2. Radiative lifetime of a square quantum well as a function of the transition energy. A constant value of the oscillator strenght of 23, corresponding

5.2 Phonon scattering

Lattice vibrations, phonons, perturb the strict periodicity of the

semiconductor crystal.For this reason the electronic states are no longer strict stationary eigenstates of the **(p.70)** Hamiltonian. In the basis of these eigenstates, phonons induce inter-level transitions and allow energy flow between the lattice and the electron system. Of all phonon scattering mechanisms, polar optical phonon scattering is the most efficient in III–V material systems.[1]

to an InGaAs/AlInAs quantum well, has been assumed

5.2.1 Optical phonons

The simplest treatement of optical phonon scattering in polar III–V semiconductors follows an approach first outlined by Price [73]. It is based on two important assumptions, usually well verified in III–V structures:

- The phonon scattering is dominated by the Frölich Hamiltonian, i.e. an interaction between the electron and the piezoelectric potential created by the local lattice deformation.
- The optical phonons have very little dispersion, and therefore can be assumed to be monoenergetic.

In this approach, an electron with an initial wavevector k_i and energy E_i in the upper subband scatters to the lower subband at k_f, therefore losing (or gaining, in the case of optical phonon absorption) an energy equal to $\pm \hbar\omega_{LO}$ in the process. The process is shown schematically in Fig. 5.3. The momentum Q exchanged in the process is

$$Q = \sqrt{k_i^2 + k_f^2 - 2k_i k_f \cos\theta}$$

(5.2.8)

The value of Q is on average larger for phonon absorption than for phonon emission because the final state is higher in the band. The scattering rate is then obtained by applying Fermi's golden rule on the envelope wavefunction. Almost all the treatments neglect the valence band part of the wavefunction (although the computation does not cause any significant difficulty) and treat the material as being a bulk material.

The scattering rate is then written as

$$w_{LO} = \frac{m^* e^2 \omega_{LO}}{2\hbar^2 \epsilon_P} \sum_f \int_0^{2\pi} d\theta \frac{I^{ij}(Q)}{Q}$$

(5.2.9)

where the effective dielectric constant entering the Frölich Hamiltonian is

$$\epsilon_P^{-1} = \epsilon_\infty^{-1} - \epsilon_s^{-1}$$

(5.2.10)

where ϵ_∞^{-1} and ϵ_s^{-1} are the mid-infrared and the static dielectric function, respectively.

The form factor entering in eqn. 5.2.9 is computed using the envelope wavefunctions:

$$I^{ij}(Q) = \int dz \int dz' \chi_i(z)\chi_j(z) e^{-Q|z-z'|} \chi_i(z')\chi_j(z')$$

(5.2.11)

The intersubband lifetime due to optical phonon emission, computed using eqn. 5.2.9, with zero initial wavevector, is plotted as a function transition energy for a squaresingle quantum well in Fig. 5.4. As shown in Fig. 5.4, a strong reduction of the lifetime **(p.71)**

is predicted when the two states are spaced resonantly with the optical phonon. Technically, this increase in scattering rate appears when Q approaches zero at resonance. Note, however, that the model used here, based on Fermi's golden rule, is not able to compute this value exactly at resonance. The global trend shown in Fig. 5.4 is rather robust, but will hold only for small ratios of the in-plane kinetic energy to the intersubband energy. The computed lifetime as a function of in-plane kinetic energy is shown in Fig. 5.5 for a square InGaAs/AlInAs, 220 Å thick quantum well. The strong increase of lifetime with initial energy can easily be understood by a study of eqn. 5.2.9 and Fig. 5.3. As the electron kinetic energy is increased, the average value of Q is increased as the integral in Eqn 5.2.9 runs on the circle of the available final k_f values. This increase in lifetime should be kept in mind when estimating population inversion in quantum cascade structures, as the electrons will exhibit a very large kinetic energy after the first optical phonon emission.
(p.72)

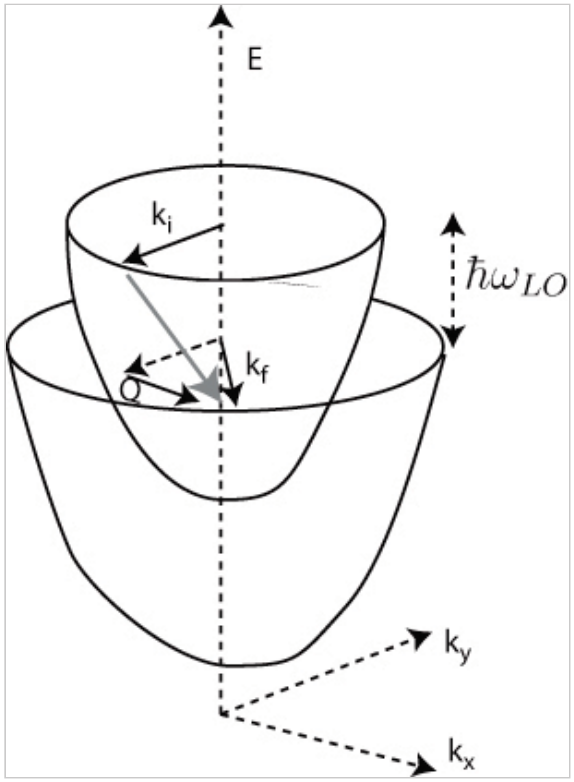


Fig. 5.3. Scattering by an optical phonon between two subbands. An electron with initial energy E_i and initial wavevector k_i is scattered to the lower subband at k_f, therefore losing (or gaining) an energy equal to $\hbar\omega_{LO}$ in the process.

Temperature dependence. The above computation was carried out for spontaneous emission of an optical phonon. When considering a system at finite temperature, absorption and stimulated emission of an optical phonon should be considered along with the spontanous emission factor. The phonon population n_{LO} is given by the Bose–Einstein factor:
(p.73)

$$n_{LO} = \frac{1}{\exp(\hbar\omega_{LO}/kT) - 1} \tag{5.2.12}$$

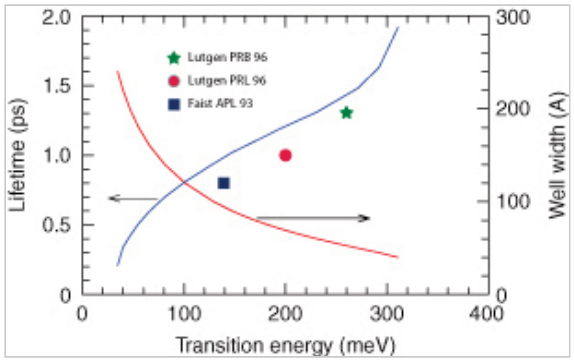


Fig. 5.4. Computed electron lifetime (left vertical axis) due to optical phonon scattering as a function of transition energy in a square quantum well, whose width is indicated in the rigth vertical axis. Zero

The total scattering rate is then obtained by separately adding the emission and absorption processes

initial kinetic energy in the upper subband, and zero temperature was assumed. Experimental measurements are reported from Faist APL 1993 [74], Lutgen PRL 96 [75], and Lutgen PRB 96 [76].

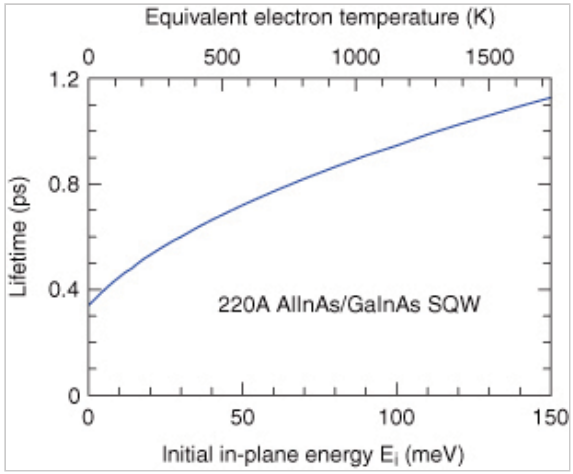


Fig. 5.5. Computed electron lifetime due to optical phonon scattering as a function of the initial kinetic energy. A square, 220 Å thick $Ga_{0.47}In_{0.53}As/Al_{0.48}In_{0.52}As$ quantum well was considered. The equivalent electron temperature is shown in the top horizontal axis

$$w_{tot} = (1 + n_{LO}) * w_{LO}^{(em)} + n_{LO} w_{LO}^{(abs)}$$

(5.2.13)

where $w_{LO}^{(em)}$ and $w_{LO}^{(abs)}$ are the optical phonon emission and absorption rate computed using eqn. 5.2.9. For transition energies larger than the optical phonon energy, as the phonon energy is larger than kT even at room temperature, the resultant dependence of the scattering time is relatively weak, as the lifetime decreases by about a factor of 2 when the temperature increases from zero to 400 K

.

5.2.2 Acoustic phonons

The computation of the acoustic phonon scattering rate is usually made using the same approximations (one sound velocity c_s, bulk phonons) as in the optical phonon case [73,77]. The computation is, however, more cumbersome as one has to integrate over the possible values of the phonon wavevector, assuming

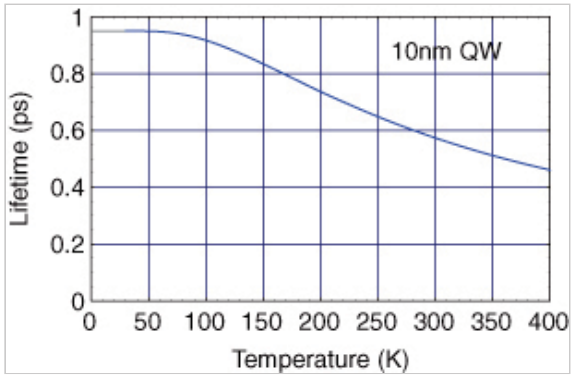


Fig. 5.6. Computed temperature dependence of the intersubband lifetime for a 100 Åthick

conservation of energy and in-plane momentum for the scattering process between the intial $\mathbf{k_i}$ and final $\mathbf{k_f}$ wavevector of the electron. The momentum of the phonon perpendicular to the interfaces is $\hbar q_{||}$. The angle between the initial and final wavevector is θ. The total scattering rate for the phonon emission $w_{if} = 1/\tau_{if}$ is then [77]

InGaAs/AlInAs quantum well

$$w_{if} = \frac{c_0}{4\pi^2\hbar}\int_{-\infty}^{\infty} dq_{||}|f(q_{||})|^2 \int_0^{2\pi} d\theta \int_0^{\infty} dk_f k_f \hbar\omega(1+n_{ac})\delta(E_i - E_f - \hbar\omega + (\hbar^2/2m^*)(k_i^2 - k_f^2))$$

(5.2.14)

(p.74) where the angular frequency of the phonon ω is obtained from the phonon wavevector $q_{||}$

$$\omega = c_s\sqrt{q_{||}^2 + |\mathbf{k_i} - \mathbf{k_f}|^2}$$

(5.2.15)

$$c_0 = D^2/2\rho c_s^2$$

(5.2.16)

and the form factor is computed

$$f(q_{||}) = \int_{-\infty}^{\infty} \chi_i(z) e^{-iq_{||}z}\chi_f(z)dz$$

(5.2.17)

D is the deformation potential for electrons, ρ the density, and c_s the longitudinal velocity of sound (8.6 eV, 5.3 g/cm^3 and 3700 m/s for GaAs respectively).

As in the case of the optical phonon, n_{ac} is the Bose–Einstein factor for the emission of acoustic phonons, given by

$$n_{ac} = \frac{1}{\exp(\hbar\omega/kT) - 1}$$

(5.2.18)

The expression for the scattering rate for the absorption of acoustic phonon is very similar to eqn. 5.2.14, replacing the term $(1 + n_{ac}) \rightarrow n_{ac}$ and changing the sign of the optical phonon energy $\hbar\omega \rightarrow -\hbar\omega$. The total scattering rate is then the sum of the scattering rates for emission and absorption. The computation is more cumbersome because the Bose–Einstein factor enters directly into the integral that must then be evaluated at each temperature individually.

The lifetime for the emission of an acoustic phonon increases with quantum well thickness, from a value of 80 ps for a 8 nm thick quantum well (E_{21} = 112 meV) to a value of 240 ps for a 20 nm thick quantum well (E_{21} = 29 meV) in the GaAsAl$_{0.3}$Ga$_{0.7}$As material system [77]. The lifetime increases further as the well width is increased, to a value of 400 ps for a 28 nm (E_{21} = 16 meV).

These long lifetimes have been very difficult to confirm experimentally. The main reason is that it is difficult to design an experiment that will measure the lifetime of a *single* electron in an excited subband and not the overall lifetime of a population. As is shown further in the text, electron–electron scattering will shorten the lifetime, as the electron density is above a few 10^7 cm^{-2}

5.3 Elastic scattering

Elastic scattering between subbands may occur whenever the translational invariance of the subband within the plane of the quantum well is lost. This may occur either through collision with an ionized impurity or through scattering at an interface step.

5.3.1 Impurity scattering

Ionised impurity scattering is the strongest scattering mechanism limiting the low-temperature mobility in clean samples based on "pure" (i.e. not alloyed) semiconductors. Its influence on the intersubband lifetime was evaluated by Ferreira and Bastard [77]. It was found that, as compared to other scattering mechanisms, a significant **(p.75)** scattering rate is found only for relatively large doping levels. In fact, in a 100 Å square well, a scattering time of 40 ps was computed for a relatively large background impurity level of $N_{imp} = 10^{16}$ cm^{-3}. Nevertheless, it was shown that doping the active region with doping levels $N_{imp} = 10^{17}$ cm^{-3} dramatically increased the linewidth of the intersubband electroluminescence [78]. Similarly, the initial quantum cascade lasers with superlattice active regions also suffered from this poor performances owing to the additional scattering mechanism introduced by the intentional doping of the active region [79]. Impurity scattering is also expected to play an important role in terahertz quantum cascade lasers.

5.3.2 Alloy scattering

Alloy disorder is another fluctuation that destroys the translational invariance of the potential seen by the electron in the plane of the layers. In fact, alloy scattering is found to be the dominant scattering mechanism limiting the mobility of high-purity samples of alloy semiconductors at low temperatures. The scattering rate between two subbands i and j in the presence of the potential of the alloy scattering V_{alloy}, describing the fluctuation from the average alloy potential caused by the individual atoms, can be written as [80,81]

$$\frac{1}{\tau_{alloy}}(\mathbf{k}_{\|}) = \frac{2\pi}{\hbar}\sum_{\mathbf{k}'_{\|}}\delta[\epsilon_{i,\mathbf{k}} - \epsilon_{j,\mathbf{k}'}]\,|\langle i,\mathbf{k}_{\|}|V_{alloy}|j,\mathbf{k}'_{\|}\rangle|^2$$

(5.3.19)

Assuming a random point contact potential and summing on the final states, one obtains finally

$$\frac{1}{\tau_{alloy}} = \frac{m^* x(1-x)\Omega_0|\langle\delta V\rangle|^2}{\hbar^3}\int_{alloy}|\chi_i(z)|^2|\chi_j(z)|^2 dz$$

(5.3.20)

where χ_i and χ_j are the envelope functions of the initial and final states, x is the alloy fraction (= 0.47 for $In_{0.53}Ga_{0.47}As$), Ω_0 is the volume of the unit cell, and $\langle\delta V\rangle$ is the spatial average of the fluctuating alloy potential over the alloy unit cell. The integration is meant to be carried over the region of the space where the alloy is present.

The alloy scattering has some interesting features. First of all, it is expected to be relatively weak in $GaAs/Al_{0.3}Ga_{0.7}As$ QCLs, as most of the wavefunction resides in the binary GaAs material. In contrast, it is expected to be much stronger for the lattice-matched $In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$/InP as the scattering is maximum for x=0.5 for a given ternary mixture. Moreover, it is also expected to *decrease* in strain-compensated $Ga_xIn1_{-x}As$/ $Al_yIn_{1-y}As$ as both x and y are departing from 0.5.

According to eqn. 5.3.20 the alloy scattering rate should increase weakly with energy. This dependence entirely derives from the integral over the wavefunction squares that have an overall

dimension of inverse distance. For this reason, a dependence of / the scattering rate in $\sqrt{E}$ is expected; as such, alloy scattering is increasingly relevant for short wavelength QCLs.

(p.76) 5.3.3 Interface roughness scattering

We showed in the previous chapter that interface roughness scattering is responsible in great part for the broadening of intersubband transitions in the mid-infrared. A powerful model describing the role of interface roughness as a scattering mechanism was introduced by Unuma *et al*. [60], based on ideas already developed by Ando [5]. This model assumes a statistical description of the interface that exhibits Gaussian fluctuations of step height Δ with a correlation length Λ, such that

$$\langle \Delta(\mathbf{r})\Delta(\mathbf{r}')\rangle = \Delta^2 \exp(-\frac{|\mathbf{r}-\mathbf{r}'|^2}{\Lambda^2})$$

(5.3.21)

where **r** = (*x, y*) is in the plane of the layer. The matrix element of scattering between an initial state in the subband *m*, at wavevector **k**' to subband *n* at wavevector **k** is

$$\langle m\mathbf{k}'|H_l|n\mathbf{k}\rangle = \int d^2r F_{mn}\Delta(\mathbf{r})e^{i\mathbf{q}\mathbf{r}}$$

(5.3.22)

where the form factors F_{mn} is proportional to the amplitude of the wavefunctions at the interface, and is defined for such an interface located at z_i as

$$F_{mn} = \delta U(z_i)\chi_m(z_i)\chi_n(z_i)$$

(5.3.23)

Equivalently, one may express F_{mn} as

$$F_{mn} = \sqrt{(\partial E_m/\partial L)(\partial E_n/\partial L)}$$

(5.3.24)

Intrasubband scattering. The total intrasubband scattering rate (neglecting screening) may be written as

$$\Gamma_{intra}^{IFR}(E) = \frac{m^*\Delta^2\Lambda^2}{\hbar^2}(F_{00}-F_{11})^2\int_0^{\pi} d\theta e^{-q^2\Lambda^2/4}$$

(5.3.25)

where the wavevector *q* exchanged during the process is expressed as

$$q^2 = 2k_i^2(1-\cos\theta)$$

(5.3.26)

Using this model, the authors were able to explain the difference between the lifetime deduced from the mobility and the one obtained from the linewidth of the intersubband absorption. Apart from an additional angular dependence, the key difference is that for the mobility the prefactor in the scattering rate is F_{00}^2, whereas it is $(F_{00}-F_{11})^2$ for the intersubband absorption and $|F_{11}\rangle$ $|F_{00}|$. Fig. 5.7 compares the measured and computed values of the linewidth for an intersubband transition in a GaAs quantum well as a function of temperature T.

This model was successfully extended to the case of many interfaces by Tsujino [61] by introducing a correlation coefficient $c_{i,j}$ between successive interfaces. For the pair of subbands

(mn), the broadening is now (in the limit of $q\Lambda \gg 1$)

$$\Gamma_{intra}^{IFR}(E) = \frac{\pi m^* \Delta^2 \Lambda^2}{\hbar^2} \sum_{i,j} c_{ij} \mathscr{F}_i^{(mn)} \mathscr{F}_j^{(mn)}$$

(5.3.27)

(p.77)

where the form factors $\mathscr{F}_i^{(mn)}$ and $\mathscr{F}_j^{(mn)}$ are expressed as a function of the amplitude of the wavefunction at the interface located a the position z_i by

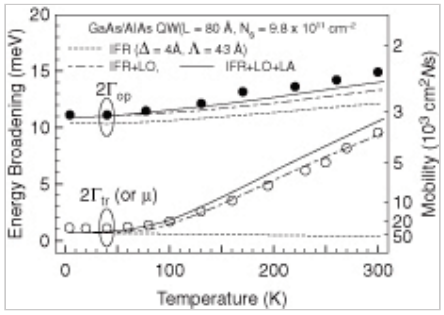


Fig. 5.7. Experimental measurement of the broadening of the intersubband transition and mobility as a function of temperature, compared with the model proposed by Unuma *et al*. [60]. Optical and acoustic phonon intra- and intersubband processes are also included in the model, as indicated.

$$\mathscr{F}_i^{(mn)} = (F_{mm} - F_{nn}) = \delta U(z_i)(|\chi_n(z_i)|^2 - |\chi_m(z_i)|^2)$$

(5.3.28)

Uncorrelated interfaces correspond to $c_{ij} = \delta_{ij}$. Using this model, the author was able to correctly predict the linewidth of various structures based on vertical and diagonal transitions and grown by various MBE reactors. Although the broadening is mostly driven by the intrasubband scattering, interface roughness is also predicted to induce intersubband scattering.

Intersubband scattering. This formalism can also be used directly to compute the contribution of this scattering phenomena to the intersubband lifetime, using the form factor F_{mn} for a transition from a subband m to n and a wavevector q given by

$$q^2 = k_i^2 + k_f^2 - 2k_i k_f \cos\theta$$

(5.3.29)

with k_i the initial wavevector of the electron in the subband. The final wavevector k_f can be expressed as a function of the wavevector $k_{isb} = \sqrt{2m^* E_{ij}}/\hbar$ corresponding to the intersubband transition as

$$k_f^2 = k_i^2 + k_{isb}^2$$

(5.3.30)

Because the correlation length Λ enters at the same time squared as a prefactor but also as an exponential cut-off in the computation of the scattering rate, the total lifetime has a non-trivial

dependence on this parameter, the initial wavevector, and intersubband transition energy. It is given by

$$\Gamma_{inter}^{IFR}(k_i) = \frac{m^* \Delta^2 \Lambda^2}{\hbar^2} F_{01}^2 \int_0^{\pi} d\theta e^{-q^2\Lambda^2/4}$$

(5.3.31)

To obtain an order-of-magnitude estimate of the lifetimes involved, the scattering rate was evaluated for a transition between the ground and first excited state at an angular frequency ω_{12} of a quantum well with infinite barriers:

(p.78)

$$\Gamma_{inter}^{IFR} = \frac{32}{27\pi\hbar} \Delta^2 \Lambda^2 m^{*2} \omega_{12}^3 e^{-\Lambda^2 k_i^2/2} e^{-\Lambda^2 k_{isb}^2/4} \int_0^{\infty} d\theta e^{-\frac{1}{2}\Lambda^2 k_i \sqrt{k_i^2 + k_{isb}^2}\cos\theta}$$

(5.3.32)

The result is plotted for two transition energies at various values of the correlation lengths in Fig. 5.8. Note that in contrast to the optical phonon scattering, and for typical values of the correlation length Λ, the scattering rate tends to grow with the transition energy, as shown by the ω_{12}^3 dependence of the scattering rate. This results, for short wavelength lasers ($\lambda \lesssim 5\ \mu$m), in a very strong dependence of the laser performance in the interface quality [82], as both step heights Δ and correlation length Λ are expected to be dependent on the growth parameters.

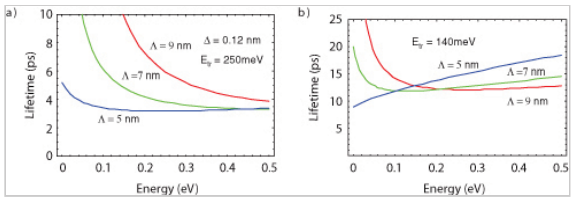


Fig. 5.8. Interface roughness scattering rate between two subbands as a function of the kinetic energy within the band for various values of the correlation length Λ, as indicated. The height of the interface step is $\Delta = 1.2 \times 10^{-10}\ m$. a) Transition energy $E_{tr} = 0.25eV$. b) Same computation, for a transition energy $E_{tr} = 0.14eV$.

As mentioned above, the relative strengths of the elastic and inelastic non-radiative channels are expected to be exchanged as a function of transition energy. This behavioris shown in Fig. 5.9, where the relative strengths of intersubband roughness scattering, alloy scattering, and optical phonon scattering are compared for a strain-compensated $Ga_{0.36}In_{0.64}As$ / $Al_{0.67}In_{0.33}As$ square quantum well as a function of the subband spacing. For the typical values of interface roughness parameter chosen (a correlation length Λ = 90 Å and a height of the interface step Δ = 1.2 Å), the interface roughness scattering becomes dominant even at 300 K for transition energies larger than 250 meV. Note the much faster increase of the interface roughness scattering rate as compared to the alloy scattering as a function of transition energy, as expected from the expressions 5.3.32 and 5.3.20.

This strong dependence of the performance of short wavelength quantum cascade lasers on the elastic scattering has been observed, using the MBE growth temperature as a way of changing the interface roughness. Fig. 5.10 shows the slope efficienciesand threshold current densities of a series of otherwise identical lasers, operating at$\lambda \approx 4.6\ \mu$m, and grown at different substrate temperatures. The large variation in device performances was interpreted as a result of the change

in the correlation length Λ of the interface roughness of the active region quantum wells [82] **(p.79)**

.

5.3.4 Electron–electron scattering

Electron–electron scattering (EE) is a process in which two electrons exchange energy and momentum. This process may also be accompanied by a change of the subband index, and is usually labeled by the index number of the intial and final electron states, i.e. process *klmn* is an interaction with carrier 1 scattering from sub-band *k* to *m* and carrier 2 from state *l* to *n*. The various processes are shown schematically in Fig. 5.11. An important feature of electron–electron scattering is the fact that *total energy and momentum are conserved*, and as a result this process can only thermalize a carrier population but not cool it.

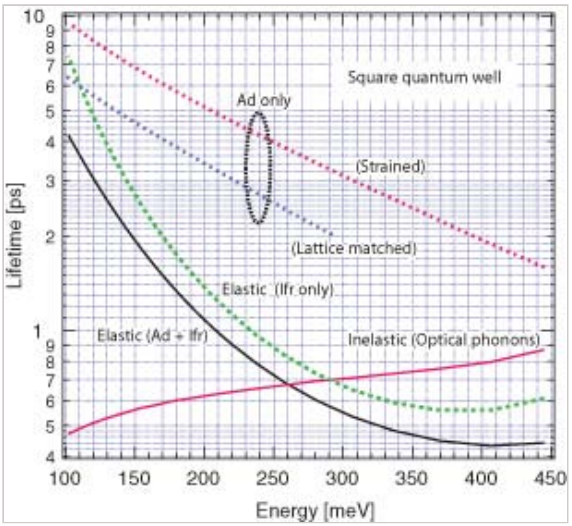


Fig. 5.9. Comparison between the elastic and inelastic non-radiative scattering channels for a single quantum well using a $Ga_{0.36}In_{0.64}As$ / $Al_{0.67}In_{0.33}As$ strain compensated material as a function of transition energy. For the interface roughness scattering (Ifr trace) a correlation length Λ = 90 Å and a height of the interface step Δ = 1.2 Å have been assumed. Note how the alloy scattering (Ad trace) is weaker in the strained material than in the lattice-matched one, shown here for comparison. Adapted from [72]

While processes 1212 and 2121 have no influence on the overall population of the subbands, processes 2211 or 2221 are responsible for decreasing the upper-state population. In summary, the influence of the various processes is

- 2222 or 1111 (scattering within a single subband) will help establish a thermal equilibrium *within a single subband.* In particular, for quantum cascade lasers with a large transition energy (200 meV) where electrons should emit a large number of optical phonons before reaching the bottom of the band, the strength of this process will determine whether the electron distribution will be a set of phonon replicas or a thermal one.

(p.80)

- Process 2121 or 1212 will tend to equilibrate the temperatures of the two sub-bands. Again, for large transition energies, this process will tend to heat the upper-state population, since the lower one is

heated by the large energy loss of the electrons.

• Process 2211 or 2221 will transfer electrons from the upper to the lower subbands. As shown by the computations presented later in this text, this process is important for low tansitions energies and medium densities.

(p.81) A theoretical treatement of electron–electron scattering between subbands was performed [83] in the limit of an infinite square well potential. Unfortunately, they focused their work on process 2221, which was shown later to be forbidden by parity in a symmetric structure [84]. Smet, Fonstad, and Hu gave a thorough treatement of electron–electron scattering and evaluated its importance as a function of temperature and intersubband transition energy [84]. They demonstrated that process 2211 is the dominant scattering mechanism. Central to the computation of the electron–electron scattering rate is the form factor for the transition:

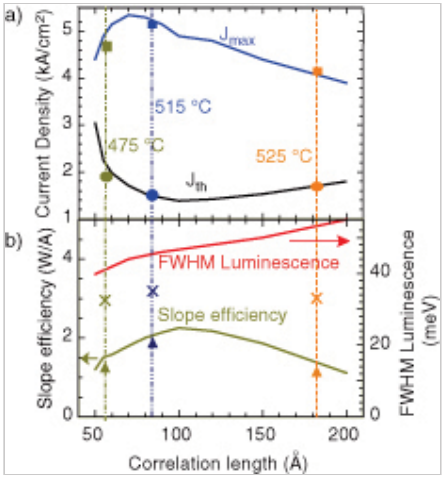


Fig. 5.10. Operation characteristics of a quantum cascade laser operating at $\lambda \approx 4.6\ \mu$m as a function of the correlation length of the interface. Solid lines are computed values using a full-density matrix theory, and symbols are experimental results. a) Maximum current density J_{max} and threshold current density J_{th}. b) Slope efficiency per facet and then full-width-athalf-maximum of the electroluminescence measured at a voltage of 14 V. The measured values for the samples grown at 475, 515 and 525°C, were compared with simulated values using the values of the correlation length as a fit parameter. Optimum performance is reached for $\Lambda \approx 100$ Å. The design is shown in Appendix A.3.2. Reprinted with permission from [82], copyright 2011, AIP

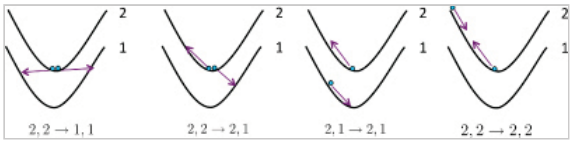


Fig. 5.11. Various electron–electron scattering processes. The 2, 2 → 1, 1 is the dominant electron–electron scattering term, as the Auger-like term 2, 2 → 2, 1 was shown to vanish in symmetric quantum wells. The 2, 1 → 2, 1 will lead to thermalization between subbands, while the 2, 2 → 2, 2 or 1, 1 → 1, 1 will tend to thermalize the subbands themselves.

$$A_{i,j\to f,g}(q_\perp) = \int_{-\infty}^{\infty} dz \int_{-\infty}^{\infty} dz' \chi_i(z)\chi_j(z) e^{-q_\perp |z-z'|} \chi_f(z')\chi_j(z') \tag{5.3.33}$$

where $q_\perp = |\mathbf{k}_i - \mathbf{k}_f|$ is the wavevector exchanged during the process. The transition rate for electron–electron scattering can then be evaluated using Fermi's golden rule, and yields the

following expression:

$$W_{i,j\to f,g}(\mathbf{k}_i) = \frac{e^4}{2\pi\hbar\epsilon_r^2\epsilon_0^2}\int d^2\mathbf{k}_j \int d^2\mathbf{k}_f \int d^2\mathbf{k}_g \frac{|A_{i,j\to f,g}(q_\perp)|^2}{\epsilon_{sc}^2(q_\perp,T)q_\perp^2} \times f_j(\mathbf{k}_j)(1-f_f(\mathbf{k}_f))(1-f_g(\mathbf{k}_g)) \times\delta(E_f(\mathbf{k}_f)+E_j(\mathbf{k}_j)-E_i(\mathbf{k}_i)-E_g(\mathbf{k}_g))\delta(\mathbf{k}_f+\mathbf{k}_g-\mathbf{k}_i-\mathbf{k}_j)$$

(5.3.34)

where ϵ_{sc}^2 is the screened dielectric function, and $f_i(\mathbf{k}_i)$ is the Fermi distribution. Even after reduction of eqn. 5.3.34 using the δ functions, the total rate will require a three-dimensional integral for each initial wavevector $\mathbf{k}_i$, causing a non-negligible computational burden. Computation of the scattering rate as a function of transition energy and density is shown in Fig. 5.12.

(p.82) In the same work the temperature dependence of the electron–electron scattering rate was also evaluated. The latter is virtually independent of temperature for reasonable physical parameters. The reason for this behavior is because the integral 5.3.34 is dominated by contributions in which there is no exchange of energy between the electrons, and for which

$$q_\perp^{elastic} = \sqrt{\frac{2m^* E_{21}}{\hbar^2}}$$

(5.3.35)

Quantum cascade lasers operate at very low densities, even in the terahertz region. As a consequence, the Fermi factor appearing in eqn. 5.3.34 can be assumed to be $f_{i,j,k,l} \ll 1$, and the rates are not expected to be limited by final state occupation. Another consequence of this fact is that the strength of the electron–electron scattering can be compared between structures by a mere computation of the form factor 5.3.33 at $q_\perp^{elastic}$. Note also that this form factor is the same (except computed for another wavevector) as the one for the optical phonon scattering.

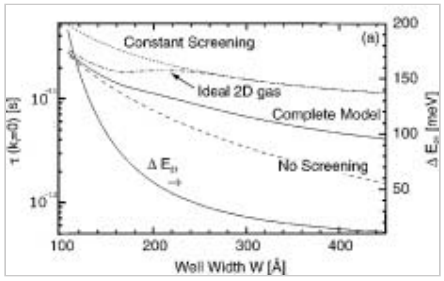


Fig. 5.12. The electron–electron scattering time for the 2, 2 → 1, 1 process presented as a function of well width for an infinite quantum well at zero temperature for different screening models: no screening, constant screening length, static single subband screening for an ideal two-dimensional system with no z-dependence (i.e. $|A_{2,2\to 1,1}| = 1$), and static single subband screening including $q_\perp$-dependence form factor. The Fermi energy of the excited state E_{f2} is equal to 10 meV and the initial electron state wavevector $\mathbf{k}_i$ is equal to zero. Adapted from Ref [84], with the correction suggested by [85], with permission.

A computation of the electron–electron scattering rate according to eqn. 5.3.34 is shown in Fig. 5.12 as a function of transition energy, for assuming various screening approximations. The vertical axis for the scattering rate was rescaled by a factor of 4 compared to the published data, following a suggestion by a subsequent work of the same group [85]. In this figure, the initial Fermi energy is assumed to be E_{f2} = 10meV, corresponding to a population $n_s = 2.6 \times 10^{11}$ cm^{-3}. Under this assumption, the computed scattering rate for a 28 nm thick quantum well, corresponding to a transition energy of 15 meV, is 8 ps. The upper-state lifetime is half of this value, as each scattering event will promote two electrons to the lower subband. However, as will

be shown later in this text, the actual density of the upper-state is closer to $n_s = 2 \times 10^9$ cm^{-3}—a value about 100 times smaller.

Hyldgaard and Wilkins [86] gave a simplified expression for the electron–electron scattering rate that, in addition to giving an order of magnitude estimate for the lifetime, has the advantage of showing the dependence of the scattering rate on the physical parameters. In their expression, the scattering rate $\Gamma_{ee} = \hbar/\tau_{ee}$ between two subband separated by an energy E_{21} is given by

$$\Gamma_{ee} = \frac{\hbar}{\tau_{ee}} = \frac{Ry^*}{\pi^2}\frac{\mu_2}{E_{21}}|U^0(q_{E_{21}})|^2 I_P(0)$$

(5.3.36)

where $U^0(qE_{21})$ is the unscreened interaction matrix element, and $I_P(0) \approx 0.785$ is a dimensionless integrated phase space factor. μ_2 is the chemical potential of the upper subband and is equal to $\mu_2 = \frac{\pi\hbar^2}{m^*}n_2$ at low temperatures, with n_2 the upper-state sheet density. The key result is the linear dependence of the scattering rate in the electron density, as well as its inverse dependence on the intersubband transition energy. Numerically, for a coupled well system, and assuming a Fermi energy of $E_{f2} = 5$ meV, their computed lifetime is $\tau = 0.94ps$ for a transition energy of about 10.6 meV. Assuming the linearity between scattering rate and density, this would extrapolate to a scattering rate of about 45 ps for a density of $n_s = 2 \times 10^9$ cm^{-3}. This value is not unreasonable for a square well at low temperatures.

From these computations it is quite clear that electron–electron scattering will play a significant role only for very clean structures at low temperatures **(p.83)** .

5.4 Comparison with experiments

The measurement of the intersubband scattering time τ_{ISB} is a very challenging task. It took a large number of experiments, spread over more than a decade, to reach a general consensus on its value. Although the rough estimate of the scattering time is not in question (τ_{ISB} 〈 1 ps) for transitions energies larger than the optical phonon, and about $\tau_{ISB} \approx 10$ ps at low temperature for transition energies below the optical phonon energy, there is still a lack of systematic measurement of its dependence with temperature, well width, electron density, and materials. The difficulty arises from the combination of short lifetime, strong dependence on the experimental conditions (temperature, density) and on the sample quality, especially for the low transition energies.

The intersubband lifetime was measured by a number of techniques that one may separate according to the different excitation and detection mechanisms used. As shown schematically in Fig. 5.13, these are:

- Interband pump and probe.
- Intersubband pump and probe.
- Intersubband saturation experiments.
- Intersubband electroluminescence.

Understanding the limitations of each of these techniques is very important when comparing the experimental results shown in Tables 5.1 and 5.2.

5.4.1 Interband pump and probe

In this approach an initial electron density is created in the excited state by a firstpump pulse. The population of the upper and lower levels is then probed by a weaker **(p.84)** probe pulse that follows with a variable delay. In general, this technique presents two key difficulties:

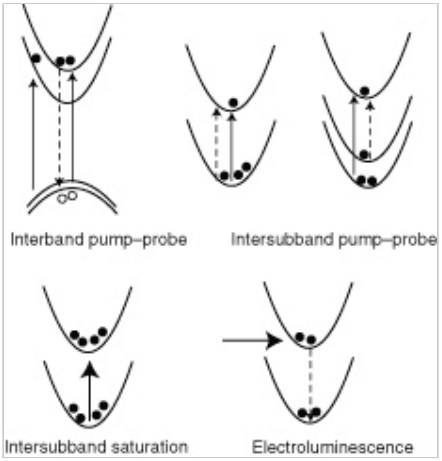


Fig. 5.13. Techniques for the measurement of the intersubband lifetime.

• Interband pumping on the second excited state will always create a concommitant population on the first subband at the same energy. This population will then interact with the upper-state population.
• Holes are also created and may affect the intersubband lifetime.
• The most popular optical source, the Ti-saphire mode-locked laser, has a repetition rate (about 80 MHz) such that the population does not have enough time to recombine totally across the gap between successive pulses. A parasitic population of the lower state is then induced. A possibility for preventing this problem is to use p-doped samples that exhibit much shorter electron radiative lifetimes (although p-doping will add additional scattering channels), or to sample the pulses with a lower repetition rate.

Compared with the early results, the generalized use of high-sensitivity detectors has allowed to greatly decrease the pump and probe pulse energies.

5.4.2 Intersubband pump and probe

In this technique a population is created in the upper-state by a pump photon tuned at the energy of the transition. The decay of the upper-state population is then measured by either a measure of the recovery of the fundamental absorption, or by the decay of the absorption from the excited state. This technique is free from many of the problem cited above and yielded the most reliable results. It is, however, a technique that is not easy to implement, as it requires the generation and detection of short pulses in the mid- or far-infrared. This was accomplished by either downconversion from the visible pulses or by direct generation by a free electron laser.

5.4.3 Intersubband saturation experiments

A simpler technique consists in measuring the intensity-dependent absorption. In a simple two-level approximation, the latter is written as

$$\alpha(I) = \frac{\alpha(0)}{1 + \frac{I}{I_{sat}}}$$

(5.4.37)

Fitting the data will yield the saturation intensity I_{sat}, which in turns will yield the lifetime T_1:

$$I_{sat} = \frac{\hbar^2 c\epsilon_0 n_{refr}}{2q_0^2 z_{12}^2 T_1 T_2}$$

(5.4.38)

where a Lorenzian line with dephasing time T_2 is assumed. The latter is usually fitted from the measurement of the lineshape for small pumping intensities. As shown by the preceding equation, a drawback of this technique, besides the strong pump powers that it requires, is the requirement to evaluate correctly the power density inside the sample as well as a knowledge of the dipole matrix element z_{12} **(p.85)** .

5.4.4 Experimental results

Because of a threshold for the onset of optical phonon emission, it was expected that a very strong contrast would be observed when measuring the intersubband lifetime below and above the optical phonon energy. Indeed, the first measurements were the time-resolved Raman experiments of Oberli *et al*. [87], which yielded, for a transition energy slightly below the optical phonon energy (28meV), a lifetime of about 300 ps, while the intersubband pump and probe measurements of Seilmeier yielded a lifetime of 10 ps for a transition energy of 150 meV, clearly above the optical phonon emission threshold. However, it was found out that both measurements did not yield the true intersubband lifetime but rather the global relaxation time of the hot electron plasma distributed thermally across the two transitions. When large populations are excited, because of Pauli blocking, the electrons could not scatter back to the ground subband because of the lack of available states.

In fact, time-resolved Raman experiments by Tatham *et al*. [88] yielded an intersubband lifetime of τ 〈 1 ps for a 14.6 nm thick GaAs active quantum well. Later measurements using various techniques performed at low excitation levels agreed with the values computed using optical phonon emission, as shown in Fig. 5.4. These values are summarized in Table 5.1.

5.1 Summary of lifetime measurements for transition energies above the optical phonon energy in GaAs quantum Early measurements were inaccurate due to the strong pumping needed. ISB: intersubband, IB: interband. Only mperature measurements (T = 10 K) are included

	E_{12} (meV)	τ (ps)	Technique	ref	Yea
SQW	150	10	ISB pump probe	[89]	198
SQW	150	0.9	ISB Saturation	[90, 91]	198
SQW	53	〈 1	IB Time Res. Raman	[88]	198
SQW	125	0.65	IB lum. and transport	[92]	199
SQW	112	0.65	ISB cw pump probe	[74]	199
As SQW	129	0.8	ISB cw pump probe	[74]	199
SQW	175	0.65, 300K	ISB/IB pump probe	[93]	200
As SQW	200	1	ISB/IB pump probe	[75]	199
As SQW	260	1.3	ISB/IB pump probe	[76]	199

In contrast, the lifetime for transition energies below the optical phonon energy, if limited by the emission of acoustic phonons, would lead to very long lifetimes of the order of more than 100 ps, raising the hope of intersubband lasers with very low threshold current densities.

Indeed, subsequent work [94,95] found scattering times spread from 20 to 750 ps. Thanks to the improvement of the detection techniques, the detection of the luminescence signal was possible even for extremely low excitation densities. In a series of experiments, Hartig *et al*. [96] showed that the intersubband lifetime first decreases with decreasing excitation density, as Pauli blocking effects are reduced. The shortest value measured of 5 ps, limited by the temporal resolution of their experiment, is achieved for densities around 10^{11} cm^{-2}. As shown in Fig. 5.14, the lifetime then increases again **(p.86)** when the density is further lowered and saturates at about 30 ps. These measurements, however, were performed in a coupled well system for technical reasons. Another aproach has been intersubband measurements with quasi-continuous-wave excitationwith a CO_2 [98] or free electron laser [99]; detection was either by intersubband absorptio [98] or capacitive measurements [99]. Both groups used modulation-doped coupled quantum wells. In the low-temperature, low-excitation limit, the measured values of lifetime range from $\tau = 300ps$ [98] to 1ns [99, 100]

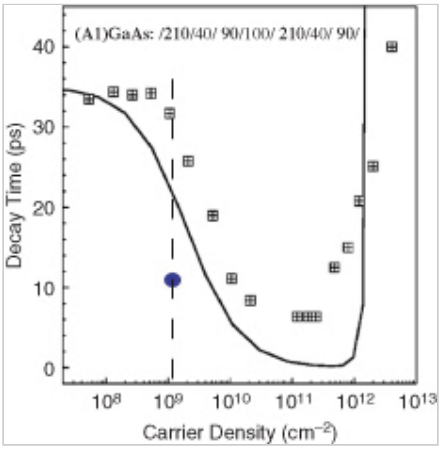


Fig. 5.14. Lifetime as a function of pumping density, measured by intersubband pump–probe spectroscopy in a coupled quantum well with an energy separation smaller than the optical phonon energy. The solid line is the result of a Monte Carlo computation. Reprinted with permission from [96]. Copyright 1999, APS. The solid disk is a measurement deduced from an electroluminescence experiment in a single-quantum-well terahertz QCL [97]

5.2 Summary of some lifetime measurements for transition energies below the optical phonon energy.

meV)	τ (ps)	Technique	Ref	Year
	325	Time-resolved PL	[87]	1987
	40	ISB pump–probe	[95]	1994
	300	Continuous wave ISB pump–probe	[98]	1994
	1200	Intersubband pump–probe	[99]	1995
	30	Interband pump–probe	[101]	1998

.

5.4.5 Intersubband electroluminescence

Pump and probe as well as saturation measurements require relatively strong pumpintensities. Another drawback is that the measurement will be performed in a model **(p.87)** system, usually a single or coupled quantum well. In contrast, quantum cascade lasers will operate in the regime of very low upper-state densities (in general with a sheet density of about $n_s \approx 10^9$ cm^{-2}) and with an active region consisting of a complex arrangements of coupled wells. For this reason, a measurement of the intersubband luminescence intensity, although by itself a very indirect measurement of the upper-state lifetime, proved a very useful technique because it directly includes some important additional parameters such as the injection efficiency in the upper-state. In particular, it is especially valuable to measure the *variation* of the lifetime with an external parameter such as the electron density, temperature, or material.

For example, the initial measurements performed in the first cascade structure [102] showed that the electroluminescence intensity decreased with temperature in agreement with an intersubband lifetime dominated by optical phonon scattering. In another set of experiments, Benveniste *et al*. compared the luminescence intensities from three similar quantum cascade structures grown from the InAs/AlSb, InGaAs/AlInAs and GaAs/AlGaAs material system [103]. The results, shown in Fig. 5.15, confirms that the lower electron effective mass of the In-containing compounds will yield larger efficiencies. In fact, for an intersubband lifetime limited by optical phonon emission, the luminescence efficiency should scale approximately with $((m^*)^{-3/2})$. As shown in Fig 5.16, this dependence is verified experimentally.

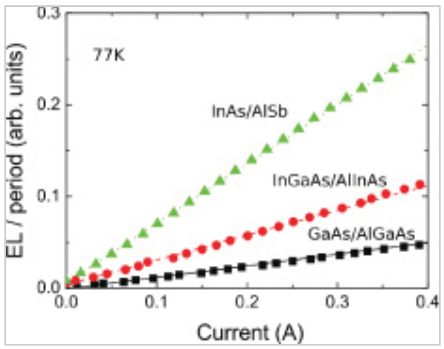


Fig. 5.15. Comparison of the normalized light–current characteristics in pulsed operation for InAs/AlSb, GaInAs/AlInAs, and GaAs/AlGaAs cascade structure. The low electron effective mass, responsible for the larger dipole matrix element, as well as the lower optical phonon scattering rate, are the reasons for the increasing efficiency with In content in the quantum well layers. Reprinted with permission from [103]. Copyright 2008, AIP.

As discussed in Chapter 7, the differential gain of a quantum cascade structure should theoretically be directly proportional to the electroluminescence quantum efficiency, provided that the line broadening remains constant, as it was indeed observed experimentally. Based on these considerations, InAs-based quantum cascadelaser should clearly outperform the other material systems. However, laser performances also depend critically on the waveguide losses as well as on the thermal conductivity of the laser structure—two criteria where the InAs-based lasers are at a clear disadvantage compared to the InGaAs/AlInAs/InP material system **(p.88)**

.

5.4.6 Magneto spectroscopy

Lifetime measurements do not, in general, provide direct information of the nature of the scattering mechanism. This information is obtained only when an additional parameter (temperature, density) can be freely adjusted while the measurement is performed. A very powerful technique is to apply a very strong magnetic field perpendicular to the layer. In this case, and considering a one-band model for simplicity, the Schrödinger equation for the envelope function in the Laudau gauge $\varphi k(x, z)$ is

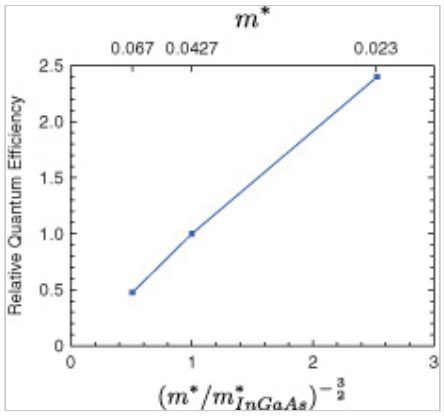


Fig. 5.16. Experimental quantum efficiency, relative to the lattice matched InGaAs/AlInAs material, compared with the $(m^*)^{-3/2}$ dependence. Data is from [103].

$$[-\frac{\hbar^2}{2m^*}\frac{\partial^2}{\partial z^2} - \frac{\hbar^2}{2m^*}\frac{\partial^2}{\partial x^2} + \frac{1}{2m^*}(\hbar k_y + eB_\perp x)^2 + V_{conf}(z)]\varphi_k(x,z) = E\varphi_k(x,z)$$

(5.4.39)

The wavefunction $\varphi k(x, z)$ can be factorized into an in-plane and an out-of-plane motion, such that $\varphi k(x, z) = \varphi_n(x) \cdot \chi_m(z)$. The equation for $\chi_m(z)$ is identical to that of the envelope function at $k = 0$ at zero field. The motion along x for $\varphi_n(x)$ can be recast into an harmonic oscillator with an energy spectrum consisting of equispaced Landau levels [104]

$$E_n^{(m)}(k_y) = \left(n + \frac{1}{2}\right)\hbar\omega_c + E_0^{(m)}; n = 0, 1, 2\ldots$$

(5.4.40)

where E_0^m is the energy of the m th subband at k = 0 in the absence of magnetic field, and the cyclotron frequency is given by $\omega_c = \frac{eB_\perp}{m^*}$. For each subband, the energy spectrum then consists of a ladder of equidistant Landau states. By tuning the magnetic field, the energy separation between these states can be continuously tuned.

Because the energy is now quantized, the intersubband scattering is now strongly constrained, especially since the key scattering mechanisms in quantum cascade lasers, optical phonon emission and elastic scattering, each require a fixed energy exchange. As described schematically in Fig. 5.17b, for well-resolved Landau levels the scattering is strongly suppressed, except when the condition

(p.89)

$$E_{21} = N\hbar\omega_c + \hbar\omega_{LO}$$

(5.4.41)

is set for a resonance with the optical phonon emission, or

$$E_{21} = N\hbar\omega_c$$

(5.4.42)
for an elastic resonance. Each time such a resonance occurs, the scattering rate betweenthe subband is enhanced, leading to a local maximum of the injected current and a minimum of the light output. As a result, the measurement of the current and lightoutput as a function of magnetic field B enables not only the measurement of E_{21} but also of the strength of the individual scattering mechanisms.

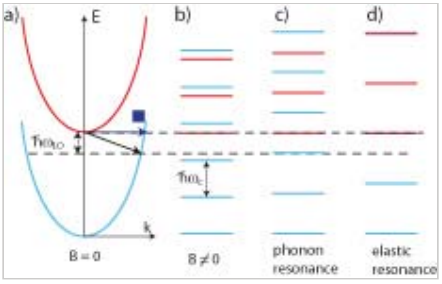


Fig. 5.17. Energy levels in a quantum well with two subbands. a) B=0, the states are characterized by a parabolic dispersion. b) When a strong perpendicular magnetic field is turned on, the subband splits into Landau levels. c) Optical phonon resonance: $E_{12} = N\,\hbar\omega_c + \hbar\omega_{LO}$ d) Elastic resonance $E_{12} = N\,\hbar\omega_c$.

Such measurements were initially performed in terahertz quantum cascade lasers [106, 107], as mid-infrared structures required a much stronger magnetic field that can only be provided by a resistive magnet. In contrast with the measurements ofterahertz quantum cascade lasers, performed at cryogenic temperatures, and where the elastic scattering was observed to dominate the scattering processes, the initial measurements on mid-infrared GaAs/$Al_{0.3}Ga_{0.7}As$ showed a dominant contribution of the LO optical phonon scattering [108,109]. Further measurements of GaAs/$Al_{0.3}Ga_{0.7}As$ lasers showed, however, small features corresponding to elastic resonances, attributed mainly to interface roughness scattering [110].

Measurements of a QCL laser operating at $\lambda \approx 8.7\ \mu$m and based on the $Ga_{0.47}In_{0.53}As/Al_{0.48}In_{0.52}As$ material system showed, however, a significantly different picture [105]. As shown in Fig. 5.18, the threshold current and optical scattering trace cannot be explained in a satisfactory manner without the presence of the elastic scattering channel that was shown to be 40% of the total scattering rate at the measurement temperature of 77 K. The computation of the different process showed that at low temperature, alloy scattering lifetime did correspond to a lifetime of 2.7 ps, while the interface roughness scattering corresponds to a lifetime of 5.3 ps. By itself, these results, although demonstrating clearly the importance of elastic scattering in InGaAs-based QCL, did not, however, contradict the general consensus attributingto LO phonon scattering the strongest non-radiative channel, as the ratio of elastic **(p.90)**

to inelastic scattering is expected to drop to 20% at 300 K. The situation is, however, expected to change for short-wavelength lasers, as the optical phonon and elasticscatterings are following opposite trends, the first one decreasing and the second one increasing as a function of transition energy.

Notes:

([1]) In contrast, scattering in non-polar materials such as group IV

(Ge, Si) semiconductors is dominated by deformation potential.

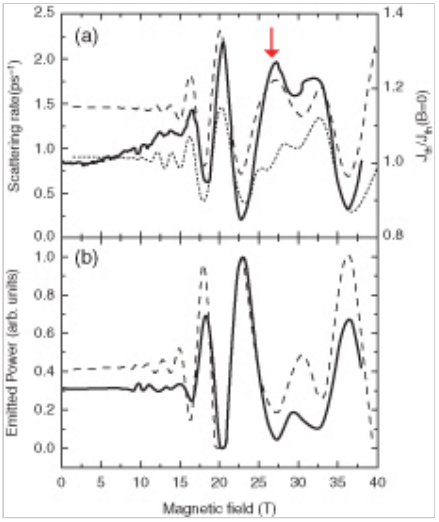
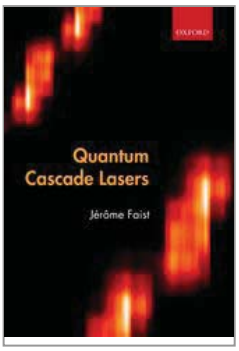


Fig. 5.18. a) Normalized threshold current density $J_{th}(B)/J_{th}(B = 0)$ (solid line) and scattering rates $1/\tau_3$ as a function of the magnetic field. The dotted curve takes into account the LO-phonon emission; the dashed curve also includes the elastic contribution (alloy and interface roughness scattering). The arrow points at the feature at 27 *T* that can be explained only by elastic scattering. b) Measured (solid line) and calculated (dashed line) emitted power at I/I_{th} = 1.25 as functions of B. Reprinted with permission from [105]. Copyright 2006, AIP.



Quantum Cascade Lasers
Jérôme Faist

Print publication date: 2013
Print ISBN-13: 9780198528241
Published to Oxford Scholarship Online: May 2013
DOI: 10.1093/acprof:oso/9780198528241.001.0001

Mid-infrared waveguides

Jérôme Faist

DOI:10.1093/acprof:oso/9780198528241.003.0006

Abstract and Keywords

In a quantum cascade laser, the light is confined both vertically, by the sequence of the epitaxial layers, as well as horizontally, by the ridge structure that also confines laterally the current. Because the epitaxial growth allows the tighter control of layers thicknesses as compared to the lateral one, optical confinement is usually stronger in the vertical direction. As a result, the simplest model of the one-dimensional dielectric slab model correctly predicts the propagation index and overlap factor of the waveguide, and is usually a good starting approximation for a more complete computation. This chapter discusses the dielectric slab waveguide, interface plasmon mode, optical properties of doped layers, two-dimensional confinement, large optical waveguides, and thermal properties.

Keywords: quantum cascade lasers, dielectric slab waveguide, interface plasmon mode, doped layers, two-dimensional confinement, large optical waveguides, thermal properties

The very short upper-state lifetime of intersubband transitions means, as a consequence, that optical gain is achieved only at the cost of a very large power dissipation in the material. For a typical operation condition of a current density of 3kA/cm^2 at an applied electric field of 100 kV/cm, the power dissipated per unit volume is 300 MW/cm^3, a factor of 10–100 larger than for interband devices and 10^8 larger than for a Nd YAG solid-state laser. As a result, the use of laser cavities based on optical waveguides is essential to keep the total electrical dissipation to a minimum.

Waveguiding can be achieved fundamentally using either metallic or dielectric confinement or a combination of both. In general, dielectric confinement has the advantage that it is essentially loss-free, whereas the use of metals enables the confinement of the light in regions much smaller than the wavelength. Fortunately, the same semiconductor materials used to achieve the quantum confinement of the electrons also exhibit, because of their different bandgaps, a concomitant difference in their refractive index that can be used to guide the light. Compared to interband devices, the longer wavelength in mid-infrared quantum cascade lasers limits the relative thickness of the cladding layers that can be used before creating growth problems. For example, the slight lattice mismatch between $Al_{0.45}Ga_{1-x}As$ and GaAs constrains the use of dielectric cladding for GaAs-based quantum cascade lasers.

In a quantum cascade laser, the light is confined vertically, by the sequence of the epitaxial layers, as well as horizontally, by the ridge structure that also confines laterally the current. Because the epitaxial growth allows the tighter control of layer thicknesses as compared to the lateral ridge width, optical confinement is usually stronger in the vertical direction. As a result, the simplest model of the one-dimensional dielectric slab model already predicts correctly the propagation index and overlap factor of the waveguide, and is usually a good starting approximation for a more complete computation.

6.1 Dielectric slab waveguide

As shown in Fig. 6.1, in the simplest dielectric optical waveguide a slab of dielectric material (core) with refractive index n_1 is sandwiched between two semi-infinite lower index cladding layers satisfying $n_1 \rangle (n_2, n_3)$. In general, the thickness of the guiding layer is of the order of a wavelength. The electromagnetic treatment of such a problem is relatively easy due to the homogeneity of the medium in each segment. We consider now the propagation of monochromatic radiation at frequency ω following Maxwell's equation for a dielectric, source-free medium:

(p.92)

$$\nabla \times \mathbf{H} = i\omega\epsilon_0 n_i^2 \mathbf{E} \tag{6.1.1}$$

$$\nabla \times \mathbf{E} = -i\omega\mu_0 \mathbf{H} \tag{6.1.2}$$

where n_i are the refractive indices and μ_0 the vacuum permittivity. As the refractive index is assumed to be constant (or even slowly varying), eliminating one of the two **H** or **E** fields, we obtain the Helmholtz wave equation for each field **E** and **H** valid in each region:

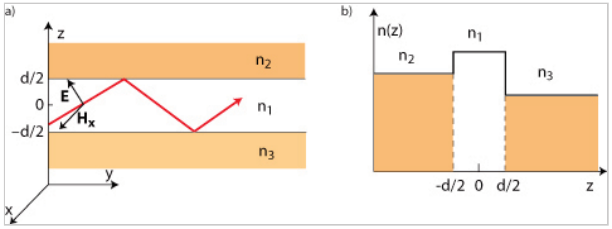


Fig. 6.1. a) Schematic drawing of a slab dielectric waveguide, showing the axes as well as the **E** and **H** fields for a TM polarized wave. b) Schematic refractive index profile as a function of distance.

$$\left(\nabla^2 + \frac{n_i^2 \omega^2}{c^2}\right)(\mathbf{H}, \mathbf{E}) = 0$$

(6.1.3)

For a refractive index constant across each layer, the solution can be constructed using, in each layer, plane waves propagating in the z-y plane. The phase factor in the y-direction is assumed to be common to all layers. We have then for the **E** and **H** fields:

$$\mathbf{E} = \mathbf{E}_m e^{-ik_z z} e^{i(\omega t - \beta_m y)}$$

(6.1.4)

$$\mathbf{H} = \mathbf{H}_m e^{-ik_z z} e^{i(\omega t - \beta_m y)}$$

(6.1.5)

where β_m is defined as the propagation constant of the mode with index m. From the Helmoltz equation 6.1.3, the wavevector in the z-direction is given by solving the eqn.

$$\beta^2 + k_z^2 = \frac{\omega^2 n_i^2}{c^2}$$

(6.1.6)

We will look for values of β where k_z is real in the guiding layer and imaginary in the claddings, i.e. satisfying

$$k_0 n_1 \rangle \beta \rangle (k_0 n_2, k_0 n_3)$$

(6.1.7)

with $k_0 = \omega/c$, as they will correspond to a guided wave that decays exponentially in the cladding layers.

Two polarization directions are possible—either transverse electric (TE) or transverse magnetic (TM) in which the electric and respectively magnetic field vector is **(p.93)** oriented along the y-direction. The polarization direction will select the continuity of the relevant fields at the interfaces:

$$\text{TE} : \mathbf{H_y}, \mathbf{E_x}$$

(6.1.8)

$$\text{TM} : \mathbf{H_x}, \mathbf{E_y}$$

(6.1.9)

Using the relations between **H** and **E** fields given by eqn. 6.1.1, one shows that these boundary conditions correspond to the continuity of the following fields:

$$\text{TE} : \mathbf{E_y}, \frac{\partial \mathbf{E_y}}{\partial z}$$

(6.1.10)

$$\mathrm{TM}: \mathbf{H_y}, \frac{1}{n_i^2}\frac{\partial \mathbf{H_y}}{\partial z}$$

(6.1.11)

In particular, the boundary conditions for the **H** field of the TM wave correspond to a continuity of

$$n_i^2 \mathbf{E_z}, \frac{\partial \mathbf{E_z}}{\partial z}$$

(6.1.12)

that we can also recognize as a continuity of the displacement field $\mathbf{D} = \epsilon\,\mathbf{E}$ and of the derivative of $\mathbf{E_z}$. As the refractive index is larger in the core of the waveguide, this boundary condition will force a lower value of $|\mathbf{E}|$ at that location. As a result, the TM modes are in general less confined than TE modes.

As discussed in Chapter 4, the polarization selection rule of intersubband transitions allow gain only for TM modes. A standard procedure is to compute first the **H** field, and to then extract the electric field from

$$E_z = \frac{i}{\omega n_i^2}\frac{\partial}{\partial y}H_x$$

(6.1.13)

$$E_y = \frac{-i}{\omega n_i^2}\frac{\partial}{\partial z}H_x$$

(6.1.14)

6.1.1 Symmetric waveguide

A simple solution exists for the case of the symmetric waveguide (i.e. $n_3 = n_2$). In that case, the field $H_m(z)$ can be written as

$$H_m(z) = \begin{array}{ll} A\sin(hz) + B\cos(hz) & \frac{1}{2}d\rangle z\rangle -\frac{1}{2}d \\ Ce^{-qz} & z\rangle \frac{1}{2}d \\ De^{qz} & z\langle -\frac{1}{2}d \end{array}$$

(6.1.15)

where $h = \sqrt{(\frac{n_2\omega}{c})^2 - \beta^2}$ and $q = \sqrt{\beta^2 - (\frac{n_1\omega}{c})^2}$. Applying the boundary conditions of the fields and their weighted derivative (we consider the TM polarization) at the interfaces ($z = \pm d/2$) yields a set of equations:

$$\begin{aligned} A\sin\left(h\tfrac{d}{2}\right) + B\cos\left(h\tfrac{d}{2}\right) &= Ce^{-q\frac{d}{2}} \\ -A\sin\left(h\tfrac{d}{2}\right) + B\cos\left(h\tfrac{d}{2}\right) &= De^{q\frac{d}{2}} \\ Ah\cos\left(h\tfrac{d}{2}\right) - Bh\sin\left(h\tfrac{d}{2}\right) &= -\bar{q}Ce^{-q\frac{d}{2}} \\ Ah\cos\left(h\tfrac{d}{2}\right) + Bh\sin\left(h\tfrac{d}{2}\right) &= \bar{q}De^{q\frac{d}{2}} \end{aligned}$$

(6.1.16)
(p.94) where $\bar{q} = qn_1^2/n_2^2$. Eliminating the coefficients A, B, C, and D leads to the transcendental equation

$$\tan(hd) = \frac{2h\bar{q}}{h^2 - \bar{q}^2}$$

(6.1.17)
which yields the possible values of the propagation constant for the mode β_m. The *effective* refractive index for the mode is defined as

$$n_{eff} = \frac{\beta}{\omega/c}$$

(6.1.18)
For confined modes, its value should lie between the refractive index of the core and of the cladding $n_1 < n_{eff} < n_2$.

6.1.2 Generalization

A nice feature of the model presented above is that it leads itself easily to a generalization of an arbitrary large number of layers with *complex* dielectric constants ϵ_i. In each layer, a pair of plane waves are assumed to propagate, with a propagation constant in the x-direction k_z given by the (complex) roots of the equation:

$$k_z^2 = \epsilon_i \left(\frac{\omega}{c}\right)^2 - \beta^2$$

(6.1.19)
Depending on the polarization direction (TE or TM), the appropriate boundary conditions are applied at the interface, given by eqn. 6.1.10.

The loss of the waveguide can be computed by seeking a solution of β in the complex plane, and extracting the loss as

$$\alpha_w = 2\Im m(\beta)$$

(6.1.20)
In this approach, the solution is a wave that has the same decay constant in the propagation direction.

In a laser, this optical loss is compensated by gain in the active region. The effect of the gain on the propagation constant can be evaluated using a perturbation approach to the Helmoltz equation, seen as an eigenvalue problem for β^2:

$$\left(\nabla^2 + \frac{n_i^2\omega^2}{c^2}\right)\mathbf{E} = \beta^2\mathbf{E}$$

(6.1.21)
The effect of the small change of the (imaginary) part of the refractive index caused by the material gain g_m is evaluated on the propagation constant using a first-order perturbation

expansion,

$$\Delta(\beta^2) = \langle \mathbf{E} | \Delta \left(\frac{n_i^2 \omega^2}{c^2} \right)^2 | \mathbf{E} \rangle$$

(6.1.22)

As the imaginary part of n_i and β are much smaller than the real part, we can immediately extract the modal gain:

$$G_m = \langle \mathbf{E} | g_m | \mathbf{E} \rangle = g_m \Gamma.$$

(6.1.23)

(p.95) The overlap factor Γ is defined for the y-component of the optical field normalized to the intensity as

$$\Gamma = \frac{\int_{active} dz dx E_x^2(z, x)}{\int_{z,x=-\infty}^{\infty} dz dx E^2(z, x)}$$

(6.1.24)

The amount of material gain g necessary in one layer to compensate the laser losses (α_m mirror and α_w waveguide) is then given by the equation

$$g\Gamma = \alpha_m + \alpha_w.$$

(6.1.25)

In this approach, however, the fact that in the final device some layers are going to be lossy and some will exhibit gain is not explicitly taken into account in the optical field profile. Rigorously, a better (but computationally more demanding) solution to the problem is to include the gain directly in the dielectric function of the active region, and iterate until the equality

$$-2\Im m(\beta) = \alpha_m$$

(6.1.26)

is verified. This approach, however, is computationally intensive and is needed only when the change of refractive index due to gain affects the shape of the guided mode. It was shown, however, to be very successful for the modeling of “cavityless” and “random” laser devices [111], and should probably be considered for some terahertz waveguides.

6.1.3 Dielectric waveguide for short wavelengths QCL

In QCLs based on the InGaAs/AlInAs/InP material system, the material combination is very favorable for dielectric waveguides, as the InP substrate has a refractive index n = 3.1 with a good contrast compared to AlInAs (n = 3.2) and especially GaInAs (n = 3.5). Fig. 6.2 shows the refractive index profile as well as the first three computed eigenmodes for a QCL emitting at 4.2 μm. In that initial device the upper cladding was AlInAs grown by MBE, displaying a lower refractive index contrast with the active region than the InP

substrate. The active region is separated from the heavily doped InP substrate by a low-doped AlInAs lower cladding layer.

6.2 Interface plasmon mode

In TM mode the boundary condition expressing the continuity of $\mathbf{H_x}$ and $\frac{1}{\epsilon_i}\frac{\partial \mathbf{H_x}}{\partial z}$ enables a particular bounded mode if the dielectric constant $\epsilon_i = n_i^2$ changes sign at the interface, i.e. if one of the two material behaves optically as a metal. As shown schematically in Fig. 6.3, in that case, the slope of $\mathbf{H_x}(z)$ changes sign at the interface, enabling a confined mode that decays exponentially in both materials.

In a simplified treatment, let us assume a metal, characterized by a negative dielectric constant ϵ_m for $z < 0$, while the regions $z > 0$ has a dielectric constant ϵ_d. We therefore assume the following field:

$$\begin{aligned} H_m(z) &= Ae^{qz} \quad z<0 \\ &= Be^{-hz} \quad z>0 \end{aligned}$$

(6.2.27)

(p.96)

where h and q are now $h = \sqrt{\beta^2 - \epsilon_d\left(\frac{\omega}{c}\right)^2}$ and $q = \sqrt{\beta^2 - \epsilon_m\left(\frac{\omega}{c}\right)^2}$ The boundary conditions at $z = 0$ are

$$A = B$$

(6.2.28)

$$A\frac{q}{\epsilon_m} = -B\frac{h}{\epsilon_d}$$

(6.2.29)

The second equation can be satisfied only if the dielectric constant changes sign at the interface. The propagation constant can be obtained after some algebra:

(p.97)

$$\beta = \frac{\omega}{c}\sqrt{\frac{\epsilon_m \epsilon_d}{\epsilon_m + \epsilon_d}}$$

(6.2.30)

In addition to the condition on the change of sign at the interface, an additional

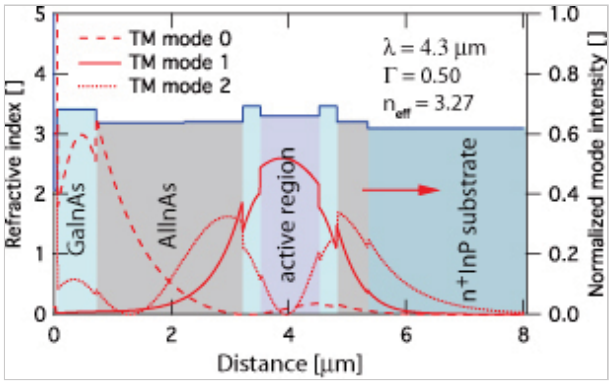


Fig. 6.2. Refractive index profile (left axis) and normalized mode intensity (right axis) of a mid-infrared dielectric waveguide of the initial QCL operating at $\lambda = 4.3$ μm. Note that the mode m=0 with the highest effective index is not the one supporting the lasing but an interface plasmon one. The parameters of the laser mode are shown in the inset.

condition is that the dielectric constant of the metal is "large enough", i.e. $|\epsilon_m| \rangle \quad |\epsilon_d|$.

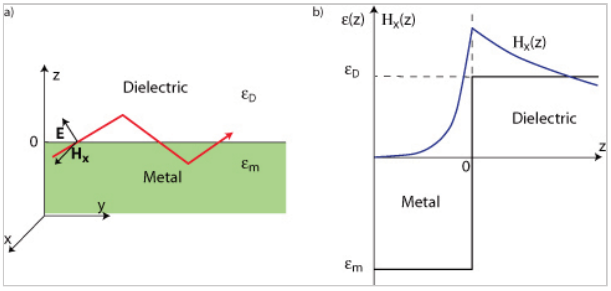


Fig. 6.3. a) Schematic drawing of a single plasmon waveguide, showing the axes as well as the **E** and **H** fields for a TM polarized wave. b) Schematic plot of the dielectric constant and of the magnetic field as a function of distance.

As QCLs operate using the TM polarization direction, interface plasmon modes play an important role in the light confinement. In fact, as the electrical contacting of QCLs is done by a surface metalization, an interface plasmon solution will in general be present in the confined mode. In the mid-infrared it is usually more lossy than the dielectric mode. For very long wavelengths ($\lambda \gtrsim 15\ \mu$m), however, modes based on single plasmons have losses and confinement factors that are competitive with the dielectric wavelength.

6.2.1 Waveguide based on an interface plasmon

As a rule, in dielectric waveguides, the cladding thickness scales linearly with wavelength. In addition, as the wavelength increases the doping needed to enable electrical transport creates a loss that grows as λ^2. As a result, for devices operating at wavelengths longer than about 10 μm, an approach based on single plasmon waveguide becomes more and more attractive. In fact, the loss of a single plasmon waveguide, derived from eqn. 6.2.30 can be rewritten as

$$\alpha = \frac{4\pi n_m n_d^3}{k_m^3 \lambda} \tag{6.2.31}$$

as a function of the real n_m, n_d, and imaginary k_m parts of the refractive index of the metal and dielectrics, given by $\epsilon_m = (n_m + ik_m)^2$ and $\epsilon_d = n_d^2$. As shown in Eqn. 6.2.31, the waveguide loss decreases with increasing wavelength as n_m and k_m grow approximately linearly with wavelength. An example of a QCL waveguide based on a plasmon waveguide [112] and operating at 11.2 μm is shown in Fig. 6.4. In that example, the confinement of the mode is further enhanced by the refractive index step between the active region and the substrate.

Compared to a dielectric waveguide, the single plasmon waveguide exhibits a much stronger confinement factor ($\Gamma = 0.7$ as compared to $\Gamma = 0.54$ for a similar thickness of the active region [113]) as well as a much reduced total thickness of the grown layer. Although the computed waveguide losses should have remained relatively low ($\alpha_w = 14\ \text{cm}^{-1}$), the measured threshold current of these initial devices indicated a much larger optical loss value.

In such waveguides the very strong coupling between the metal and the optical mode has been used to manipulate the confinement of the mode in the plane of the layers. In particular, it was used to realize the first QCL based on a photonic crystal cavity [114], and is still the subject of active research on single-mode devices [115].

6.3 Optical properties of doped layers

Because of the presence of free carriers in both the active region and the cladding layers, it is essential to develop a model that evaluates the effect of these doped regions on the **(p.98)** optical properties of the materials. The ability to control the free carriers effects is an important engineering tool to control waveguiding in the mid-infrared and terahertz regions. A simple Drude-like approximation approach that considers the motion of classical free electrons of effective mass m^* only subjected to a damping force yields usually excellent results. The susceptibility of such an electron gas is given by

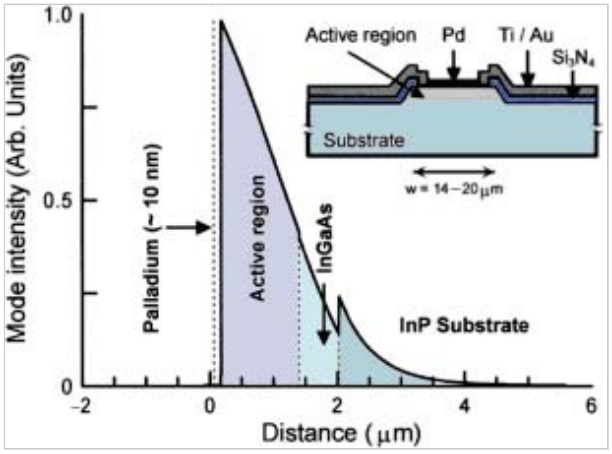


Fig. 6.4. Normalized mode intensity of a single plasmon waveguide of a QCL operating at $\lambda = 11.2$ μm. Adapted from [112].

$$\chi_{fc}(\omega) = -\frac{Nq^2}{m^*\epsilon_0}\frac{1}{\omega^2 + i\omega\gamma} \tag{6.3.32}$$

where N is the (volume) electron density. To compute the response of the system, the susceptibility of the free electrons must be added to the susceptibility of the semiconductor itself, given by $\chi_{sc} = 1 - \epsilon_{sc}$, so that the complete response can be characterized by the dielectric function:

$$\begin{aligned}\epsilon_r(\omega) &= 1 + \chi_{sc} + \chi_{fc}(\omega) \\ &= \left(\epsilon_{sc} - \frac{Nq^2}{m^*\epsilon_0}\frac{1}{\omega^2 + i\omega\gamma}\right) \\ &= \epsilon_{sc}\left(1 - \frac{\omega_p^2}{\omega^2 + i\omega\gamma}\right)\end{aligned} \tag{6.3.33}$$

where the plasma frequency ω_p is now defined as

$$\omega_p = \sqrt{\frac{Nq^2}{\epsilon_0\epsilon_{sc}m^*}} \tag{6.3.34}$$

The damping parameter γ can be seen, adopting Drude's model, as the inverse of the electron scattering time. Values of $\tau = \gamma^{-1} = 100 - 200fs$ are typically used for doped semiconductors.

(p.99)

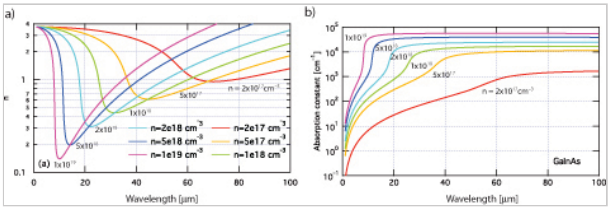


Fig. 6.5. (a) Refractive index and (b) absorption coefficient of GaInAs as a function of wavelength and for various doping densities.

The refractive index and absorption coefficient $\alpha = 4\pi k/\lambda$ of InGaAs latticematched to InP, computed using eqn. 6.3.33, is shown in Fig. 6.5 as a function of wavelength for various carrier densities. Three regimes can be clearly distinguished, dependent on the wavelength relative to the plasma frequency. For frequencies much higher than the plasma frequency, the refractive index is essentially unchanged, and the free carriers only introduce an absorption proportional to the square of the wavelength and of the electron density, which can be written from eqn. 6.3.33 as

$$\alpha_{fc} = -\frac{4\pi}{\lambda}\frac{\Im m(\epsilon_r)}{2n_{refr}} = \frac{e^2}{4\pi^2 c^3 \epsilon_0 n_{eff}}\frac{N\lambda^2}{m^*\tau}$$

(6.3.35)

The free-carrier absorption, plotted as a function of wavelength, is shown for InGaAs material for various doping concentrations in Fig. 6.6. As a comparison, the typical mirror losses of QCL are also shown. For best efficiency, the losses created by the cladding layers should remain small compared to the cavity losses. As a result, doping in the range of $N \approx 2 \times 10^{16}$ cm^{-3} are used close to the active region, where the overlap factor is large. To limit the additional series resistance, higher doping levels can be used closer to the contact regions where the optical field is small. Considering again Fig. 6.5, as the light frequency is decreased and approaches the plasma frequency, the refractive index of the semiconductor is strongly depressed. This effect can be used to create low-index confining layers that will tend to expel the optical field. If the frequency is further decreased, the semiconductor behaves optically as a metallic layer with a low electron concentration. This effect has been exploited to create plasmon waveguides for QCLs operating in the terahertz range.

6.3.1 Plasmon enhanced waveguides

We now examine some examples of how this effect can be exploited to produce efficient waveguides. In Fig.6.7 an example of a *plasmon enhanced* waveguide is reported. Here the variation of the refractive index as a function of the doping and of the frequency is exploited in order to obtain a low index (red region) in order to “push” the mode onto the core of the waveguide.

(p.100)

As shown in Fig. 6.5, care must be taken not to overdope these layers, so that the plasma resonance does not reach the

photon energy, otherwise the absorption coefficient increases by more than an order of magnitude. This technique has also been used successfully to create waveguides for QCL based on GaAs [116] or InAs substrate materials [117]. In both material systems the substrate has a refractive index that is larger than the active region and therefore is an antiguiding layer. In addition, the growth of very thick low-index ternary cladding layers is in both cases problematic because of their poor electrical and thermal conductivity. The use of a heavily doped layer as a low-index material has enabled, for both devices, relatively low-loss waveguides with good electrical and thermal properties.

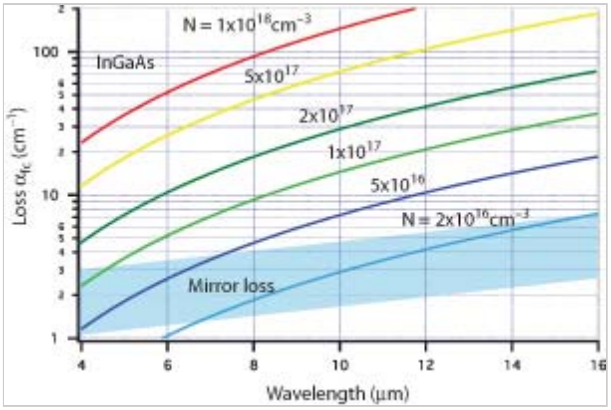


Fig. 6.6. Absorption as a function of wavelength for various free-carrier concentrations, as indicated. The shaded area represents typical mirror losses of QCL cavities.

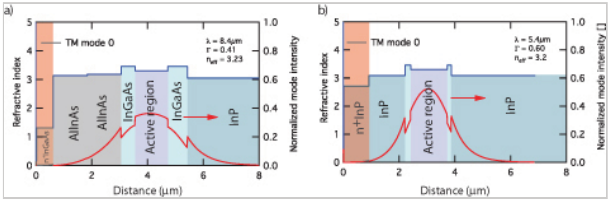


Fig. 6.7. a) Long-wavelength (λ = 8.4 μm) and b) shorter-wavelength (λ = 5.4 μm) waveguides using a low-refractive-index, highly doped top layer (regions marked n^+) to enhance the confinement.

(p.101) 6.4 Two-dimensional confinement

The fabrication of ridges creates a lateral waveguide, which is usually broader than the vertical one, as the resolution of the fabrication technique used to define the ridge is of the order of approximately 1 μm. In addition, as the optical power scales roughly with width, broad ridges are favorable for high power devices.

Fig. 6.8 shows the two common lateral waveguide geometries used in QCL devices. In the conventional ridge-etched structures Fig. 6.8a, the lateral waveguide is achieved by etching through the active region and coating the sidewalls with Si_3N_4 or SiO_2 before the metalization. As a result, the electric field vanishes at the sidewalls and the mode is completely confined inside the ridge structure. In a slightly different approach, shown in Fig. 6.8b, after the ridge etching, a layer of semi-insulating InP is regrown on the sidewalls, providing both optical and electrical confinement. Fig. 6.8 also shows

the computed electric field for both waveguide geometries. In this case, the computation was performed directly by solving Mawell's equations 6.1.1 numerically, using the

commercial package COMSOL. Excellent results can also be obtained using the effective index method, in which the problem is solved along both axes sequentially. This, of course, is especially convenient in rectangular waveguide geometries. An illustration of the effective index technique is shown in Fig. 6.9. The field is written as

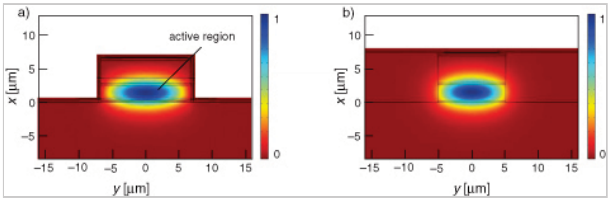


Fig. 6.8. Computed electrical field distribution of a) a 14 μm-wide conventional ridge waveguide design, and b) a 10 μm wide buried heterostucture waveguide design for an emission wavelength of 8.4 μm.

$$E(z,x) = E_x^i(z)A(x) \tag{6.4.36}$$

where the field E_x^i is the solution of the one-dimensional waveguide problem in region *i* (*i* = *I, II, II* in our example), defining an effective index for the three regions n_{eff}^I, n_{eff}^{II}, and n_{eff}^{III}. These three indices are then used to create an effective guide in the x-direction, and allows solving for the envelope function $A(x)$, as shown in Fig. 6.9b. This technique, however, is restricted in the number of geometries it can accommodate.

The optical losses α_w and overlap factor Γ for the ridge waveguide and the buried heterostructure designs are compared in Fig. 6.10 as a function of ridge width. The two waveguides have the same computed values for Γ and α_w down to a ridge width **(p.102)**

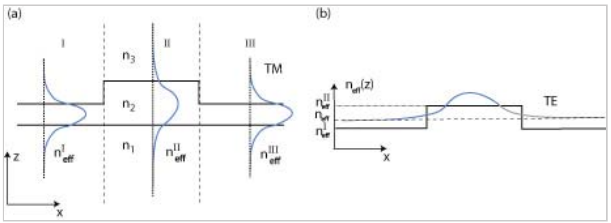


Fig. 6.9. Illustration of the effective index technique for ftwo-dimensional waveguide computation. a) A waveguide is defined by an etching of the core layers on both sides of the ridge. The one-dimensional problem is solved in the three different regions I, II, and III along the z-direction, yielding each time an effective index for the mode for a given polarization. b) The transverse part of the mode profile is now obtained by computing an effective index problem in the x-direction for the opposite polarization.

of about 10 μm. Below this value the penetration of the field in the lateral InP region yields a decrease of the overlap factor for the buried heterostructure waveguide. For the ridge waveguide, on the other side, the increasing penetration of the optical field in the metal increases the optical loss of the mode.

6.4.1 Lateral mode selection

In both ridge and buried heterostructure waveguides, because of the large refractive index step between the core and the side, these waveguides support more than one transverse mode. The multi-mode character of the waveguides is important

for its **(p.103)** spatial and spectral characteristics of the laser. In the ridge process the higher-order lateral modes will also suffer from a higher optical loss, as the intensity of the optical field is larger at the lossy metal sidewall. This mode selection has been successfully used in both Fabry–Perot and DFB lasers. In contrast, in buried heterostructure waveguides the mode discrimination arises from the change in overlap factor between the lateral modes, and is, as shown in Fig. 6.10, weaker for a given ridge width than the change in waveguide loss.

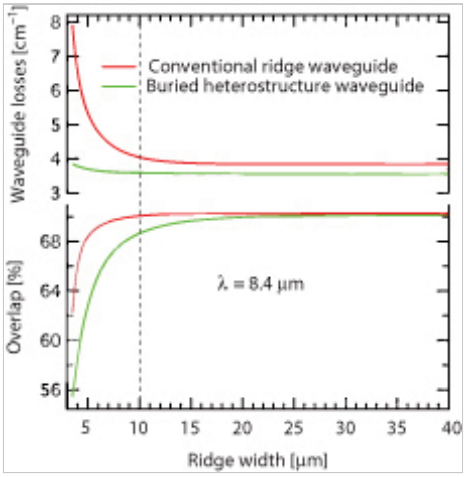


Fig. 6.10. Computed waveguide losses and overlap factor for an emission wavelength of 8.4 μm, for a ridge waveguide and buried heterostructure waveguide design.

In interband devices, lateral mode discrimination is very often achieved by a weak lateral confinement obtained, for example, by a shallow lateral etch of the active region as shown schematically in Fig. 6.9. In fact, if the difference between the effective index in the center and the side of the active region $n_{eff}^{I} - n_{eff}^{II}$ is small enough, the lateral waveguide will support only one guided mode. This technique has been attempted in QCL by a shallow etching of the cladding that induced a weak lateral confinement [118]. However, in contrast with interband devices the conductance of a QCL active region is strongly anisotropic, and is much larger in the lateral than in the vertical direction. As a result, due to current spreading, the threshold current of these devices remained very large.

In an attempt to control this problem, devices where the lateral current spreading was limited using H+ implantation were developed, as the definition of a lateral window for current injection would perform lateral mode selection [119]. Excellent results were achieved using GaAs-based materials [120]. In InP-based devices, however, the performance of the lasers was strongly limited by the poor electrical stability of the H+ implanted layers as a function of temperature.

6.5 Large optical waveguides

In pulsed operation, the optical power of a laser is limited by gain saturation of the medium. As a result, scaling the power up is achieved by increasing the mode crosssectional area. As mentioned above, this can be achieved by merely widening the laser stripe. In fact, very large optical powers, of the order of 100s of watts, have been achieved in this manner. However, the problem of this approach is the strong asymmetry in the divergence angle of the beam and high multi-mode character of the lateral mode profile.

For this reason, it is also advantageous to increase the mode size in the growth direction. This can also be combined with an increase of the number of periods of the active region.

The refractive index profile and computed mode profile of a devices where such a large optical waveguide was implemented is shown in Fig. 6.11. In this device, designed for operation at a wavelength close to $\lambda \approx 10\ \mu$m, 4 × 15 periods of the active region are separated by 1 μm thick InP interstacks that provide additional lateral heat transport [121]. The voltage and light-versus-current characteristics of one of these large optical cavity devices is shown in Fig. 6.11b. Compared to a normal buriedheterostructure waveguide, this large optical cavity device demonstrated an improvement of the peak power by a factor of × 10 while keeping a single spatial mode in both vertical and transverse directions.

A measured far-field of such a large optical waveguide is shown in Fig. 6.12 along with the computed profiles, showing a FWHM of 17.8° and 39° in the xand z- **(p.104)**

directions, respectively. These values are to be compared with the much higher ones of typical buried heterostructure devices (45× 65°).

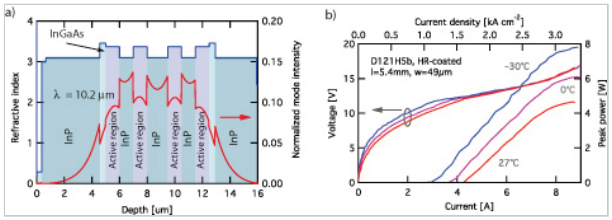


Fig. 6.11. a) Refractive index and computed mode profile of a large optical waveguide in which the active regions are separated by InP interstack spacers. b) Voltage and light-versuscurrent for such a device operating at $\lambda \approx 10\ \mu$m. Adapted with permission from [121]. Copyright 2008, AIP.

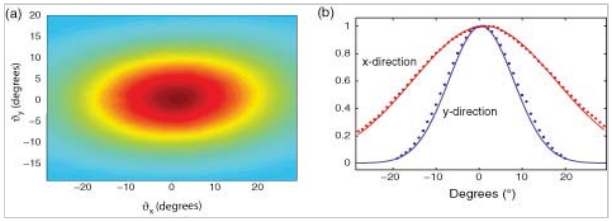


Fig. 6.12. a) Measured far-field pattern for 40 μm wide and 3 mm long device with a large optical cavity. b) Comparison between the experimental (dotted lines) and computed (solid lines) intensities along the z (growth) and x (transverse) directions.

6.6 Thermal properties

Quantum cascade laser active regions dissipate a very large amount of heat. The resultant increase in the active-region temperature, especially important in high duty cycle operation, is in general detrimental to the device performance. The only exception is the tuning of single-mode distributed feedback lasers that use exactly the refractive index change with temperature. In steady state the active-region temperature is related to that of the surmount by a linear relationship

$$T_{act} = T_{sub} + R_{th}UI \tag{6.6.37}$$

(p.105) where R_{th} is the thermal resistance, and U and I are the device operating bias and current, respectively. It is therefore important to minimize R_{th} in order to achieve the best possible continuous wave

performance. The design of optical waveguide is therefore a compromise between the requirements of achieving low optical losses and tightest possible confinement with the goal of simultaneously decreasing the thermal resistance.

Table 6.1 Thermal conductance of various materials used in waveguides and submounts.

Material	κ [W/(m K)]
InP	74
$Ga_{0.47}In_{0.53}As$	4.84
Si_3N_4	15
Au	317
Cu	384
AlN	257
Diamond	1200
In solder	81
SnAu solder	57
Active region	4.72

In the semiconductor materials used for QCLs, the heat is carried mostly by acoustical phonons, as the low doping densities used prevent efficient heat transport by the electrons. All III–V have very similar phonon density of states. In contrast, as compared to binary materials, the phonon mean free path in ternary materials is very strongly reduced by the alloy scattering. As a result, the thermal conductance of $Ga_{0.47}$ $In_{0.53}$ As and $Al_{0.48}$ $In_{0.52}$ As is about 15 times *lower* than the one of InP. For this reason, InP is a preferred material for waveguide claddings. In fact, the lack of a binary cladding material (outside pure AlAs that is prone to oxidation in ambient atmosphere) has been a consistent drawback of GaAs QCLs.
Thermal considerations were incorporated very early in the design of QCLs, and stimulated the early use of InP-based claddings, enabling average powers of 5 mW to be achieved at room temperature for the first time [32]. The use of junction-down mounting was also demonstrated relatively early to significantly improve the continuous-wave device performances [122].

The thermal resistance does depend not only on the waveguide but also on all the mounting and soldering process. For this reason, the latter have also to be included in a realistic simulation of laser devices.

Thermal modeling.The temperature T in the active region can be found by solving Fourier's heat transport equation

$$-\nabla(\kappa_{th}\nabla T) = Q$$

(6.6.38)

where κ_{th} is the thermal conductance and Q the source power density. The latter can be written as **(p.106)**

$$Q = J\mathcal{E} \tag{6.6.39}$$

where $\mathcal{E}$ is the electric field in the active region and J the current density. For a device operating with $\mathcal{E} = 80kV/cm$ and a current density of $J = 5kA/cm^2$ we obtain $Q = 400MW/cm^3$! For this reason, the heat generation outside the active region, due to Joule heating of the claddings and contacts, is usually neglected.

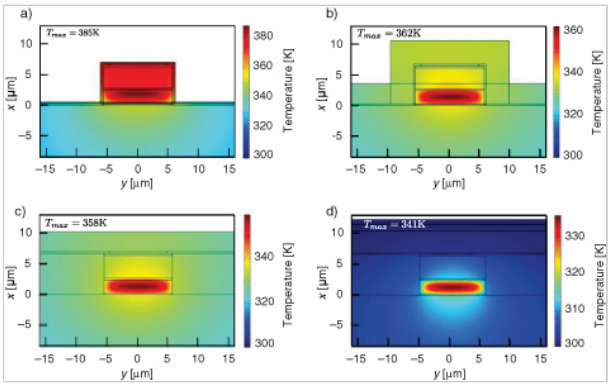


Fig. 6.13. Computed temperature plots of the front facet of an otherwise identical QC laser structure, 12 μm wide and 3 mm long, subjected to an input electrical power of 8.8 W (24.4 kW/cm^2) for various waveguide designs. The maximum temperature of the active region is indicated in each case. An isotropic thermal model was used for the active region. a–b) Conventional ridge waveguide with 0.3 μm Si_3N_4 passivation and 0.2 μm (a) and 4.0 μm (b) thick gold metallization. c–d) Buried waveguide heterostructure (BH) waveguide design for episide-up (c) and episide-down (d) mounting.

As already mentioned, the thermal conductance k_{th} varies strongly between the different materials. Scalar values commonly used are displayed in Table 6.1. As the thermal conductance in the semiconductor material is limited by the phonon mean free path, one may suspect the latter to be anisotropic in the layered active region. In particular, the thermal conductance should be lower for heat flowing perpendicular to the layer rather than the one flowing parallel to the layers, i.e. $\kappa_\perp < \kappa_\parallel$. This result was indeed found experimentally by Lops *et al*. [123], with a ratio $\kappa_\perp/\kappa_\parallel$ strongly temperature dependent, with a value of $\kappa_\perp/\kappa_\parallel = 0.1$ at 100 K, and an extrapolated value of $\kappa_\perp/\kappa_\parallel = 0.3$ at 300K, with the corresponding thermal conductance values of $\kappa_\perp = 2.5W/mK$ and $\kappa_\parallel = 8W/mK$.

A comparison of the thermal properties of various waveguide architectures is shown in Fig. 6.13, where the computation of the temperature profile across a waveguide is shown for identical thermal load conditions. The effect of a thick gold layer, as well as a buried heterostructure and a junction-down mounting, are evidenced, and translate **(p.107)**

in a reduction of the ΔT between the active region and the substrate from 90 K for the thin gold, junction-up ridge to 40 K for the buried

heterostructure, junction down device. Besides the material used, the thermal resistance of a device depends strongly on the geometry. To facilitate the comparisons, it is worth rewriting eqn. 6.6.37 as a function of the current density J:

$$T_{act} = T_{sub} + \frac{UJ}{G_{th}}$$

(6.6.40)

where G_{th} is now the area-specific thermal conductance. As shown in Chapter 11, values of G_{th} are measured between 250 W/cm^2 for wide ridges with thin evaporated gold to 2000 W/cm^2 for narrow, buried heterostructure devices mounted junctiondown. This dependence is shown in Fig. 6.14, where experimental measurements of the thermal resistance of buried heterostructure devices are compared to the computed values. The data is also shown as a function of the thermal resistivity of the active region, related to the specific heat conductance by

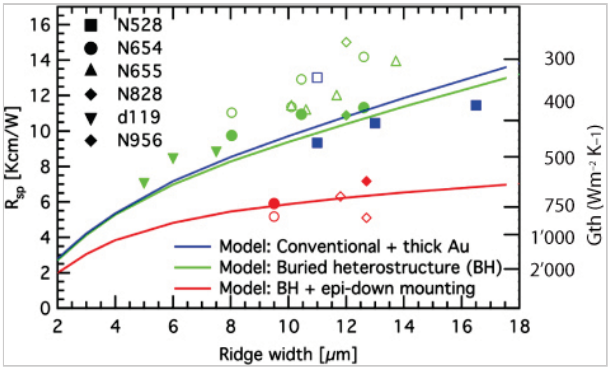
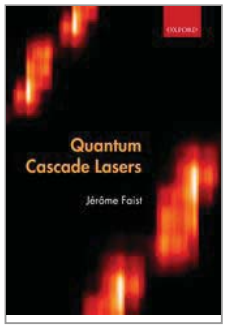


Fig. 6.14. Experimental (markers) and theoretical (lines) data of the thermal resistivity (left axis) and the area specific conductance (right axis) for different thermal waveguide designs: (upper curve) conventional waveguide with thick electroplated gold, (intermediate) buried waveguide design, and (lower curve) buried waveguide mounted epi-down on diamond. The experimental data have been deduced either from the threshold currents (open marker), or spectral characteristics (filled marker).

$$R_{sp} = \frac{1}{d_{act} G_{th}}$$

(6.6.41)

where d_{act} is the thickness of the active region.



Quantum Cascade Lasers
Jérôme Faist

Print publication date: 2013
Print ISBN-13: 9780198528241
Published to Oxford Scholarship Online: May 2013
DOI: 10.1093/acprof:oso/9780198528241.001.0001

Active region design

Jérôme Faist

DOI:10.1093/acprof:oso/9780198528241.003.0007

Abstract and Keywords

A laser based on intersubband transitions benefits from many interesting features of this system. As a detector, it is based on an optical transition with high oscillator strength that can be tuned over a wide range of frequencies in the mid and far-infrared. However, such a laser must overcome the challenge of the very short lifetime of the upper state transition. Such a short lifetime will severely limit the available gain at a given pumping power density. This chapter discusses active region design covering historical perspectives, fundamental concepts, intersubband versus interband lasers; rate equation analysis, threshold condition, slope efficiency; optimization of the active region, scaling with the number of periods, temperature dependence, doping of the active region, and wallplug efficiency.

Keywords: lasers, upper state transition, rate equations, threshold condition, slope efficient, scaling, doping, wallplug efficiency

It is clear from the preceding chapters that a laser based on intersubband transitions will benefit from many interesting features of this system. First of all, as for a detector, it is based on an optical transition that exhibits a large oscillator strength ($f \approx 20$–30 for common GaAs or InGaAs materials) that can be tuned over a wide range of frequencies in the mid- and far-infrared. However, it is also clear that such a laser has to overcome a number of inherent challenges, the most serious one being the very short lifetime of the upper-state transition. Let us emphasize that in contrast to the case of a light-emitting diode, such a short lifetime will *not* be an inherent limitation to the efficiency of such a device, but will rather severely limit the available gain at a given pumping power density.

7.1 Historical perspective

The idea to use intersubband transitions to achieve optical gain traces back to the original work of Kasarinov and Suris in the early 1970s [8]. In this paper they examined the optical response of a superlattice under the application of a strong electric field, such that the ground state of one well would be energetically situated just below the second excited state of the quantum well downstream. As compared to earlier superlattice theoretical investigations, focusing on the appearance of Bloch oscillation and negative differential conductivity in the miniband picture, they investigated a regime where the electron states would be localized by the electric field in their respective wells. The mechanism for population inversion and gain is fairly easy to understand. As shown schematically in Fig. 7.1, the transition between the ground state of one well 1 to the excited state 2 of the well downstream must proceed through a tunnel barrier. The strength of this interaction can therefore be reduced by an arbitrary amount by just increasing the thickness of the tunnel barrier. In contrast, the scattering between levels belonging to the same quantum well, such as that occurring between the states $n = 2$ and $n = 1$, is expected to be much more efficient. In such a system, as the ground state of each well is also the upper-state of the laser transition, the population inversion is a robust feature that will arise regardless of temperature for any monotonic electron distribution.

One of the key problems that has so far prevented this proposal from being demonstrated experimentally is that the situation they described could hardly be obtained experimentally, because it would mean operating the device in an unstable electric field region of negative differential resistance. In other words, as for the Bloch oscillator, the gain extends down to zero frequency, preventing a stable field distribution in **(p.109)**

the device. As a result, as shown experimentally by K. K. Choi [124], a doped superlattice under strong applied bias will break into field domains with, in each domain, energy states aligned in resonance. This configuration is more stable because it enables a larger current to flow through the structure. For example, in the case studied by Kazarinov and Suris, the real band alignment for a realistic structure would be a part of the structure that will exhibit a ground state of each well aligned with the *first* excited state of the following well, and the next part of the structure where the ground state of each well is aligned with the *second* excited state.

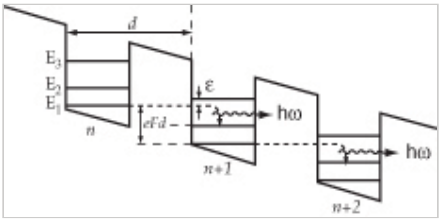


Fig. 7.1. Schematic band structure of the device considered theoretically by R. Kazarinov and R. Suris. A superlattice is under a strong applied electric field, such that the ground state of a well is energetically above the first excited state of the downstream well.

7.2 Active region: fundamental concepts

7.2.1 Injection/relaxation region

The design of a successful intersubband laser has to solve simultaneously a number of problems:

- The structure should be stable electrically at the operating point.
- Population inversion should be achieved.
- The resultant gain should overcome waveguide and other active region losses.

As compared to all the previous proposals, the quantum cascade laser was the first demonstration of an active region that fulfilled all these requirements [20]. To this end, an architecture was chosen in which the active structure is no longer a simple superlattice but now consists of a periodic arrangement of unit cells, each one consisting of a complex heterostructure potential consisting of a number of wells and barriers. As shown in Fig. 7.2, roughly speaking, each cell can be divided into a gain region and an injection/relaxation region. The gain region is the structure that will create and maintain a population inversion between the two levels of the laser transition. As will be shown later, this result can be obtained using various designs. In general, the active region contains a ladder of at least three states (or continuum of states), such that electrons are injected in the $n = 3$ state and the population inversion is maintained between the states $n = 3$ and $n = 2$. Assuming that electrons are exclusively injected in the $n = 3$ state, the population inversion requirement translates into the following requirement on the lifetimes.

$$\tau_{32} > \tau_2$$

(7.2.1)

(p.110)

i.e. the total lower-state lifetime τ_2 is shorter than the electron scattering time from the n=3 to the n=2 levels. An injection/relaxation region follows the gain region. The main purpose of this injection/relaxation region is to raise the energy of the electron compared to the band edge, thus enabling injection by resonant tunneling in the next period. This function is achieved conveniently by a sequence of alternating quantum wells and barriers with a **(p.111)** changing duty cycle, creating an effective "graded gap" that compensates the applied electric field. A portion of the injection region is also doped, providing the electron charge needed for the transport, in order to prevent the formation of space-charge domains. This region also behaves as an "electron reservoir" that will feed the carriers to the next period; the low effective applied electric field due to the bandgap grading will promote a cooling of the electron distribution towards the lattice temperature.

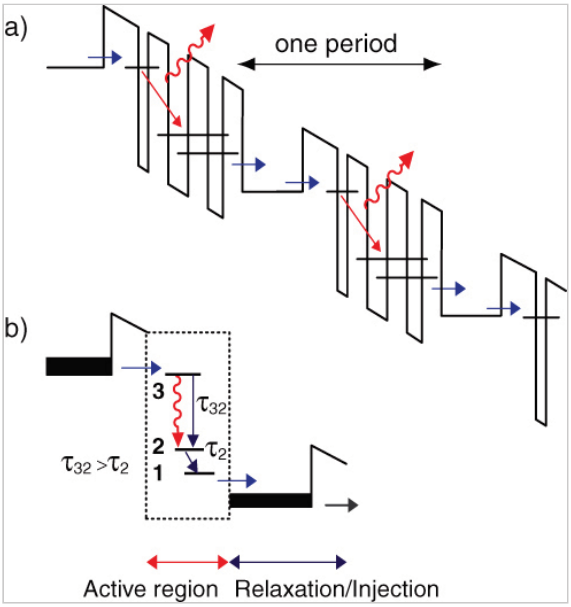


Fig. 7.2. a) Schematic conduction band diagram of a quantum cascade laser [20]. Each stage of the structure consists of an active region and a relaxation/injection region. Electrons can emit up to one photon per stage. b) General philosophy of the design. The active region is a three-level system. The lifetime of the 3 → 2 transition has to be longer than the lifetime of level 2 to obtain population inversion.

In conclusion, the presence of the injection/relaxation region will

• Prevent the formation of electrical domains.
• Reduce the applied electric field by lengthening the period.
• Block electron escape from the upper-state of the laser transition by suitable engineering.

7.2.2 Cascading

A very attractive feature of lasers based on intersubband transitions is the fact that cascading the active-region periods is easy and comes almost naturally. In a cascade laser, the gain region consists of a large (typically $N_p = 10 - 100$) number of repeated periods. Electrons are then "recycled" from period to period as they cascade down the structure. Cascading has two main advantages. First, by increasing the size of the region in which gain occurs it will decrease the population density required for each individual active region period. This allows a reduction of the threshold current density of the device. In addition, a single electron is able to potentially emit a number N_p of photons, increasing the slope efficiency. As shown in Fig. 7.3, both effects are obtained at the cost of a larger applied bias. In essence, in a cascade laser one is exchanging a lower operating current for a larger operating voltage. This enables a large reduction of the device's ohmic losses, especially in the mid- or far-infrared where the photon energy is very small. In fact, as the slope efficiency grows with N_p, the device-operating **(p.112)** current for a given power is proportional to N_p^{-1}. The ohmic losses ($= RI^2$) for this device therefore decrease with the square of the number of the period.

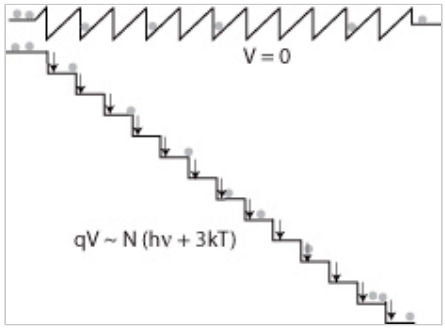


Fig. 7.3. Highly simplified picture of the effective potential in a quantum cascade laser. At zero field the current is blocked by the injection regions that prevent the current from flowing. With the application of a strong electric field, these injection regions are "flattened" and the current flows along the electronic cascade, generating the photons.

This advantage is, of course, valid for any semiconductor laser. The problem in applying such an idea to interband near-infrared structures is the necessity of inserting a Zener tunnel junction in the structure to enable reinjection of carriers across the gap. Such a tunnel junction must be heavily doped, and therefore tends to increase the optical losses. An elegant way around this problem is to work in a vertical cavity semiconductor laser, and place the Zener junction at a node of the optical field.

7.3 Intersubband versus interband lasers

Intersubband and interband lasers are radically different devices. In an interband laser:

• The population inversion is obtained naturally through a physical property of the material (its energy gap).
• A minimum current must be injected to reach material transparency. An essential element of the optimization is therefore to minimize this transparency current.
• The maximum current carried by the device is limited by either heating or catastrophic damage on the facet.

In contrast, in an intersubband laser:

• The population inversion must be established by a suitable design of the active region.
• As a general rule there is no transparency current.
• The device reaches a maximum current given by the doping of the device and an alignment of the band structure.

As a result, the optimization of the active region in both cases is very different. For interband lasers the reduction of the transparency current is achieved through a reduction of the active volume, by confining the electron-hole pairs in quantum wells or even quantum boxes. As a result, the transparency current historically fell from more than $10kA/cm^2$ for a p-n junction diode to $1-3kA/cm^2$ for a double heterostructure active region to less than $100A/cm^2$ for strained InGaAs quantum well lasers.

In contrast, in a quantum cascade laser, the gain must be maximized by engineering the lifetimes and oscillator strengths. In fact, mimimum threshold current (or even power, taking into account the applied bias) is obtained for relatively large active regions with 25–50 periods, as there is no direct penalty for increasing the active region thickness. Besides the active–region design, the other key parameter for achieving highest performance is the minimization of the optical losses coming from both the waveguide and the active region.

7.4 Rate-equation analysis, threshold condition, slope efficiency

A key understanding of the requirements for the design of a QCL may be achieved by considering a simple atomic-level analysis of the active region. For historical reasons the gain region is assumed to consist of three states, with laser action occurring between levels $n = 3$ and $n = 2$, whereas the injector is assumed to consist of only one level, with a constant population n_g, aligned with the upper level of the next period. A **(p.113)** schematic diagram of the process connecting these three different states is shown in Fig. 7.4. Electrons are injected in the $n = 3$ state from the ground state of the injector

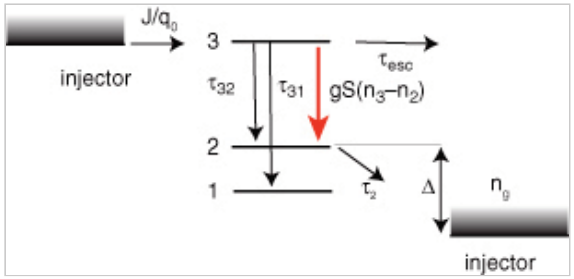


Fig. 7.4. Population inversion between the two upper levels of a ladder of three states. The relevant non-radiative process are also indicated.

of the previous period at a rate equal to J/e. Electrons may scatter from this state to the lower $n = 2$ and $n = 1$ states with rates τ_{32}^{-1} and τ_{31}^{-1} respectively. The $n = 3$ lifetime, given by the relation $\tau_3^{-1} = \tau_{32}^{-1} + \tau_{31}^{-1} + \tau_{esc}^{-1}$ may include an escape process into the continuum with a lifetime τ_{esc}. We assume also that electrons may be thermally activated to state $n = 2$ to an equilibrium population n_2^{therm}. The latter term may be approximated by an activated behavior $n_2^{therm} = n_g \exp(-\Delta/kT)$, where n_g is the sheet doping density of the injector and Δ the energy difference between the Fermi level of the injector and the $n = 2$ state of the laser transition. Therefore, the rate of change of the population of the three levels may be written as

$$\frac{dn_3}{dt} = \frac{J}{e} - \frac{n_3}{\tau_3} - Sg_c(n_3 - n_2) \tag{7.4.2}$$

$$\frac{dn_2}{dt} = \frac{n_3}{\tau_{32}} + Sg_c(n_3 - n_2) - \frac{n_2 - n_2^{therm}}{\tau_2}$$

(7.4.3)

$$\frac{dS}{dt} = \frac{c}{n_{refr}}[(g_c(n_3 - n_2) - \alpha_{tot})S + \beta\frac{n_3}{\tau_{sp}}]$$

(7.4.4)

In eqn. 7.4.4, α_{tot} is the total modal loss (sum of the waveguide and mirror losses), β the fraction of the spontaneous emission coupled in the laser mode, and τ_{sp} the spontaneous emission lifetime.

In considering eqn. 7.4.4, it should be noted that the electron n and photon S densities do not have the same units. In this work, we choose to define n_i as *sheet densities per period* and S the photon flux *per period and unit active region width*. However, it should be remembered that these equations represent only one period of the active region, whereas a quantum cascade laser will have many periods sharing the same optical waveguide. Usually, one considers the gain cross-section g_c appearing in eqn. 7.4.4 as one of all the periods in the active region put together, i.e. proportional to the total overlap factor Γ:

(p.114)

$$g_c = \Gamma\frac{4\pi e^2}{\epsilon_0 n_{refr}\lambda}\frac{z_{32}^2}{2\gamma_{32}L_p}$$

(7.4.5)

where z_{32} is the dipole matrix element, $2\gamma_{32}$ is the FWHM broadening of the transition, λ the wavelength, and L_p the length of the period. If this convention is chosen, it should be noted that the photon flux S is the one per period, not the total one for the whole active region.

The behavior below threshold is obtained by setting the time derivatives and S to zero. In this approach the amplified spontaneous emission is neglected because $\tau_{sp} \rangle\rangle \tau_3$ and only becomes significant extremely close to the threshold current. From the first equation we obtain a relation between the upper-state population and the electrical pumping: $n_3 = J\tau_3/e$. Making use of the second, one obtains for the population inversion $\Delta n = n_3 - n_2$,

$$\Delta n = \frac{J\tau_3}{e}(1 - \frac{\tau_2}{\tau_{32}}) - n_2^{therm} = \frac{J\tau_{eff}}{e} - n_2^{therm}$$

(7.4.6)

where the effective lifetime $\tau_{eff} = \tau_3(1 - \tau_2/\tau_{32})$ relates the population inversion to the electrical pumping. The effective lifetime converges to the upper-state lifetime for states displaying a large ratio of lifetimes $\tau_2/\tau_{32} \langle\langle 1$. Note also that the population inversion can only exist if $\tau_{eff} \rangle 0$, i.e. if $\tau_2 \langle \tau_{32}$.

The threshold current density is reached when the modal gain $g_c\Delta n$ compensates for the losses α_{tot},

$$J_{th} = e\frac{\alpha_{tot}/g_c + n_2^{therm}}{\tau_{eff}}$$

(7.4.7)

or, expressing the gain cross-section explicitly:

$$J_{th} = \frac{1}{\tau_{\text{eff}}}\left[\frac{\epsilon_0 n_{refr} L_p \lambda(2\gamma_{32})}{4\pi e \Gamma z_{32}^2}\alpha_{tot} + e n_2^{therm}\right]$$

(7.4.8)

Above threshold, the gain is clamped and S increases linearly. By again setting the time derivatives to zero and differentiating over J, one obtains the slope efficiency for the whole stack:

$$\frac{\mathrm{d}P}{\mathrm{d}I} = N_p h\nu \alpha_{m,1}\frac{\mathrm{d}S}{\mathrm{d}J} = \frac{N_p h\nu}{e}\frac{\alpha_{m,1}}{\alpha_{\text{tot}}}\frac{\tau_{\text{eff}}}{\tau_{\text{eff}}+\tau_2}$$

(7.4.9)

where $\alpha_{m,1}$ is the front mirror loss. The behavior of S and Δn as a function of injected current is shown schematically in Fig. 7.5.

7.5 Optimization of the active region: the intersubband toolbox

An analysis of eqns. 7.4.8 and 7.4.9 leads to the following conclusions: a low-threshold, high-efficiency laser will be obtained in a structure that achieves:

- A large ratio of upper-state to lower-state lifetimes τ_{32}/τ_2.
- A low waveguide loss α_w.
- A narrow transition linewidth γ_{32}.
- A long upper-state lifetime τ_3.

(p.115)

The multiple QCL designs that exist simply reflect the complicated trade-off that has to be made when trying to simultaneously fulfil these conflicting requirements. The required engineering of the lifetimes and dipoles is achieved by using an "intersubband toolbox", the ingredients of which are combined to construct the various active-region architectures.

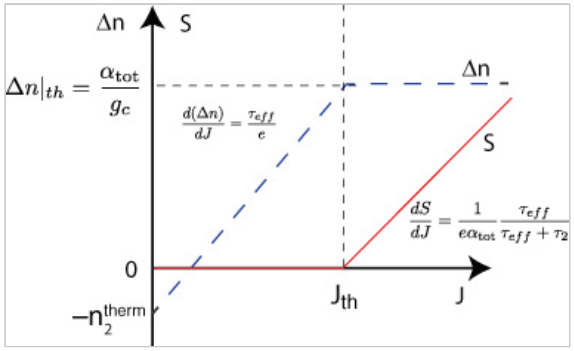


Fig. 7.5. Schematic illustration of the behavior of the population inversion Δn and the photon flux S with injected current J.

7.5.1 Tunneling

A very interesting situation occurs when, as shown in Fig. 7.6, the initial and final states of the laser transition are localized in two different quantum wells, coupled by a tunnel barrier. Mathematically, these states are localized in their respective wells when the coupling energy $\Omega_{1'1}$ between the initial state $1'$ and final state 1 is much smaller than the energy difference between the uncoupled states $|\chi_{1'}^0\rangle$ and $|\chi_1^0\rangle$, i.e. when $\Omega_{1'1} \ll E_{1'} - E_1$. In this situation, any scattering between the two states

will contain the coupling energy $\Omega_{1'1}$. The latter is proportional to the tunneling **(p.116)** probability $\exp(-\kappa_b L_b)$, where

$$\kappa_b = \sqrt{2m^*(V_b - E)}/\hbar$$

(7.5.10)

is the electron's wavevector and L_b is the thickness of the tunnel barrier. As a result, the scattering rate will be lowered by simply increasing the tunnel barrier thickness, *regardless of the scattering mechanism*.

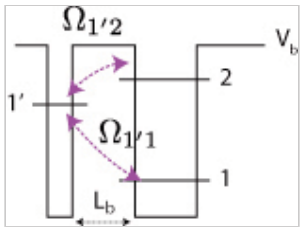


Fig. 7.6. Two states coupled by a tunnel barrier. The lifetime of the upper-state is controlled by the thickness of the coupling barrier.

This can easily be seen by applying first-order perturbation theory and computing the scattering operator *S* between the perturbed initial $|\chi_{1'}\rangle$ and final states $|\chi_1\rangle$, taking for the perturbed state

$$|\chi_{1'}\rangle = |\chi^0_{1'}\rangle + \frac{\Omega_{1'1}}{E_{1'} - E_1}|\chi^0_1\rangle$$

(7.5.11)

and a similar expression for the perturbed ground state $|\chi_1\rangle$. The result yields

$$\langle\chi_{1'}|S|\chi_1\rangle = \frac{\Omega_{1'1}}{E_{1'} - E_1}\{\langle\chi^0_1|S|\chi^0_1\rangle - \langle\chi^0_{1'}|S|\chi^0_{1'}\rangle\}$$

(7.5.12)

The above expression is indeed directly proportional to $\Omega_{1'1}$. In addition, as mentioned already by Kazarinov and Suris in their seminal paper [8], it appears immediately that a scattering potential having the same matrix elements with both states 1 and 1′ will experience a destructive quantum interference, as the right-hand side of eqn. 7.5.12 vanishes. Remote impurities—that is, located far away from both wells—would induce such a potential. However, most other scattering mechanisms, like optical phonons or interface defects, have a typically much shorter range and will act at least with a different phase on states $|\chi^0_1\rangle$ and $|\chi^0_{1'}\rangle$.

The approximation used in eqn. 7.5.12 is valid as long as the coupling to other quantum states can be neglected. In most active regions, however, the above expression must be modified to take into account the fact that the upper-state couple mainly to the *excited* state of the downstream well. A very similar model may then be used based on a set of three states with an upper-state $|\chi_{1'}\rangle$ coupled to the states of the following well.

$$|\chi_{1'}\rangle = |\chi^0_{1'}\rangle + \frac{\Omega_{1'1}}{E_{1'} - E_1}|\chi^0_1\rangle + \frac{\Omega_{1'2}}{E_{1'} - E_2}|\chi^0_2\rangle$$

(7.5.13)

We will then assume that state 1′ couples mostly to the excited state of the following well, i.e. that

$$|\frac{\Omega_{1'1}}{E_{1'} - E_1}|\langle\langle|\frac{\Omega_{1'2}}{E_{1'} - E_2}|$$

(7.5.14)

The expression for the scattering matrix element of S is then even simpler:

$$\langle\chi_{1'}|S|\chi_1\rangle = \frac{\Omega_{1'2}}{E_{1'} - E_2}\langle\chi^0_2|S|\chi^0_1\rangle$$

(7.5.15)

This approximation justifies the “tunnel-assisted” terminology used for the description of the processes occurring in such coupled wells, for the matrix elements between states 1′ and 1 may be seen as those of a single quantum well multiplied by a tunneling coupling.

(p.117) Because of the coupling induced by the tunnel barrier, the upper-state is a superposition of states 1′ and 2. The scattering matrix element is simply proportional to the amplitude of state $|\chi_2\rangle$ in the mixture of states. An important limiting case is the one where states 1′ and 2 are resonant, i.e. $E_{1'} - E_2 = 0$. In this case, the perturbation approach of eqn. 7.5.15 is not valid anymore; however, the upper-state wavefunction is composed at equal parts of state 1′ and 2 with amplitude $1/\sqrt{2}$. As a consequence, the matrix elements will all be multiplied by the same factor and the total lifetime, proportional to the square of the matrix element, will be twice that of the single quantum well.

In general, in a Fermi golden rule approach, the scattering rates are proportional to the matrix elements squared, i.e.

$$\tau^{-1} \sim |\langle \chi_{1'}|S|\chi_1\rangle|^2 = \frac{\Omega_{1'2}^2}{(E_{1'} - E_2)^2} |\langle \chi_2^0|S|\chi_1^0\rangle|^2$$

(7.5.16)

and therefore proportional to Ω^2. The latter is proportional to the intensity transmission of the tunnel barrier, i.e. to $\exp(-2\kappa_b L_b)$.

Let us take as an example a system consisting of 4 nm and 1.5 nm InGaAs wells coupled by an AlInAs barrier. The ground state of the two wells are separated by 230 meV. As shown in Fig. 7.7, both the oscillator strength and the intersubband lifetime decrease strongly with the tunnel barrier thickness; the dependence expected by eqn. 7.5.15 with the squared matrix elements varying exponentially in barrier thickness as $\exp(-2\kappa_b L_b)$ is indeed observed in the limit of thick tunnel barriers.

As is apparent in Fig. 7.7, both the oscillator strength and the intersubband scattering rate have approximately the same dependence in barrier thickness. Since the gain cross section is proportional to the ratio of these two numbers, the latter is expected **(p.118)** to remain approximately constant with barrier thickness. A closer look at Fig. 7.7, however, shows that the initial decrease of the optical phonon scattering with barrier width is slightly faster than that of the oscillator strength. As a result, a merit factor $f\tau$ of the transition, given as the product of the oscillator strength and the lifetime, increases with barrier thickness.

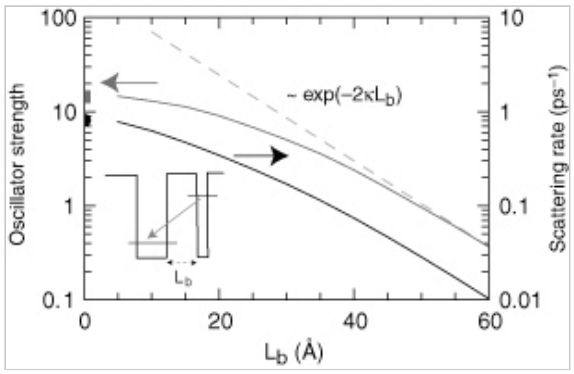


Fig. 7.7. Photon-assisted tunneling transition. Plotted are the oscillator strength $f = \frac{2m_0}{\hbar^2} z_{ij}^2 E_{ij}$ and optical phonon scattering rate as a function of barrier width for a system consisting of a 4 nm and 1.5 nm InGaAs well coupled by an AlInAs barrier. The additional points are the two values for a 6 nm

Indeed, as shown in Fig. 7.8 where this quantity is plotted as a function of the tunnel barrier width between the quantum wells for a fixed transition energy, $f\tau$ increases from 18 to 35 as the diagonality of the transition increases. This improvement of $f\tau$ was one of the justifications for the investigation of quantum cascade lasers based on a photon-assisted tunneling transition [125]

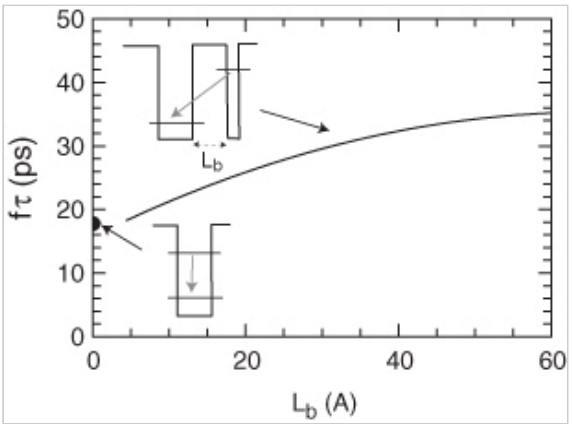


Fig. 7.8. Product of the oscillator strength and the lifetime as a function of tunnel barrier thickness. This number is a figure of merit of the differential gain cross-section. The value for $L_b = 0$ is the one computed for a single quantum well 6 nm thick, as in Fig. 7.7.

.

7.5.2 Optical phonon

As was shown in the preceding chapter, for lasers operating in the mid-infrared, the lifetimes are ultimately limited by optical phonon emission, as the other parasitic processes such as tunnel escape in the continuum can be eliminated with a good design. A *reduction* of a lifetime may be obtained by using a resonant optical phonon emission process. As shown in Chapter 5, a strong reduction of the intersubband lifetime is predicted when the transition energy approaches the optical phonon energy. Consider the structure shown in Fig. 5.4(a). In this three-level system the first two states are separated by an optical phonon energy (34 meV in InGaAs and 36 meV in GaAs). Resonant optical phonon emission between these two states will reduce the lifetime of the state $n = 2$ to about 200 fs (close to $k = 0$). Intersubband transitions between the $n = 3$ and the $n = 2$ states are less likely because they are non-resonant processes as soon as the emitted photon energy is larger than the optical phonon energy. In other words, transitions between the $n = 3$ and the $n = 2$ and $n = 1$ states will proceed **(p.119)** with large wavevectors, whereas transitions between the $n = 2$ and $n = 1$ proceed at nearly zero exchanged wavevectors.

7.5.3 Phase space

Finally, a third approach for modifying the intersubband lifetime is based on a phase space argument. Consider a superlattice formed by the periodic alternance of wells and barriers. As in its three-dimensional counterpart, this one-dimensional potential will create a set of minibands and minigaps whose width can be adjusted by modifying

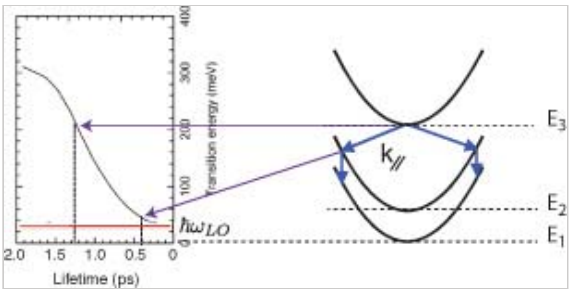


Fig. 7.9. Engineering a population inversion using an optical phonon resonance. In a ladder of three

the well and barrier thicknesses, as shown in Fig. 7.10. In an electrically pumped quantum cascade laser, the carriers are injected in the lower state of the upper miniband, and population inversion and optical gain is achieved at the edges of the minigap. Because of the nature of the electron–phonon interaction that depends on the exchanged wavevector, the intersubband lifetime in a single quantum well will be almost identical (within 20%) to that of a superlattice with a minigap of identical size. For this reason, the scattering rate between levels at the edges of the minigap will tend to *decrease* with the total number of wells in the superlattice, as the total rate must be shared between more and more lower levels. In addition, the lower state lifetime τ_2 is very short because electrons in the upper-state of the lower miniband may scatter to any point of the lower miniband very efficiently.

states, where the two lower ones are spaced by an optical phonon energy, the $n = 3$ lifetime will naturally be longer than the $n = 2$ because of the smaller exchanged wavevector for the optical phonon emission between the $n = 2$ and $n = 1$ states.

Superlattice active regions also exhibit another advantage. In these structures, since the optical transition occurs between *excited* states, its oscillator strength will be much larger than the one taking place in a single quantum well with the same transition energy.

This behavior is illustrated in Fig. 7.11, where the product of the oscillator strength and the upper-state lifetime is plotted as a function of number of period for a superlattice consisting of an alternance of 6 nm thick $Ga_{0.47}In_{0.53}As$ wells and 0.5 nm thick $Al_{0.48}In_{0.52}As$ barriers. This figure of merit increases from 22 to 50 ps as the number of periods increases from 2 to 9. Similarly, the ratio between the total upper-state lifetime and the lifetime between the edges of the minigap (called τ_3 and τ_{32} by analogy **(p.120)**

to a three-level system) decreases from 0.5 to 0.11, close to the value of $1/N$ expected for a model where the scattering rate is shared equally between lower states. Fig. 7.11 shows a continuous increase of merit $f\tau$ as a function of a number of periods. Actually, the concept of a superlattice, strictly speaking, implies an infinite number of periods. Two physical reasons prevent, in practice, increasing the number of active-region periods to a number larger than 8–10. First, as the number of periods is increased, the spacing of adjacent levels is decreasing, until the point where the states are closer than their intrinsic broadening. Upper states separated by less than kT will also decrease the injection efficiency, as the population of the upper miniband will be spread over many states. Secondly, a related issue is the electrical injection that by

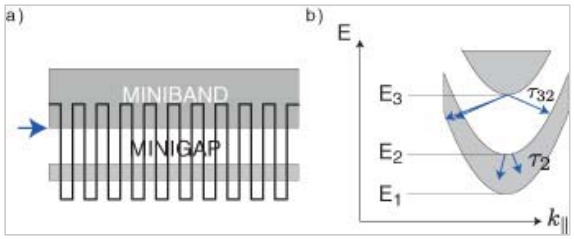


Fig. 7.10. Engineering the lifetimes. In a superlattice, a population inversion builds up at the edges of the minigap because the phase space for scattering out from the upper-state of the lower miniband is much larger than the phase space to scatter in the same state. a) Real-space and (b) reciprocal-space picture of the energy bands. The lifetime "τ_{32}#x201D;, using the notation of the three-quantum-well active region, is inefficient because the scattering from the upper-state is spread into the whole miniband.

(p.121) definition will be at an edge of the superlattice and that can only inject electrons over a maximum of a coherence length.
Intersubband lasers based on superlattice active regions have demonstrated high output powers and high operating temperatures. Some of these phase-space concepts may also be used in structures based on discrete levels.

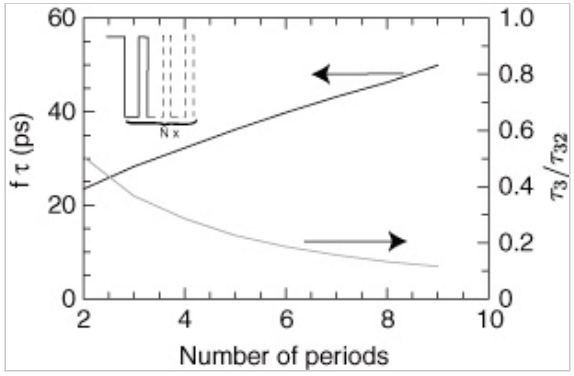


Fig. 7.11. Product of the oscillator strength and the upper-state lifetime (left axis) and branching ratio τ_3/τ_{32} (right axis) for a 60/5 Å $Ga_{0.47}In_{0.53}As/Al_{0.48}In_{0.52}As$ superlattice as a function of the number of period *N*.

A very important feature of both the phase-space argument and the tunnel effect do not depend on the exact nature of the scattering mechanism. They will therefore also be applicable to other quantum cascade laser systems such as terahertz QC lasers and QC lasers based on non-polar materials such as Si/SiGe.

7.5.4 Escape time, Bragg reflection, and upper-level confinement

Strictly speaking, under applied an electric field, all states of a quantum cascade structure are in fact resonances with a lifetime caused by the finite probability of tunneling by autoionization. This autoionization will translate into an escape time τ_{esc} from the upper-state into the continuum, reducing the total upper-state lifetime. To circumvent this problem, the confined states of the active region structure are computed by artificially bounding the structure with high barriers, transforming all the states into *real* bound states.

However, especially for states with large confinement energy, the autoionization into continuum may be relevant to the upper-state lifetime, occurring either directly or through coupling with an intermediate state further down in the epilayer stack. Preventing this escape path is especially important in the excited state of the active region, since this process will reduce artificially the upper-state lifetime below the limit set by the optical phonon scattering.

This lifetime can be computed directly by invoking Heisenberg's relationship and computing the FWHM linewidth 2Γ of the resonance, leaving the structure unbounded.

$$\tau = \hbar/\Gamma$$

(7.5.17)

The problem with this approach is that it tends to underestimate the strength of the process, since it assumes an infinite coherence length. In practice, this problem may be circumvented by computing a structure truncated to a few periods.

To design around this problem, care must be take to design the injection/relaxation region with a minigap facing the upper-state of the laser transition. Since the atten-uation of the wavefunction inside this region will be proportional to the width of the gap, it is usually a good design rule to maximize the width of the latter and to center it around the excited state of the laser transition, as shown in Fig. 7.12. In fact, it is usually not a solution to just increase the thickness of the escape

barrier, since this will reduce the escape rate of the lower state as well. When this gap has to be built in the continuum, a design rule may be used based on the Bragg reflection condition [126]. In this approach, for each well/barrier pair, the condition

$$k_w l_w + k_b l_b = \pi$$

(7.5.18)

is imposed, where $k_w = \sqrt{2m^* E}/\hbar$ and $k_b = \sqrt{2m^*(\Delta E_c - E)}/\hbar$ are the wavevector in the quantum well and the quantum barrier, respectively. The individual well and **(p.122)**

barrier thickness l_w and l_b may deviate from the exact quarter condition, since we choose to keep the effective conduction band edge of the injection region flat under the applied field. Mathematically, it translates into the requirement that the effective conduction band potential $V(x)$ of the injector at position x, approximated by

$$V(x) = \Delta E_c \frac{l_b}{l_b + l_w}$$

(7.5.19)

exhibits an effective electric field $\frac{\partial V}{\partial x} = F$ that approximatively cancels the one applied at threshold.

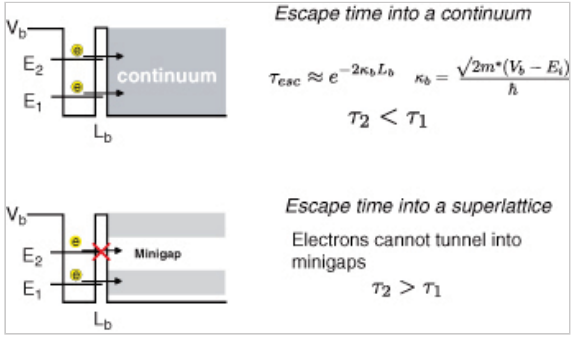


Fig. 7.12. A minigap facing the excited state of the laser transition will prevent electron escape into the continuum.

In Fig. 7.13 the electron transmission coefficient is plotted as a function of energy for an injector designed to provide Bragg reflection at the energy of the excited state E_3 = 500 meV. At this energy and given the applied electric field of $85 kV/cm$, the electron is barely confined in the well and has an energy above the barrier in the whole injection region. A transmission coefficient of 3 × 10^{-3} and 10^{-4} is found, assuming that the electron coherence is maintained over 3, respectively 5 periods, while a relatively large transmission coefficient of ≈ 10^{-1} is achieved at the energy of the lower states. The escape time therefore depends critically on the number of periods. Here we use for the escape time the expression

$$\tau_{esc} = \hbar/\Gamma = \hbar/(E_3 t)$$

(7.5.20)

where t is the transmission coefficient. For $t = 1 \times 10^{-4}$ we obtain $\tau_{esc} = 12 ps$, which is much larger than the intersubband phonon scattering time. However, if the coherence length is reduced to three well/barrier periods, the computed escape time decreases to 0.4 ps.

In Fig. 7.14 the electroluminescence intensities of a structure with and without a Bragg reflector are compared at low temperature. As demonstrated by an improvement of about a factor of 3 to 4 of the electroluminescence efficiency, it is clear that the Bragg reflector is effective in significanly reducing the electron escape time. Finally note that as the temperature is raised, the coherence length of the electron will be **(p.123)**

reduced, and the electron may also be thermally activated above the minigap,

all of which increases the electron escape rate.

7.5.5 Injection efficiencies

One could be tempted to conclude from the description above that optimization of a gain region may be reduced to the simple optimization of the computed ratio of upper- **(p.124)** to lower-state lifetimes, combined with an optimization of the intersubband oscillator strength. This picture is, however, oversimplified, because it implicitly assumes that all the electrons are initially injected in the upper-state, and, furthermore, that they are all extracted immediately from the lowest state. In fact, the situation is much more complicated by the fact that both of these processes (injection and extraction) are essentially non-ideal.

Let us assume, as shown schematically in Fig. 7.15, that a fraction η_3 of the current is injected into the upper-state, and η_2 into the lower state. Such a non-resonant injection originates from elastic or inelastic scattering between the injector state and the lower states of the active region. In this case the expression for the material gain g reads:

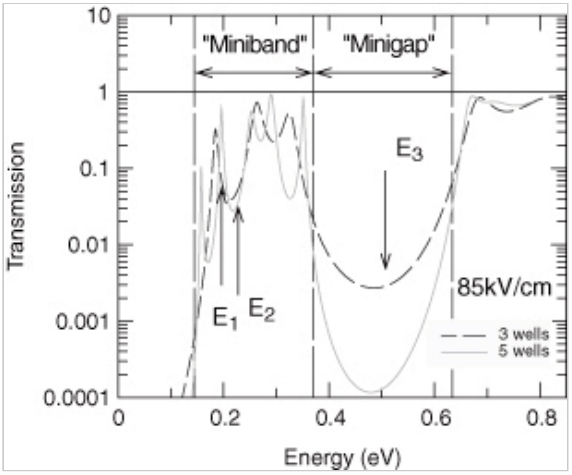


Fig. 7.13. Computed transmission of an injector designed to provide maximum reflection at the energy of the excited state.

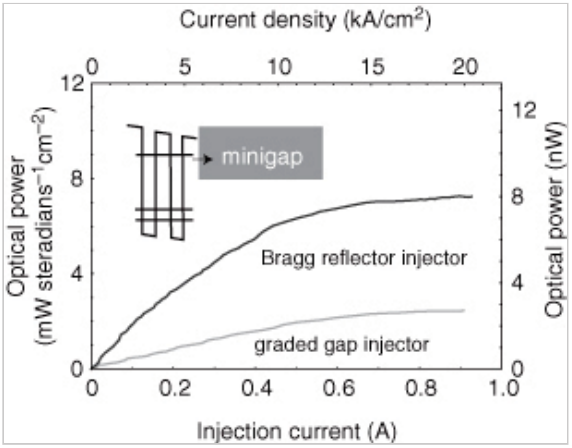


Fig. 7.14. Comparison of the electroluminescence efficiency of a structure designed with and without a Bragg reflector facing the upper-state.

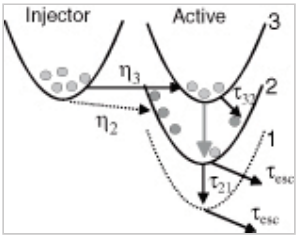


Fig. 7.15. Schematic description of the relevant levels and injection efficiencies in a cascade laser.

$$g = g_c \Delta n = g_c \frac{J}{q_0}(\tau_3 \eta_3 (1 - \frac{\tau_2}{\tau_{32}}) - \eta_2 \tau_2)$$

(7.5.21)

The influence of the injection efficiencies η_3 and η_2 are clearly not symmetric. The gain is simply scaled by the value of η_3, whereas, depending on the ratio of τ_2/τ_3, the value of η_2 can have a much more dramatic consequences of cancelling the gain altogether. Clearly, although the design should in principle aim at reaching injection efficiencies as close to the ideal values ($\eta_3 = 1$ and $\eta_2 = 0$) as possible, the detrimental effects of non-ideal injection efficiencies may be mitigated by a system with a large lifetime ratio (τ_2/τ_3). In the case of a quantum cascade laser based on a superlattice active region, eqn. 7.5.21 can still be applied with the $n = 3$ state being the lower state of the upper miniband and the $n = 2$ the upper-state of the lowest miniband. Eqn. 7.5.21 shows clearly the strong dependence of the gain on both injection efficiencies and the ratio of the lifetimes: one must design a structure which maximizes the difference $\eta_3 \tau_3 - \eta_2 \tau_2$ and minimizes the ratio τ_2/τ_{32}.

7.6 Optimization of the active region: different designs

7.6.1 Designs: general trends

As shown schematically in Fig. 7.16, after the demonstration of the first working laser, the studies first aimed at testing the validity of the different components of the intersubband toolbox individually, i.e. whether a quantum cascade laser could be built based only on tunneling, optical phonon resonance, or phase space. In later studies **(p.125)** these different strategies were combined in a single design in order to optimize the various components of the threshold current density appearing in eqn. 7.4.8, as shown in Fig. 7.17. The result of this optimization has led to the focus, for the mid-infrared on roughly two or three designs, that will be presented here.

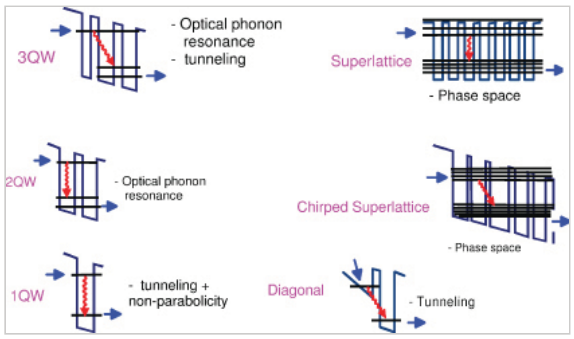


Fig. 7.16. Schematic diagram of the first generation of quantum cascade laser active region designs.

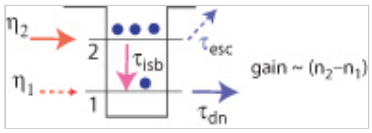


Fig. 7.17. Schematic of the parameters optimized in the active regions.

7.6.2 Three-quantum-well active region

We will start our discussion of our different QC laser designs with an evolution of the original structure designed for operation at 4.3 μm wavelength [20]. This design has some unique features: first of all it is the one which allows the shortest wavelength to be reached for a given band discontinuity; by combining the optical phonon resonance with the possibility of a diagonal transition it enables a good ratio of upper-state to lower-state lifetime. A schematic description showing how such an active region design can be decomposed is shown in Fig. 7.18, where the “bare” states, the states of the quantum well before coupling by the tunnel barrier, are shown along the “dressed” states. The active region is based on a pair of quantum wells (see Fig. 7.18a), designed in such a way that the ground state of the thinner well is brought in resonance

with the ground state of the thicker well. The tunnel barrier is then adjusted such that the energy splitting matches the optical phonon resonance (see Fig. 7.18(b)). Because the **(p.126)**

excited states gain energy faster with quantum well thickness, the excited states will not be in resonance. A third quantum well is then added upstream. Two possibilities exist. In the first one, the state of this additional quantum well is resonant with the $n = 3$ state of the quantum well, and will anticross with it as shown in Fig. 7.18c. In this case the transition is strongly diagonal and, as a result, the lifetime will be approximatively twice that of the vertical transition. If, on the contrary, the state of this thin well is clearly above the $n = 3$ state, see Fig. 7.18d, the transition will have the lifetime and oscillator strength of a vertical transition, but the extension of the wavefunction in the injection barrier will improve the injection efficiency. A schematic band diagram of both structures, designed for operation at $\lambda = 10.3\ \mu$m and $\lambda = 5.3\ \mu$m respectively, are shown in Fig. 7.19. The choice of a diagonal transition at longer wavelength allows retention of a good ratio of upper- to lower-state lifetime even for a relatively low transition energy of 120 meV. The diagonal nature of the laser transition is responsible for the lifetime (1.6 ps at low 80 K) longer by a factor of two as compared to a vertical transition at the same energy. In contrast, the lifetime of the upper-state of the vertical transition device is approximatively the same as that of the square quantum well. While the oscillator strength of the diagonal structure ($f = 16$) is smaller than that expected for a vertical transition at the same energy, it is about the same as that of the vertical transition at 5.2 μm because of non-parabolicity effects. These devices operated above room temperature, up to $T = 330$ K with 200 mW high peak power [32].

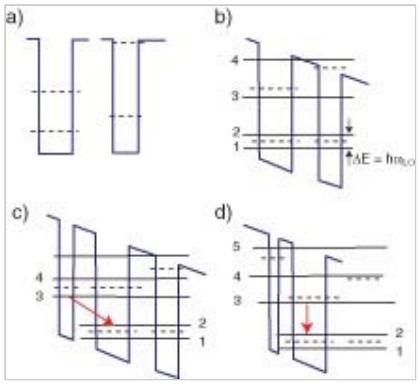
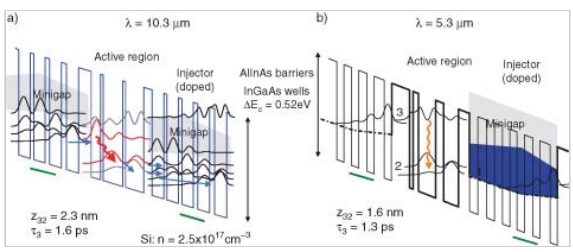


Fig. 7.18. Schematic potential profile and design of the active region of a quantum cascade laser based on three quantum wells. a) Two-quantum well before the application of the electric field. b) The electric field brings the two ground states in resonance and yields a splitting equal to the optical phonon energy. A third thinner well is added upstream. If the state of this well is resonant with the excited state of the coupled well, the resultant transition is diagonal (c), if the well is thinner and the resonance is above, the transition is vertical (d).

7.6.3 Double phonon resonance

One weakness of active regions based on three quantum wells is the extraction from the ground level of the active region. In fact, studies on diagonal transitions showed **(p.127)**

that the extraction from the active region, due to resonant tunneling, proceeded on a timescale of about 2 ps. This relatively long lifetime, compared to the fast scattering between level $n = 2$ and $n = 1$ of the active region, enabled electrons to be scattered back to the $n = 2$ state by



multiple scattering events between states. As a result, the population inversion was more difficult to achieve, degrading the high-temperature and high-power performances of these devices.

Fig. 7.19. Examples of active regions based on three quantum wells, using lattice-matched InGaAs/AlInAs material. Structure a) is designed for a wavelength of $\lambda = 10.3\ \mu$m using the concept shown in Fig. 7.18c, which leads to a diagonal transition with a longer lifetime. The structure is shown in Appendix A.1.6. Structure b) is designed for $\lambda = 5.3\ \mu$m and exhibits a vertical transition, as explained in Fig. 7.18d [32]. The structure is shown in Appendix A.1.7.

One attempt to solve this problem has been the design based on so-called two- phonon extraction [22]. In the two-phonon resonance design, the extraction is obtained by a ladder of three states separated by an optical phonon each. According to our computation, the addition of this extra level decreases significantly the population of the lower state. For a structure emitting at $\lambda \approx 9\ \mu$m, the computed distribution show a much reduced population of the lower laser state. A schematic conduction band diagram of one period of such an active region is shown in Fig. 7.20. The upper and

(p.128) lower lasing states are the wavefunctions with numbers 4 and 3, respectively; and the ground state of the injector is denominated with the letter "g". The active region was composed of four QWs, which results in three coupled lower states (levels 1, 2, and 3). Each two of the latter (i.e. level 3 and 2, and level 2 and 1) were separated by one phonon energy. This double phonon resonance yielded a short intersubband electron scattering time, and therefore an efficient extraction of the electrons into the injector region. The upper lasing state exhibits a much longer intersubband electron scattering time of $\tau_{up} = (\tau_{43}^{-1} + \tau_{42}^{-1} + \tau_{41}^{-1})^{-1} = 0.54$ ps (at $T = 300$ K) where the lifetimes contain both emission and absorption processes. The relatively large dipole matrix element, $\langle z_{43} \rangle = 3.0$ nm, confirms that the lasing transition is mainly a vertical one. Due to the thin first well, which reduces the overlap of the injector ground state g with the lower lasing state wavefunctions 1, 2, and 3, the injection efficiency was kept similarly high as in the classical three-quantum-well design. The design thus takes advantage of the good properties of both the three-quantum-well design (high injection efficiency) and the superlattice design (short lifetime of the lower lasing state).

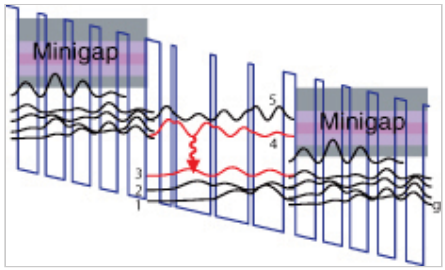


Fig. 7.20. Schematic band diagram of a two-phonon resonance gain region designed for operation at $\lambda \approx 9\ \mu$m. The structure is shown in Appendix A.2.1.

Low-temperature ($T = 80$ K) luminescence spectra were taken from a device in which the light was extracted from a 45-degree wedge polished in the substrate. These spectra, taken for various injection currents, are shown in Fig. 7.21 along with the calculated transition energies. All the transitions in the active region can be easily

identified. Our computation show that the dipole matrix elements between states 4–1 and 5–1 are always very small ($z_{51}, z_{41} < 0.1nm$), regardless of the applied electric field. Peaks related to transitions from level 5 can

be observed only experimentally in the strong injection regime. The computed dipole matrix element z_{42} decreases strongly with applied field, as observed in Fig. 7.21. In general, the luminescence peaks shown in Fig 7.21 confirms the location and strength of the respective intersubband transitions **(p.129)** .

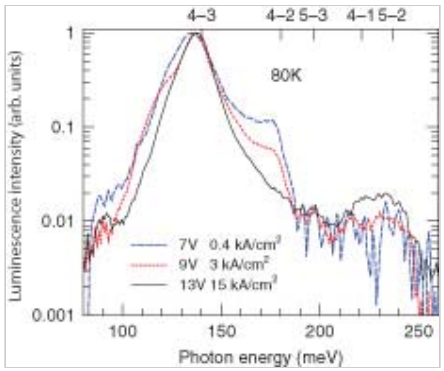


Fig. 7.21. Low-temperature ($T = 80$ K) luminescence spectra of the active region based on a two-phonon resonance. Calculated transition energies are shown on the top horizontal axis.

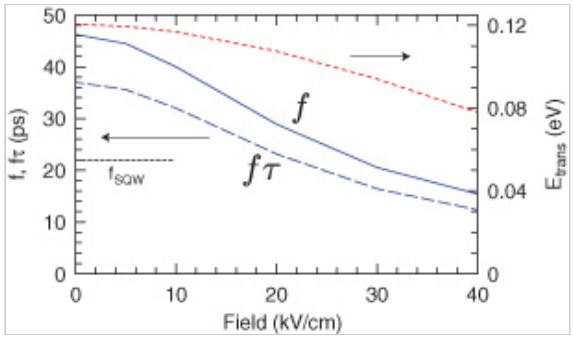


Fig. 7.22. Oscillator strength *f*, product of oscillator strength and upper-state lifetime $f\tau$ (right axis), and transition energy *Etrans* (right axis) as a function of applied electric field for a superlattice consisting of five periods of a 59 Å InGaAs quantum well and a 9 Å AlInAs barrier. The value of oscillator strength for an unbiased single quantum well is shown for comparison.

7.6.4 Bound-to-continuum active regions

Besides the two-phonon resonance approach, the bound-to-continuum approach provides another route for maintaining a high population inversion and low threshold current densities even at high temperatures. The bound-to-continuum active region can be seen as an evolution of the active regions based on superlattices. As described above, a superlattice exhibits two very attractive features: a very large oscillator strength and a favorable ratio of lifetimes at the edge of the minigap.

However, active regions of quantum cascade lasers operate under a strong applied electric field. The latter, applied on a periodic superlattice, would break the miniband of the superlattice into a set of localized states, losing all the favorable features of the extended states. An example of the effect of an applied field on the product of upper-state lifetime and oscillator strength is shown in Fig. 7.22.

For this reason, the first quantum cascade lasers based on a superlattice were doped homogeneously inside the active region to screen the field. This technique, however, presented the disadvantage of introducing additional broadening and loss caused by the ionized dopants. As a result, operation at room temperature of these devices was strongly limited. A very efficient technique to solve this problem was proposed by Tredicucci and consisted in “chirping” the superlattice to artificially compensate the applied electric field. Numerically, it consisted in changing the barrier and well widths of the superlattice such as to keep the energy of the miniband edges constant even in the presence of the applied electric field. This procedure can be

performed numerically by solving iteratively eqn. 3.4.74 for a well and barrier pair of thickness L_i and h_i, solving for the first minigap edges occurring at $qd = \pi$:

$$\cos(k_w L)\cos(k_b h) - \frac{1}{2}(\xi - \frac{1}{\xi})\sin(k_w L)\sin(k_b h) = 1$$

(7.6.22)

(p.130) where k_w and k_b are computed at an energy $E(z) = E_0 + \sum_j \langle_i e(L_j + h_j)\mathcal{E}$. The procedure must be stopped when the excess energy E at the upper band edge approaches close to the discontinuity. The result of this procedure is shown for the chirped superlattice demonstrated by A. Tredicucci [127] in Fig. 7.23 and compared to the same computation performed for a bound-to-continuum structure.

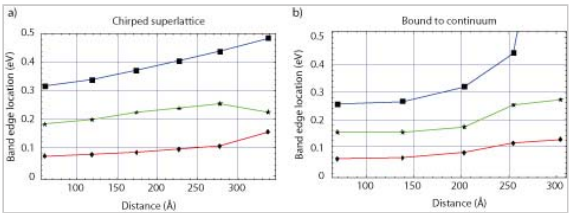


Fig. 7.23. Three first miniband edges as a function of position for a chirped superlattice (a) and for a bound-to-continuum structure (b).

In the latter design, shown schematically in Fig. 7.24, the active region spans the whole period and consists of a chirped superlattice presenting a tilted lower miniband whose width is maximum in the center and decreases on both sides close to the injection barriers. The upper-state is created in the first minigap by a small well adjacent to the injection barrier. Its wavefunction has a maximum close to the injection barrier and decreases smoothly in the active region. This upper-state is well-separated from the higher-lying states of the superlattice, lying in its first minigap. It therefore does not need to be confined by separating the structure into an active region and an injection/relaxation region. Because of this large energy separation (60 meV), the injection efficiency η_3 is not reduced by electron injection into higher energy states of the superlattice [128, 129]. In addition, this configuration of the wavefunction should enable an injection efficiency comparable to the one of the three-quantum-well design.

As a result of the localization of the upper-state, the miniband selection rules are somewhat relaxed, and the oscillator strength is now spread between the upper-state and a number of lower-states. This feature is apparent in the electroluminescence spectrum show in Fig. 7.24 along with the computed matrix elements. This design architecture was found to offer many interesting features:

High-temperature performance The devices shown in Fig. 7.24 operated up to 150° C with a very low temperature dependence on the threshold current, translated in a T_0 parameter as high as 190 K.

Long-wavelength operation Because of its good lifetime ratio, this design performs very well at long wavelengths. Above room temperature, operation at wavelengths as long as $\lambda = 16\ \mu$m were achieved.

(p.131)

Broadband gain Because a single upper-state shares a number of lower states, this design is very well adapted

for the achievement of gain over a broad wavelength range.

Non-linear generation Bound-to-continuum active regions have been shown to have a combination of mid-infrared gain and terahertz susceptibility, enabling terahertz generation by intracavity non-linear mixing up to room temperature [131].

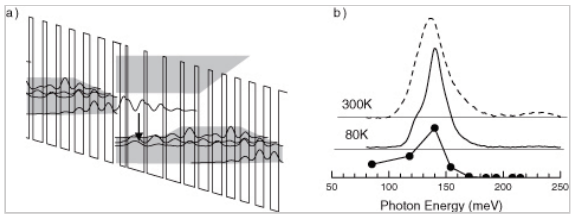


Fig. 7.24. a) Schematic conduction band diagram of one stage of a bound-to-continuum active region under an applied electric field of $3.5 \cdot 10^4$ V/cm. The moduli squared of the relevant wavefunctions are shown. b) Luminescence spectrum of the active region at 300 and 80 K, as indicated. The applied bias is 9 V. Lower curve: computed oscillator strength of the various transitions from the upper-state. The design is shown in Appendix A.2.2. Reprinted with permission from [130]. Copyright 2001, AIP.

7.7 Cascading: scaling with the number of periods

As discussed above, the possibility of cascading the active regions is a key advantage of intersubband lasers. In the simple rate equation model, the number of periods appear in both the threshold current density and slope efficiency.

7.7.1 Threshold current density

The total overlap factor between the optical mode and the gain region, Γ, is a function of the number of periods of the structure,

$$\Gamma = \sum_{p}^{N_p} \Gamma_p$$

(7.7.23)

This expression can be simplified to $\Gamma = \Gamma_p N_p$, when Γ covers only a central part of the optical mode, i.e. when the optical intensity is constant over the active region periods and therefore the Γ_p are almost identical.

As shown in Fig. 7.25 where the overlap factor is plotted as a function of the number of periods for a device operating at $\lambda = 9\ \mu$m with an active region period length of 60 nm, this approximation is valid as long as the active region thickness is much smaller than the mode width. For typical InP-based waveguides, it is verified **(p.132)** for $N_p < 20$ for lasers operating at $\lambda \approx 5\ \mu$m, and $N_p < 30$ for lasers operating at $\lambda \approx 9\ \mu$m. In this regime the threshold current density is simply inversely proportional to the number of periods, as the reduction of the current density thus relies simply on a geometrical increase of the gain region interacting with the optical mode. As the laser active regions are electrically pumped in a series configuration, the amount of carriers injected per period therefore do not vary with their number. Adding more period does not affect the population inversion in each single stage, but it increases the applied voltage. This situation is very different from that of conventional diode lasers, where by increasing the number of quantum wells in the active region one has also to increase the current to reach threshold because a larger active region volume must reach transparency.

When more stages are added, Γ_p decreases, as the total overlap factor is naturally constrained by the condition $\Gamma < 1$. In practical applications a good practice is to keep the number of periods

such that the total overlap factor is in the range 0.5 〈 Γ 〈 0.8.

7.7.2 Slope efficiency

In a simple picture the electrons are simply recycled from period to period, emitting a photon at each step. As a result, and as predicted by eqn. 7.4.9, the slope efficiency of a quantum cascade laser is proportional to the number of stages. In a simplified model this remains true even when the optical field is not constant over the various periods, i.e. when the Γ_p are not constant. Intuitively, this can be understood by the fact that the slope efficiency inherently does not depend on Γ.

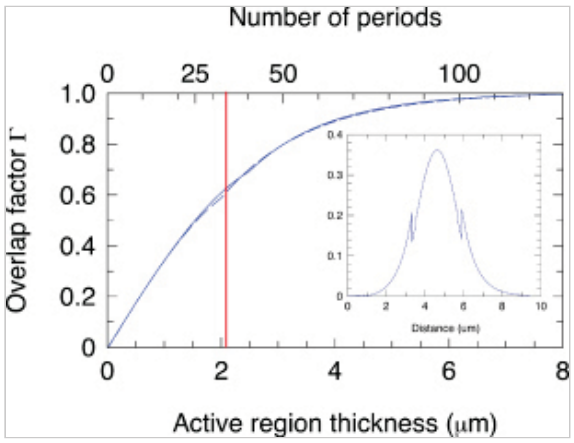


Fig. 7.25. Overlap Γ as a function of the number of periods for a device operating at $\lambda = 9$ μm. Inset: mode profile for an active region thickness of $t = 2$ μm.

However, a more detailed analysis shows that in the latter case a hole burning occurs in the center of the waveguide and the slope efficiency is renormalised by a factor η_v equal to

(p.133)

$$\eta_v = \frac{\left(\sum_{i=1}^{N_p} \Gamma_i\right)^2}{N_p \sum_{i=1}^{N_p} \Gamma_i^2}$$

(7.7.24)

This factor is η_v 〉 0.9 in common waveguides.

The dependence of the threshold current density, operating voltage and slope efficiency has been verified in experiments in which quantum cascade lasers were grown and tested with a number of stages varying between 3 and 75 [132]. As shown in Fig. 7.26, a very good agreement between the simple theory shown here and the experiments was found

.

7.8 Temperature dependence

A key feature of mid-infrared quantum cascade lasers, especially when compared to semiconductor lasers operating at similar wavelengths, is the weak dependence of the threshold current density as a function of temperature. As in conventional semiconductor lasers, the threshold current density $J_{th}(T)$ is a temperature-dependent parameter

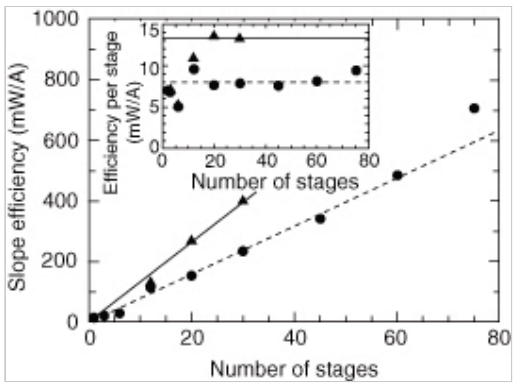


Fig. 7.26. Slope efficiency as a function of the number of periods in a series of otherwise identical

that can be conveniently fitted by the expression

$$J_{\mathrm{th}}(T) = J_0 \exp(\frac{T}{T_0})$$

(7.8.25)

where T_0 is a parameter that expresses the temperature dependence of the threshold current density.

quantum cascade lasers. Triangles correspond to data obtained in continuous wave mode, the circles to pulsed data. The lines are the best fit to the data. The discrepancy between the two lines is due to the different collection efficiencies of the two measurements: 100% for continuous wave, and 60% for pulsed. Inset: slope efficiency per stage. Copyright 1999 IEEE, Reprinted, with permission, from [132].

As shown in Fig. 7.27 where the light-versus-current plot of a quantum cascade laser based on a bound-to-continuum transition operating at $\lambda = 9\ \mu$m is shown as a function of temperature, the threshold current density varies according to eqn. 7.8.25 **(p.134)** with a value of T_0 as large as 190 K. Both the large value of T_0 as well as the fact that the device operates up to such a high temperature of 150° C is unique to quantum cascade lasers. An interband laser operating at similar wavelengths, based on lead salts, exhibited maximum operating temperatures of 200 K with T_0 values around 30 K

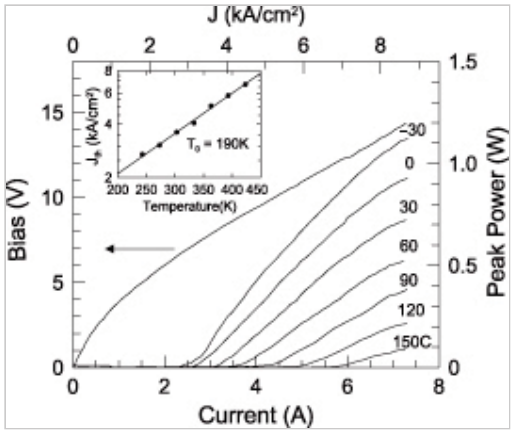


Fig. 7.27. Temperature dependence of the light-versus-current characteristics of a quantum cascade laser based on a bound-to-continuum transition operating at $\lambda = 9\ \mu$m. The design is described in Appendix A.2.2. Inset: Threshold current density as a function of temperature, fitted by eqn. 7.8.25 with a T_0 of 190 K. Reprinted with permission from [133]. Copyright 2001, AIP.

7.8.1 Intersubband physics

Theoretically, the weak temperature dependence of the threshold current density is expected because of two inherent features of intersubband transitions. First, unlike interband devices, the joint density of states of intersubband transition is atomic-like. For this reason and as discussed in detail in Chapter 5, if one neglects non-parabolicity, the temperature dependence of the broadening of the gain curve originates only from the dependence of the in-plane scattering mechanisms. As a result, the broadening parameter γ in the equation of the threshold current 7.4.8 exhibits a weak temperature dependence. Experimentally, this can be observed by the small change in the linewidth of the luminescence shown in Fig. 7.24 that increases only from about 20 meV at 80 K to 24 meV at 300 K.

The second reason, also discussed in Chapter 5, is the fact that the main intersubband scattering mechanism is the optical phonon scattering. For this reason, the upper-state lifetime is expected to decrease by a factor of approximatively two between cryogenic temperature and 400K. If a design

displaying a large ratio of upper to lower lifetime is choosen, the threshold is expected to display approximately the same temperature dependence as the τ_{eff} appearing in eqn. 7.4.8.

(p.135) 7.8.2 Backfilling

Another important term in the expression of the threshold current density is the thermal population of the lower state n_2^{therm}. As shown schematically in Fig. 7.28a, this population originates from carriers thermally excited to the lower state from the injection region. The population n_2^{therm} must be compensated by the same upper-state population to achieve transparency, and therefore should be minimized in a given design. In a simplified model the thermal population can be approximated by a simple thermal activation term at an electronic temperature T_{el}

$$n_2^{\text{therm}} = n_g \exp(-\Delta_{inj}/kT_{el}) \tag{7.8.26}$$

where n_g is the sheet density of the injector and Δ_{inj} the energy difference between the lower state and chemical potential of the injector. The effect of varying the value of Δ_{inj} for a given design is shown in Fig. 7.28b, where the threshold current density versus temperature for a series of quantum cascade lasers with injectors of various lengths show different values of Δ_{inj}. The detrimental effect of a low value of Δ_{inj} on the high-temperature operation is especially apparent. Considered in detail, the computation of n_2^{therm} is complex because the electronic temperature T_{el} is in general larger than the lattice temperature and is a strong function of injected current. Therefore, assuming a constant value of n_2^{therm} as a function of injected current density J is a rather crude approximation.

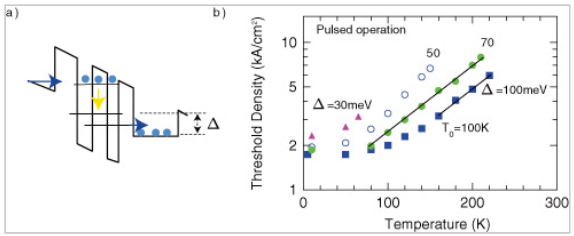


Fig. 7.28. a) Electrons thermally excited from the injector populate the lower state. The critical energy is the difference Δ between the energy of the injector and the lower state. b) Threshold current density versus temperature for a series of quantum cascade lasers with an identical active region based on a vertical transition operating at $\lambda = 4.6\ \mu$m (Design described in Appendix A.1.2) but with injectors of various lengths showing different values of Δ_{inj}.

7.8.3 Self-heating

To reach threshold of operation, quantum cascade devices must dissipate a power density amounting to 20–50 kW/cm^2. For this reason, especially when operating in **(p.136)** continuous wave, one cannot assume an active-region temperature T_{act} equal to the temperature of the submount T_{sub}. In a simple model the active-region temperature is assumed to be constant and related to the submount temperature by a single thermal resistance R_{th}:

$$T_{act} = T_{sub} + R_{th} U I \tag{7.8.27}$$

where U and I are the operation voltage and current. Expressing the above relation at threshold, and assuming that the threshold current density has the usual exponential dependence in temperature with a characteristic temperature T_0 given by eqn. 7.8.25, we obtain

$$T_{act} = T_{sub} + \frac{U_{th} J_0 \exp(T_{act}/T_0)}{G_{th}}$$

(7.8.28)

where the specific thermal conductance G_{th} in $\mathrm{Wcm^{-2}K^{-1}}$ has been used. The above equation defines naturally a maximum for T_{sub}, which can be obtained by setting $dT_{sub}/dT_{act} = 0$:

$$T_{sub,max} = T_0[\ln(\frac{T_0 G_{th}}{J_0 U_{th}}) - 1]$$

(7.8.29)

The above equation defines the maximum operating temperature in a continuous wave. A large value of the latter requires the combination of:

- A large value of T_0.
- A low intrinsic threshold power $U^{th} J^0$.
- A high specific thermal conductance G^{th}.

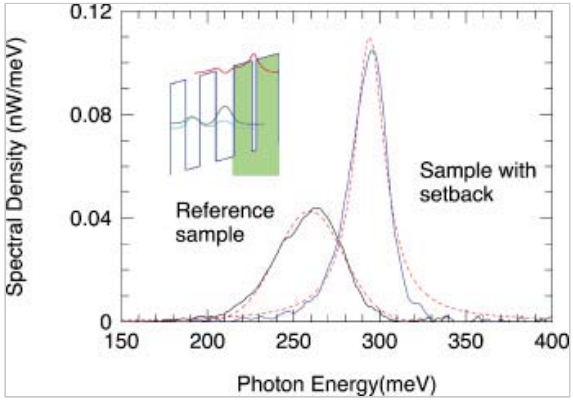
The above criteria have driven the research in quantum cascade lasers. While the first two items drove the development of new active region designs and low-loss waveguides, the last point is concerned with thermal engineering of the active region.

7.9 Doping of the active region

Active region doping is a "necessary evil" in a quantum cascade laser, as to a large extent the electrical stability of the system depends on having a structure displaying global electrical neutrality in each period. In order to minimize the scattering induced by the presence of ionized impurities, the latter are inserted as far as possible from the active region. The dramatic broadening effect of doping impurities inserted directly into the active region is shown in Fig. 7.29. However, when the dopants are remote from the active region (in practice about 20 nm away from it), no direct relationship between injector doping density and linewidth is observed, in agreement with a linewidth dominated by the interface roughness effect. We saw in Chapter 6 that the main loss mechanism of a waveguide is the free carrier absorption. In a simple model where the optical losses are dominated by the active region one, we therefore expect the waveguide losses to scale with the doping of the active region. Similarly, to the extent that the lifetimes are not limited by electron–electron scattering, and therefore are independent of electron density, we expect the current–voltage characteristic to scale linearly with doping.

Such a simple picture is well supported by the experiments. A device with an active region based on a two-phonon resonance was grown, and the doping of the injector was **(p.137)**

varied between a sheet carrier density of $n_s = 1 \times 10^{11}$ $\mathrm{cm^{-1}}$ and $n_s = 2.5 \times 10^{11}$ $\mathrm{cm^{-1}}$. As expected from this simple model and shown in Fig. 7.30, both maximum operating current *and* the threshold current increased linearly with doping. As the maximum current is linear in the sheet carrier density, we can write it as characterized by a single transport time at resonance τ_{trans}:



$$J_{max} = \frac{n_s e}{\tau_{trans}}$$

(7.9.30)

Therefore the slope of the J_{max} versus n_s yields the inverse transport lifetime. The transport lifetime extrapolated from the data shown in Fig. 7.30 yields a transport time of τ_{trans} = 4.3 ps. This time takes into account the sum of the time in the active region and the one of the transport in the injector. Similarly, one can express the threshold current density 7.4.7 in a simple form:

Fig. 7.29. Comparison of the electroluminescence from two otherwise identical samples with different doping profiles. In the sample with the setback, only the injector region is doped. The design is described in Appendix A.1.1. In the reference sample the area marked in grey in the schematic band structure in the inset is also doped. The dotted lines are line fits, with a Gaussian profile for the reference sample and a Lorentzian profile for the device with setback.

$$J_{th} = \frac{e\alpha^*_{tot}}{g_c \tau_{eff}}$$

(7.9.31)

where we now included the ground state reabsorption into $\alpha^*_{tot} = \alpha_{tot} + g_c n_2^{\text{therm}}$. Expressing the total losses α^*_{tot} as a linear expression in n^s:

$$\alpha^*_{tot} = \alpha_m + \alpha_w + g_{fc} n_s$$

(7.9.32)

where α_m are the mirror losses, α_w are the losses of the empty waveguide, and g_{fc} is the "free carrier" cross-section. As in the case of the maximum current, the slope of the J_{th} versus n_s characteristics yields

(p.138)

$$\frac{dJ_{th}}{dn_s} = \frac{e g_{fc}}{g_c \tau_{eff}}$$

(7.9.33)

Assuming the effective upper-state lifetime can be computed, the above expression enables the computation of the ratio of the loss-to-gain cross-section g_{fc}/g_c. As we will see in the next paragraph, this number and the ratio of τ_{eff}/τ_{trans} are the key numbers characterizing the fundamental performance limitations of an active region. Applying this analysis on the data of Fig. 7.30 and using τ_{eff} = 0.55 ps, one obtains a ratio of g_{fc}/g_c = 1.9×10^{-2}.

Of course, this simple picture of the effect of doping breaks down if

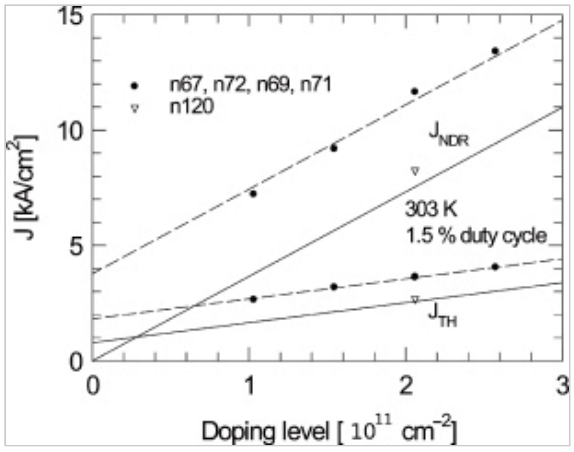


Fig. 7.30. Pulsed, room temperature threshold current density J_{th} (lower curve) and maximum current density J_{NDR} (upper curve) as a function of nominal active region sheet doping for a series of lasers grown in sequence. Solid disks, experimental measurements; dashed lines, fitted curves corresponding to an extrapolated doping offset of

extremely large doping levels are used. In that case, the Hartree potential introduces large distortions in the conduction band of the injector and significantly alters the alignment between injector and active region levels.

$n_{offset} = 1.0 \times 10^{11}$ cm^{-2} added to the active region doping. Open triangles represent J_{th} and J_{NDR} of a laser grown after a longer MBE running time. Reprinted with permission from [40]. Copyright 2006, AIP.

7.10 Wallplug efficiency

The optical power of a laser can be scaled, up to a certain point, by merely increasing the volume of the gain medium. At the same time, the electrical power needed to drive the laser will change accordingly. For this reason, the wallplug efficiency of a laser η_{wp}, defined simply as the ratio of the extracted optical power to the injected electrical power

$$\eta_{wp} = \frac{P}{UI}$$

(7.10.34)

(p.139) is a number that characterizes the scale-invariant properties of a quantum cascade laser. It is also a very useful quantity for the design of systems involving QCLs.

7.10.1 Fundamental dependencies

The simple atomic model developed in this chapter, combined with considerations of the doping dependence developed above, enables us to express the fundamental limitations on the wallplug efficiency of quantum cascade lasers. Neglecting gain saturation and assuming a constant slope efficiency, we have

$$\eta_{wp} = \frac{(J_{max} - J_{th})}{J_{max}} \frac{dP/dI}{U}$$

(7.10.35)

As in the preceding paragraph, we will assume that the maximum operating current can be defined using a global transport time across one period τ_{trans} defined by eqn. 7.9.30. The operating voltage U at resonance can be easily written as

$$U = N_p(\frac{\hbar\omega}{e} + \Delta_{inj})$$

(7.10.36)

Using for the slope efficiency eqn. 7.4.9, and for the threshold eqn. 7.4.7, we can finally combine these expressions and obtain

$$\eta_{wp} = \eta_{tr}\frac{\alpha_m}{\alpha_{tot}}\frac{1}{1+\Delta_{inj}/(\hbar\omega)}[1 - \frac{\tau_{trans}}{\tau_{eff}}(\frac{\alpha_{tot}}{n_s N_p g_c} + \frac{n_{therm}}{n_s})]$$

(7.10.37)

where η_{tr} is the differential efficiency of the laser transition, given in a three-level model of the active region by

$$\eta_{tr} = \frac{\tau_{eff}}{\tau_{eff} + \tau_2}$$

(7.10.38)

For any given material and design, the parameters on which the design will be optimized are the doping level n_s, the number of periods N_p, and the mirror loss α_m. We are obviously interested in the maximum efficiency achieved for optimized values of these parameters. As in the previous paragraph, we assume that the waveguide losses are dominated by the doping of the active region, i.e. by the term $g_{fc}\, n_s$ in eqn. 7.9.32. We assume also that Δ_{inj} is sufficiently large that the thermal population of the lower laser state n_2^{therm} is negligible. Under those assumptions, the maximum value of the wallplug efficiency, i.e. computed for the value of the mirror loss α_m for which Eqn. 7.10.37 is maximum, takes an especially simple form:

$$\eta_{wp,max} = \eta_{tr} \frac{1}{1 + \Delta_{inj}/(\hbar\omega)} \left[\frac{\sqrt{g^* \tau^*} - 1}{\sqrt{g^* \tau^*}} \right]^2$$

(7.10.39)

In this equation 7.10.39, we have introduced the reduced dimensionless upper-state lifetime $\tau^* = \tau_{up}/\tau_{tr}$ and cross-section $g^* = g_c/g_{fc}$. The fact that eqn. 7.10.39 does not depend on either the doping level or number of periods justifies a posteriori our assumption that losses from the active region could be assumed to dominate.

In eqn. 7.10.39 the product $\sqrt{g^* \tau^*}$ is a key parameter determining the wallplug efficiency: a value lower than 1 indicates a device that will never reach threshold, as **(p.140)** the free carrier absorption will always be larger than the intersubband gain. In fact, the optimization of an active region should, of course, try to maximize this parameter.

A large value of τ^* is achieved in an active region where transport is very fast in the injector and most of the carriers are remaining on the upper-state of the laser transition. An inspection of the parameter g^* demonstrates the very strong wavelength dependence of the wallplug efficiency. In fact, assuming a Lorenzian lineshape for the gain cross-section characterized by an in-plane lifetime broadening $\tau_{\square}$ and a normalized oscillator strength of the laser transition:

$$f' = f \frac{m^*}{m_0}$$

(7.10.40)

(where m^* is the carrier's effective mass). Taking similarly the free carrier absorption to originate from intersubband absorption within the injector states (assumed at a vanishing transition energy) characterized by a broadening lifetime τ_{fc}, the dimensionless cross-section is

$$g^* = \frac{1}{2} \omega^2 \tau_{\|} \tau_{fc} f'$$

(7.10.41)

The intrinsic ω^2 dependence of eqn. 7.10.41 stresses the increasing difficulty of successfully fabricating quantum cascade lasers operating at low frequencies. Expressing the normalized cross-section g^* using eqn. 7.10.41 removes all explicit dependence on the wallplug efficiency from eqn. 7.10.39 in the effective mass. Mathematically, it is simply because both the gain cross-section *and* the free carrier absorption are proportional to the inverse effective mass. In the mid-infrared there is, however, a very strong indirect dependence of the wallplug efficiency in the effective mass through the dependence of both dephasing times $\tau_{\square}$ and τ_{fc} in m^*. The linewidth of an intersubband transition is limited by the interface roughness [59, 134]. Thus, for any given interface structure (described by the step height Δ and the correlation length Λ in Ref [134]) and band discontinuity ΔE_c, the lifetime $\tau_{\square}$ is proportional to $(m^*)^{-2}$. The effective mass dependence

of the free carrier absorption is not as easy to express because it depends on the dominant mechanism of scattering (interfaces, phonons, or impurities). As shown in Fig. 7.31, the simplifed expression given by eqn. 7.10.39 predicts the wavelength dependence of the wallplug efficiency. In the latter plot, an efficiency of the transition $\eta_{trans} = 0.7$ and a reduced value of the lifetime $\tau^* = 0.11$ were used. The value $\Delta_{inj} = 150$ meV, obtained by an optimization of the full eqn. 7.10.37 was used. In the spirit of the discussion of the interpretation of the free carrier absorption as originating mainly from intersubband transitions in the active region (see Section 9.3), $\tau_{fc} = \tau$ □ was assumed, and the value of the wallplug efficiency was plotted as function of photon energy for three values of τ □.

One limitation of these studies is that the most relevant physical number, the wallplug efficiency measured in pulsed mode, is also the one most difficult to measure experimentally, as discussed in Chapter 11. Continuous wave data, usually more accurate, are further limited by self-heating effects. In a continuous wave the maximum efficiency must be achieved at the minimum possible current density in order to minimize self-heating effects. In that case, both doping and mirror losses are minimized, and the optimum value depends strongly on the residual waveguide loss α_{scatt}.

(p.141)

The model above is overly simplified, and a few factors not taken into account could allow devices with higher wallplug efficiencies than that predicted by Fig. 7.31. First of all, whereas the upper-state lifetime is clearly limited by optical phonon scattering in the long-wavelength limit, elastic scattering plays an increasingly important role as the wavelength is shortened. The latter is critically dependent on the growth conditions and in particular the growth temperature [82], and on the choice of the interface construction. Secondly, the effect of photon-driven transport on the maximum current was not taken into account; in a high performance device the maximum current is *increased* by the photon-driven transport, and both effects will improve the value of g^*. Finally, engineering the free carrier absorption by redistributing the oscillator strength, as described in Section 9.3, will increase the value of the normalized gain coefficient g^*.

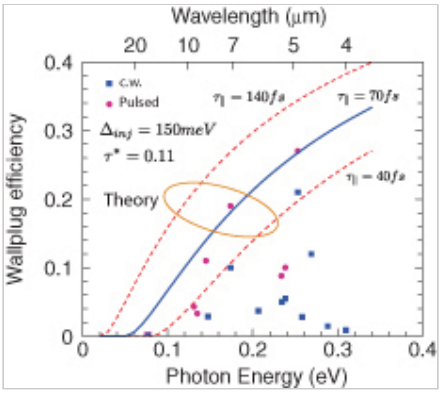


Fig. 7.31. Wallplug efficiency as a function of photon energy. Filled squares, experimental values in continuous wave, from references [28, 135–142]. Solid disk: pulsed values from references[40, 121, 139, 141–145]. The solid line represents the prediction from the simplified formula with the key parameters derived from [40]. The influence of a shorter or longer in-plane dephasing time τ □ is shown by the two dotted curves.

7.10.2 High-wallplug-efficiency devices

The practical optimization of these parameters in order to achieve the best possible wallplug efficiency has been the subject of intense research activity. In a relatively early study performed using $\lambda \approx 5$ μm wavelength device, an initial optical power level of 0.7 W with an 8% wallplug

efficiency was achieved at cryogenic temperatures [146]. These early results showed the potential of these devices and helped promote further research in the field. The most recent results report up to 25% wallplug efficiency [142] and multi-watt output power in a continuous wave at room temperature [147, 142].

As described in the preceding paragraph, these designs are attempts to enhance the values of g^*, τ^*, η_{tr} while minimizing the effect of backfilling by choosing a well adapted value of Δ_{inj}.

In order to maximize τ^*, a strong reduction of the transport time across one period was attempted by the use of very short active regions with a reduced number of active **(p.142)** quantum wells [148–151]. As in terahertz devices, the use of very short active regions with large applied electric field introduced problems with the overall electrical stability of the device and backfilling was difficult to avoid, limiting the performances at high temperatures.

Another approach has been the use of "strong coupling" devices, where the use of a very thin injection barriers [152] should increase the maximum operating current for a given doping level. However, in the short wavelength range, the injection barrier is not the only limitation to the current flow as the need to cool the electron distribution is a significant portion of the transport time. For a device with a potential drop of 380 meV, the fastest possible cooling occurs when each electron emits an optical phonon approximately every 0.2 ps. The electron cooling time for the non-radiative transport, in this limit, is already 2.2 ps, yielding a maximum current density of 7.2 kA/cm^2 for a sheet density of 1×10^{11} cm^{-2}.

Fig. 7.32 shows the computed band diagrams of three active regions that recently reported high values of wallplug efficiencies [147, 153, 154]. The band structure was computed using a Schrödinger–Poisson solver. As discussed in the next chapter, these devices operating in the short wavelength band use strain-compensated materials to achieve a large enough conduction band discontinuity. The active region thicknesses and discontinuities were obtained directly from the relevant publications, sometimes by measuring directly on the graphs. The average electric field chosen corresponds to the one at which the maximum wallplug efficiency was computed.

The active region was separated in two sub-units after the fifth well, and this basis was chosen to compute the light and voltage–current characteristics of the devices using a model described in section 12.3. Some key numbers computed using this model are compared in Table 7.1. In contrast to the numbers reported in the original literature, the lifetime reported here includes all non-radiative channels, including interface roughness and alloy scattering.

7.1 Key parameters of the three high-wallplug-efficiency designs described in Fig. 7.32. The energies as given in and the lifetimes in ps.

	λ(μm)	$\eta_{wp}^{(exp)}(\%)$	$\eta_{wp}^{(th)}(\%)$	E_{exc}	Δ_{inj}	τ_{up}	τ_{dn}	f	f′
esonant	4.6	15.3	15.5	66	117	0.29	0.07	19	0.8
diagonal	4.9	15	15	71	125	0.39	0.11	11.2	0.4
ow well	4.9	27	29	100	124	0.48	0.06	6.8	0.2

The first one (Fig. 7.32a) is described as a non-resonant extraction design in which one first level is spaced one phonon energy below the lower laser state, and a further doublet 60 meV below. In fact, this structure can also be seen as a bound-to-continuum active region with a relatively thick (100 meV) lower miniband. This is especially clear when one considers the shape of the wavefunction obtained by solving the potential in one unseparated block. They spread relatively uniformly over the whole length of the device. Of the three devices, it is the one that has the largest oscillator strength ($f' = 0.81$) and shortest upper-state lifetime. As in the other two devices the energy **(p.143)** separation with the next excited state E_{exc} is kept large ($E_{exc} \rangle\rangle kT$) to prevent leakage into through this state. The second one (Fig. 7.32b) is a diagonal three-quantum-well active region design, as discussed in Section 7.6.2, designed in a very similar way as that shown in Figs. 7.18c and 7.19a. Actually, in the article presenting this design, this device was compared to a more vertical version of the same active region (similar to Fig. 7.19b) and shown to exhibit much higher performances [153]. Compared to the non-resonant design, it exhibits a lower oscillator strength $f' = 0.48$ but a longer upper-state lifetime (0.39 ps). Moreover, this one does not have a "bound-to-continuum" character, as **(p.144)**

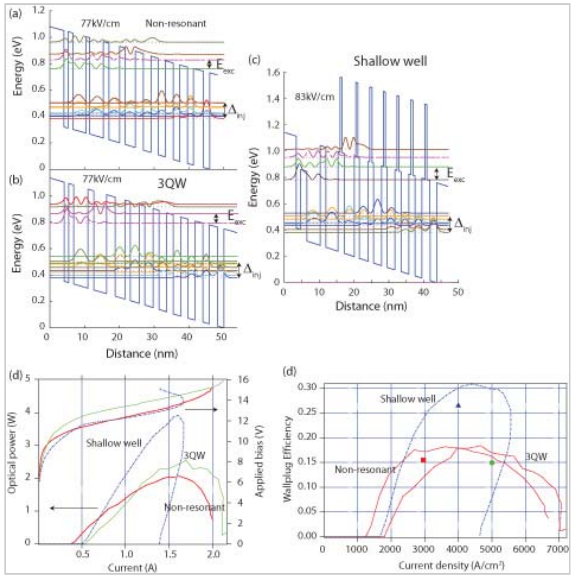


Fig. 7.32. Schematic diagram of the active region diagram with high wallplug efficiency a) non-resonant extraction [147], b) three quantum well [153], and c) shallow well designs [154]. d) Computed light and voltage versus current for a 10 μm wide, 3mm long HR coated device with 35 periods of the active region. e) Comparison of the computed wallplug efficiencies as a function of injected current. The symbols represent the experimental reported values.

the wavefunctions of the active region and of the injector are clearly separated. This feature is also responsible for the longer lower-state lifetime of the device (0.11 ps). The third one (Fig. 7.32c)is described as a shallow quantum well active region. The general design, as for the non-resonant device, can again be seen as a bound-to-continuum architecture with a relatively diagonal transition ($f' = 0.29$) and a narrower lower miniband (60 meV) at the location of the active laser transition, widening in the injection/relaxation region. The active region uses five different compositions of the In-Ga-Al-As material; the use of a reduced discontinuity at the location of the upper-state minimizes the negative

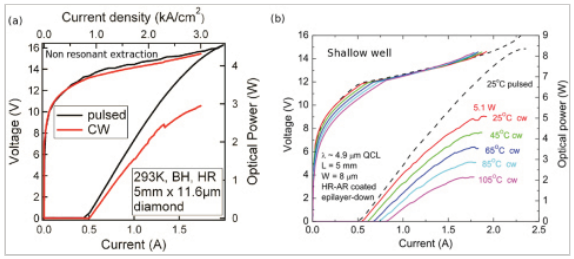


Fig. 7.33. Lightand bias-versus current characteristics for two high-performance devices at room temperature. a) Active region based on a non-resonant design, and b) on a shallow well. Reprinted with permission from [142] and [147]. Copyright 20109, 2011, AIP.

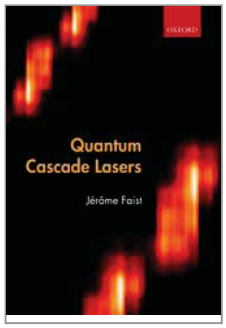
influence of the interface roughness scattering. Consequently, it is a design that combines the longest upper-state lifetime (0.48 ps) and the shortest lower-state lifetime (60 fs).

Fig. 7.32d shows the computed light-versus-current and bias-versus-current characteristics of the three devices for a 3 mm long, 10 μm wide cavity with a backside high-reflectivity coating. Each device was assumed to consists of 35 periods of the active region with an overlap factor of Γ = 0.7. The facet reflectivity was assumed to be R = 0.22, and an additional waveguide loss due to the cladding and fabrication of 0.3 cm^{-1} was assumed. As discussed in Sections 9.3 and 11.5 the free carrier absorption associated with the active region is fully taken into account by the model directly. The model predicts large optical powers in the multi-watt range already for these relatively short devices. The very large ratio of upper to lower lifetimes enables a very strong photon-driven transport of the shallow well design, and enables the very high optical power and wallplug efficiency achieved by this device.

The wallplug efficiency of these three devices (this time neglecting the additional 0.3 cm^{-1} loss) as a function of current are compared in Fig. 7.32e. Compared to the three-quantum-well active region, the non-resonant design has the advantage of achieving its highest wallplug efficiency at a lower current, enabling much better continuous wave performances. The shallow well device demonstrates the best operating performance. **(p.145)** Note that the reported experimental wallplug efficiencies correspond well to those predicted by the model.

In addition to the intrinsic performances of the active region, the wallplug efficiency in a continuous wave depends in a critical way on the thermal resistance of the device. For this reason, sophisticated processing and mounting techniques have to be used. Fig. 7.33 shows two light and bias characteristics as a function of injected current, in a continuous wave, for both the non-resonant extraction device (Fig. 7.33a) and for the shallow well device (Fig. 7.33b). A continuous wave power up to 5 W is achieved in continuous wave for a single-ridge device.

The very strong photon-driven transport is evident from the shape of the current– voltage curve that displays a very strong discontinuity at threshold.



Quantum Cascade Lasers
Jérôme Faist

Print publication date: 2013
Print ISBN-13: 9780198528241
Published to Oxford Scholarship Online: May 2013
DOI: 10.1093/acprof:oso/9780198528241.001.0001

Short-wavelength QCLs

Jérôme Faist

DOI:10.1093/acprof:oso/9780198528241.003.0008

Abstract and Keywords

This chapter discusses short wavelength quantum cascade lasers covering conduction band discontinuity and performance, heterostructure materials, and strain-compensated $In_xGa_{1-x}As/Al_yIn_{1-y}As/InP$ material system.

Keywords: quantum cascade lasers, conduction band discontinuity, heterostructure materials, quantum wells

8.1 Conduction band discontinuity and performance

In quantum cascade lasers the ability to tailor the photon energy is limited, on the high-energy side, by the size of the conduction band discontinuity between the two semi-conductor materials. In the case of the $In_{0.53}Ga_{0.47}As/In_{0.52}Al_{0.48}As$ material which is lattice matched to InP and has a conduction band discontinuity $\Delta E_c = 520$meV, lasers with photon energies up to $h\nu = 290$ meV ($h\nu = 0.56 \cdot \Delta E_c$) were realized [155], although high-temperature operation ($T \rangle$ 300 K) was achieved typically for photon energies up to $h\nu = 250$ meV ($h\nu = 0.48 \cdot \Delta E_c$). The reason for this limitation is shown schematically in Fig. 8.1. Electrons from the upper-state of the laser transition can be thermally activated to continuum states. This process is exponential in temperature and will severely limit the high-temperature operation. A rough estimate of this parameter can be obtained by assuming an escape time τ_{esc} of the form

$$\tau_{esc}^{-1} = \tau_{cont}^{-1} \exp(-E_{act}/kT)$$

(8.1.1)

where the escape rate in the continuum $\tau_{cont} \approx 100fs$ is estimated from the width of a typical continuum resonance. These quantities are very difficult to estimate because

(p.147) they depend in a critical manner on the coherence length chosen. Under these assumptions, an activation energy of E_{act} = 85meV, usually chosen as the minimum value needed for a good room-temperature operation, yields an escape time of τ_{esc} = 2.7 ps.

An experiment [156] using devices fabricated from the $Al_xGa_{(1-x)}As$/GaAs material system illustrates this effect very nicely, and uses the fact that GaAs and $Al_xGa_{(1-x)}As$ are lattice matched for all Al contents x. In this work, two otherwise identical quantum cascade lasers are compared—one using a barrier with an aluminium content of 33%, and the other one with 45%. In Fig. 8.2 are reported the threshold current density versus temperature for both devices. As expected, both devices operate with identical threshold currents at low temperature. However, as the temperature is raised above about 100 K, the device with the lower conduction band discontinuity of about 300 meV exhibits a much worse temperature performance characterized by a T_0 of 75 K, while the device with the larger discontinuity of 390 meV exhibited a T_0 of 166 K, close to the theoretical value.

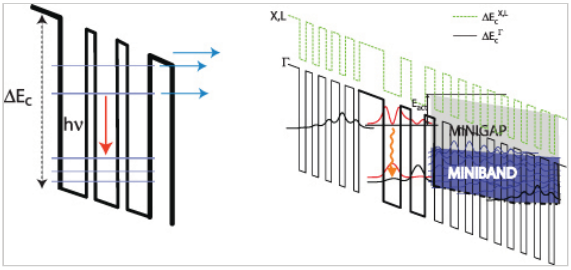


Fig. 8.1. Schematic band structure of a lattice-matched InGaAs/AlInAs QCL based on a vertical transition operating at $\lambda \approx 4.6\ \mu$m. The design is described in Appendix A.1.2. The high-temperature operation of such a device is limited by the thermal escape from an electron either to continuum states of the Γ point or states in the lateral X and L valleys.

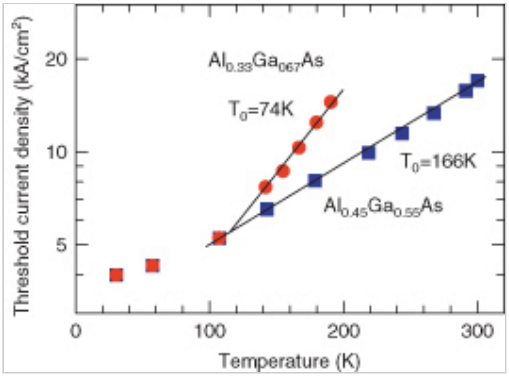


Fig. 8.2. Threshold current density versus temperature for two GaAs-based quantum cascade lasers, operating at $\lambda \approx 9\ \mu$m, that use two different aluminum mole fractions (x = 0.33 and x = 0.45). A clear relationship between discontinuity and high temperature behavior is apparent. (Redrawn using data from [156].)

These considerations are well supported by experimental evidence. In Fig. 8.3 and Table 8.1, the characteristic temperature T_0 is shown as a function of the emitted photon energy for quantum cascade lasers fabricated using materials with various conduction band discontinuities.

For devices limited by optical phonon emission and exhibiting a good population inversion, the characteristic temperature should lie between 400 K (assuming a broadening independent with temperature) and 170 K (for typical measured values of the broadening of the electroluminescence peak). From the data shown in Fig. 8.3, it is clear that the theoretical value of $T_0 \approx 170$ K, limited by optical phonon emission and the broadening of the intersubband

transition, can be achieved for large enough discontinuities, though a limited discontinuity rapidly limits high-temperature operation. **(p.148)**

8.1 Characteristic temperature T_0 for operation around room temperature for quantum cascade lasers fabricated various material systems.

	ΔE_c^{Γ} (meV)	E_{phot} (meV)	λ (μm)	T_0 (K)	reference
/$Al_{0.33}Ga_{0.67}As$	300	107	11.5	96	[157]
/$Al_{0.45}Ga_{0.55}As$	390	120	9.3	166	[156]
As/AlInAs	520	108	11.5	170	[113]
As/AlInAs	520	138	9	190	[133]
As/AlInAs	520	159	7.8	164	[158]
As/AlInAs	520	238	5.2	114	[32]
$Ga_{0.3}As/Al_{0.6}In_{0.4}As$	800	288	4.3	188	[136]
$Ga_{0.3}As/Al_{0.6}In_{0.4}As$	800	355	3.49	85	[159]
$Ga_{0.27}As/AlAs$	1000	326	3.8	143	[160]
$Ga_{0.3}As/AlAs(Sb)$	1600	375	3.3	101	[161]
$Ga_{0.3}As/AlAs(Sb)$	1600	344	3.6	116	[161]
AlSb	2100	375	3.3	200	[162]

8.2 Heterostructure materials

From the above considerations it is clear that quantum cascade lasers operating at wavelengths shorter than $\lambda \approx 5 - 6\ \mu$m require the use of a material system with a larger discontinuity as the one (ΔE_c = 520meV) provided by the lattice-matched $In_{0.53}Ga_{0.47}As/Al_{0.48}In_{0.52}As$ material. For quantum cascade lasers based on GaAs, the limit is much more severe since, as shown in the previous paragraph, lack of conduction band discontinuity starts to hamper the operation of quantum cascade lasers with wavelengths shorter than $\lambda \approx 9\ \mu$m. For this reason a good deal of interest has been carried over other material systems that would provide a larger discontinuity. The conduction band discontinuity of the main III–V semiconductor materials (to the exception of group III-nitrides) are shown in Fig. 8.4.

Another important parameter in the choice of the material system is the position of the lateral valleys. In the III–V materials the lateral valleys are characterized by a large electron effective mass and therefore large density of states. As a result, the upper-state of the laser transition must remain energetically far enough below the lateral valley to prevent signifficant current leakage and a degradation of the laser performance. The detrimental effect of the lateral valley current leakage in $GaAs/Al_{0.33}Ga_{0.67}As$ QCLs was demonstrated in high-pressure experiments by Jin and coworkers [163].

Fig. 8.5 shows the position of the Γ, X, and L points as a function of Ga and Al composition in pseudomorphic $Ga_xIn_{1-x}As$ and $Al_xIn_{1-x}As$ respectively. It is clear from this graph that large indium and aluminum mole fractions in the $Ga_xIn_{1-x}As$ and $Al_xIn_{1-x}As$ quantum well and barrier materials are desired, and allow large Γ-(X, L) separation.

From the above considerations, one very good candidate for short-wavelength application is the InAs/AlSb material system grown on InAs substrates. Both the very large conduction-band discontinuity (ΔE_c = 2.1 eV and the very light effective conduction band mass m^*/m_0 = 0.023 should enable both very large optical gain and short-wavelength operation. In addition, the lateral L and X valleys are also very high **(p.149)**

(p.150) in energy (about 0.8eV) from the bottom of the conduction band of InAs. Indeed, devices grown using this material system hold the present record for short wavelength operation (λ = 2.9 μm) as well as very high performance at λ = 3.6 μm. Compared to the device based on the InP system, this material system suffers from less favorable refractive index and thermal conduction and the epitaxial growth is more difficult to control. For this reason, another very successful material system has been the straincompensated $In_xGa_{1-x}As/Al_yIn_{1-y}As/InP$.

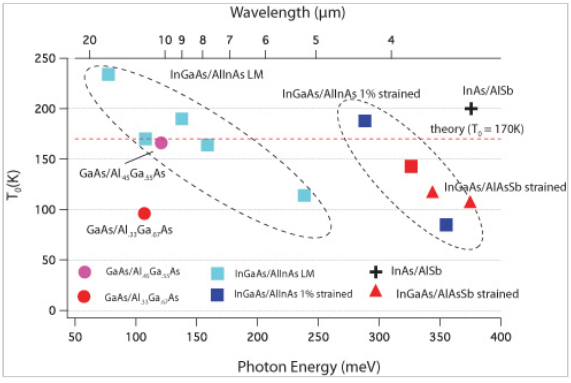


Fig. 8.3. Characteristic temperature T_0 for various quantum cascade lasers, as a function of the photon energy and for various material systems. A clear relationship between discontinuity and temperature behavior is apparent.

8.3 Strain-compensated $In_xGa_{1-x}As/Al_yIn_{1-y}As/InP$ material system

The initial research in optoelectronics demonstrated that high crystalline quality, therefore the absence of dislocations, was a prerequisite for the achievement of semiconductor lasers with high performances and long lifetimes. Such a requirement also implied that the heterostructure materials had to be carefully lattice-matched so that bulk layers would not develop dislocations due to strain between epitaxial layers. It therefore came as a surprise that thin layers of strained material could be grown with *no detrimental effect on material quality* as long as the cumulated strain remained below the level needed to destabilize the two-dimensional growth. Using this approach, strained InGaAs/GaAs quantum wells were used to achieve interband lasers at wavelengths close to $\lambda = 1\ \mu$m with excellent performance. These results could be achieved because the active region material thickness was kept very small—a few quantum wells of typical 10 nm thickness. The active region of a quantum cascade laser is also based on thin quantum wells, and therefore is also compatible with such an approach. If strain with one sign is used only, for example, for compressive strain in a InGaAs/GaAs/AlGaAs material system, it is the cumulated strain required by the large number of active region periods that will strongly limit the amount of strain that can be introduced in each layer. If a material system is chosen where strain of both **(p.151)** signs is possible—one for the barriers and the other for the quantum wells—imposing the condition of zero accumulated strain on one active region periods implies that an arbitrary number of active region periods can be grown.

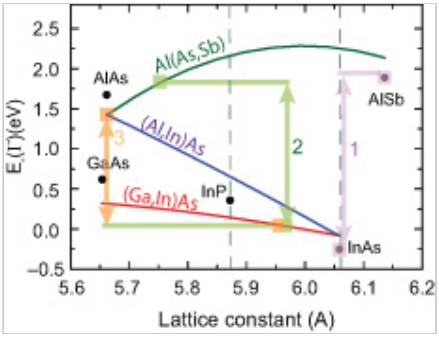


Fig. 8.4. Γ-valley conduction band edges for the three most used material systems for short-wavelength QCLs. (**1**) InAs/AlSb, (**2**) Strain-compensated InGaAs/AlAsSb, (**3**) straincompensated InGaAs/InAlAs. The position of the band edges as function of the lattice parameter are considered for a material grown pseudomorphically on InP, to the exception of the material system (**1**) InAs/AlSb that is considered as grown on InAs.

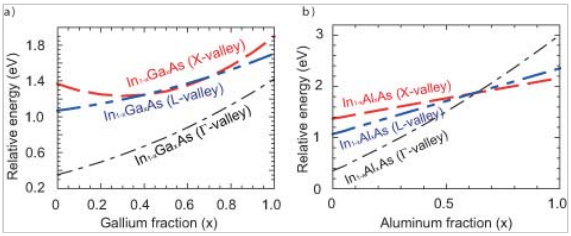


Fig. 8.5. Γ, X, L conduction band edges for $Gax In_{1-x}As$ and $Al_x In_{1-x}As$ as function of the Gallium and Aluminum mole fraction, computed using the parameters from [164, 165].

Using compressively strained $In_x Ga_{1-x}As$ and tensile-strained $Al_x In_{1-x}As$, such that the total strain on an InP substrate is zero, a material combination with a number of attractive features is realized.

- The conduction band discontinuity can be increased by about 200–250 meV as compared to the lattice-matched case.
- The Γ–X discontinuity also increases with compressive strain in the InGaAs.

- InP-based waveguides can still be used, with their low losses and high thermal
- conductivity.
- The amount of strain can be chosen according to the target wavelength.

Using this material system, very-high-performance quantum cascade lasers have been realized from about $\lambda \approx 3.8\ \mu$m to $\lambda \approx 7\ \mu$m [141]. The conduction-band discontinuity can be tailored by changing the In content in the well and barrier material in a concomitant way. The conduction-band discontinuity can then be computed using van de Walle model solid theory [164]. In this approach an absolute energy scale is rigorously defined from pseudopotential band computation. Deformation potentials are then introduced to compute the new location of the bands. Taking the total thickness of the well material h_w and barrier material h_b for one period, imposing lattice-matching to the substrate means that the lattice constant of the heterostructure at mechanical equilibrium, given by

$$a_{\parallel} = \frac{a_w G_w h_w + a_b G_b h_b}{G_w h_w + G_b h_b}$$

(8.3.2)

is equal to the lattice constant of InP (a_0= 5.87 A). In eqn. 8.3.2 the indices w and b refer to the well and barrier materials and will be referred to as $i = w, b$ in the following equations. As each layer now under biaxial strain, its lattice constant in the perpendicular direction also changes and is now given by

$$a_{i\perp} = a_i[1 - D_i(a_{\parallel}/a_i - 1)]$$

(8.3.3)

As a result, in-plane strain components are given by

$$\epsilon_{i\perp} = \frac{a_{i\perp}}{a_i} - 1$$

(8.3.4)

and an out-of-plane component is given by

$$\epsilon_{i\parallel} = \frac{a_{\parallel}}{a_i} - 1$$

(8.3.5)

In eqn. 8.3.2 the shear modulus G_i is related to the elastic constants c_{ij} by

$$G_i = 2(c_{11}^i + 2c_{12}^i)(1 - D_i/2)$$

(8.3.6)

where, for a growth in the [001] direction, the values of D_i are given by

(p.152)

$$D_i^{001} = 2\frac{c_{12}}{c_{11}}$$

(8.3.7)

As the values of elastic constants are not tabulated for the alloys, the usual approach is to interpolate those linearly from the values of the binaries. The key values needed for the model presented here are listed in Table 8.2.

8.2 Some elastic constants, energies and deformation potentials relevant for In-GaAs/AlInAs material system. [164]).

a_0(Å)	G^{001} (M bar)	$E_{v,av}$(eV)	E_c(eV)	Δ_0(eV)	a_v(eV)	a_c^{dir} (eV)
5.65	2.522	–6.92	–5.29	0.34	1.16	–7.17
5.65	2.656	–7.49	–4.27	0.28	2.47	–5.67
6.08	1.587	–6.67	–6.13	0.38	1.00	–5.08
5.87	1.897	–7.04	–5.58	0.11	1.27	–5.04

The biaxial strain has an hydrostatic component, responsible for the change of the crystal volume Ω. The fractional change in volume ΔΩ/Ω is given by the trace of the strain tensor

$$\frac{\Delta\Omega}{\Omega} = Tr(\overleftrightarrow{\epsilon}) = (\epsilon_{xx} + \epsilon_{yy} + \epsilon_{zz}) = 2\epsilon_{i\parallel} + \epsilon_{i\perp}$$

(8.3.8)

Because of its symmetry, the minimum of the conduction band is only sensitive to the hydrostatic component of the strain, and the change in the conduction-band energy is given by the deformation potential a_c:

$$\delta E_c = a_c \frac{\Delta\Omega}{\Omega}$$

(8.3.9)

To compute the discontinuity, the conduction band positions are first interpolated from the binaries for the unstrained material, the values then corrected by the shift caused by the strain using the above equation:

$$\Delta E_c = E_c^b + \delta E_c^b - E_c^w - \delta E_c^w$$

(8.3.10)

The effect of strain is shown schematically in Fig. 8.6, along with the effect on the various bands of the InGaAs/AlInAs material system. The conduction band discontinuity for AlInAs/GaInAs heterostructures for In concentrations between 0 and 1 and pseudomorphically grown on InP substrates is shown in Fig 8.7a as contour plot. As can be seen in the graph, the conduction band discontinuity is much more sensitive to a variation of the In concentration in AlInAs than in GaInAs. The design of a quantum cascade laser structure should therefore be compensated by setting the composition of $Al_xIn_{1-x}As$ and choosing a composition of $Ga_xIn_{1-x}As$ that compensates the strain and not vice versa. The compositions that yield materials lattice-matched to the InP substrate are indicated as black dashed lines, and the resultant conduction band discontinuity is $\Delta E_c \approx 0.52$ eV. The compositions used in [159] are indicated as dashed grey lines, and the resultiant ΔE_c is 0.78 eV. As shown in Fig. 8.7 the discontinuity between the two materials can be increased from 520 meV to about 900 meV, depending on the amount of strain added to each layer. The latter cannot be increased **(p.153)**

indefinitely, as the critical thickness above which the crystalline coherence cannot be maintained is strongly decreasing with increasing strain. Another important consequence of strain is its strong effect on the effective mass of the material. The hydrostatic component will reduce the bandgap, while the biaxial one will split the heavy and light-holes degeneracy at the Γ point. As the effective gap of both conduction light-hole and conduction split-off changes, so does the effective **(p.154)**

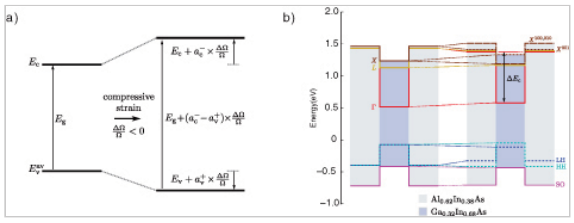


Fig. 8.6. a) Illustration of the energy shift of the band edges due to compressive strain. b) $Ga_{0.32}In_{0.68}As$/ $Al_{0.62}In_{0.38}$ As heterojunction showing the band energies when strain effects are neglected (left) and taken into account (right).

mass of the conduction band as a result of the **k** · **p** interaction. Fig. 8.8 shows the effective mass of the $In_xGa_{1-x}As$ material as a function of Indium content x, taken from the work of Sugawara *et al*. [166]. The mass relevant for the computation of the eignenstates, as well as for the computation of the dipole matrix elements, is perpendicular mass $m_\perp$. For increasing indium content, the latter is actually increasing before decreasing again for indium concentrations larger than about 80%. Fig. 8.9 shows the schematic band diagram of a quantum cascade laser designed with $In_{0.7}Ga_{0.3}As/Al_{0.6}In_{0.4}As$ strained compensated on InP. The band structure was computed using effective masses of unstrained material. As a result, the computed emission wavelength $\lambda = 3.16\ \mu$m is signifficantly shorter than the one measured in the actual devices, $\lambda =$ 3.4–3.6 μm. Using the masses from Suwagara *et al*, a value of $\lambda \approx 3.4$ μm—much closer to the experimental one—is predicted. The laser emission spectrum of these devices is shown in Fig. 8.9. The relatively low $T_0 = 85$ K measured on these devices demonstrated that the upper-state exhibited a lifetime limited by thermoionic emission above the barrier.

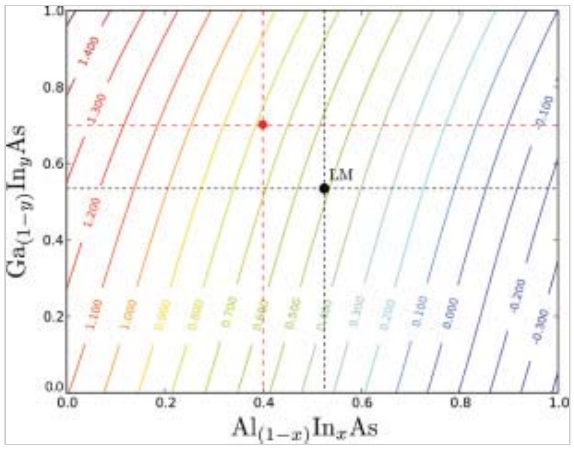


Fig. 8.7. Contour plot of the conduction band discontinuity of the GaInAs/AlInAs material system for In concentrations between 0 and 1. Effects of strain are taken into account assuming that thin layers of material are grown on a InP substrate. The black dot represents the latticematched composition resulting in $\Delta E_c \approx 0.52$ eV, and the grey dot represents the materials used in the first strain-compensated quantum cascade laser.

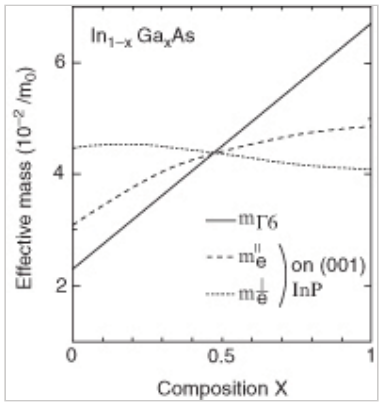


Fig. 8.8. Electron effective mass of $In_{1-x}Ga_xAs$ material as a function of the Ga content x (in contrast to the text where x labels the indium content). Full line: unstrained material. Dotted, dashed lines: in-plane, respectively out of plane mass for material grown on InP. Reprinted with permission from [166]. Copyright 1993, APS.

8.3.1 Lattice-matched and strain-compensated $In_xGa_{1-x}As/Al_yAsSb/InP$

Another possibility is to substitute the $Al_{0.48}In_{0.52}As$ material with AlAsSb. As shown in Fig. 8.4, this pair of materials enables large conduction band discontinuities even with lattice-matched InGaAs quantum wells, simplifying greatly the design of devices. In fact, this approach was pioneered by the group led by J. Cockburn [167], achieving emission down to a wavelength of 3.05 μm [168]. However, for short-wavelength operation, the best results were achieved by using strain-compensated materials. In

particular, above-room-temperature operation was achieved between 3.3 and 3.5 μm using a $Ga_{0.7}In_{0.3}As$/AlAsSb/InP strain-compensated material system [161]. Further- **(p.155)**

more, it was also found that the Sb-containing interfaces were inferiors to the As ones; as a result, AlAs barriers were used instead of AlAsSb in the active region of the latter device.

8.3.2 Strain-compensated $In_xGa_{1-x}As/Al_yIn_{1-y}As$/AlAs/InP

In the $Ga_xIn_{1-x}As/Al_yIn_{1-y}As$ strain-compensated material system, the necessity to keep the barrier thickness below the critical thickness limits the maximum aluminum fraction that can be used. A technique introduced by the group led by T. Masselink[169] is the growth of composite $Al_{0.48}In_{0.52}As$/AlAs barriers combined with a relatively high Indium content quantum well $Ga_{0.27}In_{0.73}As$. In this way, the benefit of the high barrier can be conserved and the strain-compensation can be handled in a way independent of the heterostructure design. Lasers operating at cryogenic temperatures down to wavelengths as short as 3.05 μm at 80 K were achieved [170]. For such short-wavelength devices, the limitation in performance was attributed to electron leakage into the lateral valley X of the InGaAs material.

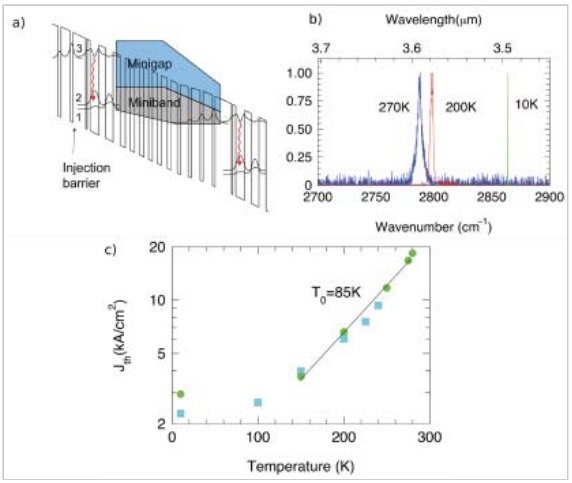


Fig. 8.9. Conduction band structure of a strain-compensated quantum cascade laser designed for short-wavelength operation. The predicted emission wavelength was $\lambda = 3.16$ μm using the effective mass of the unstrained material. b) The measured wavelength was between 3.4 μm and 3.6 μm, compatible with a much heavier mass as predicted by theory. c) Threshold current density versus temperature for the device operating at $\lambda = 3.6$ μm. The relatively low T_0 observed (85 K) is an indication of leakage paths above the barriers [159]. The design is described in Appendix A.3.1.

(p.156)

Using a similar material system but a different active region architecture, Bismuto *et al*. achieved room-temperature operation in the important 3.3 μm band [31]. The conduction-band diagram of the structure, under an applied field of 120 kV/cm, is shown in Fig. 8.10. The active region design is based on a bound-to-continuum architecture, where particular care has been taken to widen the lower miniband in order to obtain an efficient carrier extraction. Moreover, to reduce the resonant

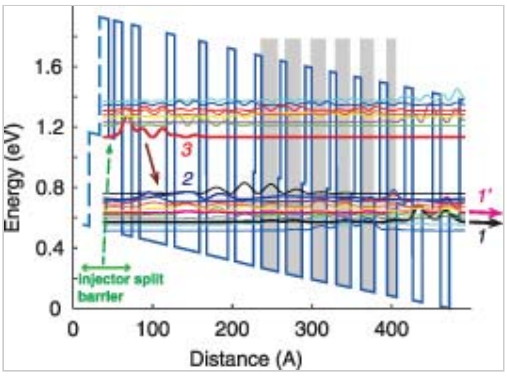


Fig. 8.10. Conduction-band diagram of a period of the active region at an average field of 120 kV/cm.

reabsorption from the extractor levels, the energy of the upper injector miniband was maximized. In Fig. 8.11a the emitted optical power and the applied voltage as functions of the current are shown for a typical 18 μm wide and 4.75 mm long laser. A high-reflection (HR) coating, consisting of 250 nm of Al_2O_3, 10 nm of Ti, and 100 nm of Au, was applied on the laser back facet. Watt-level emission was observed at room temperature and the laser performance is comparable to what was observed for Sb-containing QCLs [161]. The laser emission spectra for temperatures between 250 K and 350 K are shown in Fig. 8.11b. A tuning coefficient of 0.98 nm/K was found by linear fitting of the peak emission wavelength as a function of the temperature.

Moduli squared of the relevant envelope wavefunctions are shown. Light gray is used to show doped wells (6×10^{17} cm^{-3}). The layer sequence is described in Appendix A.3.3.

In order to better explore the influence of thermally activated carrier leakage on laser behavior the dependence of the threshold current density and of the slope efficiency are plotted as function of the heat sink temperature in figures 8.12)a and b respectively, along with the predictions of a density-matrix model. Both characteristics were fitted with the phenomenological relationship eqn. 7.8.25 as well as the equivalent for the slope efficiency:

$$\frac{dP}{dI} = \frac{dP}{dI}|_{T=0} \exp(-T/T_1).$$

(8.3.11)

In both cases the measured values of the characteristic temperatures T_0 = 100 K and T_1 = 70 K are signifficantly lower than the predicted one from a density matrix-based simulation code T_0 = 170 K and T_1 = 600 K. As the latter code includes electron leak- age into the continuum but not the effect of the lateral valley, we see this discrepancy as a strong indication that the lateral valley starts to have a non-negligeable effect on **(p.157)**

transport. In addition, the experimental characteristic temperatures are comparable with what is observed in Sb-containing QCLs that display a very large conduction band discontinuity for Γ as well as X and L points. This further indicates that the most likely leakage path occurs in the X and L valleys of $Ga_{0.27}In_{0.73}As$ materials that are believed to lie, for this indium mole fraction, ~ 0.6 eV above the Γ-valley.

As a consequence, the achievement of even shorter wavelengths could be feasible by relying on higher indium mole fractions or other active region architectures.

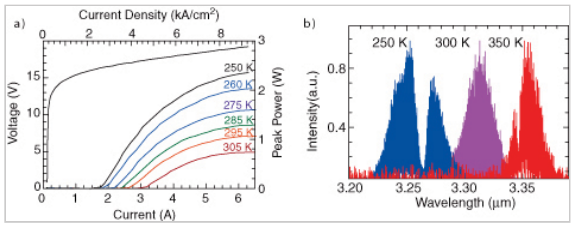


Fig. 8.11. a) Pulsed power–voltage–current characteristics of a HR coated 4.75 mm long and 18 μm wide laser for different heat-sink temperatures. b) Fabry–Perot laser emission spectra.

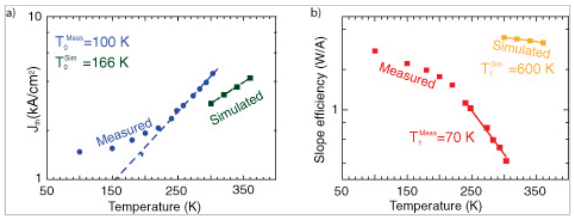
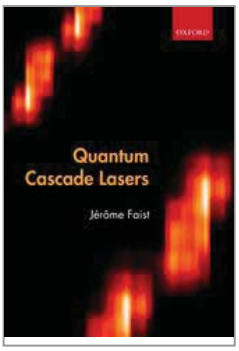


Fig. 8.12. a) Measured and simulated threshold current density as a function the laser temperature. b) Same comparison for the slope efficiency. In both

cases, the measured characteristic temperature is signifficantly lower than the predicted one.



Quantum Cascade Lasers
Jérôme Faist

Print publication date: 2013
Print ISBN-13: 9780198528241
Published to Oxford Scholarship Online: May 2013
DOI: 10.1093/acprof:oso/9780198528241.001.0001

Terahertz QCL

Jérôme Faist

DOI:10.1093/acprof:oso/9780198528241.003.0009

Abstract and Keywords

This chapter discusses Terahertz quantum cascade lasers covering Terahertz waveguides, active region designs, free carrier absorption, and key operation characteristics.

Keywords: quantum cascade lasers, waveguides, active region design, free carrier absorption

One very attractive feature of the intersubband transition is that in contrast to the interband system, the optical transition can be tuned to arbitrarily low energies with no obvious a priori limit. In fact, the first observation of intersubband absorption in an inversion layer as well as the first observation of electroluminescence from intersubband transition were both first demonstrated in what is now referred to as the terahertz frequency range.

In fact, in the original intersubband laser proposals [9–13] or failed experimental attempts,[1] it had been assumed that achieving a laser threshold would be easier in the far-infrared (i.e. for photon energies below the optical phonon energy) than in the midinfrared because of the absence of optical phonon emission at low temperatures. The longer upper-state lifetime would lead to lower threshold current densities. As shown in Chapter 7, however, it happens that the optical phonon process is a key ingredient in the design of a quantum cascade laser. Resonant optical phonon emission is used in many designs to reduce the lower-state lifetime. Even in designs in which it is not explicitly used to engineer the upperand lower-state lifetimes such as in laser based on diagonal transitions and on superlattices active regions, the optical phonon emission is the dominant electron cooling

mechanism. As a result, one expects operation in this wavelength range to require new active regions architectures.

The second and even more serious problem is to find a strategy to solve the waveguiding problem. In fact, as shown in Chapter 6, the quantum cascade lasers use a waveguide based on dielectric confinement, in many ways very similar to that used in interband lasers. In the terahertz, such an architecture would be impractical, due to the prohibitively thick cladding layers it would require (about 100 μm thick), as well as the very high optical losses caused by the doping of the cladding layers, needed to feed in the current in the layers. For these reasons the optical waveguides used in terahertz QCL use various combination of metallic or quasi-metallic confinement.

This chapter will cover only a few salient aspects of the topic. Readers are referred to more thorough review articles for a more in-depth description of these devices [171, 172] **(p.159)**

.

9.1 Terahertz waveguides

9.1.1 Single plasmon

This waveguide is based on the realization that in this frequency range a heavily doped semiconductor layer behaves as a "low density metal". As already discussed in Section 6.3, the plasma frequency for a GaAs layer doped to $N = 1 \times 10^{-18}$ cm^{-3} is about 10 THz. As a result, it will exhibit an overall negative dielectric constant in all the terahertz frequency ranges. Unfortunately, because of the relatively short lifetime of the carriers, as well as the limited densities that can be reached by doping, such doped layers cannot directly substitute a metal without creating unacceptably high losses.

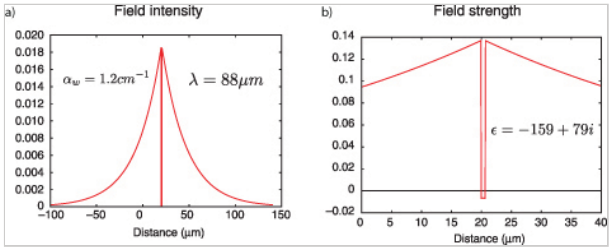


Fig. 9.1. Intensity (a) and normal component of the electric field (b) for the single plasmon mode trapped by single GaAs layer doped with $N = 2 \times 10^{18}$ cm^{-3} at $\lambda = 88$ μm.

However, a thin doped layer can create a waveguide with relatively low losses. Fig. 9.1 shows the intensity and electric profile of the plasmon mode trapped by a 1 μm thick GaAs doped layer with $N = 2 \times 10^{18}$ cm^{-3} at $\lambda = 88$ μm. At this wavelength, the doped GaAs has a dielectic constant of $\epsilon = -159 + 79i$. As for the interface plasmon polariton mode, see Section 6.2, the electric field is confined by the boundary condition of the continuity of the displacement field $\mathcal{D} = \epsilon\mathcal{E}$ that forces the electric field to change sign in the doped layer (see Fig. 9.1 b). Furthermore, for the same reason, the optical mode is strongly depressed in the doped region. As, in addition, the doped layer is thin, the optical losses are limited to the very low value of $\alpha_w = 1.2$ cm^{-1}.

A slightly modified version of this waveguide has been used to demonstrate the first terahertz quantum cascade lasers [23], and was shown to exhibit very low losses [173]. In the latter, the confinement of the electromagnetic wave is based on a metallic reflection at

the top metallization and the quasimetallic confinement provided by a thin, heavily doped buried contact placed below the heterostructure.

A two-dimensional computation of the mode intensity is shown in Fig.9.2 (a). The heterostructure has to be grown on a semi-insulating substrate, in order that the large part of the mode which overlaps with the substrate brings a very small contribution to the waveguide losses. The heavily doped (typically 0.8-2 × 10^{18} cm^{-3}) buried layer **(p.160)**

has the double function of optically confining the mode and acting as an electrical contact. As in the previous example, the overlap factor between the field and this heavily doped region is very small. The main advantage of this waveguide resides in its quite high figure of merit (overlap divided by losses) in the 2–4 THz region associated to a facet reflectivity still close to the Fresnel reflectivity calculated via index mismatch [174], which eases the coupling to free space modes and yields a good far-field. This waveguide loses its efficiency as the photon energy is lowered: the longer wavelength has to accommodate the same epilayer thickness, and the reduction of the dielectric constant of the active region pushes the mode towards the substrate, reducing the overlap. Another general disadvantage of this waveguide is the non-optimal and non-uniform ridge pumping due to the injection via the buried layer which shows significant in-plane resistance [175].

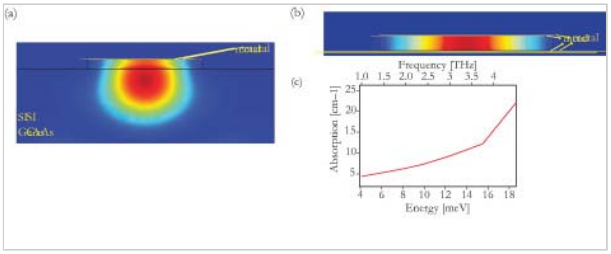


Fig. 9.2. (a) Computed mode intensity for a single plasmon waveguide at 3 THz (200 μm wide, 17 μm thick). (b) Computed mode intensity for an "empty" double-metal waveguide 200 μm wide and 17 μm thick at 950 GHz. (c) Calculated absorption for the empty waveguide of point (b) as a function of photon energy. (Adapted from[172].)

9.1.2 Double-metal waveguide

The double-metal waveguide is derived directly from the microstrip resonator widely employed in the microwave range. In this waveguide the active region is sandwiched between two metallic layers, typically Au, yielding an almost unity overlap factor almost independently of frequency. In Fig. 9.2b is plotted the two-dimensional simulation of the mode intensity for a double metal resonator at 950 GHz.

This waveguide was first demonstrated in QCL at mid-infrared frequencies [176] and then widely employed in combination of resonant-phonon designs in devices that have shown the highest temperature performance to date [171]. Recently, Cu has been proficiently used to enhance the figure of merit of this waveguide, setting the new limit in pulsed high-temperature operation [177]. The main disadvantage of this waveguide is represented by the patterned far-field [178] due to the impedance mismatch at the laser facet, and solutions based on micro-lenses [179] or horn antennas [180], which show substantial improvement in the collected powers have been developed. As shown in Fig. 9.2c, as the frequency is

lowered the waveguide loss decreases as the metals behave more and more as perfect conductors.

(p.161) 9.2 Active-region designs

The design of quantum cascade lasers for the terahertz also implies a change in the energy scale, as the photon energy is now smaller than the energy of the optical phonon and also much closer to the value of the thermal energy, even at cryogenic temperatures. These constraints yielded new approaches for the active region architecture, which will be reviewed briefly here.

9.2.1 Superlattice and bound to continuum

Gain between subband depends critically on the population inversion. Using the notations shown schematically in Fig. 9.3, the population inversion is maximized by achieving the largest ratio of upper to lower lifetime τ_{32}/τ_2, a large injection efficiency in the upper-state η, and a long upper-state lifetime τ_3. The standard approach used in mid-infrared quantum cascade lasers has been to use the properties of the intersubband non-radiative transitions to achieve a large ratio of lifetimes, and employ a resonant tunneling injection of the electrons from the injector into the upper-state to obtain a high value of η. As show in Fig. 9.3b, a similar approach has worked very well in the terahertz at cryogenic temperatures. The active region architecture of the first quantum cascade laser operating in the terahertz [23] is essentially a scaled version of a mid-infrared chirped superlattice active region [127], in which the population inversion between the two states at the edge of the miniband is obtained by a phase-space argument, as discussed in Section 7.5.3. The scattering rate from the lower state τ_2^{-1} is very large, the latter being the upper-state of the lower miniband while the scattering from the upper-state is similarly spread among the whole lower miniband, keeping the scattering rate into the lower state small, i.e. $\tau_{32}^{-1} \langle \langle \tau_2^{-1}$. Similarly, the bound-to-continuum approach reported in [181] is a transposed version of the same approach demonstrated at shorter wavelengths [133, 143]. In the bound-to-continuum approach, the population inversion, achieved by a combination of the phase-space used in the superlattice active region with a diagonal transition between states with a reduced spatial overlap, further enhancing the population inversion.

Chirped superlattice and bound-to-continuum were also used very successfully to extend the wavelength range of mid-infrared to longer wavelength. The terahertz bound-to-continuum approach was further developed [175, 182], and yielded devices with a very high slope efficiency and power at temperatures of about 20 K.

However, the maximum operating temperature of these devices remained limited to about 100 K. The problem originates from the scaling of the miniband width Δ with photon energy hv, considering the effect of temperature kT and the broadening of the individual levels γ. For the miniband and bound-to-continuum approach to succeed, one should maintain the miniband width much larger than the broadening of the individual levels $\Delta \rangle \rangle \gamma$ and larger than the thermal energy $\Delta \rangle kT$. In the mid-infrared, at room temperature ($kT = 26$ meV) the conditions are satisfied because $\Delta = 100 - 150 meV \rangle \rangle \gamma \approx 5 - 10$ meV and $\Delta \rangle kT$. In the terahertz, however, the same inequality holds only at cryogenic

temperatures, because the miniband width is only 15 meV, while the broadening at room temperature remains of the order of a few meV. Similarly, one cannot keep a very wide miniband Δ = 100 meV with a photon energy of $h\nu$ = 15 meV, as the latter should be much larger than the individual level **(p.162)**

spacing of the miniband. In addition, as soon as the sum of the miniband width and the photon energy becomes equal or larger than the optical phonon energy $\Delta + h\nu \rangle \hbar\omega_{LO}$ = 36 meV, the lifetime of the upper-state drops very strongly even at low temperatures. The schematic band structure diagram, computed at an average electric field of 2.5 kV/cm, of such a bound-to-continuum design is shown in Fig. 9.4a. The lightand voltage-versus-current characteristics, measured in a continuous wave from a 1.5 mm long and 200 μm wide device processed into a single plasmon, are shown in Fig. 9.4b. The large lifetime ratio between upper and lower states at low temperature enables high optical powers, and also translates into a strong change in the differential conductivity at threshold **(p.163)**

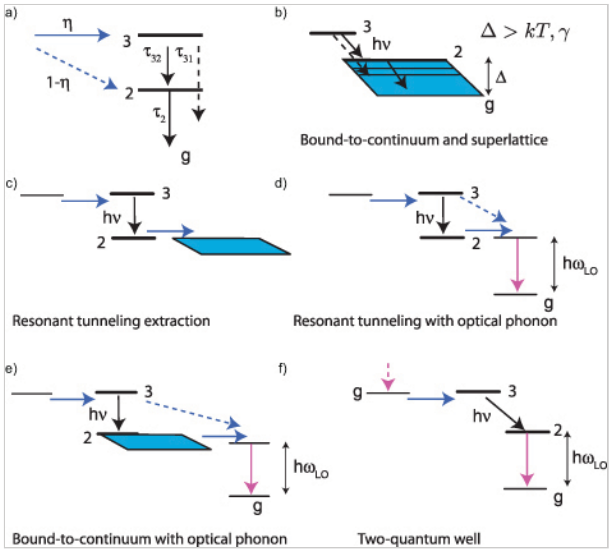


Fig. 9.3. a) Schematic levels of a generic quantum cascade laser. By tradition, the upper level is labeled 3. b) Bound-to-continuum or superlattice design. c) Design based on a resonant tunneling extraction. d) Design combiningresonant tunneling extraction with an optical phonon resonant stage. e) Bound-to-continuum with optical phonon resonant extraction stage. f) Two-quantum well: photon–phonon cascade.

.

9.2.2 Resonant tunneling extraction

As shown schematically in Fig. 9.3b, a more selective depopulation of the lower state can be achieved by coupling the lower state by resonant tunneling to a miniband. Such a laser, based on a single quantum all-active region, achieved

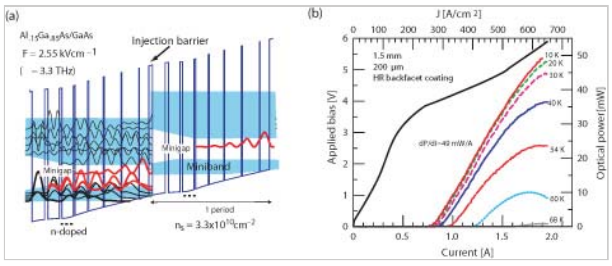


Fig. 9.4. Band structure (a) and light-versus-current characteristics (b), measured in a continuous wave, of a bound-to-continuum design operating at λ = 88 μm wavelength.

record low threshold current densities of 30 A/cm^2 [183]. In terms of high-temperature operation, however, this laser suffered from the same limitations as the one described above because of the small width of the injection miniband.

Coupling by resonant tunneling the lower state of the laser transition to a very short-lived upper-state of a nearby well, as shown in Fig. 9.3c allows a significant reduction of the lower-state lifetime while preserving a long upper-state lifetime. This idea is the basis of the so-called resonant phonon design first demonstrated by the MIT group [184]. This design has seen many variations [185, 186], and is now the one that has demonstrated the highest operating temperature of a terahertz quantum cascade laser so far [187]. The low lifetime of the upper level of the well following the active region is usually achieved by spacing the levels resonantly with optical phonon energy. In this way, the condition $\Delta \rangle\ kT$ is satisfied. This architecture has the additional benefit of shifting the intersubband absorption of the electrons from the ground state of the quantum well to a photon energy larger than the laser transition. The transfert of some or all of the oscillator strength from the ground state from very low frequencies to frequencies above the transition energy was one of the key ingredients for the achievement of intersubband lasers of 1.2–2 THz [188]. In that case, the pair of states was spaced by less than the optical phonon energy. A general feature of these designs is the requirement for a simultaneous alignment of the injector and extraction stage at the same electric field [183]. Such a double-resonance condition, however, increases the sensitivity of the design to growth inaccuracies.

Another class of design aims at combining the advantages of both approaches. As shown schematically in Fig. 9.3e, the active stage is based on a bound-to-continuum transition, while the miniband is coupled to an extractor quantum well [189]. This **(p.164)**

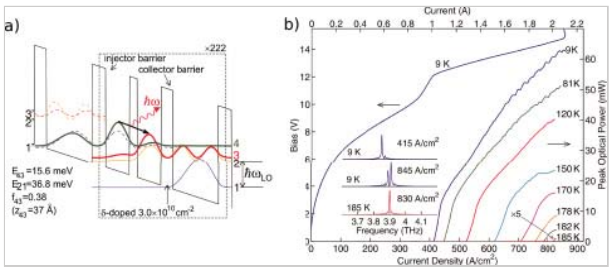


Fig. 9.5. Band structure (a) and light versus current (b) of a three quantum wells active region operating up to a maximum temperature of 186K. Reprinted with permission from [187]. Copyright 2009, AIP.

architecture has the advantage of reducing the direct coupling between the upper state and the extractor well, because they are physically separated by the length of the miniband region. The alignment condition on the extraction stage is also somewhat relaxed by the fact that it has to be satisfied over the width of the miniband and no longer over that of a single state. The original design following this architecture [189] was based on a five-quantum-well active region and a two-quantum-well injector region. A shortened version of this design, where the total number of quantum wells has been reduced to four, with a three-quantum-well active region coupled to a one-quantumwell injector, is discussed extensively in [190]. Recently, structures with three-quantumwell active regions [186] have demonstrated operating temperatures up to ≈ 200 K [191]. As shown in Fig. 9.5, in such a device the lower state of a

diagonal transition between two coupled wells is in resonance with the upper-state of a thicker phonon well.

Finally, a very interesting structure is the two-quantum-well active region shown schematically in Fig. 9.3f. In this structure the lower state of a diagonal laser transition is also the upper-state of an optical phonon transition. This structure should in principle exhibit population inversion up to room temperature, because the lower-state lifetime remains very short. The first demonstrations, however, showed operation up to 125 K [192–194]. The origins of the limitation in temperature are not completely elucidated, and are further discussed in the context of resonant tunneling diodes in Section 12.2.5. Extremely wide gain bandwidths (1 THz) have been demonstrated using this device [194]

9.3 Free carrier absorption

As discussed in Section 7.10.1, free carrier absorption imposes some fundamental limits on the operation of quantum cascade lasers. In the classical Drude model, the free carrier absorption grows with the square of the wavelength. As shown in Fig. 9.6, the **(p.165)** experimentally measured values for mid-infrared quantum cascade lasers, representing the state of the art in 1997, indeed showed the expected λ^2 dependence, and the extrapolation of this trend to the terahertz yielded unacceptably high losses in the range of 1000 cm^{-1}.

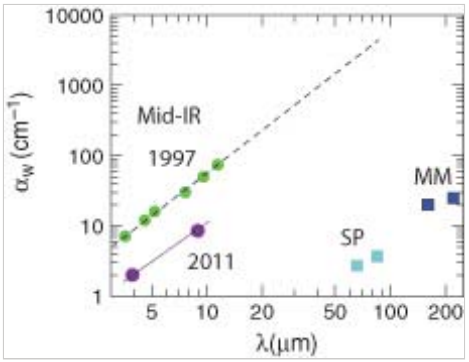


Fig. 9.6. Waveguide losses as a function of wavelength. Solid disks, mid-infrared devices, with the upper curve representing the existing literature in 1997, while the two lower points are more recent high performance devices. Solid squares, terahertz devices. SP, single plasmon and MM metal–metal waveguides.

Interestingly, the values measured for terahertz QCL and shown in the same figure lie about three orders of magnitude below these extrapolated values. One reason for these lower values is the use of a single plasmon or a metal–metal waveguide that removed the need for a doped cladding, as well as by the use of a lower average doping of the active region, reducing from $\approx 8 \times 10^{16}$ cm^{-3} to $\approx 5 \times 10^{15}$ cm^{-3}.

However, a closer look shows that treating the active region as a bulk is a crude approximation, especially in quantum cascade laser waveguides where the electric field is normal to the layers. A better approach consists in treating all the absorption in the active and injector regions as resulting from resonant or non-resonant intersubband absorption. In fact, this approach is obtained naturally for transport and gain models of the whole active region based on density matrix or on the non-equilibrium Green's function, as discussed in Chapter 12. In particular, the second-order absorption formula (eqn. 12.2.58) allows the

description of the intersubband absorption as a function of temperature in the presence of an in-plane elastic scattering.

Considerations based on sum rules are very useful, because, as shown theoretically [195], free carrier absorption can be seen as the limit of intersubband absorption for vanishing intersubband transition energy. Therefore, there is a transfer of oscillator strength between the Drude-like free carrier absorption within a miniband and interminiband absorption [196]. As a result, for devices that operate at low frequencies, injectors can be designed to transfer the oscillator strength of the free carrier absorption into an intersubband absorption at a higher frequency. This approach is most beneficial for terahertz QCLs, but is also applied successfully in mid-infrared devices. A very rough approximation for oscillator strength carried by the intraminiband conduction can be obtained by considering the dipole matrix element between the first two states of a coupled quantum well, which is in the tight-binding approximation:

(p.166)

$$z_{12} = \frac{1}{2}(z_{22} - z_{11}) = \frac{L}{2}$$

(9.3.1)

where L is the thickness of an individual quantum well. A similar result can be established with a series of N coupled wells [196]. Using the infinite barrier approximation for the transition energy, the oscillator strength is then

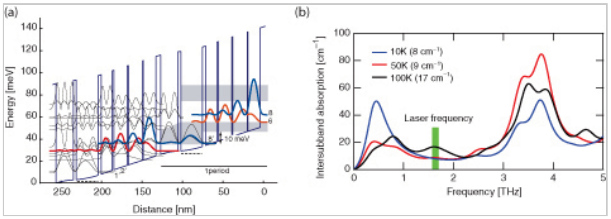


Fig. 9.7. a) Band structure for a 1.7 THz design. b) Calculated intersubband absorption spectrum of the band structure, assuming a thermal equilibrium distribution of the carriers. Adapted with permission from [197]. Copyright 2006, AIP.

$$f_{intra} = \frac{3\pi^2}{4}\frac{2\Omega}{E_{inter}}\frac{m_0}{m^*}$$

(9.3.2)

where 2Ω is the splitting between the two levels and E_{inter} the transition energy to the next excited state of the quantum well. Note that the numerical prefactor in eqn. 9.3.2 is much larger than unity, and therefore a significant fraction of the oscillator strength remains in the miniband even for a ratio $2\Omega/E_{inter}$ below unity.

Fig. 9.7a shows the design of a device operating at 1.7 THz. The injector consists of a double well, with an energy splitting (2 meV) much smaller than the design photon energy (6 meV), preceded by a wider miniband located at a higher energy (10 meV) [197]. As a result, the reabsorption of the injector, shown in Fig. 9.7b, is such that the low-frequency component is much weaker than the high-frequency component, and that laser operation can happen in a minimum between the lowand high-frequency peaks. Note that because the two states of the double well have a splitting smaller than kT, the contribution of this

doublet to the absorption *decreases* with increasing temperature. The rise of intersubband absorption at 100 K originates from resonant absorption of the laser transition due to backfilling.

Note also that this feature arises naturally in single-quantum-well injector devices with a phonon extraction stage.

9.4 Key operation characteristics

As for their mid-infrared counterparts, terahertz quantum cascade lasers are able to deliver large optical powers. Up to 250 mW in pulsed operation was achieved for multimode devices, while single-mode devices have shown powers up to 50 mW, due to a **(p.167)**

photonics crystal structure. Because of the low frequency, as well as the low lattice operating temperature, the temperature and the current dependence of the laser is weak. As a result, very narrow linewidths, down to a Schawlow–Townes limit of 60 Hz (unpublished) have been measured. However, for the same reason the tuning range of terahertz QCL is very limited, and an electrical tuning range of a few gigahertz is usually achieved. The maximum values, achieved by cavity pulling, are of the order of 30 GHz, as shown in Fig. 10.17.

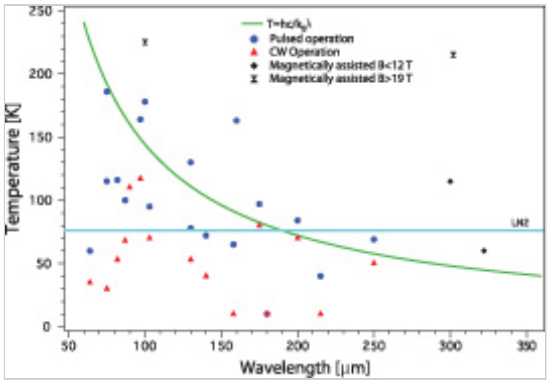


Fig. 9.8. Maximum operating temperature as a function of wavelength, summarizing devices reported in the literature.

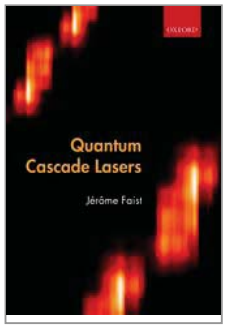
A summary of the maximum temperature as a function of frequency is shown in Fig. 9.8. As mentioned, the maximum operating temperature reported is 186 K for a device operating at 3 THz. It can also be seen that the majority of the reported values follow a general trend of having the photon energy $h\nu \approx kT_{max}$.

The temperature dependence of the threshold current of terahertz quantum cascade lasers was shown to be dominated by optical phonon emission and can be described using simple models, as shown already for designs based on bound-to-continuum [181, 172], bound-to-continuum with phonon extraction [190], or resonant phonon extraction [198]. Although these models predict the temperature dependence of the threshold current density with various degrees of accuracy, they do not predict the maximum operating temperature. In fact, the latter is reached when the threshold current density is equal to the maximum current, at which the gains rolls over because of the onset of a negative differential resistance. Prediction of the device characteristics in this regime are very difficult, because they require a good model of the gain saturation mechanisms.

Operation at room temperature remains an elusive goal of terahertz QCL research.

Notes:

([1]) Such as pumping optically with a CO_2 laser in a coupled quantum well system, trying to reach population inversion between excited levels.



Quantum Cascade Lasers
Jérôme Faist

Print publication date: 2013
Print ISBN-13: 9780198528241
Published to Oxford Scholarship Online: May 2013
DOI: 10.1093/acprof:oso/9780198528241.001.0001

Mode control

Jérôme Faist

DOI:10.1093/acprof:oso/9780198528241.003.0010

Abstract and Keywords

This chapter discusses mode control for quantum cascade lasers covering the Fabry Perot cavity, distributed feedback cavity, external cavities, and DFB arrays.

Keywords: quantum cascade lasers, Fabry-Perot cavity, distributed feedback cavity, external cavities, dfb arrays

10.1 Fabry–Perot cavity

The active region of a quantum cascade laser is placed in the center of a mid-infrared waveguide forming a long ridge. By cleaving the latter at a length of typically 1–3 mm, a Fabry–Perot cavity is formed. The cleaved facets exhibit an amplitude reflectivity $r = \sqrt{R}$. Imposing a stationary mode running in the cavity requires the field to close on itself after a round-trip in the cavity:

$$r^2 \exp(2i\beta L) = 1$$

(10.1.1)

where

$$\beta = n_{eff} k_0 - i\frac{g - \alpha_w}{2}$$

(10.1.2)

is the propagation constant for a mode with a vacuum propagation constant $k_0 = 2\pi/\lambda$ and a net gain $g - \alpha_w$. Imposing the condition set by eqn. 10.1.1 for both the amplitude and phase imposes the well-known conditions for the gain:

$$g = -\frac{1}{L}\ln(R) + \alpha_w$$

(10.1.3)
and the condition on the wavelength

$$\lambda = \frac{2n_{eff}L}{N}$$

(10.1.4)
where N is any positive integer. Taking into account the fact that the refractive index has in general a wavelength dependence, the mode spacing between adjacent cavity modes is

$$\Delta\lambda = \frac{\lambda^2}{2L(n_{eff} - \lambda\frac{\partial n_{eff}}{\partial\lambda})}$$

(10.1.5)
Expressed in wavenumbers, the expression reads

$$\Delta(\frac{1}{\lambda}) = \frac{1}{2L(n_{eff} + \frac{1}{\lambda}\frac{\partial n_{eff}}{\delta(\frac{1}{\lambda})})}$$

(10.1.6)
where the group refractive index n_g is

$$n_g = n_{eff} - \lambda\frac{\partial n_{eff}}{\partial\lambda} = n_{eff} + \frac{1}{\lambda}\frac{\partial n_{eff}}{\delta(\frac{1}{\lambda})}$$

(10.1.7)
(p.169)

As shown in Fig. 10.1, where a series of spectra of short-wavelength quantum cascade lasers is shown as a function of injected current, the Fabry–Perot modes of a laser cavity are clearly apparent below threshold. Their contrast (which can be directly evaluated using the logarithmic scale of the figure) increases with injected current; their location shifts with the change of n_{eff} induced by the temperature change with injected current. The latter effect can be analyzed using the expression for the transmission of a Fabry–Perot cavity:

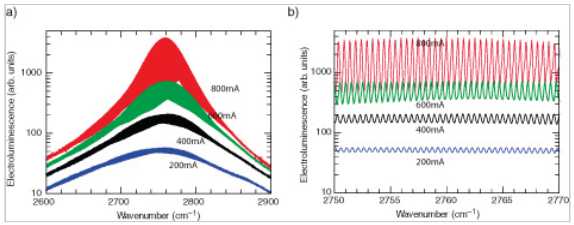


Fig. 10.1. Subthreshold luminescence of a short-wavelength quantum cascade laser as a function of increasing current, from 200 mA to 800 mA by steps of 200 mA. The temperature is 20 K and the threshold current is at 950 mA. The spectra are shown with two different horizontal scales a) and b). The shift of the individual Fabry–Perot resonances is caused by the device heating.

$$I(1/\lambda) = (\frac{1-R}{\sqrt{R}})\frac{b}{1+b^2-2b\cos(4\pi n_{eff}L/\lambda)}$$

(10.1.8)
where the parameter controlling the finesse of the Fabry–Perot cavity b is given by

$$b = R\exp((g-\alpha_w)L)$$

(10.1.9)

For this reason, the measurement of the fringe contrast as a function of injected current enables a measurement of the gain, as will be discussed in detail in the next chapter.

10.2 Distributed feedback cavity

As is apparent in Fig. 10.1, for usual Fabry–Perot cavity lengths the mode spacing (of the order of 1 cm^{-1}) is much narrower than the gain bandwidth of 100–200 cm^{-1}. For this reason, even if because of mode competition, the laser operates on a single mode above threshold, the exact location of the latter is almost impossible to predict with enough accuracy in order to target a specific molecular vibration line.

10.2.1 Multi-layer approach, Bragg reflection condition

One solution used to solve this problem is to incorporate a distributed feedback inside the active region of the device by creating a small, periodic modulation of the **(p.170)** waveguide effective index. In a first approach, let us consider two regions of effective refractive index n_a and n_b of lengths a and b repeated with a periodicity $\Lambda = a+b$. The coupling between forward and backward propagation of the plane wave at frequency ω can be described by a 2x2 complex matrix M describing the coefficients of the plane waves at the nth layer with the one at the $n-1$th layer:

$$\begin{pmatrix} a_{n-1} \\ b_{n-1} \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} a_n \\ b_n \end{pmatrix}$$

(10.2.10)

where the coefficients of the matrix M are given for a TE mode at normal incidence:

$$A = e^{i\frac{2\pi n_a}{\lambda}a}[\cos(\frac{2\pi n_b}{\lambda}b) + \frac{i}{2}(\frac{n_b}{n_a} + \frac{n_a}{n_b})\sin(\frac{2\pi n_b}{\lambda}b)]$$

(10.2.11)

$$B = e^{-i\frac{2\pi n_a}{\lambda}a}[\frac{i}{2}(\frac{n_b}{n_a} - \frac{n_a}{n_b})\sin(\frac{2\pi n_b}{\lambda}b)]$$

(10.2.12)

$$C = e^{i\frac{2\pi n_a}{\lambda}a}[-\frac{i}{2}(\frac{n_b}{n_a} - \frac{n_a}{n_b})\sin(\frac{2\pi n_b}{\lambda}b)]$$

(10.2.13)

$$D = e^{-i\frac{2\pi n_a}{\lambda}a}[\cos(\frac{2\pi n_b}{\lambda}b) - \frac{i}{2}(\frac{n_b}{n_a} + \frac{n_a}{n_b})\sin(\frac{2\pi n_b}{\lambda}b)]$$

(10.2.14)

where λ is the wavelength in vacuum. As the structure is periodic, the solutions must satisfy Bloch's theorem:

$$\begin{pmatrix} a_n \\ b_n \end{pmatrix} = e^{-iK\Lambda} \begin{pmatrix} a_{n-1} \\ b_{n-1} \end{pmatrix}$$

(10.2.15)

which finally yields the dispersion relation

$$\cos K\Lambda = \cos k_a a \cos k_b b - \frac{1}{2}(\frac{n_b}{n_a} + \frac{n_a}{n_b})\sin k_a a \sin k_b b$$

(10.2.16)

where we have introduced the wavevectors $k_a = \frac{2\pi n_a}{\lambda}$ and $k_b = \frac{2\pi n_b}{\lambda}$ to emphasize the similarity with the electronic dispersion relation obtained for the superlattice, in particular when considering the two states in the continuum (eqn. 3.4.74). As in the electronic case, gaps will open for frequencies corresponding to the Bragg condition

$$\lambda_B = 2\bar{n}\Lambda/N$$

(10.2.17)

where N is an integer that characterizes the order of the grating and $\bar{n}$ is the average refractive index, assuming a weak modulation of the latter. The last equation can also be written in terms of the propagation constant $\beta = \bar{n}\frac{2\pi}{\lambda}$ of the mode as

$$2\beta = N\mathbf{G}$$

(10.2.18)

where $\mathbf{G} = \frac{2\pi}{\Lambda}$ is the reciprocal vector of the grating. The Bragg condition can be seen as a conservation of the in-plane wavevector with the grating providing an integer number of wavevectors $N\,\mathbf{G}$, and can be graphically described schematically by the vector diagrams shown in Fig. 10.2. Of course, as the order of the grating increases, so does the number of different scattering directions available for the light. Of particular **(p.171)**

interest are the first-order grating, as only backscattering is allowed. The second-order grating enables both the frequency selection *and* the surface emission.

To use again the analogy with solid-state systems, the Bragg reflection creates a gap for which the propagation is not possible. If the modulation is purely of the real part of the refractive index, the gap is opened in the frequency domain; if the modulation is only on the gain, the gap opens in the allowed propagation wavevector. Assuming again a small refractive index change

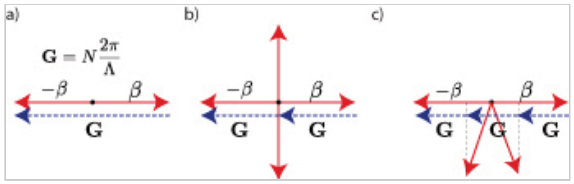


Fig. 10.2. The Bragg reflection condition for gratings of various order N. a) In the first-order grating, only backscattering is possible. b) For $N = 2$, either backscattering or vertical emission is possible. c) In the third-order grating ($N = 3$), emission is only possible inside the substrate because of the large refractive index difference between the guided mode ($n_{eff} = 3.2 \rangle\ 3$) and the vacuum. Third-order DFB lasers using strong gratings have been successfully developed in the terahertz, however [199].

$$\left|\frac{n_a - n_b}{n_a + n_b}\right| = \frac{1}{2}\left|\frac{\Delta n}{n}\right| \langle\langle 1$$

(10.2.19)

and equal thicknesses ($a = b$) the width of the gap in angular frequency is shown to be

$$\frac{\Delta\omega}{\omega_B} = \frac{2}{\pi}\frac{\Delta n}{n}$$

(10.2.20)

by Taylor expansion of the eqn. 10.2.16 to the second order around the Bragg frequency

$$\omega_B = c\beta_B = \frac{2\pi\bar{n}c}{\Lambda}$$

(10.2.21)

and solving for cos $K\Lambda = -1$. The computed transmission for a finite quarter wave stack with a periodicity of 1.3 μm of average refractive index $\bar{n} = 3.2$ and a refractive index step of $\Delta n = 1 \times 10^{-2}$ is shown in Fig. 10.3. As expected, a gap appears centered at the Bragg wavelength. The value of the stopband width, 2.76 cm^{-1}, obtained by computing the distance between the two transmission maxima, agrees with the value obtained from eqn. 10.2.20 (2.52 cm^{-1}). The asymmetry of the transmission characteristic, as well as the small difference between the two computed values for the bandgap width, originate from a high-reflectivity coating assumed on one end of the device and providing an additional reflectivity with a phase shift.

The coupling constant of the grating κ describes the attenuation of the propagating wave in the gap center. It is obtained as the solution for the imaginary part of K of eqn. 10.2.16 at the frequency corresponding to the Bragg condition. The quantity $1/\kappa$ is the attenuation length of the field inside the periodic grating. This quantity plays a major role in the evaluation of the performances of distributed feedback lasers. Again **(p.172)**

using a series expansion of eqn. 10.2.16 for small values of $\Delta n/n$, the coupling constant is written as

$$\kappa = \frac{1}{\Lambda}\frac{\Delta n}{n}$$

(10.2.22)

A laser with a periodic grating can only operate at the edge of this gap. In fact, the gap opened in the ω versus k dispersion curve will force the derivative to be zero at the edge. As a consequence, the group velocity

$$v_g = \frac{\partial\omega}{\partial k}$$

(10.2.23)

is zero at the band edge. For this reason, light at this band edge is stationary and therefore a mode can be confined.

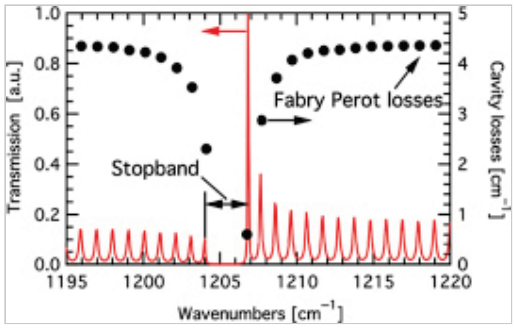


Fig. 10.3. Computed transmission (left axis) of a 1.5 mm thick stack of layers with alternating high and low refractive index. The periodicity is $\Lambda = 1.3$ μm, the average refractive index is 3.2, and a refractive index step of $\Delta n = 1 \times 10^{-2}$ has been assumed. A stop band of width about 2.76 cm^{-1} appears in the transmission. The high-reflectivity coating assumed on one end of the device is responsible for the asymmetry of the transmission characteristic. The computed loss of the modes corresponding to the transmission maxima is plotted on the right axis.

Using the multi-stack approach, one is also able to compute the threshold for the modes of a distributed feedback laser. The transmission of the finite cavity is obtained by the multiplication of the transmission matrices for each layer pair given by the elements 10.2.14 but with a complex refractive index $\tilde{n}$ given by

$$\tilde{n}_{a,b} = n_{a,b} - ig\frac{\lambda}{4\pi}$$

(10.2.24)

where g is the gain. The laser modes are defined by the couples of values (λ, g) for which the transmission diverges.

The result of such a computation is shown in Fig. 10.3. As expected, the threshold corresponds to the values predicted for a Fabry–Perot device ($g = -\frac{1}{2L}\ln R_1 R_2 = 4.3\text{cm}^{-1}$, assuming a 100% high reflectivity coating on one end) for the modes away from the grating resonance. The modes on the side of the grating resonance have their threshold gain reduced to about 2.2 and 0.4 cm^{-1}, respectively.

(p.173) 10.2.2 Coupled mode analysis

Although the previous analysis enabled the computation of the threshold for realistic devices, it is based on purely numerical simuations and therefore hides the physics of the device. Kogelnick and Shank [200] have derived a coupled-wave analysis of distributed feedback lasers that shows the physics of the system. Using their notation, the refractive index is assumed to be modulated sinusoidally for both its real part

$$n(z) = n + n_1 \cos(2\beta_0 z)$$

(10.2.25)

as well as for the loss

$$\alpha(z) = \alpha + \alpha_1 \cos(2\beta_0 z)$$

(10.2.26)

where $\beta_0 = \frac{\pi}{\Lambda} = \mathbf{G}/2$ is the propagation constant of the grating. The field inside the cavity is assumed to be the sum of a right-going wave and a left-going wave:

$$E(z) = R(z)\exp(-i\beta_0 z) + S(z)\exp(i\beta_0 z)$$

(10.2.27)

where both functions $R(z)$ and $S(z)$ are assumed to vary slowly. Using the latter assumption, a set of coupled differential equations can be derived for R and S:

$$-\frac{\delta R}{\partial z} + (\alpha - i\delta)R = i\kappa S$$
$$\frac{\delta S}{\partial z} + (\alpha - i\delta)R = i\kappa R$$

(10.2.28)

where the coupling constant κ is given by:

$$\kappa = \frac{\pi n_1}{2n\Lambda} + i\frac{\alpha_1}{2}$$

(10.2.29)

and the normalized detuning δ is given as

$$\delta = \frac{n\Delta\omega}{c}$$

(10.2.30)

Note that the ratio of the coupling constant defined by eqn. 10.2.22 for a rectangular refractive index profile and the one defined above is equal to $\frac{4}{\pi}$ (taking $\Delta n = 2n_1$), which is exactly the Fourier component of a rectangular wave of unit amplitude. This model enables the computation of the threshold gain in general, however, without the effect of the facets. The key results are the

position of the modes for modulation of both the real part and the imaginary part of the refractive index. It also enables us to study the effect of the product κL on the internal optical field along the cavity. The solution of eqn. 10.2.28, using the appropriate boundary condition for a device spanning symmetrically between $-L/2$ to $L/2$, is given by the solution of the transcendental equation:

$$\kappa = \frac{i\gamma}{\sinh \gamma L}$$

(10.2.31)

where the parameter γ is given as a function of the loss α and detuning δ as

(p.174)

$$\gamma^2 = \sqrt{\kappa^2 + (\alpha - i\delta)^2}$$

(10.2.32)

Once eqn. 10.2.31 is solved, the functions $R(z)$ and $S(z)$ are given by

$$R(z) = \sinh \gamma (z + L/2)$$

(10.2.33)

$$S(z) = \pm \sinh \gamma (z \mp L/2)$$

(10.2.34)

The solutions for the bandgap edge mode of a 1.5mm long laser operating at $\lambda = 7\ \mu$m for various values of the coupling constant κ are shown in Fig. 10.4. In general, a larger value of κL lowers the device threshold and improves the mode selectivity but at the cost of a lower power efficiency.

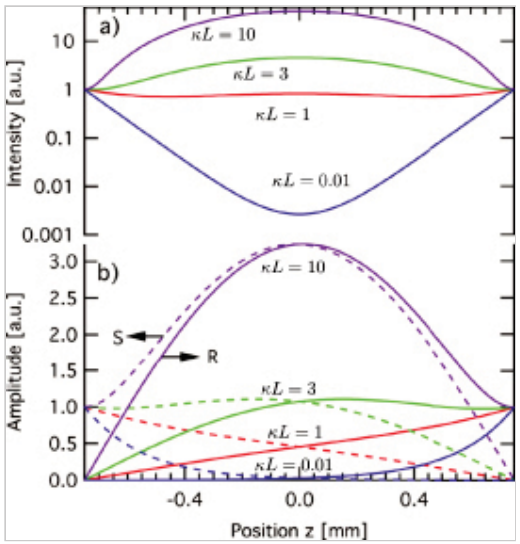


Fig. 10.4. a) Intensity distribution and b) amplitudes of the right-going R(z) (solid) and left-going S(z) (dashed) in a 1.5 mm-long DFB-QC laser for different values of the product κL, as indicated. Non-reflecting facets are assumed.

10.2.3 Fabrication geometries

Surface grating.The first distributed feedback quantum cascade lasers were fabricated by etching the grating directly onto the surface of the device onto a waveguide specially designed with a thin AlInAs cladding. The interaction between the mode and the surface plasmon propagating at the semiconductor–contact interface enabled significant modification of both refractive index and loss, and therefore yielded complex coupled DFBs. Single mode operation was achieved at both 5.7 and 7.2 μm [21]. The modulation of the effective index can be modeled by the effective index technique, in which the effective index of a propagating mode is computed in both regions of the grating (i.e. in the etched and unetched parts). These numbers are then used in a one-dimensional simulation of the structure as described above. For the structure **(p.175)**

shown in Fig. 10.5, a refractive index step of $\Delta\tilde{n} = 7 \times 10^{-3} + 3 \times 10^{-4}\ i$ is found for an etch depth of 0.5 μm. The distributed feedback laser is then

predominantly index coupled with a non-negligible loss component. Computation techniques based on the effective index method tend to predict everincreasing losses for stronger (deeper) grating because of the increasing interaction with the surface metal. A more refined analysis [201], however, shows that with the correct set of geometrical parameters, a combination of strong coupling and low losses could also be achieved for deep etched gratings. In fact, very high-performance single-mode operation was achieved for devices based on both GaAs [202] and InP [203].

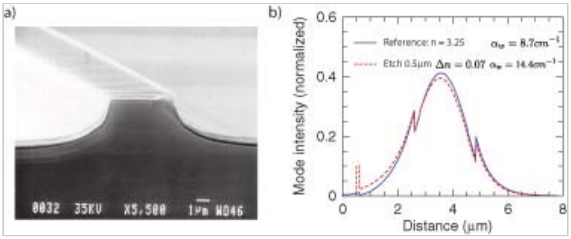


Fig. 10.5. a) Scanning electron micrograph of a distributed feedback quantum cascade laser with a grating on the surface. b) Computed effective index, losses, and mode profile of the reference waveguide and the one for an etch depth of 0.5 μm. For an etch depth of 0.3 μm, a value of $\Delta n = 0.03$ and $\alpha = 10.2$ cm^{-1} are computed. [21]

Another successful technique has been to use a device where air is used as the top cladding layer and the current injection proceeds laterally through a doped ($n = 2 \times 10^{18}$ cm^{-3}) InGaAs conducting layer which is simultaneously the layer host to the grating [204]. A schematic drawing of the waveguide, together with a scanning electron micrograph of a processed device, is shown in Fig. 10.6. Devices fabricated in this way exhibit a number of advantages. As in the case for the surface-etched devices, the fabrication does not require a regrowth etch step, and the technique allows strong **(p.176)**

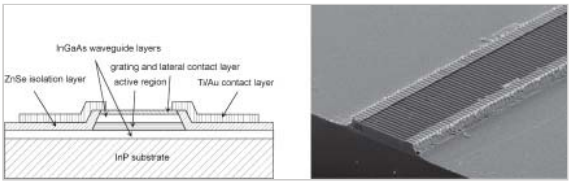


Fig. 10.6. a) Schematic cut through a waveguide design that uses air as the top cladding and a lateral injection through a doped InGaAs layer. b) Scanning electron micrograph of a finished device. Reprinted with permission from [204]. Copyright 1999, AIP.

complex grating coupling. Wide devices can achieve large peak powers (1W) while maintaining single transverse mode operation due to the lateral losses brought by the lateral contacts. Finally, the top surface of the grating could be used as an access to the optical field circulating inside the device for intracavity detection. However, the relatively large width of the lasers, which enables the large peak powers, is also responsible for the high drive currents of these devices and prevents their use in a continuous wave. The light-versus-current, as

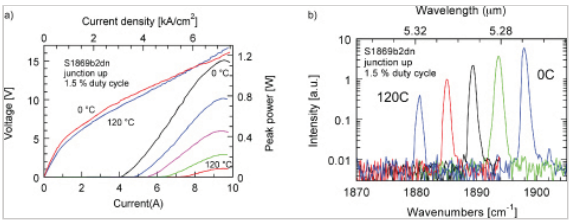


Fig. 10.7. a) Light- and bias-versus-current characteristic of a distributed feedback quantum cascade laser with a air/semiconductor top cladding and a lateral injection. b) Representative spectra of the device at various operation temperatures. Reprinted with permission from [205]. Copyright 2001, AIP.

well as some representative spectra of such a device, is shown in Fig. 10.7. Single-mode peak optical powers as high as 1.2 W at 0° C and operation up to 120° C were achieved [205].

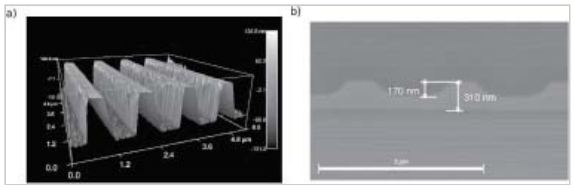


Fig. 10.8. a) Atomic force microscopy of a buried grating prior to the regrowth step. b) Scanning electron micrograph of the active region of a device after the regrowth was performed.

Buried grating. As for the telecom devices, a very effective architecture is to etch the grating in the InGaAs guiding layer above the active region, followed by an InP regrowth. Atomic-force microscopy of a grating prior to the etching step, as well as a scanning electron micrograph of the active region of such a device, is shown in Fig. 10.8. In such a device the refractive index step between InGaAs and InP ($\Delta n_{etch} \approx 0.4$) is close to the maximum of the waveguide. Actually, using the effective-index approach, **(p.177)**

the change in effective index Δn_{eff} as a function of etching depth w can be immediately estimated:

$$\Delta n_{eff} = \Delta n_{etch} w I_{norm}(z_{grating}) \quad (10.2.35)$$

where $I_{norm}(z_{grating})$ is the normalized intensity of the mode in the waveguide at the location of the grating, i.e. satisfying

$$\int_{-\infty}^{\infty} I_{norm}(z)dz = 1 \quad (10.2.36)$$

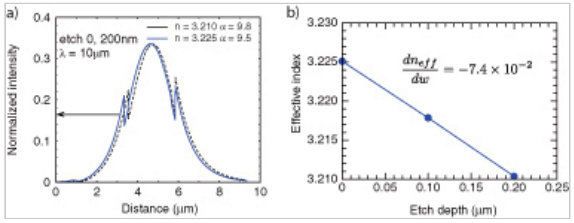


Fig. 10.9. a) Computed mode for a 10 μm wavelength with (dashed) and without (solid) 200 nm thick etched grating. The computed change of refractive index is $\Delta n = 1.5 \times 10^{-2}$ while the losses are not affected. b) Real part of the mode effective index as a function of etched depth. As expected from eqn. 10.2.35, the relationship is linear.

This approach is verified by effective-index computation of an example shown in Fig. 10.9, where the mode profile and effective index are shown as a function of etch depth. As expected from eqn. 10.2.35, a linear relationship is found between the effective index and the etch depth, where the slope of this characteristic $\frac{dn_{eff}}{dw} = -7.4 \times 10^{-2}$ is reasonably close to the one (6.8×10^{-2}) found by eqn. 10.2.35 assuming an average normalized intensity at the interface (from Fig. 10.9) $I_{norm} = 0.17$.

Globally, the buried grating technology allows the fabrication of distributed feedback quantum cascade lasers with the highest performance levels, at least in terms of low dissipation and continuous-wave operation. This is especially true when this technology is combined with a buried heterostructure waveguide. Examples of performance achieved with devices based on this technology are shown in Figs. 10.10 and 10.11, demonstrating true uncooled operation with tens of milliwatts of single-mode operating power.

10.2.4 Tuning and linewidth

*Temperature.*The operating frequency of a distributed feedback quantum cascade laser is fixed by the Bragg condition:

$$\lambda_B(T) = 2n_{eff}(T)\Lambda(T)$$

(10.2.37)

As a result, the tuning with temperature has two terms, and can be conveniently written as:

(p.178)

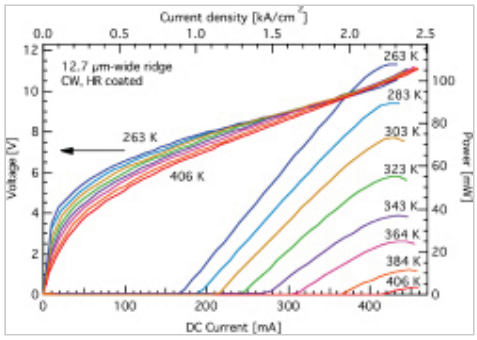


Fig. 10.10. Single-mode optical power versus injected current for a distributed feedback quantum cascade laser combining a buried grating and a buried heterostructure waveguide. Reprinted, with permission, from [206]. Copyright 2009 IEEE.

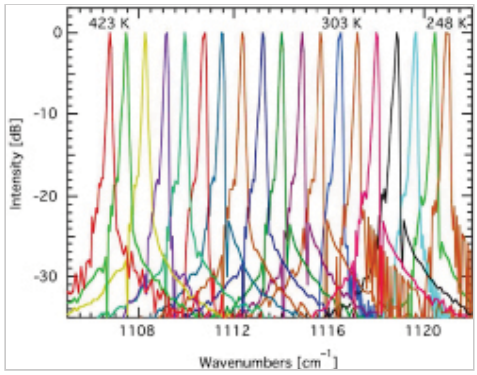


Fig. 10.11. Temperature tuning of the a buried grating device operating in continuous wave above room temperature. Reprinted, with permission, from [206]. Copyright 2009 IEEE.

$$\beta_{tun} = \frac{1}{\lambda_B}\frac{d\lambda_B}{dT} = \frac{1}{n_{eff}}\frac{dn_{eff}}{dT} + \frac{1}{\Lambda}\frac{d\Lambda}{dT}$$

(10.2.38)

Experimentally, this tuning coefficient is $\beta_{tun} = -(6.5 - 9) \times 10^{-5}$ K^{-1}) and is dominated by the temperature tuning of the refractive index, as the second term is the thermal expansion of InP and is an order of magnitude smaller (4.6×10^{-6}). As shown in Table 10.1, this coefficient is

wavelength independent but increases with temperature. It is the reason why the tuning coefficient of continuous-wave devices is larger than that of pulsed ones at the same holder temperature, as the active-region temperature is significantly increased by self-heating in devices that are operating in a continuous wave.

Current.The same temperature coefficient is also responsible for the tuning of the devices with injected current. For slow variations of the thermal load in the device, **(p.179)**

10.1 Measured temperature tuning coefficients of distributed feedback quantum cascade lasers, as reported in ıs references. The tuning coefficient increases with the active region temperature but remains in a very good xımation independent of the wavelength.

ence	$\beta_{tun}(K^{-1})$	λ(μm)	Temperature (K)	Operation
	-6.8×10^{-5}	7.8	220	pulsed
	-6.5×10^{-5}	5.3	220	pulsed
	-6.5×10^{-5}	10.16	200	pulsed
	-7.1×10^{-5}	4.9	260	pulsed
	-7.8×10^{-5}	5.3	330	pulsed
	-8.9×10^{-5}	7.8	310	continuous wave
	-9.5×10^{-5}	5.4	253	continuous wave
	-7.9×10^{-5}	9.0	400	continuous wave

the average temperature of the waveguide T_{wav} is related to that of the holder by a thermal resistance:

$$T_{wav} = T_{sub} + R_{th} U I.$$

(10.2.39)

Note that the thermal resistance defined by the above equation is slightly lower than the one defined by the similar equation for the active region 7.8.27, as the optical mode spreads over the claddings that are cooler than the active region. As a result, the tuning coefficient can therefore be written as

$$\frac{1}{\lambda}\frac{d\lambda}{dP_{elec}} = \frac{1}{\lambda}\frac{d\lambda}{dT}\frac{dT}{dP_{elec}} = \beta_{tun} R_{th}$$

(10.2.40)

As expected, a strong tuning is achieved for a device with a large thermal resistance R_{th}. In short, a device that will provide a large current tuning exhibits a equally large difference between the active region and holder temperatures. In this respect, the opimization for maximum tuning is opposite to the one aiming at reaching the largest continuous wave operation temperature. An example of tuning with both submount temperature and dissipated power is shown in Fig. 10.12.

Dynamical behavior and linewidth. During short current pulses, the waveguide temperature of a distributed feedback quantum cascade laser will undergo a change in temperature and, as a result, induce a chirp of the ouput frequency with time. The example of such a behavior is shown in Fig. 10.13, where a series of time-resolved spectra of a distributed feedback quantum cascade laser were performed on a device driven by a 100 ns long pulse at room temperature. The spectra were aquired using a fast, room-temperature mercury–cadmium–telluride detector and a gated averager with a 3 ns long gate. Because for such short timescales the heat has no time to diffuse out of the active region, the thermal behavior of the device is dominated by the specific heat of the active region, and therefore the mode will shift linearly with time. This is indeed the behavior observed in Fig. 10.13. The device remains single mode but tunes over 2 cm^{-1} during the 100 ns long pulse. This rapid chirp enables an “intrapulse” spectroscopy technique in which absorption features are detected using the very fast chirp of the laser across the gas absorption line.

(p.180)

As a result of this thermal chirp, the integrated linewidth of a distributed feedback laser is expected to grow linearly with pulse length. This is exactly the behavior observed experimentally, as shown in Fig. 10.14. For this device driven at 1.25x the threshold, the linewidth is found to grow linearly with time at a rate $\frac{d(1/\lambda)}{dt} = 0.0075 \mathrm{cm}^{-1} ns^{-1}$. In a very rough approximation, assuming little significant heat flow occurs out of the active region, the chirping rate can be written as

$$\frac{d\nu}{dt} = \frac{V_{op} J_{op}}{t_{act} C_v} \nu \beta_{tun}$$

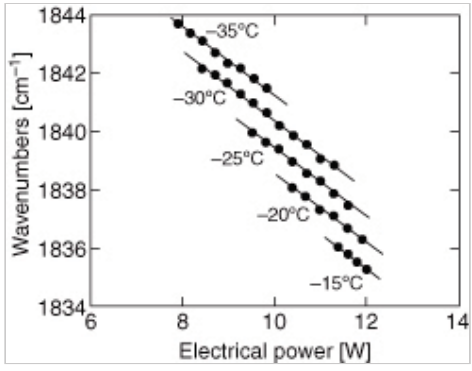


Fig. 10.12. Mode position of a distributed feedback quantum cascade laser as a function of both the submount temperature and the dissipated electrical power. Reprinted with permission from [209]. Copyright 2005, AIP.

(10.2.41)

(p.181)

where J_{op} = 5.4kA/cm^2, is the operating current, V_{op} = 11 V is the operating voltage, t_{act} = 1.75 μm, is the active region's thickness, C_v = 1.64 J K^{-1} cm^{-3} is the specific heat and v is the operation frequency. The value computed by the above equation yields a tuning rate of $\frac{d(1/\lambda)}{dt} = 0.013\text{cm}^{-1}ns^{-1}$, larger than the value measured over the maximum 100 ns pulse length but close to the value measured for the initial 20 ns, where no significant heat flow is expected out of the active region.

Obviously, for very short pulse lengths, the linewidth is expected to grow again because of the time–bandwidth uncertainty relations. These can be expressed for the FWHM of the intensity of both the pulse $\Delta\tau$ and its Fourier transform, in frequency units $\Delta\nu$ as $\Delta\nu\,\Delta\tau$ = 0.44 for a Gaussian pulse and $\Delta\nu\,\Delta\tau$ = 0.88 for a rectangular pulse. In a simple approach the two contributions are simply summed, and, assuming rectangular pulses,

$$\delta\nu = 0.88\frac{1}{\Delta\tau} + \frac{d\nu}{dt}\Delta\tau$$

(10.2.42)

This equation fits relatively well with the experimental data shown in Fig. 10.14, and predict a minimum linewidth of about $\delta(1/\lambda) = \delta\nu/c$ = 0.03cm^{-1} for a pulse length of 2 ns. The minimum pulse width achievable for a given pulse shape (characterized by its time-bandwidth product $\Delta\nu\,\Delta\tau$) and a given tuning rate $\frac{d\nu}{dt}$, is given then by:

$$\delta\nu_{min} = 2\sqrt{\Delta\nu\Delta\tau\frac{d\nu}{dt}}$$

(10.2.43)

Distributed feedback quantum cascade lasers driven by short pulses and tuned by a subthreshold ramp have been used to perform high resolution spectroscopy, an approach pioneered by the group at Aerodyne Research Inc. A measurement of the laser linewidth can be obtained fitting the measured absorption of a known absorption line. Using this technique, a narrow linewidth of

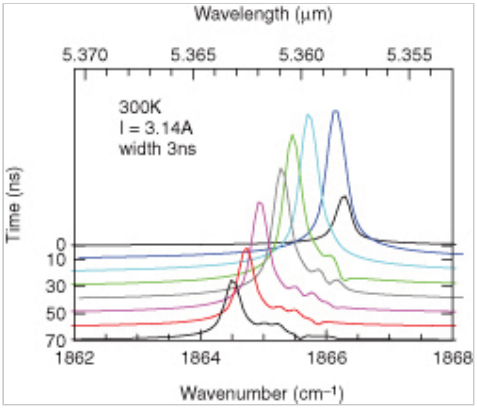


Fig. 10.13. Single-mode spectra of a distributed feedback quantum cascade laser as a function of time during a current pulse. The chirp is thermally driven. Reprinted with permission from [21]. Copyright 1997, AIP.

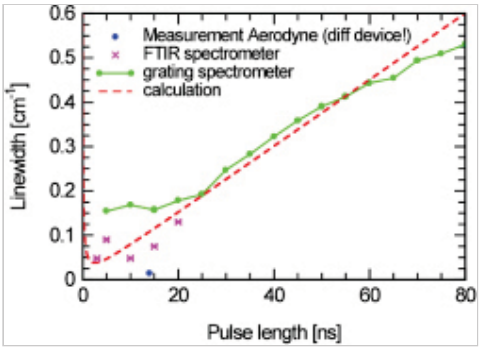


Fig. 10.14. Measured linewidth as a function of pulse length. The point labeled "Aerodyne" is the result of a spectral fit of the absorption line of NH_3 at 967 cm^{-1} that yielded a linewidth of 0.012 cm^{-1}. Adapted from [210].

0.012 cm^{-1} has been measured on a device similar to that used to measure the data shown in Fig. 10.15. Indeed, assuming a Gaussian pulse shape and operating very close to threshold, a minimum linewidth of ≈ **(p.182)**

0.02 cm^{-1} is inferred from eqn. 10.2.43, showing that simply summing the contributions from the uncertainty and chirp tends to overestimate the final linewidth.

In fact, measurements on high-performance devices yielded very narrow linewidths. An example of such a measurement is shown in Fig 10.15, where the measured linewidth is shown as function of the voltage on the pulse, the latter acting as a surrogate for the pulse length [211]. In the latter measurement the devices used exhibited much lower threshold current densities (about $1 kA/cm^2$) and, therefore, yield a predicted linewidth of only 0.01 cm^{-1}, still above the measured experimental value of 0.0072 cm^{-1}. The latter value is obviously larger than the one (0.005 cm^{-1}) given by the time–bandwidth product.

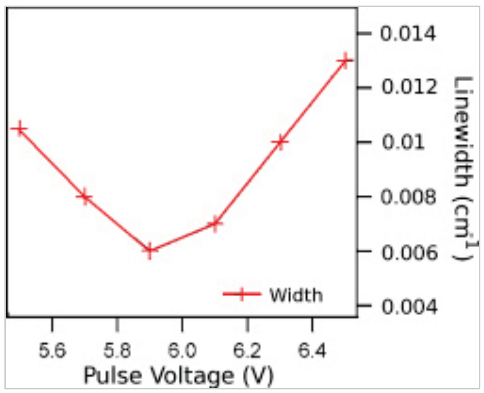


Fig. 10.15. Measured half-width at half-maximum of the emission spectra of a quantum cascade laser as a function of pulse length. The linewidth is obtained by fitting a vibrational molecular line, and the pulse length is an increasing function of the voltage on the pulser, indicated in the horizontal axis. Adapted from [212].

The fact that experimentally measured linewidth values close to the limit imposed by the time–bandwidth product can be achieved is far from obvious. In fact, assuming a perfect rectangular current pulse, much *broader* linewidths than the values given by the time–bandwidth limit are predicted, because the mode competition does not have time to efficiently narrow the emission. The experimental results have been achieved using a current pulse that exhibits a cosine shape where the device is brought just above threshold for only a fraction of the pulse length. In effect, this corresponds to an amplification of an already narrow amplified spontaneous emission.

Linewidth.It was shown by Schawlow and Townes in their celebrated paper [213] that the linewidth of a laser is limited by the ratio between the spontanous and stimulated photon emission. This result was adapted later by Henry [214], who showed that the result from Schawlow and Townes must be multiplied by a term $(1+\alpha_{LEF}^2)$, where the linewidth enhancement factor α_{LEF}, defined as

$$\alpha_{LEF} = \frac{\partial n'/\partial N}{\partial n''/\partial N}$$

(10.2.44)

(p.183) expresses the ratio of the change of the real n' to the imaginary n'' part of the refractive index as a function of the upper-state population N. The interpretation is that each random spontaneous emission event, besides changing the phase of the optical field, also induces a phase

modulation of the total field through the change of the refractive index as a function of the population N. The laser linewidth $\Delta\nu$ is then expressed as

$$\Delta\nu = \frac{\pi h\nu\Delta\nu_c^2}{P}\frac{\alpha_m}{\alpha_{tot}}\frac{n_3}{n_{3t}-n_{2t}}(1+\alpha_{LEF}^2)$$

(10.2.45)

where ν is the lasing frequency, α_{tot} the total losses of the cavity, α_m the outcoupling losses of the output facet, P the output power, $\Delta\nu_c = \alpha_{tot}\, c/(2\pi\, n_g)$ the full width at half maximum of the cavity resonance, and n_{3t} and n_{2t} the populations of the upper and lower states of the laser transition at threshold. Because it expresses the fact that the linewidth of a laser is limited by the ratio of the spontaneous to the stimulated photon emission, eqn. 10.2.45 remains valid for quantum cascade lasers, although it was originally derived in the context of interband semiconductor lasers. As will be discussed below, the very narrow linewidths measured in quantum cascade lasers are *only indirectly* a result of their unique physics.

Using the QCL rate equations presented in Section 7.4, and remaining sufficiently close to threshold so that the upper-state population is close to the threshold value, one can show that

$$\frac{n_3}{n_{3t}-n_{2t}} = \left(1-\frac{\tau_2}{\tau_{32}}-\frac{q}{\tau_3 J_{th}}n_2^{\text{therm}}\right)^{-1}$$

(10.2.46)

where τ_3 and τ_2 are the lifetimes of the upper and lower states, τ_{32}^{-1} is the scattering rate from level 3 to level 2, and n_2^{therm} is the thermal population of level 2.

In an interband semiconductor laser the linewidth enhancement factor is usually large because the gain and the change in absorption do not peak at the same frequency. In contrast, for quantum cascade lasers, it was noted that the symmetric gain profile would yield a small to vanishing value for α_{LEF} for devices operated. Indeed, assuming a Lorentzian lineshape with an FWHM equal to $2\gamma_{32}$ for the intersubband transition, the linewidth enhancement factor of a QCL can be expressed as

$$\alpha_{LEF}(\nu) = -(\nu-\nu_{32})/\gamma_{32}$$

(10.2.47)

and thus is equal to zero at the maximum of the gain curve. Distributed feedback lasers *LEF* will usually operate within a range of $\nu_{32}\pm\gamma_{32}/2$, and as a result $\alpha_{LEF}^2 \langle 0.25$. In fact, a value of $|\alpha_{LEF}| \langle 0.5$ was measured directly through a high-frequency modulation experiment where the ratio of the amplitude to the phase modulation was measured for a device under high frequency modulation [215].

Evaluation of eqn. 10.2.45 with typical values for a mid-infrared distributed feedback quantum cascade laser yields extremely low values for the linewidth. For a device operating at $\lambda \approx 5\ \mu$m with an output power of 10 mW, with a waveguide loss of $\alpha = 5\ \text{cm}^{-1}$ and an outcoupling loss of $\alpha_m = 2\ \text{cm}^{-1}$, a linewidth of $\delta\nu = 400Hz$ is found. For a device of length L and group index n_g inserted in an external cavity with a passive section of length l, this value is reduced further by a factor $F = 1 + l/n_gL$.

(p.184)

Theoretical values in the 1–10 Hz range are easily obtained for a 10 cm long passive section.

Using a rate equation model, Yamanishi [216] derived an expression mathematically equivalent to eqn. 10.2.45, but where the linewidth is expressed as a function of the pumping ratio $I/I_{th} - 1$ instead of the optical power. This expression enabled the authors to show that at a given pumping ratio the low radiative efficiency of intersubband transitions, caused by strong optical phonon scattering, is indirectly responsible for the very narrow predicted line, as it naturally yields a very large ratio of the stimulated emission rate to the spontaneous rate in the laser mode. This result has been sometimes wrongly interpreted as implying that the linewidth of a QCL is *narrower* than the one predicted by the Schawlow–Townes limit of eqn. 10.2.45. This is, of course, nonsense: QCL have narrow linewidth because they are *powerful, and operate with low mirror losses and at long wavelengths*—all features that help reduce the linewidth in eqn. 10.2.45.

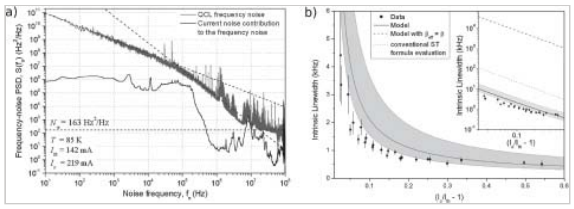


Fig. 10.16. a) Noise spectrum of a quantum cascade laser operated in a continuous wave. b) Measured linewidth as a function of current. Reprinted with permission from [217]. Copyright 2010, APS.

The noise properties of a single-mode distributed feedback quantum cascade laser operating at 4.3 μm were studied by Bartalini *et al*. [217]. In this measurement the light from the quantum cascade laser was passed through a cell filled with CO_2. The laser was tuned to the side of a strong absorption line; the frequency noise of the laser was therefore converted into an amplitude noise, recorded with a fast detector over a bandwidth up to 100 MHz. The results of such a measurement is shown in Fig. 10.16a.

In the low-frequency region the noise exhibit first a $1/f$, followed by an $1/f^2$ frequency dependence. This noise is attributed to the conversion of the amplitude noise of the current source to frequency noise through the temperature dependence of the refractive index, the $1/f$ and $1/f^2$ reflecting the thermal response roll-over of the device. At higher frequencies ($\rangle\ 10^7$ Hz) the noise reaches a plateau where a spectral density of the noise $N_w = 163$ Hz^2/Hz corresponds to the Lorenzian natural linewidth of the laser $\delta\nu = \pi N_w$ [218] that yields a value of $\delta\nu = 513$ Hz in this case. This value, achieved for an optical power of 6 mW, is in good agreement with the result ($\delta\nu = 530$ Hz) predicted by eqn. 10.2.45, assuming rather typical values for the **(p.185)**

waveguide $\alpha_w = 4$ cm^{-1} and mirror $\alpha_m = 1.7$ cm^{-1} losses of these strongly coupled distributed feedback quantum cascade lasers. Fig. 10.16b shows the comparison between the experimental results for various injection currents to the Schawlow-Townes linewidth expressed in the form given by Yamanishi, showing an excellent agreement.

Cavity pulling. An important difference between distributed

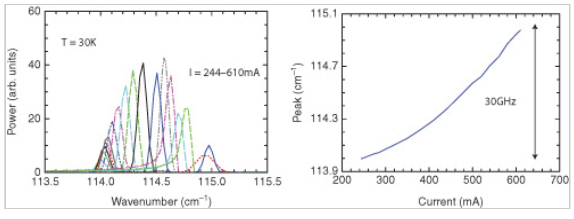


Fig. 10.17. a) High-resolution spectra of a single-frequency terahertz quantum cascade laser based on a photonic bandgap structure as a function of current. b) Peak position as a function of current. The tuning originates from a cavity pulling effect.

feedback quantum cascade lasers and their interband counterparts is the almost total absence of carrier density-driven changes of refractive index. In fact, the requirement of preserving global neutrality in a unipolar device forces the number of electrons to remain constant. As a result, the only changes in refractive index that can be expected are related to shifts of the carrier distribution or transition energies as a function of applied voltage or temperature. In particular, a blueshift of the gain curve with applied bias is expected to induce a concomitant blueshift of the mode. As shown in Fig. 10.17, this phenomenon, called "cavity pulling" in the literature of gas lasers, has been observed in single-mode quantum cascade laser-operating in the terahertz. By assuming the laser to operate close to the peak of a Lorenzian gain curve whose peak frequency shifts at a rate $\partial \nu_G/\partial J$, the cavity pulling of the lasing mode $\partial \nu_L/\partial J$, obtained by ratioing the imaginary to the real part of the susceptibility, can be rewritten as

$$\frac{\partial \nu_L}{\partial J} = \frac{\hbar c G_M}{n_{eff} 2\gamma} \frac{\partial \nu_G}{\partial J}$$

(10.2.48)

where G_M is the modal gain and γ is the half-width at half-maximum of the gain curve.

It has never been observed in mid-infrared devices, for two reasons. First, the large ratio of upperto lower-state lifetimes necessary to produce high-performance devices implies that the upper-state density is practically "locked" to a constant value above threshold. As a result, the differential resistance of the device decrease abruptly and the Stark tuning of the transition above threshold is minimum. Secondly, high-performance operation is achieved by employing low-loss cavities while the half-width **(p.186)**

at half-maximum remains large at room temperature. As a result, the prefactor in eqn. 10.2.48 remains much lower than unity, and the concomitant change in refractive index is therefore small.

10.3 External cavities

The tuning range of distributed feeback quantum cascade lasers is not limited by the width of the gain curve but by the amount that the mode effective index can be tuned, usually by temperature. As a result, for quantum cascade lasers, the maximum tuning range that has been achieved is at most ≈ 10 cm^{-1} for temperature tuning and ≈ 1 cm^{-1} for current tuning, much less than the width of a gain curve ≈ 120 cm^{-1}. Quantum cascade lasers with large tuning capabilities would be desirable for spectroscopy of multiple gas components or liquids with broad absorption lines. One technique to achieve this goal is to use a tunable wavelength filter outside the laser cavity. Such an external cavity configuration is achieved by using a combination of a quantum cascade active gain medium with an antireflection coating, followed by a lens and a grating.

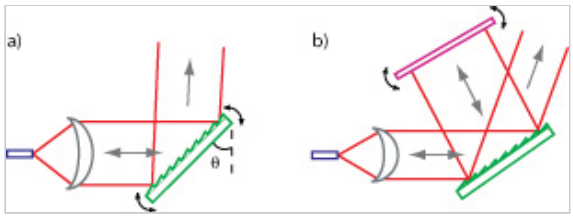


Fig. 10.18. a) Littrow and b) Littman external cavity configurations. The large gray arrows indicate the directions of propagation of light, and the small black arrows show how coarse tuning is achieved.

The most common configurations are the Littrow and Littman–Metcalf [219] configurations illustrated in Fig. 10.18. Tuning is achieved by the rotation of the grating for the Littrow, and the rotation of a mirror for the Littman–Metcalf configuration. Whereas the Littmann configuration

offers a stronger wavelength selectivity by a double pass through the grating, the Littrow arrangement maximizes the back-coupling into the active region chip. It was for this reason selected for the majority of QCL-based experiments.

10.3.1 Mode selection in the Littrow configuration

As shown in Fig. 10.19, tuning of an external cavity quantum cascade laser in the Littrow configuration can be analyzed by a wavelength-dependent amplitude reflection coefficient of the external cavity $\tilde{r}_1(\theta_G, \lambda)$. Neglecting the reflectivity of the AR coated lens and in the limit of small amplitude reflectivity from the chip facet r_1, we obtain

$$\tilde{r}_1(\theta_G, \lambda) \cong r_1 + T_1 r_{\text{ext}}(\theta_G, \lambda) e^{2ikl}$$

(10.3.49)

where θ_G is the grating angle, T_1 the intensity transmittivity of the facet, $k = 2\pi/\lambda$, l the distance from the facet to the grating, and $r_{\text{ext}}(\theta_G, \lambda)$ describes the reflection of **(p.187)** the external part of the cavity. The latter can be written as

$$r_{\text{ext}}(\theta_G, \lambda) = T_L \sqrt{R_G} \eta$$

where T_L is the transmission of the lens, R_G the first-order efficiency of the grating, and $\eta(\theta_G, \lambda)$ the coupling of the beam reflected by the grating and lens combination onto the initial waveguide mode. The latter can be computed as the overlap integral between the normalized amplitude profiles of these two beams on the coupling facet:

$$\eta = \int \mathrm{d}x \mathrm{d}y u_i^*(x, y) u_r(x, y)$$

(10.3.50)

The mode will be assumed to have an elliptical Gaussian beam profile [220] whose waists in the xand y-directions $W_{0x,\,y}$ are located on the facet ($z = 0$). The values of $W_{0x,\,y}$ are inferred from a measurement of the far-field of the device. The normalized amplitude profile of the beam on the facet reads

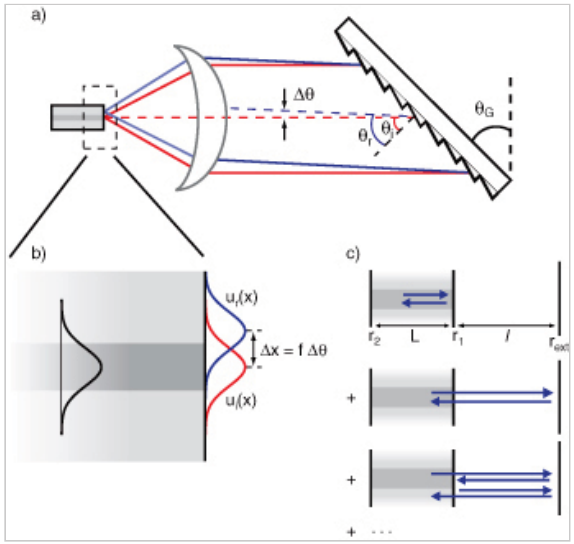


Fig. 10.19. a) Schematic drawing of the Littrow configuration. The incident horizontal beam is diffracted by the grating at an angle $\Delta\theta$ depending on the wavelength λ. b) The angle $\Delta\theta$ introduces a shift of the reflected beam by a distance $\Delta x = f\,\Delta\theta$. c)

Schematic representation of all the terms for the calculation of $\tilde{r}_1(\theta_G, \lambda)$.

$$U(x,y) = \left(\frac{2}{\pi W_{0x} W_{0y}}\right)^{\frac{1}{2}} \exp\left[-\left(\frac{x^2}{W_{0x}^2} + \frac{y^2}{W_{0y}^2}\right)\right]$$

(10.3.51)

(p.188) and inserting in eqn. 10.3.50, we obtain

$$\eta = \int dx dy U(x,y) U(x - \Delta x, y) = \exp\left(-\frac{\Delta x^2}{2W_{0x}^2}\right)$$

(10.3.52)

The displacement Δx between the incident and the reflected beams is

$$\Delta x = f \Delta \Theta$$

(10.3.53)

where f is the focal length of the lens and $\Delta\Theta$ is the difference between the angle of the incident and the reflected beams. Significant feedback will only occur when the incident and diffracted beams are nearly collinear, i.e. near the Bragg reflection condition

$$\lambda_G = 2d \sin\theta_G$$

for a grating of periodicity d, at wavelength λ_G and an angle θ_G. For wavelengths close to λ_G, we write $\theta_{r,\,1} = \theta_G + \Delta\theta$ and linearize the grating equation in the vicinity of λ_G:

$$\Delta\theta = \frac{\lambda - \lambda_G}{d \cos\theta_G}$$

(10.3.54)

We finally obtain for the effective reflectivity of the external cavity

$$r_{\text{ext}}(\theta_G, \lambda) = T_L \sqrt{R_G} \exp\left[-\frac{f^2(\lambda - \lambda_G)^2}{2W_{0x}^2 d^2 \cos^2\theta_G}\right]$$

(10.3.55)

Note that this expression has been derived for an ideal (diffraction limited) lens. In the case of an imperfect lens, the size of the reflected beam on the facet is larger, and consequently r_{ext} has a broader and less intense peak at λ_G.

The allowed modes of the external cavity and the corresponding threshold gains are determined by the stationary condition

$$\tilde{r}_1(\theta_G, \lambda) r_2 e^{2inkL} e^{(G_M - \alpha_w)L} = 1$$

(10.3.56)

(the wave should have the same amplitude and phase after one round-trip in the cavity), where n is the refractive index of the chip and L its length. The condition on the amplitude readily gives the threshold modal gain as a function of wavelength:

$$G_{M,th}(\lambda) = \alpha_{\text{tot}}(\lambda) = \alpha_w + \frac{1}{L} \ln \frac{1}{|\tilde{r}_1(\theta_G, \lambda) r_2|}$$

(10.3.57)
and the condition on the phase gives the allowed modes:

$$\arg[\tilde{r}_1(\theta_G, \lambda)] + \frac{4\pi nL}{\lambda} = N \cdot 2\pi$$

(10.3.58)
For wavelengths far from λ_G, the first term in $\tilde{r}_1(\theta_G, \lambda)$ dominates and the usual results for a Fabry–Perot chip are recovered. The modes are spaced by $\Delta\nu_{FP} = \frac{c}{2nL}$ and the threshold gain equals

(p.189)

$$G^{FP}_{M,th}(\lambda) = \alpha_{FP} = \alpha_w + \frac{1}{2L}\ln\frac{1}{R_1 R_2}$$

(10.3.59)
On the contrary, for wavelengths close to λ_G the second term dominates. In this case, $\arg(\tilde{r}_1) \cong = 2ikl$, resulting in a mode spacing equal to $\Delta\nu_{EC} = \frac{c}{2(nL+l)}$. The minimum threshold gain is

$$G^{EC}_{M,th}(\lambda_G) = \alpha_{EC} = \alpha_w + \frac{1}{2L}\ln\frac{1}{R_{EC}R_2}$$

(10.3.60)
with $R_{EC} = (r_1 + T_1 T_L\sqrt{R_G})^2$.

The coarse tuning range of the external cavity laser can be readily obtained by requiring that the threshold condition for the external cavity mode $g(\lambda) = G^{EC}_{M,th}$ is met, and that before the laser operates on the parasitic modes of the Fabry–Perot cavity $g(\lambda_{max}) = G^{FP}_{M,th}$. As a result, we obtain immediately:

$$\frac{g(\lambda)}{\alpha_{EC}} \geq 1\rangle\frac{g(\lambda_{max})}{\alpha_{FP}}$$

(10.3.61)
For a perfect antireflection coating, α_{FP} is very large and the tuning is limited only by the available gain. Conversely, large coarse tuning ranges could be achieved even with relatively poor AR coatings if the gain is very flat spectrally, i.e. if $g(\lambda) \approx const$.

The fine-tuning behavior can be predicted numerically by using eqn. 10.3.56 to determine the external cavity mode with the lowest losses. The results of this computation for the case in which only the grating angle is varied is shown in Fig. 10.20. The simulation is done for a 3 mm long HR coated chip with waveguide losses of 10 cm^{-1}. The lasing frequency as a function of the grating angle is plotted for two values of the front facet reflectivity, 2% and 0.2%. In the first case, the tuning happens only on the external cavity modes that are close to the Fabry–Perot modes of the chip. In the second case, on the contrary, tuning between the chip FP modes is possible, and mode hops occur only on the EC modes. These mode hops can be suppressed by varying the EC length together with the grating angle.

Fig. 10.21 shows the computed losses and allowed modes of the external cavity in the two cases for one given grating angle. The angle is chosen so that the grating selected wavelength λ_G is roughly mid-way between two chip FP modes. In this case the waves reflected by the grating and the facet interfere destructively. For R_1 = 2%, the amplitude of these partial waves are of the

same order of magnitude so that the modes near λ_G have a larger threshold than the modes near the next FP mode, for which r_{ext} is smaller but the partial waves are in phase. For $R_1 = 0.2\%$ the contribution from the grating dominates, and consequently the modes which are the closest to λ_G have the lowest threshold despite the destructive interferences.

10.3.2 Experiments

The first realization of an external cavity quantum cascade laser was reported by Luo *et al*. in 2001 [221]. A tuning range of 34 cm^{-1} was achieved in these initial experiments. Much broader tunings were achieved by active regions based on bound-to-continuum transitions operating close to room temperature. These active regions **(p.190)**

(p.191) exhibited both a better gain margin and an inherently broader gain spectrum enabling a tuning of 150 cm^{-1} at $\lambda \cong 10\ \mu m$ using an active region design with a gain full width at half maximum equal to 297 cm^{-1} [222]. An important step was achieved by the demonstration, by Maulini *et al*. [223] in 2005, of an external cavity device operated in a continuous wave and tunable over more than 170 cm^{-1} on a Peltier cooler. Subsequent progress in the performance levels of devices based on single active regions were the increase in the operation temperature, extension to wavelength $\approx 8-10\ \mu m$ using buried heterostructure waveguides [224,225] and the increase in the output power [226].

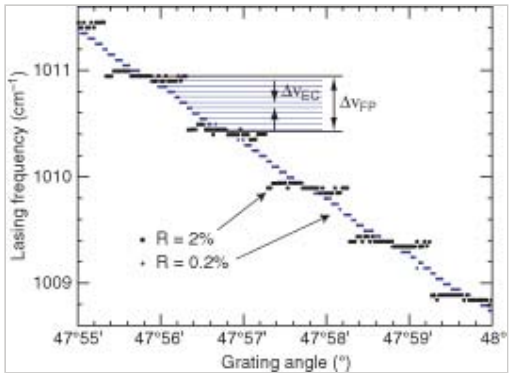
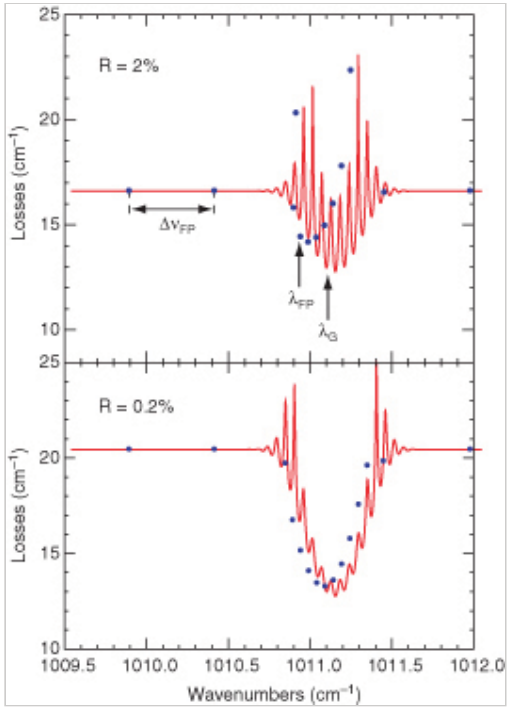


Fig. 10.20. Fine-tuning behavior of an external cavity QCL for two different values of the front-facet reflectivity. With $R_1 = 2\%$ tuning is possible only on EC modes that are very close to the FP modes of the chip. With $R_1 = 0.2\%$ tuning is possible on all the EC modes.

Fig. 10.22a shows the applied bias voltage and continuous-wave output power of a gain element, based on a bound-to-continuum active region operating at a wavelength close to 5 μm as a function of the injection current [223]. The heat-sink temperature is –30° C. With both facets as cleaved, the continuous-wave threshold current density J_{th} at this temperature was 1.73 kA/cm^2, and the maximum continuous-wave operating temperature was 0° C. After deposition of an Al_2O_3/Au (300 nm/100 nm) high-reflection coating on the back facet, J_{th} decreased to 1.35 kA/cm^2 at –30° C, and the



maximum operating temperature reached 35° C, with still more than 20 mW of output power at 25° C. As AR coating, a quarter-wave layer of Al_2O_3 on the front facet was evaporated. The computed (Fresnel) reflectivity is 1%. After deposition, a threshold $J_{th} = 2.03$ kA/cm^2 at –30° C was measured. Using the value of $\alpha_w = 5$ cm^{-1}, deduced from the threshold current measurement before and after HR coating, we computed a residual reflectivity $R_{AR} = 3\%$.

Fig. 10.21. External cavity losses as a function of the wavenumber for a given grating angle and two different values of the front facet reflectivity. The disks mark the allowed modes. The angle is chosen so that the reflections from the facet and the grating interfere destructively.

The gain element was then inserted in an external cavity configuration based on a Littrow arrangement. Fig. 10.22a shows an output power-versus-current characteristic of an ECQCL tuned near the maximum of the gain curve. The addition of the grating feedback lowered the threshold current density of the chip down to 1.26 kA/cm^2 at –30° C. The ratio of this value to that of the HR coated chip allows one to calculate the equivalent reflectivity of the external part of the cavity $REC = 36\%$.

By changing the grating orientation, the laser could be tuned over 174 cm^{-1} (0.46 μm) from 4.94 to 5.4 μm, i.e. over 9% of the center frequency. The side-mode suppression ration was larger than 25 dB over 169 cm^{-1} (see Fig. 10.22b, and the laser was single-mode with a SMSR ≥ 30 dB (the noise level of the FTIR) over 142.5 cm^{-1} (0.37 μm) from 4.95 to 5.32 μm, i.e. over 7% of the center frequency.

10.3.3 Dynamical properties of an external cavity

The spectrum of an external cavity quantum cascade laser, formed by a bound-to-continuum gain chip operating at $\lambda \approx 9.5$ μm in a Littrow configuration, was studied in pulsed operation. Some resultant spectra, taken at 2 ns intervals with a 2 ns long gate, are shown in Fig. 10.23. They clearly display the mode competition occurring at the beginning of the pulse. The Fabry–Perot cavity modes lase first because of their shorter round-trip time, but then vanish when the grating-selected mode, having a lower threshold gain, starts to lase. The linewidth of these spectra is smaller than the resolution of our spectrometer (3.75 GHz), proving that the relatively large timeaveraged linewidth is due to the thermal drift of the wavelength during the pulse. The inset is a spectrum recorded using a gate covering the end of the pulse, from 12 ns to 50 ns. **(p.192)**

During this time interval the laser is single-mode in the limit of our spectrometer resolution (~30 dB). These results show clearly the effect of mode competition between two well-separated spectral frequencies inside the gain curve and the homogeneous nature of the intersubband broadening in these bound-to-continuum structures. To illustrate this point, the quantum cascade laser rate equations were

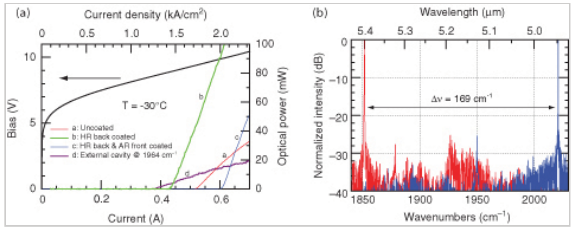


Fig. 10.22. a) Bias voltage and continuous-wave optical power of the gain element and the external cavity laser as functions of the injection current at a heat-sink temperature of – 30° C. b) Spectra taken at the extreme tuning range of an external cavity laser,

rewritten for two photon flux densities S_{FP} and S_{EC} describing the chip Fabry–Perot modes and the external cavity modes, respectively. To take into account the longer round-trip time of the lat-*nL* ter, its speed was multiplied by $1/\rho_{cav}$, where $\rho_{cav} = 1 + \frac{l}{nL}$ is the ratio of the (optical) lengths of the two cavities, *L* being the chip length, *n* its refractive index, and *l* the

(p.193)

distance from its front facet to the external grating. The resultant equation is

displaying nevertheless an SMSR ≥ 25 dB. The optical power was more than 5 mW over most of the tuning range.

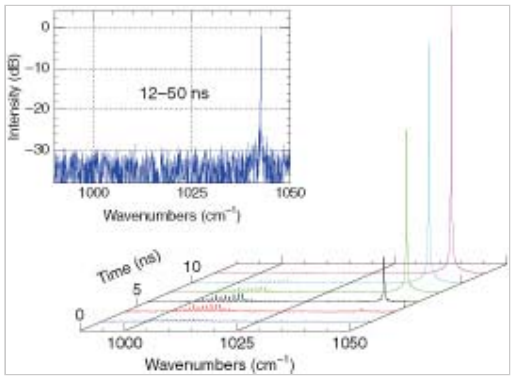


Fig. 10.23. Time-resolved spectra of an external cavity laser. The spectra are taken at 2 ns intervals with a 2 ns long gate. Inset: logarithmic scale spectrum of the end of the pulse (from 12 ns to 50 ns).

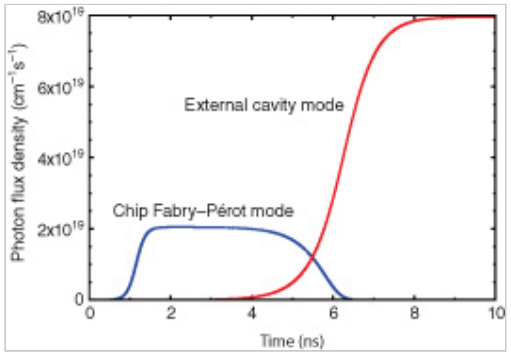


Fig. 10.24. Simulated photon flux densities as functions of time. The time origin corresponds to the begining of the current pulse.

$$\frac{dn_3}{dt} = \frac{J}{q} - [S_{FP}g_{FP} + S_{EC}g_{EC}](n_3 - n_2) - \frac{n_3}{\tau_3}$$

(10.3.62)

$$\frac{dn_2}{dt} = \frac{n_3}{\tau_{32}} + [S_{FP}g_{FP} + S_{EC}g_{EC}](n_3 - n_2) - \frac{n_2}{\tau_2}$$

(10.3.63)

$$\frac{dS_{FP}}{dt} = \frac{c}{n}[(g_{FP}(n_3 - n_2) - \alpha_{FP})S_{FP} + \frac{\beta n_3}{\tau_{sp}}]$$

(10.3.64)

$$\frac{dS_{\mathrm{FC}}}{dt} = \frac{c}{\rho_{\mathrm{cav}} n}[(g_{\mathrm{EC}}(n_3 - n_2) - \alpha_{\mathrm{EC}})S_{\mathrm{FC}} + \frac{\beta n_3}{\tau_{sp}}]$$

(10.3.65)

where g_{FP} and g_{EC} are the values of the gain cross-section $g_c(\nu)$ at its maximum and at the grating-selected frequency ν_{EC}, respectively. α_{FP} and α_{EC} are the total losses of the AR coated chip and the external cavity. β, the fraction of the spontaneous emission emitted in the mode, is assumed to be identical for both modes, for simplicity. The numerically computed electronic populations and photon flux densities (per unit length, per period) are shown in Fig. 10.24 where the experimentally observed behavior is reproduced. After a few picoseconds the population inversion Δn reaches the value $J\tau_{\mathrm{eff}}/q$. The photon flux densities grow as e^{Gvt} where $G = g_c\Delta n - \alpha$ is the net modal gain and $v = c/n$ the speed of light. Despite a higher net gain due to smaller losses, S_{EC} increases more slowly because of its ≈ 10 times smaller effective speed $c/(\rho_{\mathrm{cav}}\, n)$. As S_{FP} increases, Δn decreases because of stimulated emission until it reaches the value $\alpha_{\mathrm{FP}}/g_{\mathrm{FP}}$. Then S_{FP} is amplified no longer, and saturates, but S_{EC} continues to grow because its net gain is still positive. Doing this decreases further the population inversion, which becomes too small to compensate the losses for S_{FP}. The FP modes thus vanish, and the external cavity mode continues to grow until the population inversion reaches the equilibrium value $\alpha_{\mathrm{EC}}/g_{\mathrm{EC}}$.

(p.194) 10.3.4 Heterogeneous stacks

As predicted by eqn. 10.3.61, the part of the gain curve $g(\nu)$ over which an external cavity quantum cascade laser can be tuned is given by the condition $g(\nu)/g_{\mathrm{max}} \geq \xi$, where g_{max} is the maximum gain and ξ is the ratio between the threshold gain of the EC–QCL and the threshold gain of the AR coated chip. In a typical Littrow configuration, measured values of ξ between 0.65 and 0.8 are found, depending on the characteristics of the chip. In order to realize a more broadly tunable EC-QCL, it is thus necessary to design a structure with gain variation of less than ~20% over a larger wavelength range. A way to achieve this is to make a cascade containing dissimilar stages emitting at different wavelengths.

A similar concept already been used in the field of optical communications; nearinfrared external cavity diode lasers with a tuning range of 240 nm (17% of center wavelength) have been realized using non-identical multiple quantum-wells [227]. QCLs, however, are more adapted to multi-wavelength lasing than interband lasers, for two reasons. First, the cascaded geometry, in which the current flows through the active regions arranged in series and not in parallel as in multiple-quantum-well diode lasers, insures the same injection efficiency in all the active wells independently of the number of active regions. Secondly, in opposition to interband lasers in which the gain is always accompanied by absorption at higher frequencies, there is no reabsorption between active regions emitting at different wavelengths. As a result, an active region tuned to a specific transition frequency can be designed to be transparent over a wide range around this frequency. By combining active regions with a gain peaking a different frequencies, much larger tunings were then achieved.

This concept was first demonstrated at Bell Laboratories by Gmachl *et al*., who realized a two-wavelength laser emitting simultaneously at 5.2 and 8 μm [228] and an ultra-broadband laser [229]. These first devices exhibited relatively limited performances, especially when operated at room temperature.

Significant progress was achieved by combining active regions based on bound-to-continuum transitions. This active region architecture displays an intrinsically wide gain spectrum. By

combining two structures separated by a spacing corresponding approximately to their half-width, a resultant gain characterized by a very small frequency dependence over a wide frequency range could be obtained. As shown in Fig. 10.25, devices based on such a combination achieved a tuning range of 250 cm^{-1} [230, 231] in pulsed, and 200 cm^{-1} [231] in continuous-wave operation. By combining more stacks more widely spaced, an even wider gain curve could be designed.

In [232] a structure based on a combination of five different active region stacks was used. To prevent cross-reabsorption, a model that computed the gain over a wide frequency range was used at the design phase. As shown in Fig. 10.26 a), the computed gain profile of the individual gain regions were combined to provide a smooth gain profile that matched the expected wavelength dependence of the waveguide losses.

The device was operated in an external cavity formed using a Littrow configuration. As shown in Fig. 10.26b, a very wide tuning range of 430 cm^{-1} was achieved with a peak power larger than 0.5 W for most of the tuning range **(p.195)**

.

10.4 DFB arrays

The external-cavity QCLs demonstrate very large tuning characteristics, and are able to exploit fully the broad-gain bandwidth of QCLs and operate with a narrow linewidth. They are not, however, monolithic devices, and therefore are in general more bulky and fragile than DFB lasers. In addition, the wavelength stability depends crucially on the mechanical stability of the device. The gain chip is in general easier to fabricate than a corresponding DFB laser, though it requires the deposition on its facet of a very finely tuned anti-reflection coating to bring the residual reflectivity below 10^{-3}.

One possibility of exploiting the width of the gain curve of a QCL using a DFB laser is to fabricate an ensemble of devices at predefined wavelengths. For a general purpose spectrometer, those would be equally spaced by the spacing corresponding to the tuning of an

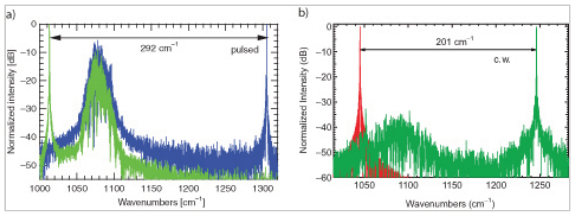


Fig. 10.25. Laser spectra, taken at the border of the tuning range, for a device based on two active regions designed at $\lambda = 8.4\ \mu$m and $\lambda \approx 9.6\ \mu$m. a) Pulsed operation. b) Continuous wave operation. Reprinted, with permission, from [231]. Copyright 2008 IEEE.

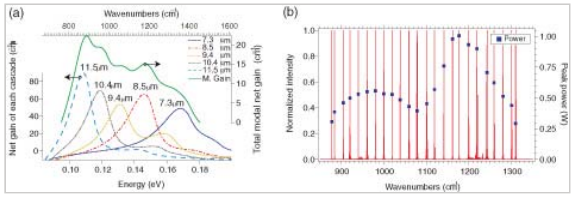


Fig. 10.26. Very broadband operation of a quantum cascade laser based on a combination of five different active regions combined in the same waveguide [232]. a) Computed gain profiles of the individual active regions close to the maximum current. b) Left axis: spectra measured in pulsed mode for various orientation of the grating, demonstrating the very large total tuning range of 430 cm^{-1}. Right axis: corresponding peak power.

individual device; for a dedicated instrument the wavelengths **(p.196)**

could be selected to discriminate a specific chemical against an ensemble of interfering substances.

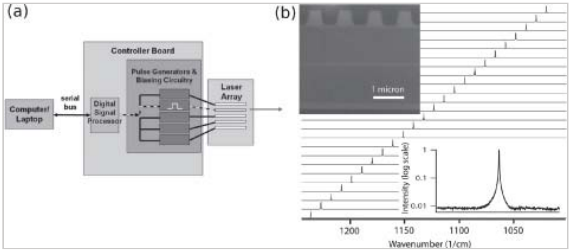


Fig. 10.27. a) Schematic block diagram of a spectrometer based on a DFB array. Reprinted with permission from [234]. Copyright 2007, AIP. b) Emitted spectrum of an array of 24 a DFB laser, spanning a frequency range of 220 cm^{-1}. Upper inset: scanning electron micrograph of the buried heterostructure grating. Lower inset: spectrum of a single device, shown in a log scale, displaying a high side-mode suppression ratio. Reprinted, with permission, from [233], copyright 2009 IEEE.

10.4.1 Linear array

Fig. 10.27 shows the spectra of a linear array of such DFB lasers, fabricated side by side on a single chip [233]. The epitaxial layer is an heterogeneous stack covering the wavelength range 8–10 μm, used originally for external cavity devices as displayed in Fig. 10.25. The devices are spaced by a physical distance of 75 μm, so that the full chip is only 5 mm long.

Although the back facet of the device was obtained by dry etching, the accuracy of the process is not high enough, so that the phase of the back facet reflectivity could be controlled to a constant value. For this reason, in order to keep the necessary high single mode yield, the devices used an extremely strong coupling constant. To this end, a strong buried etched grating with an etch depth of 500 nm was used, providing a very strong coupling product $\kappa L \approx 11$. The disadvantage of such an approach is that a large device-to-device fluctuation of the slope efficiency was observed. Further work will involve beam-combining as well as other ways to control the phase problem. Such devices hold great promise for integrated spectrometers with dedicated purposes.

10.4.2 Array of DFB ring lasers

An elegant way to solve the problem of the back facet phase of a DFB QCL is to fold the latter on itself and fabricate a DFB ring laser. The light extraction is then achieved through the surface or through the substrate by a second-order grating. The optical-loss component of the grating also discriminates between the two modes on each side of the stop band. In addition, the bending losses on the ring provide a strong **(p.197)**

mode-selection mechanism. Research in this direction has proceeded very actively in both mid-infrared [235] and terahertz [236] frequency regions. Fig. 10.28 shows the performance of a two-dimensional 16-element QCL array based on ring-cavity surface-emitting lasers. As shown in Fig. 10.28a, the entire array occupies an area 4 × 3 mm^2 in size. The inset in the same figure shows dry-etched second-order gratings to

allow for vertical light emission, and provides the wavelength control.

As shown in Fig. 10.28b, each element of the two-dimensional array operates in single mode, providing a comb of individual wavelengths separated by 12.3 cm^{-1} and covering an entire spectral range of 180 cm^{-1} from 1140 to 1320cm^{-1}. Because the **(p.198)** boundary conditions are now completely defined by the periodicity of the ring, the device not only offers a theoretical 100% single-mode yield, but also provides optical powers that vary only through the wavelength dependence of the material gain. Indeed, as shown in Fig. 10.28c, the light-versus-current characteristics of devices located side by side spectrally are now differing by less than 50%, the overall power dependence originating from the position of the single-mode emission compared to the gain curve.

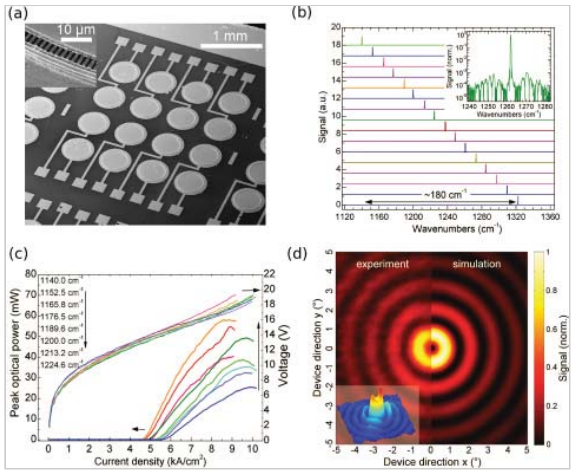
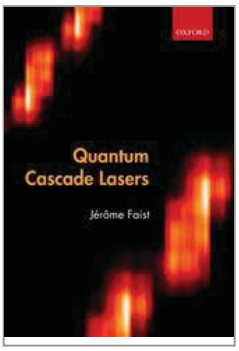


Fig. 10.28. a) Scanning electron microscopy picture of a two-dimensional sixteen-element QCL array based on ring-cavity surface-emitting lasers. The inset shows a waveguide-section illustrating the etched second-order gratings to allow for vertical light emission. b) Spectra of sixteen-surface emitting DFB lasers of the two-dimensional array. Individual wavelengths are separated by 12.3 cm^{-1} and cover an entire spectral range of 180 cm^{-1}. The inset shows a single-mode spectrum at 1261 cm^{-1} with a side mode suppression ratio of 30 dB. c) Peak optical power and voltage-versus-current density characteristics for eight individual lasers, located at the low-frequency side (frequencies shown as inset). d) Measured (left half) and simulated (right half) surface emission far-field pattern of a ring laser. Reprinted with permission from [237] and [238]. Copyright 2010, 2011, AIP.

One particularity of ring-cavity surface-emitting lasers is the shape of the far-field, illustrated in Fig. 10.28d. As predicted by a theoretical model, the latter exhibits a set of concentric rings with a low divergence. The polarization is azimuthal. As the beam is already strongly collimated, one could build a simple sensing system by using a concomitant array of detectors facing that of the lasers. The circular beam, however, would require a phase plate to be converted into a Gaussian beam shape.



Quantum Cascade Lasers
Jérôme Faist

Print publication date: 2013
Print ISBN-13: 9780198528241
Published to Oxford Scholarship Online: May 2013
DOI: 10.1093/acprof:oso/9780198528241.001.0001

Device properties and characterization

Jérôme Faist

DOI:10.1093/acprof:oso/9780198528241.003.0011

Abstract and Keywords

The performance of quantum cascade lasers can be measured after a long sequence of steps, starting from the design, epitaxial growth, process, mounting, and finally measurement. Ideally, one should find ways to characterize each step separately to shorten the optimization cycle, but this proves difficult for quantum cascade lasers. The usual proxies — optical inspection, X-ray, SIMS and TEM characterization of the epitaxial layers, and SEM characterization of the ridges for the process — provide valuable information on these respective steps. However, it is difficult to provide more direct proxies for the performance of the epitaxial layers, similar to the broad area device fabrications performed for interband lasers. This lack of reliable proxies has raised the need for careful measurements of the various spectral and electrical characteristics of the final laser devices, which are reviewed in this chapter. It discusses basic electrical and optical characterization, electroluminescence and spectral measurements, far field measurements, active region temperature, and gain and loss measurements.

Keywords: quantum cascade lasers, performance measurement, epitaxial layers, electroluminescence, spectral measurements, far field measurements, active region temperature, gain and loss measurements

Quantum cascade laser performances are measured after a relatively long sequence of steps, starting from the design, epitaxial growth, process, mounting and finally measurements. Ideally, one should find ways to characterize each step separately in order to shorten the optimization cycle. For quantum cascade lasers, this proved to be difficult. Obviously, the

usual proxy such as optical inspection, X-ray, SIMS and TEM characterization of the epitaxial layers, SEM characterization of the ridges for the process provide valuable informations on these respective steps. It has, however, been difficult to provide more direct proxies for the performance of the epitaxial layers, similar to the broad area device fabrications performed for interband lasers.

As a result, this lack of reliable proxies has stressed the need for careful measurement of the various spectral and electrical characteristics of the final laser devices. In this chapter we will review these.

11.1 Basic electrical and optical characterization

The most important and basic characterization is the measurement of the voltage and light-current characteristics as a function of temperature. For most devices it is important to be able to perform this characterization in pulsed mode using relatively short (50–100 ns) electrical pulses with a low (1-2%) duty cycle to remove the effects of self-heating from the characteristics. An example of such a measurement set-up is displayed in Fig. 11.1.

Measurement of the optical power. In this example the power is measured using a calibrated thermopile power meter, with the light collected from the device using a high (NA = 0.8) broadband antireffection coated aspheric lens. By comparing the power collected using this arrangement with that measured by placing the power meter directly facing the laser, an overall collection efficiency of 70–80% is deduced. It corresponds to the losses from the finite reflectivity of the lenses and optical windows (about 5% each), as well as the remaining clipping of the optical field. Thermopile detectors have a response time of about 1 sec, and as such, the peak power is deduced from the average power by dividing by the electrical duty cycle. The latter must be kept large enough (typically 1%) to keep a good signal-over-noise ratio on the power meter. The use of a fast detector (for example, a room-temperature MCT) enables direct measurement of the instantaneous optical power, but at the cost of a more **(p.200)**

difficult optical alignment and an additional calibration procedure, since the optical response is strongly wavelength dependent.

Current and voltage. In the set-up shown in Fig. 11.1 the current is fed from a pulse generator adapted to a 50 Ω source, with the current measurement performed using a current probe and the voltage measured directly (via another 50 Ω cable and a 50 Ω resistor) onto the high-impedance input of an

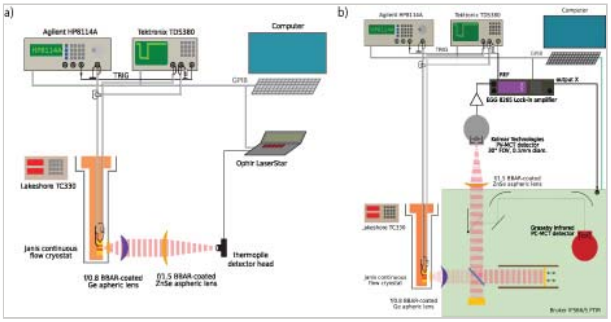


Fig. 11.1. Set-up for the electrical and optical characterization as a function of temperature. The samples are mounted on a continuous He-flow cryostat that can be operated between 5 and 325K. a) LIV measurements in pulsed operation. The average power of the emitted light is measured with a thermopile detector head. b) Schematic drawing of a

set-up used for electroluminescence measurements. The spectrum is measured with a Fourier-transform spectrometer and 77 K cooled detectors using lock-in signal filtering.

oscilloscope. This arrangement has the advantage of providing, after a 5–10 ns transient due to the imperfect impedance-matching, a reasonably flat and clean electrical pulse when using reasonably short cables (1.5 m). A better electrical pulse can be achieved by matching the dynamical resistance of the QCL to 50 Ω (by adding a 42–45 Ω resistor in series) directly in the cryostat. However, it also has the major disadvantage of increasing locally the electrical dissipation. Using commercially available 50 Ω pulse generators [239, 240], maximum currents of 1–2 A at maximum duty cycles of 95–40% and 5–7 A with $\leq 5 \times 10^{-4}$ duty cycles can be achieved.

The major disadvantage of a 50 Ω system is the large instantaneous electrical dissipation required for the operation (〉 50W at 1 A), preventing the use of such systems for high-duty-cycle measurements. For high-current, large-duty-cycle measurements, a better option is to switch the current close to the device using a power MOSFET. Such circuits can be home-built or can be found commercially [241], and an example is shown schematically in Fig. 11.2. In this circuit the current can again be measured **(p.201)**

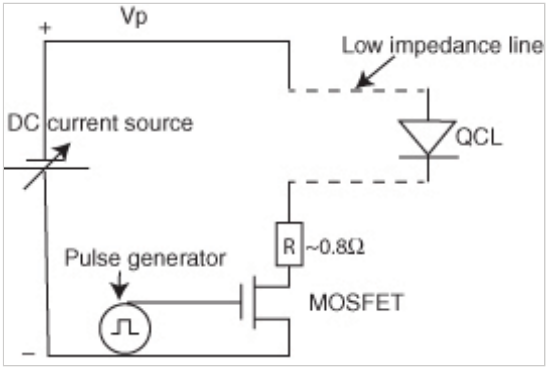


Fig. 11.2. Schematic drawing of the electrical circuit for a high duty cycle characterization.

using a current probe, taking into account that the latter acts as a transformer with a high-pass frequency or by measuring the average current on the DC current source and dividing by the duty cycle. Accurate measurements of the voltage bias are more difficult, and are best achieved by measuring the difference between the potential of the two electrodes of the laser using a high impedance, DC coupled voltage probe.

Temperature. In Fig. 11.1 the device is mounted in a cryostat, enabling the temperature to be varied between 5 and 325 K, and the temperature measured with an Si temperature sensor. This, however, restricts the implementation of a low-impedance line feeding the laser. When measurements on a more restricted temperature range are to be performed, a laser holder using a Peltier cooler is usually preferred, as it is for “real-world” applications. In such holders the temperature is best measured using a thermistor element because it provides a very large temperature coefficient enabling a good temperature stabilization.

11.1.1 Continuous wave measurements

Continuous wave measurements are conducted in a very similar way as the pulsed ones. By driving the device in continuous wave and using a four-probe arrangement, much better

accuracies can be achieved for the current and voltages ($\delta I \ \langle \ 100\mu A, \delta V \ \langle \ 1mV$) compared to the pulsed measurements. The problem, however, is that in most cases the temperature is not well defined, because self-heating effects are now very important.

In fact, comparing continuous-wave characteristics from the literature must be done with extreme care. In the older literature the temperature value was usually that of the copper block on which the device was mounted, as it is the value that is the most relevant for the end-user. However, recent reports use thermistors mounted very close to the device that, depending on the mounting technique, report a temperature close to the active-region temperature. This will, of course, yfield much better-looking curves! As these two different temperatures can differ by as much as 30–50 K, care must be taken when comparing data. Of course, the problem is most important in high-power and high-dissipation devices.

(p.202) 11.1.2 Wallplug efficiency

A simple look at the error propagation shows that one of the most important experimental parameter for a quantum cascade laser, its wallplug efficiency, is very difficult to measure accurately. NIST quotes an accuracy of 1.5–5% for calibrated optical detectors in the mid-infrared. In pulsed mode, current and voltage have easily an error of 3–5%, while the electrical pulse length will suffer from a similar error level. As a result, a very careful measurement in a continuous wave could reach an error of 5%, while pulsed measurements are limited to 10%. For this reason, measurements that are based on calorimetry, as shown later in this text, are especially interesting because they do not rely on the external measurement of the optical power.

11.2 Electroluminescence and spectral measurements

One important characterization tool is the measurement of the electroluminescence. In fact, in devices with a good population inversion the shape of the electroluminescence spectrum is very close to the true gain profile, and therefore enables checking of the gain bandwidth of an active region. For this measurement to work, it is, however, important to use a geometry where a multi-pass through the active region is avoided, to prevent optical gain to distort the true gain profile. Light extraction through a 45-degree wedge is a common approach. When buried heterostructure devices are processed, good results are also achieved by extracting the light through a side of the device.

At any rate the intensity of the luminescence is very low because of the low collection efficiency (typically below 1%) and the low radiative efficiency of the transitions (about 10^{-6}). Assuming an average dissipation of 1 W on the sample (to prevent heating), the expected average power on the detector amounts to 1–10nW. This is to be compared to the emitted power by the simple blackbody radiation of the sample that amounts, at room temperature, to tens of μW. For this reason, a good signal-over-noise requires a small active area (200 × 200 μm), a 77 K cooled detector such as a mercury–cadmium–telluride with an NEP of a few picowatts, and an integration time of hundreds of milliseconds. For these reasons, a grating spectrometer that will spread this electroluminescence intensity spectrally does not allow acquisition of good electroluminescence spectra. Conversely, the Fourier

Transform InfraRed spectrometer or FTIR, operating in step-scan mode, is well adapted to this task.

A schematic diagram of the experimental measurement is shown in Fig. 11.1b. The electrical injected current is modulated at 10–100k Hz and the signal detected synchronously by a lock-in amplifier. The moving mirror of the FTIR is stepped with a delay corresponding to a few integration-time constants (100 ms to 10 s) of the lock-in amplifier. In this manner, the very strong blackbody background is eliminated; the necessity not to overload the input of the lock-in limits in practice the stepping rate of the FTIR. High signal-over-noise spectra can be acquired within 10–30 min with a resolution of 16 cm^{-1}.

In contrast, measurements of the spectra of lasing devices do not suffer from signal-over-noise limitations, because of the large power delivered by quantum cascade lasers. A FTIR operated in rapid scan, with a room temperature thermal detector (a DTGS for example) usually provide the best combination of spectral accuracy (guaranteed **(p.203)** by the internal reference of the He-Ne laser of the FTIR), resolution (down to about 0.1 cm^{-1} for high-resolution models), linearity, and measurement speed. While they typically require longer measurements, grating spectrometers provide a better noise floor off the peak (〉 45 dB, in contrast to 30 dB for an FTIR-based measurement).

11.3 Far-field

Beside the power and the spectrum, the far-field of the devices is another characteristic that affects the use of QCLs in applications. Because of their low gain and their long operation wavelength, quantum cascade lasers require very strongly confined optical waveguides to minimize the thermal dissipation during operation. As a result, the optical beam exiting the facet exhibits a very large divergence angle, at least along the growth direction (x). The latter can be estimated by assuming an intensity with Gaussian dependence

$$I(x,z) = \exp\left(-\frac{2x^2}{w_x^2(z)} - \frac{2y^2}{w_y^2(z)}\right)$$

(11.3.1)

Then the asymptotic divergence half-angle Θ_x at e^{-2} is

$$\tan(\theta_x) = \frac{\lambda}{\pi w_x}$$

(11.3.2)

where w_x is the Gaussian broadening at the facet. Taking a typical QCL operating at $\lambda = 9$ μm, fitting the computed mode intensity in the growth direction yfields $w_x = 2.2$ μm and therefore $\theta_x = 53°$. This is a large angle, corresponding to a numerical aperture of $NA = \sin \theta_x = 0.82$ in the growth direction. Because of these large angles, measurements using two-dimensional imagers should be performed with the detector very close to the laser, and the data correctly converted to angle using the tangent that naturally arises. For this reason, the best results are usually obtained by scanning a pyroelectric detector on a two-axis goniometer in which the laser to be measured is placed in the center.

Fig. 11.3 shows an example of a far-field measurement set-up based on a two-axis goniometer, along with the measurement of a buried heterostructure quantum cascade laser operating at λ = 8.4 μm. The FWHM of the far-field in the growth direction is about 65°, corresponding to $\theta_x = \frac{65}{\sqrt{2\ln 2}} = 53^c$, close to our estimate above. The divergence in the plane of the layer is smaller, as the width of this waveguide is 8 μm, much wider than its thickness in the growth direction. In general, it is found, especially for buried-heterostructure devices, that the measured far-fields correspond very well to the computed ones.

11.4 Active-region temperature

11.4.1 Average temperature

As discussed in Chapter 6, the large electrical dissipation in a quantum cascade laser heats the active region, changing its temperature. This effect is most pronounced for **(p.204)** long pulses and continuous wave operation. It is therefore important to measure the temperature of the active region experimentally.

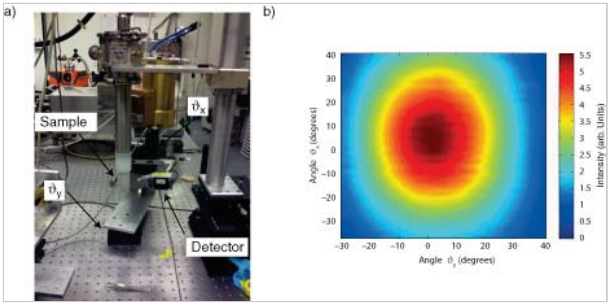


Fig. 11.3. a) Photograph of a two-axis goniometer far-field measurement set-up. The te?on window around the cryostat is adapted for measurements in the terahertz and is removed for room temperature mid-infrared devices. b) Measured far-field of a λ = 8.4 μm buried heterostructure quantum cascade laser.

The average active-region temperature can be measured by comparing the values in low-duty-cycle, pulsed mode and in continuous wave operation of a temperature-dependent parameter, such as the threshold current density or the emission wavelength of a single-mode laser. In this approach, one takes advantage of the fact that for short enough pulses, the active region lattice temperature remains the same as that of the holder.

The starting point is a measurement of the threshold current as a function of temperature using short pulses to obtain I_{th}^{p} (T_{sub}). The laser is then measured in continuous wave at the temperature of interest T_{sub}, yfielding a threshold current I_{th}^{cw} (T_{sub}. The thermal resistance is then extracted, using eqn. 6.6.37 and assuming that in pulsed operation, $T_{act} = T_{sub}$, yfielding

$$R_{th} = \frac{\Delta T_{sub}\big|_{I_{th}^{cw} = I_{th}^{p}}}{U_{th}\, I_{th}^{cw}}$$

(11.4.3)

where $T\left(I_{th}^{cw}\right)$ is the temperature at which the pulsed threshold current is equal to the one measured in a continuous wave. The latter temperature is obviously an weighted average of the temperatures in the active region. The stronger-than-linear dependence of the threshold current in temperature, combined with the shape of the optical mode that usually peaks in the center of the active region, means that $T\left(I_{th}^{cw}\right)$ is actually close to the value computed for the center of the active region.

If the threshold characteristic can be fitted using the usual exponential form $I_{th}^{p} = I_0^{p} \exp\left(T/T_0^{p}\right)$ then the thermal resistance becomes

(p.205)

$$R_{th} = \frac{T_0 \ln\left(I_{th}^{cw}/I_0^{p}\right) - T_{sub}}{U_{th}\, I_{th}^{cw}}$$

(11.4.4)

Obviously the speciffic thermal conductance can be obtained instead of the thermal resistance by substituting current by current densities and R_{th} by $1/G_{th}$ in the above equations. An illustration of the technique is shown in Fig. 11.4, where a very large thermal conductance of $G_{th} = 1840 W\,\mathrm{K}^{-1}\,\mathrm{cm}^{-2}$ is deduced for these narrow ridge buried heterostructure devices.

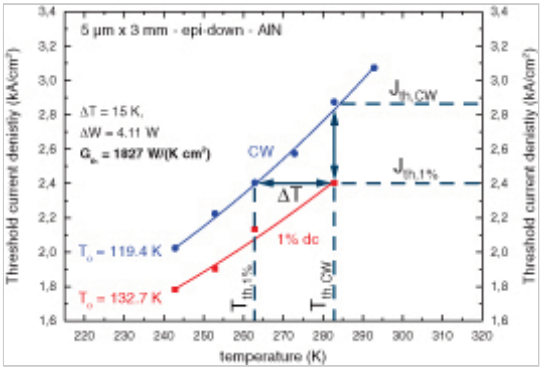


Fig. 11.4. Threshold current density as a function of temperature in pulsed mode (lower curve) and in continuous wave (upper curve). The temperature difference between the two curves yfields the thermal conductance.

As discussed in Section 10.2.4, in devices with single-frequency emission, a similar procedure can be performed using the dependence of the emission frequency as a function of temperature $\nu(T)$, first measured in pulsed mode. A completely analogous derivation can then be performed as above, substituting the function $T(I_{th})$ by $T(\nu)$, yfielding eqn. 10.2.40. In general, the thermal resistance measured using this second technique is significantly lower than that measured by the threshold current dependence. This difference completely stems from the temperature sensed by the two phenomena, as the emission frequency depends on the weighted average temperature by the mode intensity (that also includes the cladding regions) whereas, as discussed above, the threshold current density is mainly dependent on the temperature of the center of the active region. The thermal resistance measured using the threshold current data is therefore relevant to describe the optical power behavior, whereas the other one will describe the spectral characteristics, such as the mode maps, as a function of temperature and current.

11.4.2 Spatially resolved measurements

The position and shape of the interband photoluminescence depends not only on the heterostructure and applied electric field, but also strongly on the temperature. As a result, in a microphotoluminescence experiment the local lattice and electron temperature can be determined separately as a function of temperature and applied field. Using a similar set-up, a Raman microprobe is able to measure the non-equilibrium **(p.206)**

optical phonon distribution. These techniques were applied successfully in various quantum cascade structures, and yfielded spatially resolved data on the lattice [242], electron [243, 244], and phonon temperatures [245].

The variation of the photoluminescence peak energy with lattice temperature is to a very good approximation attributed to the change of the energy gap with temperature. As shown in Fig. 11.5a, the latter change is then well fitted by Varshni's semi-empirical law [246]:

$$E_g(T) = E_g(0) - \frac{\alpha T^2}{T + \beta} \tag{11.4.5}$$

The two parameters describing the temperature dependence (α and β) are then fitted by a measurement of the photoluminescence as a function of temperature at fixed dissipated power. In a further experiment, measurement of the photoluminescence peak position as a function of electrical dissipation then yfields the local temperature and allows measurement of the thermal resistance. Using this tool, a systematic study of the thermal properties of quantum cascade lasers was carried over [123]. The key result was a validation of the theoretical thermal models used, and confirmed previous results obtained by a sub-gap interferometric technique [247] that the cross-layer thermal conductivity $k_{\perp}$ was about an order of magnitude lower than the in-plane one $k_{\square}$. This **(p.207)** ratio, however, was found to depend on material and system, being extrapolated to about a factor of 3 for the $Ga_{0.47}In_{0.53}As/Al_{0.48}In_{0.52}As$ system at room temperature.

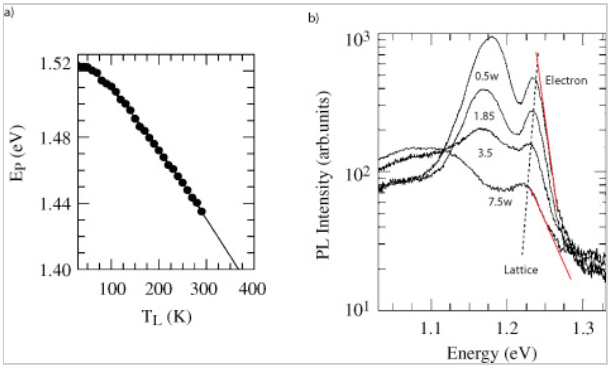


Fig. 11.5. a) Position of the photoluminescence peak of a GaAs/AlGaAs quantum cascade laser operating in the terahertz as a function of lattice temperature. The solid line is a fit using Varshin's law. b) Low-temperature photoluminescence of the facet of a strain-compensated $Ga_xIn1_{-x}As/Al_xIn_{1-x}As$ quantum cascade laser as a function of dissipated power. The peaks shown correspond to the photoluminescence from the injection/relaxation region. The dashed line shows schematically the redshift of the main features due to lattice heating, whereas the solid lines indicate the increase of the photoluminescence tail due to an increase in the electron temperature. Reprinted with permission from [244]. Copyright 2007, AIP.

Another interesting use of the temperature measurements is the possibility of improving the measurement of the wallplug efficiency of the devices in the regime of large duty-cycle. In fact, rewriting the equation relating the active-region temperature to that of the substrate 6.6.37, but taking explicitly into account the optical power generated P_{opt} [248] and inverting the equation, yfields:

$$P_{opt} = UI - \frac{T_L - T_{sub}}{R_{th}}$$

(11.4.6)

The thermal resistance can be easily fitted from the sub-threshold measurement of the active-region lattice temperature T_L versus electrical power UI as the luminescence efficiency is effectively negligible. The knowledge of the thermal resistance then enables the computation of the optical power directly. The optical power P_{opt} derived in this manner corresponds to the total optical power extracted from the device by all facets. The wallplug efficiency of QCL devices was measured in the mid-infrared [244] and terahertz spectral regions [249] using this technique.

In fact, a very attractive aspect of this technique is that it does not require an accurate temperature measurement of the active region, as the measurement of a quantity proportional to the temperature increase in the active region $\Delta T' = \alpha(T_L - T_{sub})$ is still sufficient to measure the correct value of the optical power. In the latter case, the fit of the thermal resistance below threshold yfields only the scaled value R_{th}/α.

The photoluminescence measurements are performed in the low excitation limit in a n-doped material. Because of the heavy mass of the holes and the n-doping, the shape of the photoluminescence emission peak $\mathscr{L}(E)$ is dominated by the electron distribution, and can be written as

$$\mathscr{L}(h\nu) = \mathscr{D}_j(h\nu)\exp(-h\nu/k_B T_e)$$

(11.4.7)

where $\mathscr{D}_j(E)$ is the joint density of state and T_e is the electron temperature. The joint density of state is constant for two-dimensional systems and, as a result, the slope of the high-energy tail of the luminescence should yfield directly the electron temperature.

The effect of electron heating is shown in Fig. 11.5 where the high energy tail of the photoluminescence emission clearly increases as a function of electrical dissipation. This property has been used to estimate the electron temperature during operation for GaAs-based [243] and InP-based QCLs [244] in the mid-infrared, as well as in the terahertz [250]. Extracted from these measurements were the electron-lattice coupling constant $\alpha = (T_e - T_L)/J$ for various materials and active regions dopings. An example of such a measurement performed on a strain-compensated $Ga_xIn_{1-x}As/Al_xIn_{1-x}As$ QCL is shown in Fig. 11.6. A thermal resistance of $R_{th} = 11.5K/W$ and an electron– lattice coupling value $\alpha =$ 34.8K cm^2/kA were deduced from these measurements. A wallplug efficiency of about 5% was extracted from the lattice temperature measurement as a function of power.

Measurements of the electron temperature as a function of electrical power have also shown the differential cooling effect of the optical emission, as the photon emission **(p.208)**

efficiently removes heat from the active region. An approach similar to that shown in eqn. 11.4.6 can be followed, but substituting the lattice by the electron temperature. An example of such a differential cooling is shown in Fig. 11.6c for a terahertz QCL operating at 1.8 THz. One limitation of these measurements is the fact that they are most easily interpreted by assuming a single electron temperature describing all the subbands simultaneously. This situation is most likely oversimplified, and more realistic models will yfield a small part of the electrons with a non-thermal distribution.

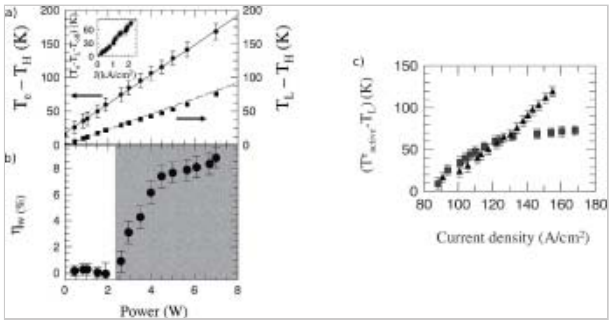


Fig. 11.6. a) Electron (left axis) and lattice (right axis) temperatures measured as a function of electrical dissipation for a strain-compensated $Ga_xIn_{1-x}As/Al_xIn_{1-x}As$ device at 60 K. A thermal resistance of R_{th} = 11.5 K/W and an electron–lattice coupling value α = 34.8 K cm^2/kA were deduced from these measurements [244]. Inset: temperature difference between the electron and lattice as a function of injected current. b) The wallplug efficiency deduced from the lattice temperature measurement using eqn. 11.4.6. c) Electron temperature as a function of injected power in a lasing (solid squares) and non-lasing (solid triangles) terahertz QCL [251]. The high internal efficiency of the device prevents a further increase of the electron temperature above threshold. Reprinted with permission from [244] and [251]. Copyright 2007, 2009, AIP.

11.5 Gain and loss measurements

What are waveguide losses after all?. The gain and loss measurements are very important tools to characterize and improve active regions. In particular, the value of the waveguide losses, because it limits both the threshold current density *and* the slope efficiency, is a very important parameter. However, because these two quantities relate to potentially a different definition of the waveguide losses, a certain confusion exists in the literature.

The waveguide loss *αtot* appears in the rate equations (eqn. 7.4.4) and in the expression for the threshold current J_{th} (eqn. 7.4.7). The sum of all optical loss mechanisms, to the exception of resonant losses of the laser transition, appears in the slope efficiency *dP/dI* (eqn. 7.4.9). As shown schematically in Fig. 11.7, the total waveguide loss can be written as:

$$\alpha_{tot} = (1 - \Gamma)\alpha_{fc} + \alpha_{scatt} + \Gamma\alpha_{nr}$$

(11.5.8)

where Γ is the overlap factor of the optical mode with the active region.

(p.209)

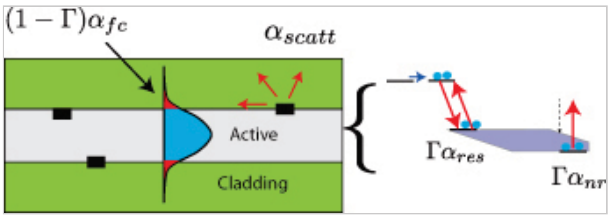


Fig. 11.7. Schematic description of the various contributions to the waveguide losses.

The cladding losses mostly originate from free carrier absorption α_{fc}, usually modeled by the Drude formula 6.3.35. The resonant absorption $\alpha_{res} = g_c n_2^{therm}$ arises from electrons thermally excited to the lower state of the laser transition. In principle, this loss affects only the threshold current density and not the slope efficiency, and is already taken into account in the rate equation model by the term en_2^{therm}/τ_{eff} in eqn. 7.4.7. The non-resonant losses of the active region arise from all the other transitions in the active region. In the early work they were estimated using a Drude formula, because the injection region was treated as a having a three-dimensional character.
It was then realized that such a three-dimensional approximation was too crude, as a quantization effect should appear, the absorption in the active being closer to that of a superlattice. As a result, a much better approach is to simply compute the total loss as the sum of intersubband transition comprising all the possible combination of subbands of the active region and injector. Note that holistic transport models that compute all the electron distribution and gain in the active region include the loss naturally in the computation of the active region gain.

The scattering losses α_{scatt} originate from the waveguide sidewall roughness. Such roughness has typically a scale much smaller than the wavelength. D. Marcuse presented a perturbative treatment of such mode conversion loss [252] for a slab dielectric waveguide. In many ways, his treatment is conceptually similar to the effect of the interface roughness to the electron lifetime in a quantum well. He derived a full expression where the loss has the following behavior:

$$\alpha_{scatt} \sim \frac{a^2 (\Delta n^2)^2}{\lambda^4} \tag{11.5.9}$$

where a is the amplitude of the roughness scattering, λ the wavelength, and Δn^2 the step of dielectric constant responsible for the waveguide confinement. Compared to interband devices operating in the near-infrared, quantum cascade lasers using etched ridges have both a much longer wavelength but at the same time a much larger value of Δn^2, making them roughly as sensitive to sidewall roughness, especially in the shorter wavelength range 3–5 μm. The effect of deliberate sidewall roughness was tested experimentally by Toor *et al.*[253], verifying the above overall conclusions.

Experimentally, it is observed that scattering loss affects more strongly the slope efficiency rather than the threshold current. Such a behavior was for example observed when considering dry-etched waveguides in devices operating at short wavelength. In fact, for a

narrow waveguide with dry-etched vertical sidewalls covered by a reffective **(p.210)** metallic coating, the mode coupling to continuum modes is extremely limited, as most scattering will then couple the mode to a counter-propagating mode. As a result, the sidewall roughness acts basically as a distributed reffector, and should rather be seen as a reduction of the mirror loss rather than an increase of the waveguide loss.

11.5.1 Length dependence

The threshold current density and slope efficiency can be measured as a function of the device length by fabricating laser devices with a varying cavity length. Assuming a simple rate equation approach where the net gain is a linear function of current and rewriting eqn. 7.4.7, writing the optical losses explicitly as

$$J_{th} = \frac{e}{g_c \tau_{eff}} \left(-\frac{1}{2L} \ln(R_1 R_2) + \alpha_w \right) + \frac{e n_2^{therm}}{\tau_{eff}}.$$

(11.5.10)

The above equation is usually rewritten in terms of a transparency current $J_{tr} = e n_2^{therm} / \tau_{eff}$ and a differential gain coefficient $g'\Gamma$:

$$J_{th} - J_{tr} = \frac{1}{L} \frac{-\ln(R_1 R_2)}{2 g' \Gamma} + \frac{\alpha_w}{g' \Gamma}$$

(11.5.11)

As a result, a plot of the threshold current density as a function of device length allows the determination of the gain coefficient and waveguide losses if the transparency current can be neglected. A more direct measurement of the waveguide loss can be achieved by measuring the inverse quantum efficiency $\eta_Q^{-1} = N_w (h\nu/e)/(dP/dI)$ as a function of cavity length L:

$$\eta_Q^{-1} = \eta_{tr} \left(1 + \frac{\ln R_2}{\ln R_1} \right) + L \frac{2 \eta_{tr} \alpha_w}{\ln R_1}$$

(11.5.12)

where R_1 and R_2 are the front and back facet reffection coefficients respectively, and $\eta_{tr} = \tau_{eff}/(\tau_{eff} + \tau_2)$ the (internal) efficiency of the transition.

A weakness of these techniques is the necessity of measuring precisely the ridge width and the slope efficiency for devices that can have variations in their geometry or optical far-field originating from their fabrication and mounting. One possibility for alleviating such variations is to measure the threshold current density and slope efficiency of the *same* device before and after a back facet coating, i.e. with R_2 taking two well-defined values. In this way, much of the uncertainty in the device parameters are ratioed away, since the reffection coefficient of a metallic back-facet coating is relatively well known.

11.5.2 Hakki–Paoli

Threshold current density and slope efficiency measurements yfield the waveguide loss (and therefore the gain) at the laser frequency only. A measurement of the gain as a function of photon energy allows a much more detailed understanding of the processes in

the active region. One of the most natural measurement of the sub-threshold gain is provided by the Hakki and Paoli technique, first demonstrated in semiconductor lasers [254]. In this technique the contrast of the Fabry–Perot optical fringes formed **(p.211)**

by the cavity itself is measured as a function of the injected current. As discussed in Chapter 10, the Airy function characterizing the contrast of the Fabry–Perot fringes (eqn. 10.1.8) depends on the quantity $b(\nu) = R\exp(-(\alpha_w - g(\nu))L)$. One notes that, writing the transmission $I(\varphi)$ as a function of the phase $\varphi = 4\pi n_{eff}L/\lambda$, we have the result

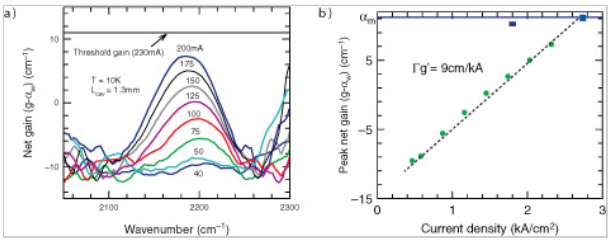


Fig. 11.8. a) Gain as a function of photon energy for various injected currents, for a 1.3 mm long device based on a vertical transition. b) Solid disks: peak gain as a function of current. *αm* are the calculated mirror losses. The actual threshold current density is indicated by a filled square. Reprinted with permission from [255]. Copyright 1995, AIP.

$$b = \frac{\tilde{I}^{(1)}}{\tilde{I}^{(0)}} = \frac{\int_{-\pi}^{\pi} I(\phi)\cos(\phi)d\phi}{\int_{-\pi}^{\pi} I(\phi)d\phi}$$

(11.5.13)

i.e. b is the ratio of the first Fourier component of the fringe spectrum to the average (DC) value [255,256]. This "Fourier variant" of the Hakki–Paoli technique is especially convenient since the Fourier transform of the spectrum is the interferogram itself! As a result, a spectrum of the gain can be obtained by performing the ratio of the Fourier transform of the interferogram around the time delay corresponding to the roundtrip time and the one around the zero path difference. This technique also allows the measurement of the group index dispersion, with, however, insufficient sensitivity for typical QCL devices[257]. The gain is then simply retrieved as

$$g(\nu) - \alpha_w = \frac{1}{L}\ln\left(\frac{\tilde{I}^{(1)}}{\tilde{I}^{(0)}}\right)$$

(11.5.14)

An example of such a measurement, performed at low temperature on a device based on a vertical transition, is shown in Fig. 11.8a. The peak gain, extracted from these measurements, is then plotted as a function of injected current in Fig. 11.8b. The waveguide losses $\alpha w = 12$ cm^{-1} can be either extrapolated from the baseline in Fig. 11.8a or from $J = 0$ in Fig. 11.8b.

The key limitation of the Hakki–Paoli technique is the assumption of the strict single monomodicity of the waveguide. In fact, a second transverse mode, never reaching

threshold, will significantly perturb the measurement, as the emission from this mode, modulated by a different effective refractive index, will interfere with the fringes **(p.212)** generated by the fundamental mode. To minimize this effect, Hakki–Paoli measurements should be performed with relatively low numerical aperture optics, in order to collect preferentially the light coming from the fundamental mode. An underestimate of the gain due to transverse modes appears clearly in a plot of the peak gain as a function of injected current (such as in Fig. 11.8b), because the extrapolated threshold current, given by the intersection between the loss line (gain = mirror losses) and the extrapolated gain (dashed line in Fig. 11.8b) would be higher than the true threshold (solid square in Fig. 11.8b). Another limitation is the fact that pulsed measurements are very difficult because thermal drifts during the pulses "wash out" the Fabry–Perot fringes. As a result, only extremely short pulses can be used, considerably reducing the optical signal that has to be acquired with a high spectral resolution.

11.5.3 Single pass gain measurements

Another limitation of the Hakki–Paoli technique is the fact that the gain can be measured only in a relatively narrow range around the laser frequency. This limitation can be overcome by coupling an external light source in the laser waveguide and then measuring the single pass gain. Two possible techniques use either a single-frequency device (possibly tunable) or a broadband source.

When a single mode source is used, such as a gas laser (3.39 μm He-Ne) [258] or a DFB QCL [259], the gain is still measured using the contrast of the fringes of the Fabry–Perot resonator. The fringes can either be scanned by a thermal chirp of the QCL DFB or of the device under test. This technique enables accurate measurement of the losses in both TE and TM at a fixed frequency. The use of a broadly tunable QCL as the source should in principle enable the use of this technique over a wide frequency range.

By coupling a broadband light such as a blackbody source into the waveguide, a direct spectrum of the transmission can be achieved [260]. As shown in Fig. 11.9, this spectrum clearly shows the evolution of the gain and absorption in the device as the carriers are redistributed inside the active region by the applied field. In particular, the dips at high energy, marked by arrows, correspond to the absorption transitions between the ground states of the active region and excited states in the two-phonon active region. The latter disappear as the carriers are transferred out of the active region into the injector states.

In contrast to the previous techniques, a direct quantitative assessment of the gain from the transmission is difficult, but the transmission can be written as

$$T = c^2 \exp(-(\alpha_w - g)L)$$

(11.5.15)

where c is the coupling efficiency in the waveguide. The latter exhibits usually a slow wavelength dependence, and must be determined using a separate measurement. To improve the signal-over-noise and enable measurements in narrow waveguides, the broadband emission from synchrotron has also been used instead of the blackbody source.

Finally, using as a source the pulses generated by a non-linear conversion of ultrashort pulses from a femtosecond laser, the advantages of both techniques can be combined, as the source is both broadband and coherent. Using this approach, Parz *et* **(p.213)**

al. [261] have measured the gain *and* the dispersion of a quantum cascade laser. Such measurements, performed on a two-phonon QCL operating at $\lambda \approx 12\ \mu$m, are shown in Fig. 11.10. To unravel the refractive index from the measurement of the transmitted electric pulse, an iterative fixed-point algorithm was used to find the real and imaginary part of the refractive index n and k that will satisfy the equation for the transmission of the optical pulse, given by

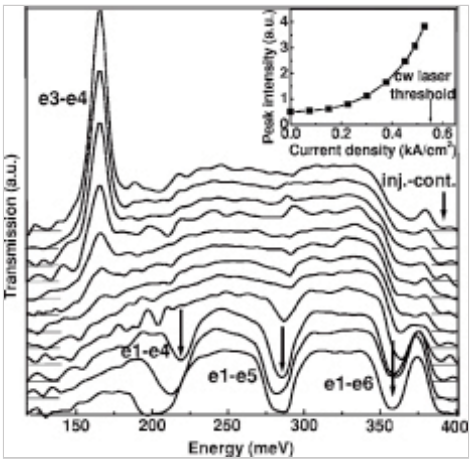


Fig. 11.9. Broadband measurement of the waveguide transmission of a QCL operating at 7.4 μm as a function of injected current, measured at low temperature. Reprinted with permission from [260]. Copyright 2006, AIP.

$$T = 4\frac{n - ik}{n + 1 - ik}\exp(-ik_0(n - ik)L),$$

(11.5.16)

where $k_0 = 2\pi/\lambda$. As shown in Fig. 11.10, the gain introduces a strong dispersion around the laser frequency.

Single-pass measurements of the gain using ultrafast lasers have been widely used to probe the gain in terahertz QCL [262], as at the same time the generation of terahertz using photoconductive antenna is in general easier than in the mid-infrared, and as the other techniques based on electroluminescence are plagued by sensitivity issues and the difficulty to discriminate between the intersubband electroluminescence and the thermal blackbody emission.

11.5.4 Multi-section

A variant of the single-pass gain measurement is a technique where the laser is separated into multiple sections, and the electroluminescence from a section in the back is amplified by passing it through a section in front. This technique, first introduced for interband devices, was used for gain [263] and waveguide loss [264] measurements. When using a two-section device, this technique naturally yfields the change in gain Δg; if the waveguide loss at zero current is known and has little wavelength dependence, the resultant data will yfield the absolute gain. When applied at room temperature this **(p.214)**

technique no longer works, as the device suffers from strong resonant reabsorption losses from the active region when not biased. For this reason, as shown schematically in the inset of Fig. 11.11a, a three-section device is preferable. The light-bulb (A) and amplifier (B) sections are of equal length (typically L = 375–700 μm) while the third section, in the back of section A, is used as an "absorber" in order to remove any feedback from the back facet. During the measurements the device is never turned off completely, and is always operated with a baseline current J_{bl} that flattens the band structure enough to remove most of the reabsorption, and therefore yfields a flat baseline. Assuming the two sections A and B are symmetric, the absorption α_{bl} at this injected current can be obtained from the measurement of the luminescence upon injection of a small probe current J_p(assumed not to create any significant gain) as:

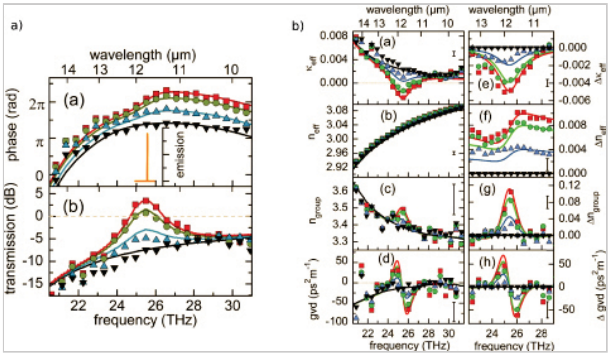


Fig. 11.10. a) Broadband measurement of the real and imaginary part of the transmission of a QCL operating at 12 μm as a function of injected current, measured at low temperature. b) Real and imaginary part of the effective index and of the group index as a function of frequency. Displayed are full (100%) bias (squares), 80% (circles), 40% (triangles up) and 0% (triangles down) [261].

$$\alpha_{bl}(\lambda) = \frac{1}{L}\ln\left(\frac{S_{A=J_{bl}}^{B=J_p}(\lambda)}{S_{A=J_p}^{B=J_{bl}}(\lambda)}\right)$$

(11.5.17)

where we have used the notation $S_{A=J_{bl}}^{B=J_p}$ as the luminescence signal obtained with section A biased with a current J_{bl}, and section B biased with current J_p.

Once this baseline $\alpha_{bl}(\lambda)$ is obtained, the device is measured at the final test current J_p, and the gain is

(p.215)

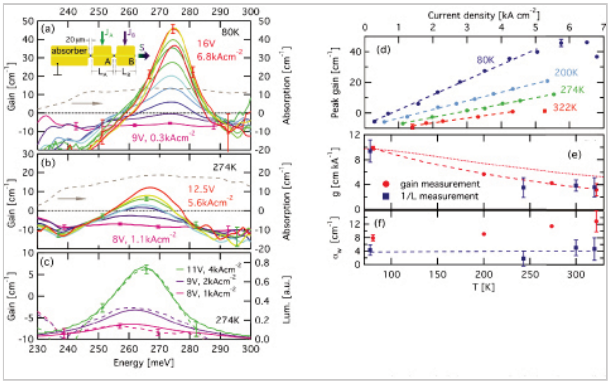


Fig. 11.11. (a) Modal gain as a function of injected current between 0.3 and 6.8 kA cm^{-2} at $T = 80$ K. Inset: sample geometry. (b) The same measurements, taken between 1.1 and 5.6 kA cm^{-2} at 274 K. The baseline absorption measurements are shown as dashed lines. (c) Comparison between gain (dashed) and electroluminescence (solid lines) spectra at the same bias at 274 K. (d) Peak values of the gain as a function of injected current and for temperatures between 80 and 322 K. Dashed lines represent linear fits that yfield the gain coefficient and waveguide losses that are shown in graphs (d) and (e) as a function of temperature (dots). Squares represent the corresponding values obtained by 1/L-measurements. The dashed line in (e) shows an exponential fit, and the dotted line the expected dependence on temperature assuming that optical phonon scattering is the only non-radiative channel. The device is described in Appendix A.3.4. Reprinted with permission from [267]. Copyright 2009, AIP.

$$g(J_p,\lambda) = \frac{1}{L}\ln\left(\frac{S^{B=J_h}_{A=J_p}(\lambda) - S^{B=J_h}_{A=J_{h1}}(\lambda)}{S^{B=J_{h1}(\lambda)}_{A=J_p}}\right) - \alpha_{ll}(\lambda)$$

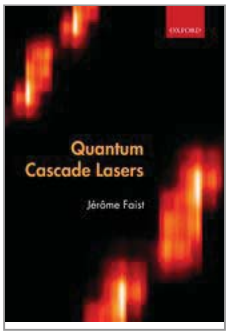
(11.5.18)

This multi-section technique is more cumbersome to implement because it requires the fabrication of devices with sections that have good electrical insulation. However, it has the advantage of providing calibrated measurements, work in pulsed mode and can measure the unclamped gain up to saturation. For these reasons it has been used to study the dispersive nature of the gain in QCL with poor population inversion [265, 266].

Using this technique, the measurement of the gain in a strain-compensated laser operating at $\lambda = 4.8$ μm, as a function of temperature, is shown in Fig. 11.11. These measurements

point out some important physical results. As expected, the peak gain is reduced from about 45cm^{-1} down to 12cm^{-1} as the temperature is increased from 80 to 274 K. In Fig. 11.11c, the measured gain at 274 K is compared to the electroluminescence lineshape. The two curves overlap within experimental errors, confirming that **(p.216)** for devices with a good population inversion, the unamplified electroluminescence is a good indicator of the shape of the gain profile. When plotted as a function of injected current (Fig. 11.11d), the peak gain exhibits a linear dependence with current. This dependence persists up to the onset of the negative differential resistance, confirming that the gain is not compressed at high injection by electron–electron scattering or hot carrier leakage. In fact, when the gain coefficient g' is plotted against temperature (Fig 11.11e, its dependence is fully explained by the expected temperature dependence (dashed line) of the optical phonon emission (eqn. 5.2.13) and the broadening of transition. As shown in the same figure, the gain coefficient found by the multisection technique agreed well with the one measured using the length dependence of the threshold current density "1/L technique".

As shown in Fig. 11.11f the measurement of the waveguide losses were found to be independent of temperature, but the value obtained using the multi-section technique (10 cm^{-1}) was found to be significantly higher than those obtained by the threshold current density measurements (4 cm^{-1}), the latter result being more consistent with the slope efficiencies measured.



Quantum Cascade Lasers
Jérôme Faist

Print publication date: 2013
Print ISBN-13: 9780198528241
Published to Oxford Scholarship Online: May 2013
DOI: 10.1093/acprof:oso/9780198528241.001.0001

Transport models

Jérôme Faist

DOI:10.1093/acprof:oso/9780198528241.003.0012

Abstract and Keywords

Unlike other optoelectronic devices, quantum cascade lasers (QCLs) cannot be described by simple drift-diffusion equation. Whereas in interband semiconductor lasers the assumption of separate quasi-Fermi levels in the conduction and valence bands allows the transport problem to be reduced to a quasi-equilibrium situation, such approximations completely break down in QCLs. Because of the subpicosecond long intersubband lifetime, the transport scattering time is never much shorter than the intersubband transition time, preventing the use of such approximation. For this reason, the development of predictive transport models requires a more sophisticated approach. This chapter discusses rate equation models, density matrix, full density matrix models, Monte-Carlo techniques, and non-equilibrium Green's function.

Keywords: quantum cascade lasers, rate equation models, density matrix, full density matrix models, Monte-Carlo techniques, non-equilibrium green's function

In contrast to other optoelectronic devices, quantum cascade lasers cannot be described by simple drift-diffusion equation. Whereas in interband semiconductor lasers the assumption of separate quasi-Fermi levels in the conduction and valence bands conveniently allow reduction of the transport problem to a quasi-equilibrium situation, such approximations completely break down in QCL. Because of the subpicosecond long intersubband lifetime, the transport scattering time is never much shorter than the intersubband transition time, preventing the use of such approximation. For this reason, the development of predictive transport models requires a more sophisticated approach.

Formally, one can see the transport problem as one of solving the equation of motion of an electron population subjected to various interactions represented by a Hamiltonian

$$H_{tot} = \frac{p^2}{2m} + V_{crystal} + V_{heterostructure} + V_{field} + H_{e-phonon} + H_{e-photon} + H_{scatt}$$

(12.0.1)

where $V_{crystal}$ is the crystal potential, $V_{heterostructure}$ the one created by the heterostructure, V_{field} the applied field, $H_{e-phonon}$ the electron–phonon interaction, $H_{e-photon}$ the interaction with the cavity field and H_{scatt} the other scattering mechanisms (interface roughness, electron–electron). Fortunately, only quantities such as a current density or a photon population, which relate to the *average* electron populations are measured experimentally.

As shown schematically in Fig. 12.1, one diffculty is that the electrons are strongly interacting with the bath of photons in the cavity on one hand, and the bath of phonons in the crystal. In addition, under laser operation, the photon population will be by nature out of equilibrium and therefore cannot be treated as a thermal bath. Although the phonons are usually treated as in thermal equilibrium, a number of experiments and theoretical prediction have hinted that this is probably too strong an approximation [245]. To keep the system tractable, the dynamics of individual subsystems (electrons, photons and phonons) are solved either quantum-mechanically or classically, and their cross-interactions treated in a mean-field approach.

As mentioned above, the strong quantum confinement of the electron in the direction of the current flow, responsible for the observation of intersubband transitions, precludes the use of a purely classical Boltzmann equation approach for the electron transport dynamics, and therefore requires a treatment of quantum effects.

A general discussion of transport theories is outside the scope of this work, and the reader is referred, for example, to the work of Wacker [268] and Rossi [269–271], where **(p.218)**

the merits of various approaches are compared. These fall roughly into four classes of model:

Rate equation In this simple approach the Schrödinger equation is solved for the complete heterostructure potential over one period. A rate-equation dynamics is then solved between these basis states.

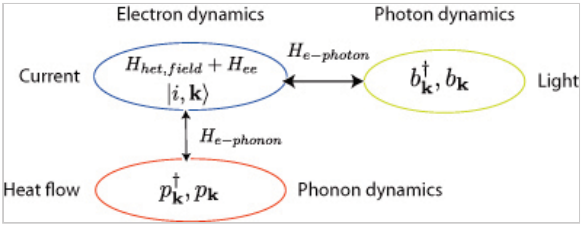


Fig. 12.1. Schematic description of the relation between the microscopic quantities, electron, photon, and phonon states, and their interactions, as well as the global observables (current, optical field, and temperature).

Density matrix In contrast to the rate equation, the density matrix approach enables simultaneous coherent and incoherent couplings between states, and is therefore able to handle resonant tunneling between states in a more accurate way.

Monte Carlo Ensemble and quantum Monte Carlo techniques compute the time evolution of the carrier distribution function. They have been used to follow the in-plane distribution

function and can include coherent terms.

Non-equilibrium Green's function The non-equilibrium Green's function is a very powerful diagrammatic technique that allows the treatment of interactions in a very clean way.

These approaches will be discussed individually, and usually represent various tradeoffs between their accuracy, numerical overhead, and completeness.

12.1 Rate-equation models

As shown in Fig. 12.2, the approach implicit to the rate equations presented in Chapt. 7 was to diagonalize the first four terms of the Hamiltonian in the stationary state using the $\mathbf{k} \cdot \mathbf{p}$ approach and the envelope function approximation. The in-plane dispersion is then treated as an additive term to the energy. The last three interaction terms are then treated in a perturbative approach. In this manner the whole problem can be reduced to a set of rate equations for the whole structure. For each state *i*, a rate equation can be written as

$$\frac{dn_i}{dt} = \sum_{j \rangle i} \frac{n_j}{\tau_{ji}} - \sum_{j \langle i} \frac{n_i}{\tau_{ij}} - \delta_{i,u} g_{ud} S(n_u - n_d) + \delta_{i,d} g_{ud} S(n_u - n_d)$$

(12.1.2)

where the optical transition is assumed to occur between the upper state *u* and the lower state *d*. The time τ_{ij} corresponds to the outgoing intersubband scattering from **(p.219)**

subband i, ($i \rangle j$), and τ_{ji} the incoming scattering ($i \langle j$). To be able to separate the scattering neatly between uniquely incoming and uniquely outgoing channels imply that the thermal energy *kT* must be assumed to be small enough compared to the intersubband energy. This model has been used with relatively good success to describe the population in the active region at low temperatures. However, this assumption breaks down as soon as the in-plane kinetic energy of the electrons becomes comparable to the intersubband spacing. On can circumvent this problem by expanding the equation to include all the outgoing $\tau_{ij}^{\rangle}$ and incoming $\tau_{ij}^{\langle}$ scattering between the two subbands *i* and *j* (of course, now $\tau_{ij}^{\langle} = \tau_{ij}^{\rangle}$ in this notation). These terms are effective scattering rates between two subbands, weighted by the occupancy $f_i(k)$ and the final-state availability $(1 - f_j(k))$. The subband population is the sum of the *k*-dependent occupancy $n_i = \sum_k f_i(k)$. The rate equation can then be written as

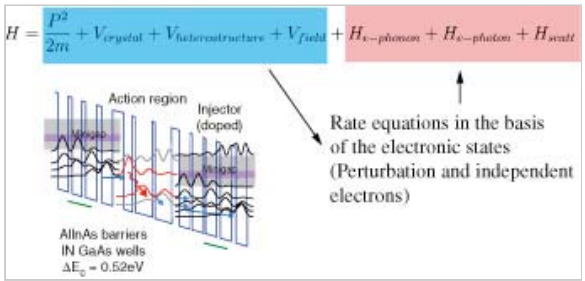


Fig. 12.2. Schematic description of the rate-equation approach.

$$\frac{dn_i}{dt} = \sum_{j \neq i} \frac{n_j}{\tau_{ij}^{\langle}} - n_i \sum_{j \neq i} \frac{1}{\tau_{ij}^{\rangle}} - \delta_{i,u} g_{ud} S(n_u - n_d) + \delta_{i,d} g_{ud} S(n_u - n_d)$$

(12.1.3)

where the weighted terms in terms of the population distribution are, for the outgoing scattering

$$\frac{n_i}{\tau_{ij}^{\rangle}} = \sum_{k_i} \sum_{k_j} f(k_i)(1 - f(k_j))\tau_{ij}^{-1}(k_i, k_j)$$

(12.1.4)

and for the incoming scattering

$$\frac{n_i}{\tau_{ij}^{\langle}} = \sum_{k_j} \sum_{k_j} f(k_j)(1 - f(k_i))\tau_{ji}^{-1}(k_i, k_j).$$

(12.1.5)

Such an approach has been used to predict transport and gain in various quantum cascade laser devices in the mid-infrared and terahertz frequency domain. Fig. 12.3 shows the voltage–current and gain–current curves, computed for various temperatures, of two GaAs-based three quantum well active region lasers differing by the aluminum content of the tunnel barriers. Specifically, it is shown that the devices with the $Al_{0.33}Ga_{0.67}$ As barrier exhibit a significant amount of carrier leakage at elevated temperatures, as compared to devices fabricated with $Al_{0.45}Ga_{0.55}As$ barriers. This **(p.220)**

effect of carrier leakage was indeed observed experimentally and is discussed in Chapter 7. The overall shape of the voltage–current, as well as the magnitude of the gain and its temperature dependence, are in good general agreement with the experimental data [272]. However, the features seen in the voltage–current curves, for example at 45 kV/cm at 300 K, do not correspond to the experiments. They probably originate in resonances between states that produce artefacts inherent to the technique used. A solution to this problem is to introduce a density-matrix framework in which dephasing, related to the in-plane scattering, can be explicitly taken into account. This will be discussed later in this chapter.

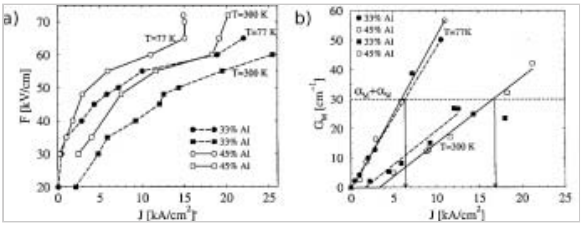


Fig. 12.3. Computation of the transport using a rate-equation approach for a GaAs-based quantum cascade laser. The goal was to compare the current leakage for the Al_xGa_{1-x} As barriers with $x = 0.33$ and $x = 0.45$. a) Voltage versus current characteristics. b) Gain versus injected current. Reprinted with permission from [272]. Copyright 2002, AIP.

12.1.1 Lasers based on photon-assisted tunneling transition

As shown schematically in Fig. 7.16, a class of quantum cascade lasers are based on a very diagonal transition, whereas the ground state of the injector of one period is also the upper-state for the next period. In these devices, because the transition must proceed through a tunnel barrier, the lifetime of the upper-state can be relatively long, up to 20 ps at cryogenic temperatures. Because upper and lower states are, to a very large extend, localized in different quantum wells, the radiative transition can be described aptly as proceeding by photon-assisted tunneling through the coupling barrier, in the manner first described by Kazarinov and Suris [8]. In such a laser, modeling the transport using the approach described above can yield very good results. In fact, because we are in a situation in which the scattering within the injector is much faster than the active region transition, we can safely describe the occupation of the injector states by a thermal distribution. The whole transport is then driven by the intersubband

scattering between periods. Such a design was first demonstrated at an operating wavelength of $\lambda = 6\ \mu$m [125]. A schematic conduction-band diagram of a device, based also on a very diagonal transition, and operating at a longer wavelength $\lambda \approx 10\ \mu$m, is shown in Fig. 12.4a. In the same figure the voltage–current characteristics of the same structure is shown and compared to the theoretical prediction, computed using the eqn. 12.1.3. An excellent agreement is found, showing that the parameters used to compute the intersubband scattering rates are correct [273] **(p.221)**

.

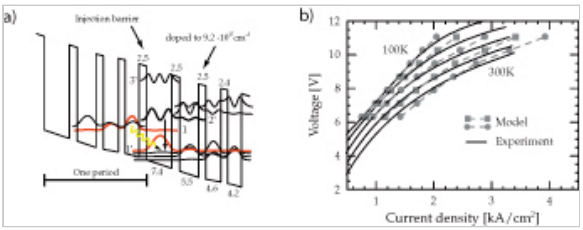


Fig. 12.4. Modeling the transport in a quantum cascade laser based on a photon-assisted tunneling transition. a) Self-consistent computation of the energy-band diagram of two periods of the InGaAs/AlInAs structure under an applied electric field of 55 kV/cm. The moduli squared of the relevant wavefunctions are shown. The thickness of the different layers is indicated in nanometers. b) Comparison between the measured (solid lines) and computed (dashed lines) voltage–current characteristics at 100, 150, 200, 250, and 300 K. The design is shown in Appendix A.1.4. Reprinted, with permission, from [273]. Copyright 2001 IEEE.

12.1.2 Photon-driven transport

The relative simplicity of the structures based on a photon-assisted tunneling transition enabled the study of two effects usually neglected in the first approaches: conservation of the total charge over one period, and the influence of the photon-driven transport. In these structures where the intersubband scattering within each period can be assumed to be much faster than period-to-period transport, a simple model can be written where each period consists of the state 1, simultaneously a ground state of the injector and upper state of the laser transition, and 1′, lower state of the laser transition, with densities n_1 and $n_{1'}$ satisfying the charge conservation condition $n_1 + n_{1'} = n_{tot}$. The rate equation in the presence of the optical field is

$$-\frac{n_1}{\tau_1} - g_c S(n_1 - n_{1'}) + \frac{n_{1'}}{\tau_{1'}} = 0$$

(12.1.6)

While the lower-state lifetime $\tau_{1'}$ is assumed to be constant, the upper-state lifetime τ_1, as well as the gain cross-section g_c, varies with the applied electric field and controls the transport. For typical structures, the product $g_c \cdot \tau_1$ is constant:

$$g_c \cdot \tau_1 = A$$

(12.1.7)

where A [cm · s] is a constant, characteristic of the heterostructure. In addition, the cross-section gain g_c is, with good accuracy, proportional to the applied voltage $V \propto g_c$. It is convenient to express the lifetime τ_1 (g_c) and $\tau_{1'}$ as a function of the cross-section gain g_0 needed by an (idealized) device with zero lower-state lifetime and total loss α_{tot}, i.e.

$$g_0 \cdot n_{tot} = \alpha_{tot}$$

(12.1.8)
and

$$\tilde{g} \cdot \tau_{1'} = A$$

(12.1.9)
so that the ratio of $\tau_{1'}/\tau_1 (g_0)$ is the inverse of the effective cross-section gains $g\tilde{}/g_c$

(p.222)

The rate equation below and above threshold can be solved analytically [273], and the result is shown in Fig. 12.5, where the normalized voltage–current and light– current are displayed for various values of the upper-state to lower-state lifetime ratio. A very strong dependence is found, with photon-driven transport for $\tau_1/\tau_{1'} = 20$ but the laser not reaching threshold for $\tau_1/\tau_{1'} \lesssim 6$. Even if the results shown in Fig. 12.5 are computed using a very simple model, meant to be applied for a very specific class of structure, they actually describe a generic feature of photon-driven transport: the abrupt change of the differential resistance at threshold. As shown in Fig. 12.5b, such change in conductance is very apparent for these devices based on photon-assisted tunneling. In this particular model, the ratio of the differential conductances just below G_b and just above G_a threshold is approximately half of the ratio of the lifetimes:

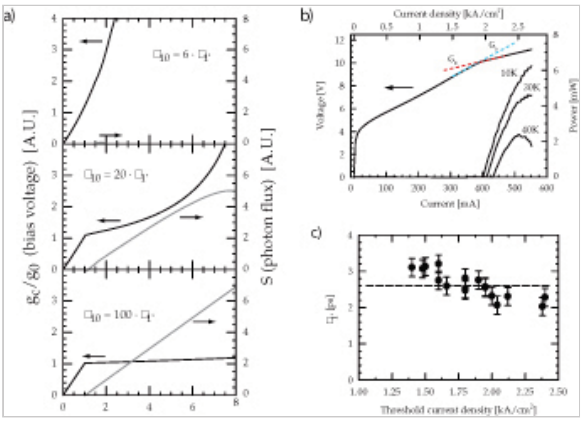


Fig. 12.5. a) Computed voltage–current and light–current characteristics for three different values of $\tau_{10}/\tau_{1'}$, i.e. of $g\tilde{}/g_0$ for the simple analytical model of transport that includes charge conservation and photon-driven transport (see text). b) Lightand voltage-versus-current characteristics, measured in continuous wave, for the device whose bandstructure is shown in Fig. 12.4. c) Value of the lower state lifetime extracted from the measurement of the ratio of conductance below and above threshold. Reprinted, with permission, from [273]. Copyright 2001 IEEE.

$$\frac{G_b}{G_a} \approx \frac{1}{2}\frac{\tau_{10}}{\tau_{1'}}$$

(12.1.10)
This model was applied to the structure shown in Fig. 12.4a. By studying the transport in devices of varying cavity lengths, the lifetime of the lower state was inferred as a function of the drive current using the computed value of the upper-state lifetime and the expression 12.1.10. As shown in Fig. 12.5c, this value, dominated by the tunneling out of the lower state of the quantum well into the injector, is roughly field-independent and is about 2.5 ps. The relatively long value of this tunneling time stimulated the search for active regions with more efficient extraction mechanisms, **(p.223)** such as the bound-to-continuum or two-phonon extraction schemes as discussed in Sections 7.6.3 and 7.6.4.

12.1.3 In-plane electron distribution

Implicit to the models presented above is the assumption that the in-plane electron distribution is thermal, characterized by a single temperature. Implicit to this hypothesis is the existence of thermalization processes happening at a much shorter timescale than the intersubband transitions. At high drive current and room temperature, this assumption is expected to break down first in the active region, where the electrons must lose a large amount of energy after being injected in the upper-state.

In-plane dynamics. A simple model of the electron dynamics in the active region can be derived by expanding the rate-equation model shown in eqn. 12.1.3 in *k*-space, by assuming that all the scattering mechanisms are dominated by optical phonon emission. A schematic description of this model is shown in Fig. 12.6a for a two-phonon active region: electrons are interacting with a bath of optical phonons in thermal equilibrium at a temperature T_L. Since we are neglecting electron–electron interactions, the electron may gain or lose energy only in quanta of the optical phonon energy. As a result, the energy of an electron in the subband i is given by $E_{i,j} = -j\hbar\omega_{LO}$, with $j = 0$, corresponding to the energy of the electron injection.

For an active region containing N states (with N being the upper laser state), the rate equation for the lower states (i, j) (with $i < N$) is

$$\frac{dn_{i,j}}{dt} = \sum_{k=1}^{N-1}\left(\frac{n_{k,j-1}}{\tau_{ik}^{e}} + \frac{n_{k,j+1}}{\tau_{ik}^{a}}\right) + \delta_{j,1}\frac{n_{N,0}}{\tau_{Ni}^{e}} + \delta_{j,-1}\frac{n_{N,0}}{\tau_{Ni}^{a}} - n_{i,j}\left[\sum_{k=1}^{N-1}\left(\frac{1}{\tau_{ik}^{a}} + \frac{1}{\tau_{ik}^{e}}\right) + \frac{1}{\tau_{esc}} + \delta_{j,1}\frac{1}{\tau_{Ni}^{a}} + \delta_{j,-1}\frac{1}{\tau_{Ni}^{e}}\right] \quad (12.1.11)$$

We assume that a cool electron distribution is injected in the upper laser level. Therefore, all the electrons are concentrated in $j = 0$ in the upper level N. The rate equation for that state is

$$\frac{dn_{N,0}}{dt} = \frac{J}{e} + \sum_{k=1}^{N-1}\frac{n_{k,1}}{\tau_{N,k}^{a}} - n_{N,0}\sum_{k=1}^{N-1}\left(\frac{1}{\tau_{N,k}^{e}} + \frac{1}{\tau_{N,k}^{a}}\right) \quad (12.1.12)$$

Solving eqns. 12.1.11 and 12.1.12 for a given injection current and temperature yields the densities at each relevant point of the *k*-space. The result of such a computation, comparing the densities for a single-phonon, three-quantum-well active region and two-phonon, four-quantum-well active region is shown in Fig. 12.6, where an extraction time from the active region of 2 ps was assumed. This model nicely demonstrates the “bottleneck” effect of the long tunneling escape time τ_{esc} in single-phonon resonance structures. At room temperature ($T = 300$ K), the ratio of upperto lower-state populations is $n_3/n_2 = 1.7$ — a value much lower than the one anticipated by considering the lifetimes ratio $\tau_{32}/\tau_2 = 5$. The reason is a a significant “reabsorption” probability: electrons scattered out into the state $n = 1$ at high kinetic energy have a high **(p.224)**

probability of scattering back to $n = 2$. In other words, the two groups of terms in eqn. 12.1.11 tend to compensate each other, slowing the scattering out of level $n = 2$.
In the two-phonon resonance structure, a larger phase-space is

available for scattering from the lower state, reducing its population. As a result, the population inversion is significantly larger, $n_4/n_3 = 2.85$. As shown in Fig. 12.6b the electron distribution in the subband, for small in-plane kinetic energies, is close to a thermal electron distribution, characterized by an electron temperature very close to the lattice one. A larger electron temperature ($T_e = 325$ K) and higher non-thermal tail characterize the electron distribution of the single phonon design.

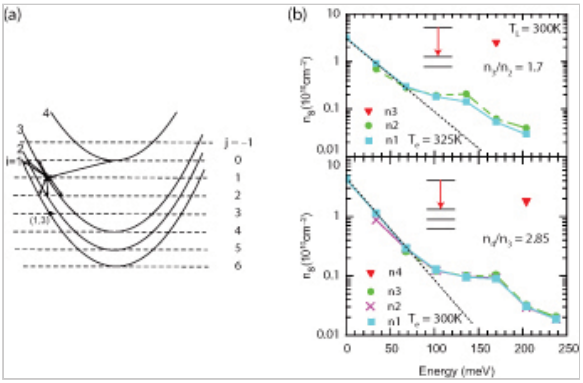


Fig. 12.6. a) Schematic description of the rate-equation model in *k*-space for an active region based on two-phonon resonance [274]. b) Electron sheet density (in units of 10^{10} cm^{-2}) as a function of energy in the relevant subbands of the active region of a three-quantumwell structure [275] (top) or a two-phonon, four-quantum-well structure [22] (bottom), each operating at $\lambda \approx 9 - 10\ \mu$m. For both devices the lattice temperature is $TL = 300$ K and the escape time from the subbands is $\tau_{esc} = 2$ ps. The injected current corresponds approximately to the threshold value, and is $J = 5$ kA/cm2 for the three-quantum-well active region, and $J = 3$ kA/cm^2 for the two-phonon resonance device. The design used are in Appendices A.2.1 and A.1.6. Reprinted, with permission, from [274]. Copyright 2002 IEEE.

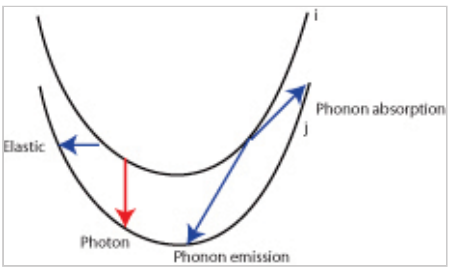
Electron temperature. Another approach, first proposed by Harrison [276], has been to assume that the populations are characterized by a common electronic temperature T_e to all subbands, different from the lattice temperature T_L. The implicit assumption is that electron–electron scattering, while not affecting the overall energy loss, will thermalize the electron distributions in each subband to a thermal distribution characterized by a single temperature. An important quantity is the average kinetic energy of all subbands, defined as

$$\langle E^k \rangle = \sum_i \langle E_i^k \rangle = \sum_i \int d^2k f(E_i^k) E_i^k$$

(12.1.13)

As shown in Fig. 12.7, in steady state the average kinetic energy will be given by the balance between the processes that tend to increase the average kinetic energy **(p.225)**

and those that, in contrast, will tend to decrease. In general, elastic scattering or intersubband optical phonon absorption increases the average in-plane energy, while optical phonon decreases it. Note that emission of photons do not modify $\langle Ek \rangle$.



The electronic temperature is found by globally balancing the energy input from all the scattering processes that increase the in-plane energy with those that decrease it, i.e. solving the equation

Fig. 12.7. Scattering processes between two subbands. While elastic scattering and phonon absorption automatically increase the average in-plane kinetic energy of the subbands, intra-subband optical phonon emission reduces it.

$$\sum_{process} \sum_{i,f} \frac{n_i(T)}{\tau_{if}} (E_i - E_f + \delta E_p) = 0$$

(12.1.14)

where E_p is the energy exchanged during the process and is equal to $\mp\hbar\omega_{LO}$ for the emission, respectively absorption of an optical phonon, zero for elastic processes, and $-\hbar\omega$ for the emission of a photon.

12.2 Density matrix

12.2.1 Limitations of the rate-equation model

In the rate-equation model, transport is completely described as a scattering mechanism between eigenstates of the unperturbed Hamiltonian. This simple approach, in which the states are not broadened, leads to unphysical results in certain situations.

Assume, as shown schematically in Fig. 12.8, two states coupled by a tunnel barrier. Often, these states would be the ground state of the injector $|g\rangle$ and the upper state $|3\rangle$ of the laser transition. However it can also be a lower state of the active region coupled to a state of the injector. When computing the transport using the rateequation approach, the Hamiltonian must be diagonalized in the extended basis where the coupling introduced by the barrier is taken into account. When the two "bare" states are in resonance, the wavefunctions of the new "dressed" states are

$$|\pm\rangle = \frac{1}{\sqrt{2}}|3\rangle \pm \frac{1}{\sqrt{2}}|g\rangle$$

(12.2.15)

where the energies of these two states are split by an energy difference $\Delta E = 2\hbar\Omega$. We assume for simplicity that the state $|g\rangle$ has an infinite lifetime and the upper state $|3\rangle$ **(p.226)**

has, due to scattering via an interaction Hamiltonian H_{int}, a lifetime τ_3. Because their wavefunction is equally splitt between the two states, the "dressed" states will have a lifetime equal to $2\tau_3$, *independent of the coupling strength* Ω. In fact, the scattering rate $\tau_{\pm}^{-1}$ is given by

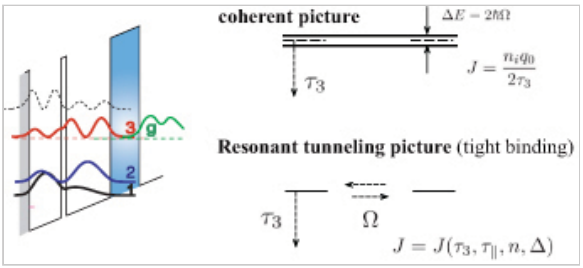


Fig. 12.8. Transport between the injector and the upper state of the laser transition (left). In the strong coupling case, the two states split in a doublet of

energies and the maximum current depends only on the upper-state lifetime and not on the coupling strength Ω. In incoherent tunneling the current depends, in addition, on the strength of in-plane scattering.

$$\tau_{\pm}^{-1} = \frac{2\pi}{\hbar}\sum_i |\langle \pm | H_{int} | i \rangle|^2 = \frac{1}{2}\frac{2\pi}{\hbar}\sum_i |\langle 3 | H_{int} | i \rangle|^2 = \frac{1}{2}\tau_3^{-1}$$

(12.2.16)

As a result, the current carried at resonance by the device will be equal to $J = \frac{ne}{2\tau_3}$, independent of the barrier thickness. This result is unphysical and arises as soon as the separation of two states is small compared to their lifetime, i.e. when the condition $\Omega\tau \gg 1$ is violated. In fact, when the coupling is small one would expect the current to decrease exponentially in the barrier thickness, as expressed already in eqn. 7.5.16. To solve this issue, a model based on density matrix theory has already been developed in Kazarinov and Suris's original work [8].

12.2.2 A two-level model system

Let us consider first a simple two-level model, described schematically in Fig. 12.9. In this model, each period has only two levels. The upper level is $n = 2$ and the injection level is $n = 1$. The coupling across the barrier between two periods is described by the Hamiltonian

$$H_0 = \begin{pmatrix} E_1 & \hbar\Omega \\ \hbar\Omega & E_2 \end{pmatrix}$$

(12.2.17)

In the Hamiltonian described by eqn. 12.2.17 we took advantage of the periodicity of the system to identify the states of each following period. For this reason, the off-diagonal element in eqn. 12.2.17 is really one coupling the state with the previous period, which is then identified with the state of the same period. While the lifetime τ_1 describes the backfilling time for electrons inside the same period, the lifetime τ_2 is the scattering time inside the period.

(p.227)

The equation of motion for the carriers, written as

$$i\hbar\dot{\rho} = [H_0, \rho] + \frac{\partial\rho}{\partial t})_{coll}$$

(12.2.18)

where the last term represents the loss of coherence due to collisions, can be rewritten in a matrix form using the Liouville operators L_0 and L':

$$i\hbar\dot{\rho} = (\mathscr{L}_0 + \mathscr{L}')\rho$$

(12.2.19)

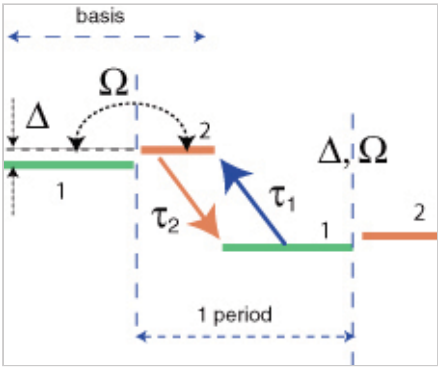


Fig. 12.9. Schematic description of a density matrix model of transport in quantum cascade lasers based on two levels. The injection level is number 1, and the upper level of the laser transition is number 2.

The density matrix elements ρ are written in vector form

$$\rho = (\rho_{11}, \rho_{12}, \rho_{21}, \rho_{22})$$

(12.2.20)
and the Liouville operators are

$$\frac{\mathscr{L}_0}{i\hbar} = \begin{pmatrix} 0 & i\Omega & -i\Omega & 0 \\ i\Omega & -i\Delta & 0 & -i\Omega \\ -i\Omega & 0 & i\Delta & i\Omega \\ 0 & -i\Omega & i\Omega & 0 \end{pmatrix}$$

(12.2.21)
where the detuning $\Delta = (E_1 - E_2)/\hbar$ has been used and

$$\frac{\mathscr{L}'}{i\hbar} = \begin{pmatrix} -\frac{1}{\tau_1} & 0 & 0 & \frac{1}{\tau_2} \\ 0 & \frac{-1}{\tau_\parallel} & 0 & 0 \\ 0 & 0 & \frac{-1}{\tau_\parallel} & 0 \\ \frac{1}{\tau_1} & 0 & 0 & -\frac{1}{\tau_2} \end{pmatrix}$$

(12.2.22)
In fact, the decay of the population is given, for the population of state 1 (represented by ρ_{11}), by a feeding term τ_2 and an emptying term τ_1, while the situation is the opposite for the second level ρ_{22}. The lifetime $\tau_\parallel$ describes the decay of the coherence between the two states by elastic or inelastic processes.

(p.228) The first quantity of interest in our model is to compute the current in the structure. It can be defined in a quantum-mechanical way by the current operator:

$$\underline{J} = e\dot{Z} = \frac{e}{i\hbar}[H, Z]$$

(12.2.23)
The expectation of the current is given by the trace of the current operator:

$$\mathscr{J} = Tr(\rho \underline{J}) = e(z_{11} - z_{22})i\Omega(\rho_{21} - \rho_{12})$$

(12.2.24)
In eqn. 12.2.24, we assumed that the states, within one period, are weakly coupled by incoherent processes, and assumed therefore that the position operator is diagonal (i.e. $z_{12} = z_{21} = 0$). In addition, in the definition of the current (eqn. 12.2.23), we should multiply the operant by a volume density; as the elements of the density matrix are thought of as areal densities we therefore divide $\mathscr{J}$ by the difference between the centroid of the wavefunctions $z_{11} - z_{22}$ to finally obtain the current density:

$$j = ei\Omega(\rho_{21} - \rho_{12}).$$

(12.2.25)
The equation system formed by the Liouville eqn. 12.2.19,

$$\dot{\rho}_{11} = -i\Omega(\rho_{21} - \rho_{12}) - \frac{\rho_{11}}{\tau_1} + \frac{\rho_{22}}{\tau_2}$$
$$\dot{\rho}_{22} = i\Omega(\rho_{21} - \rho_{12}) + \frac{\rho_{11}}{\tau_1} - \frac{\rho_{22}}{\tau_2}$$
$$\dot{\rho}_{12} = -i\Omega(\rho_{22} - \rho_{11}) - i\Delta\rho_{12} - \frac{\rho_{12}}{\tau_{\|}}$$
$$\dot{\rho}_{21} = i\Omega(\rho_{22} - \rho_{11}) + i\Delta\rho_{21} - \frac{\rho_{21}}{\tau_{\|}}$$

(12.2.26)

completed by the charge conservation condition

$$\rho_{11} + \rho_{22} = 1$$

(12.2.27)

can be solved in the steady-state ($\dot{\rho} = 0$) for $\rho_{21} - \rho_{12}$. The result, inserted into the expression for the current eqn. 12.2.25, and multiplied by the total sheet density n_s, yields

$$j = \frac{e n_s \Omega^2 \tau_{\|} \frac{\tau_1 - \tau_2}{\tau_1 + \tau_2}}{1 + \Delta^2 \tau_{\|}^2 + 4\Omega^2 \tau_{\|} \frac{\tau_1 \tau_2}{\tau_1 + \tau_2}}$$

(12.2.28)

In a situation where the backscattering is negligible and the injection state has an infinite lifetime $\tau_1 \to \infty$, the above expression reduces to the well-known Kazarinov and Suris formula for the current:

$$j = \frac{e n_s \Omega^2 \tau_{\|}}{1 + \Delta^2 \tau_{\|}^2 + 4\Omega^2 \tau_{\|} \tau_2}$$

(12.2.29)

This expression, derived first in a slightly modified form by Kazarinov and Suris [7, 8], has often been the basis used to explain the transport in quantum cascade lasers, **(p.229)**

and in particular the injection process [277]. It describes the transition between the coherent and incoherent tunneling. In an injection process, τ_2 represents the lifetime of the upper state and $\tau_{\|}$ the dephasing between the upper state of the laser transition and the ground state of the injector. The detuning Δ is given by the device's operating voltage. The coupling strength Ω is an important design parameter, because it is mainly defined at the design stage by the thickness of the injection barrier.

The condition $\Omega^2 \tau_{\|} \tau_2 \gg 1$ defines the strong coupling regime. In this case the maximum current, always obtained at resonance (Δ = 0), is given by

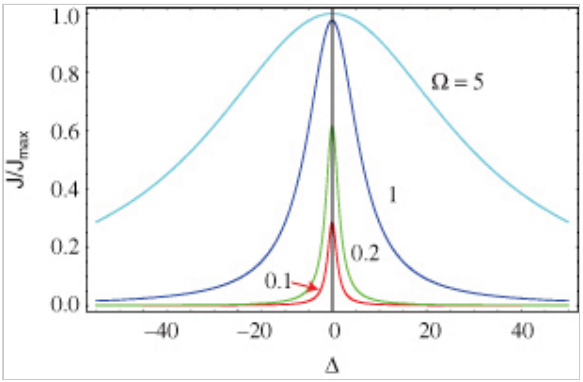


Fig. 12.10. Normalized current J/J_{max} as a function of normalized detuning $\Delta\tau_{\|}$. A value $\tau_2 = 10\tau_{\|}$ was assumed. The values of the coupling strength are Ω = 0.1, 0.2, 1, and 5, as indicated in the figure. The values of the term $4\Omega\tau_{\|}\tau_2$ are then 0.4, 1.6, 4, and 1000 respectively.

$$J = \frac{en_s}{2\tau_2}$$

(12.2.30)

and does not depend on either the dephasing $\tau_{\parallel}$ or the coupling strength Ω.

In the opposite case $\Omega\tau_{//}\tau_2 \ll 1$ we are in the weak coupling regime: the maximum current, still achieved at resonance, is now given by

$$J = en_s\Omega^2\tau_{\parallel}$$

(12.2.31)

which is, as expected, proportional to Ω^2. The transition between these two regimes is described in Fig. 12.10, where the normalized current J/J_{max} is plotted as a function of normalized detuning $\Delta\tau_{\parallel}$ for increasing values of the coupling strength Ω. The upperstate lifetime was assumed to be 10 times longer than the in-plane dephasing time, a typical value for mid-infrared quantum cascade lasers, and the maximum current available for the doping density and upper-state lifetime J_{max} is given by eqn. 12.2.30. The data shown in Fig. 12.10 illustrates the compromise needed in the design of the injection barrier. A too-low value of Ω will limit the maximum output current, while a too-thin injection barrier will unduly broaden the voltage-current curve as well as the laser transition and enhance non-resonant injection in other states. In fact, the nonresonant injection into the lower state of the laser transition is a major loss mechanism **(p.230)**

that will negatively influence the threshold current and slope efficiency. For this reason, assuming the linewidth of the luminescence is a good measure of $\tau_{\parallel}$, a rule of thumb has long been to adjust the coupling strength such that the splitting between the injector state and the upper state 2 Ω is equal to the width of the luminescence, assumed to be equal to $2/\tau_{\parallel}$, i.e. $\Omega\tau_{\parallel} = 1$. As shown in Fig. 12.10, for such a value of Ω the maximum current is almost given by eqn. 12.2.30 while the additional broadening is still kept to a reasonable value.

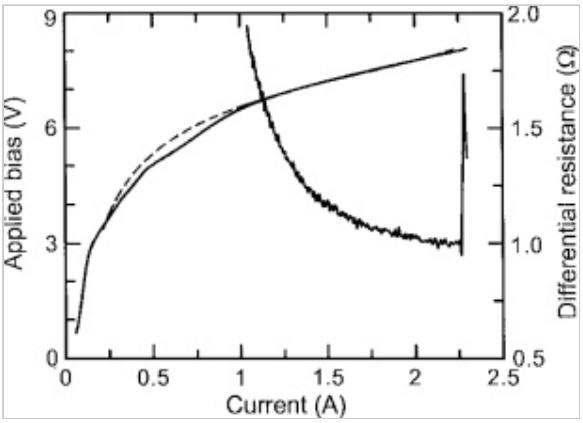


Fig. 12.11. Voltage–current (upper curves, left axis) and differential resistance-versuscurrent characteristics of a quantum cascade laser (full lines) operating at $\lambda = 8.5\ \mu$ m at 300 K. Dashed line, the same characteristic computed using expression 12.2.29, and where only the sheet density n_s and the in-plane scattering $\tau_{//}$ have been fitted. Adapted, with permission, from [277]. Copyright 1998 IEEE.

Expression 12.2.29, even if based on a very simplified model, has nevertheless been successfully used to model the current–voltage curves of quantum cascade lasers. Fig. 12.11 shows the voltage–current characteristics of a quantum cascade laser operating at operating at $\lambda = 8.5\ \mu$ m at 300 K, compared to the result of expression 12.2.29 where only the sheet density n_s and the in-plane scattering $\tau_{\parallel}$ have been fitted, the other parameters being computed from the knowledge of the

band structure. A very good agreement between the fitted and experimental values of the sheet density ($n_s = 1.6 \times 10^{11}$ cm^{-2} and $n_s = 1.5 \times 10^{11}$ cm^{-2}), as well as between the fitted value of the in-plane scattering time $\tau_{\parallel} = 40fs$ and the one derived from the linewidth of the luminescence $\tau_{\parallel} = 38fs$, were obtained [277]. In contrast to this result, it has been argued recently [152] that the luminescence linewidth is a poor estimate of $\tau_{\parallel}$ as the latter describes the loss of coherence in the injection process between the injector and the upper-state of the laser transition, while the luminescence linewidth depends on the dephasing between the upper and lower state of the laser transition. In addition, the computation of the broadening contains interference terms. Namely, a fluctuation of the layer thickness will induce a modification of the energy of both lower and upperstates by a quantity of the same sign, and therefore the linewidth of the transition will be in general narrower than the broadening of the upper-state. Mathematically, it **(p.231)**

arises from the presence of the substraction between terms $(F_{00} - F_{11})^2$ in eqn. 5.3.25, and the strength of this effect obviously depends on the active region design and the detail of the form factors F_{ii}.

Optimization of the injection barrier. The thickness of the injection barrier has to be optimized for each design. On one hand a stronger injection reduces the injection efficiency, since, at resonance, the injection into the upper-state saturates while the injection in the other states increases. A too-strong injection barrier also broadens the laser transition. A too-weak injection, provided by a thick barrier, on the other hand, limits the maximum dynamical range of the laser by reducing the maximum current that can be driven before the alignment between injector and upper-state breaks down. In [278] the authors have explored experimentally the optimization of an injection barrier in a GaAs-based quantum cascade laser using quantitative electroluminescence. As shown in Fig. 12.12, the linewidth of the electroluminescence increases, and its intensity decreases, as the thickness of the barrier is reduced. On the other hand, a too-thick injection barrier (in this case 〉 7 nm) reduces the dynamical range. Their result yielded an optimum barrier thickness (7 nm thick) corresponding to $\Omega \approx 0.2\gamma$, where 2γ is the FWHM of the luminescence. One expects that such optimization, performed at room temperature, yielding stronger couplings (even in relative terms) because of the shorter upper-state lifetime.

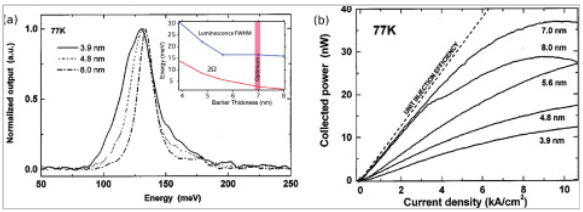


Fig. 12.12. Optimization of the injection barrier thickness in a GaAs/$Al_{0.33}Ga_{0.67}$ As quantum cascade laser. a) Normalized electroluminescence spectra at a current density $J = 6kA/cm^2$ for three barrier thicknesses, as indicated. Inset: measured FWHM and computed anticrossing energy (2 Ω) as a function of the barrier thickness. b) Calibrated LI curves for samples at 77 K. The dashed line represents the calculated collected output power for unit injection efficiency. Adapted with permission from [278]. Copyright 2001, AIP.

The numerical optimization of the injection barrier is a difficult task because it implies a model that works well in a regime between the”strong coupling” and “weak coupling” limits. Some

very encouraging results have been obtained by Dupont *et al*. in the terahertz, using a density matrix model [198].

12.2.3 Bloch gain and second-order current

The preceding derivations assumed that the states could be treated in an atomiclevel fashion, i.e. that the in-plane dispersion could be factored out. Eqn. 12.2.29, **(p.232)** while extremely useful to interpret the transport data, has still some weaknesses. For example, it is symmetric with Δ, and therefore predicts the same current for an injector misalignement below or above the resonance. For low temperatures, $kT \langle \Delta$, this approximation is a poor one.

This behavior can be corrected by explicitly introducing the in-plane dispersion in the density matrix model, and was first derived by Willenberg *et al*. [279]. This model takes advantage of the fact that the coupling Ω between two states preserves the in-plane momentum and is therefore diagonal in *k*, while the dominant intrasubband scattering is interface roughness scattering. The latter is energy-conserving and is maximum between states belonging to the same subband, and therefore will be diagonal in the state number *i, j*. As a result, starting with the same two levels as in the previous paragraph, and using the same state numbering as [279], the matrix elements of the Hamiltonian in this basis are given by

$$\mathscr{H}_{kk'}^{ij} = \langle ik|\mathscr{H}|jk'\rangle = H_{k'}^{ij}\delta_{kk'} + V_{kk'}^{ij},$$

(12.2.32)

where the respective contributions *H* and *V* take the form

$$H_k^{ij} = \begin{pmatrix} E_{2k} & \hbar\Omega_{21} \\ \hbar\Omega_{12} & E_{1k} \end{pmatrix}^{ij}, V_{kk'}^{ij} = \begin{pmatrix} V_{kk'}^{22} & 0 \\ 0 & V_{kk'}^{11} \end{pmatrix}^{ij}$$

Thus, electrons are allowed to tunnel between the subband state *i* and *j* by means of the momentum-conserving matrix elements $\hbar\Omega_{ij}$, in each of which they are possibly scattered out of a virtual intermediate state by an intrawell relaxation process *V ii* as depicted in Fig. 12.13.

For simplicity we assume a parabolic dispersion relation parallel to the layers in the effective mass approximation

$$E_{ik} = E_i + \frac{\hbar^2 k^2}{2m^*}$$

(12.2.33)

where E_i denotes the lower subband edge and m^* is the effective mass of the electron averaged over the extension of the wavefunction in well and barrier. Extension to subbands with a non-parabolic dispersion relation $E_{i,k} = E_i(k)$ is, however, straightforward, as shown later.

The time evolution of the system is driven by the equation of motion of the density matrix (eqn. 12.2.18), where, however, the collision term is now explicitly contained in the Hamiltonian $\mathscr{H}$. Eqn. 12.2.18 is first rewritten, separating the diagonal and non-diagonal part of the density matrix $\hat{\rho}_{kk'}^{ij}$ with respect to the parallel momentum *k, k′* according to[1]

$$\hat{\rho}_{kk'}^{ij} = \delta_{kk'}\rho_k^{ij} + (1 - \delta_{kk'})\rho_{kk'}^{ij}$$

(12.2.34)

We obtain two coupled equations with four terms each, which determine the time evolution of the system. Since the coherent term H is diagonal with respect to the inplane momentum and the scattering term V is purely non-diagonal, the commutators **(p.233)**

that determine the diagonal and non-diagonal part of the density matrix are evaluated as

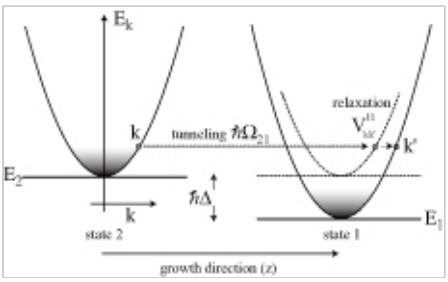


Fig. 12.13. Mixed momentum and real-space picture of a two-level system where the inplane scattering is taken into account. Tunneling into a virtual intermediate state (dotted) is expressed by a transfer matrix element $\hbar\Omega_{ij}$. Scattering is assumed to take place within each subband only.

$$i\hbar\dot{\rho}_k^{ij} = \sum_m (H_k^{im}\rho_k^{mj} - \rho_k^{im}H_k^{mj}) + \sum_{m,k'} (V_{kk'}^{im}\rho_{k'k}^{mj} - \rho_{kk'}^{im}V_{k'k}^{mj})$$

(12.2.35)

$$i\hbar\dot{\rho}_{kk'}^{ij} \approx \sum_m (V_{kk'}^{im}\rho_{k'}^{mj} - \rho_k^{im}V_{kk'}^{mj}) + \sum_m (H_k^{im}\rho_{kk'}^{mj} - \rho_{kk'}^{im}H_{k'}^{mj})$$

(12.2.36)

where the commutator of the scattering potential with the non-diagonal part of ρ has been neglected in the second equation (Born approximation). The steady-state values of the coherences of the density matrix $f\,ij$, which determine the transitions $|ik\rangle \rightarrow |jk\rangle$, the current and the absorption, are obtained from a Laplace average [280] defined by

$$f(s) = s\int_0^\infty dt e^{-st}\rho(t)$$

(12.2.37)

and performing the Laplace limit $s \rightarrow 0$ using the relation $\lim_{s\rightarrow 0_+} (\omega - is)^{-1} = P(1/\omega) + i\pi\delta(\omega)$ at the appropriate stage of the calculation. In this approach the populations f_k^{ii} of the density matrix are not accessible and appear in the resultant expression as external quantities. The Laplace average gives

$$i\hbar s f_k^{ij} =$$
$$i\hbar s\rho_k^{ij}(0) + \sum_m (H_k^{im}f_k^{mj} - f_k^{im}H_k^{mj}) + \sum_{m,k'} (V_{kk'}^{im}f_{k'k}^{mj} - f_{kk'}^{im}V_{k'k}^{mj})$$

(12.2.38)

$$i\hbar s f_{kk'}^{ij} \approx$$
$$i\hbar s\rho_{kk'}^{ij}(0) + \sum_m (V_{kk'}^{im}f_{k'}^{mj} - f_k^{im}V_{kk'}^{mj}) + \sum_m (H_k^{im}f_{kk'}^{mj} - f_{kk'}^{im}H_{k'}^{mj}).$$

(12.2.39)

In a first step, the non-diagonal part in eqn. (12.2.39) is approximated (see [279]). Specifying the assumptions of a two-level-system with intrawell scattering only, and neglecting terms of higher order in Ω_{ij} corresponding to multiple tunneling processes, gives

$$(f)^{ij}_{kk'} \approx i\pi\delta(E_{ik} - E_{jk'})[V^{ii}_{kk'}f^{ij}_{k'} - f^{ij}_{k}V^{jj}_{kk'} + \\ + \hbar\Omega_{ij}\left(\frac{V^{jj}_{kk'}(f^{jj}_{k} - f^{jj}_{k'})}{E_{jk} - E_{jk'}} - \frac{V^{ii}_{kk'}(f^{ii}_{k} - f^{ii}_{k'})}{E_{ik} - E_{ik'}}\right)]$$

(12.2.40)

(p.234) which has to be placed in eqn. (12.2.38) for the diagonal part. Simplifying for intrawell scattering here, and taking the Laplace limit, yields (see [279])

$$(E_{ik} - E_{jk})f^{ij}_{k} = \hbar\Omega_{ij}(f^{ii}_{k} - f^{jj}_{k}) - \sum_{k'}(V^{ii}_{kk'}(f)^{ij}_{k'k} - (f)^{ij}_{kk'}V^{jj}_{k'k})$$

where *f* denotes the approximated expression for the non-diagonal part in eqn. 12.2.40. Performing an ensemble average, i.e. dropping terms related to correlation effects, we obtain

$$(E_{ik} - E_{jk})f^{ij}_{k} - i\pi f^{ij}_{k}\sum_{k'}\delta(E_{ik'} - E_{jk})|V^{ii}_{kk'}|^2 + \delta(E_{ik} - E_{jk'})|V^{jj}_{kk'}|^2 \\ = \hbar\Omega_{ij}(f^{ii}_{k} - f^{jj}_{k}) + i\pi\sum_{k'}[\delta(E_{ik'} - E_{jk})|V^{ii}_{kk'}|^2\frac{\hbar\Omega_{ij}}{E_{ik'} - E_{ik}}(f^{ii}_{k'} - f^{ii}_{k}) \\ +\delta(E_{ik} - E_{jk'})|V^{jj}_{kk'}|^2\frac{\hbar\Omega_{ij}}{E_{jk} - E_{jk'}}(f^{jj}_{k} - f^{jj}_{k'})]$$

(12.2.41)

Which agrees with the previous result [8]. In contrast to the original treatment, we continue by neither neglecting the difference of the arguments in the δ functions on the LHS nor omitting the second term on the RHS. The coherence associated with the transition $|2k\rangle \rightarrow |1k\rangle$ is obtained from

$$\hbar\Delta f^{21}_{k} - \overbrace{i(\gamma^2_k + \gamma^1_k)}^{\text{broadening}}f^{21}_{k} = \overbrace{\hbar\Omega_{21}(f^{22}_{k} - f^{11}_{k})}^{\text{population difference}} + \underbrace{\frac{i\hbar\Omega_{21}}{\hbar\Delta}(\gamma^2_k(f^{22}_{q_-} - f^{22}_{k}) - \gamma^1_k(f^{11}_{q_+} - f^{11}_{k}))}_{\text{Bloch-type contribution}}$$

(12.2.42)

where we have used abbreviations for the scattering-induced broadening of the transition

$$\gamma^i_k = \sum_{k'}\delta(E_{ik'} - E_{jk})|V^{ii}_{kk'}|^2$$

(12.2.43)

the subband separation $\hbar\Delta = E_{2k} - E_{1k}$ and the in-plane momentum of the final state

$$q_{\pm} = \hbar^{-1}\sqrt{2m^*(E_k \pm \hbar\Delta)}.$$

(12.2.44)

The first term on the RHS of eqn. 12.2.42, which contains the difference of populations between the two states, corresponds to the central result of [8]. The second term, first introduced in [279], contains differences in population within a subband. When applied to the transport problem, it modifies the expression for the current by introducing an asymmetry in the dependence of the current density as a function of detuning Δ. In addition, it is also responsible

for the second-order type of gain, leading to the characteristic negative differential conductivity and the dispersive gain profile in a superlattice, and a modified spectral shape of the gain in a quantum cascade laser.

(p.235) *Second-order current*. We use for the current an expression similar eqn. 12.2.25, but where we keep explicitly the spatial term $z_{11} - z_{22}$, and we interpret the density matrix as a volume density:

$$j = \epsilon d \sum_k i(\Omega_{21} f_k^{12} - \Omega_{12} f_k^{21})$$

(12.2.45)
where the notation $d = z_{22} - z_{11}$ has been used. Using the expressions 12.2.42 for the coherences f_k^{12} and f_k^{21}, the current yields

$$j = \frac{2\epsilon d |\hbar\Omega_{21}|^2}{\hbar} \sum_k \frac{\gamma_k^1 (f_k^{22} - f_{q+}^{11}) + \gamma_k^2 (f_{q-}^{22} - f_k^{11})}{(\hbar\Delta)^2 + (\gamma_k^1 + \gamma_k^2)^2}$$

(12.2.46)
The current results from differences in population evaluated *for non-equivalentkstates in the respective subbands*. The sum runs only on allowed transitions, i.e. those for which the final state $q_{\pm}$ exists.

A possible formulation of Kazarinov and Suri's expressions 12.2.29 for the current can be recovered by setting $q_{\pm}$ equal to k, and assuming a constant broadening γ:

$$j = \frac{2e|\hbar\Omega_{21}|^2}{\hbar} \frac{\gamma}{(\hbar\Delta)^2 + \gamma^2} \Delta n$$

(12.2.47)
where $\Delta n = n_s (\rho_{22} - \rho_{11})$ is the population difference. The above expression can be directly obtained by combining the third and fourth line of eqn. 12.2.26.

The general expression for the current, 12.2.46, is a priori valid for any population distribution within the subbands. A simpler form can be derived if one assumes a classical distribution characterized by the same electron temperature T in each subband. When, in addition, a uniform scattering potential is considered ($\gamma_k^1 = \gamma_k^2 = \gamma$), the k k current density can be integrated and simply rewritten as

$$j = \frac{2\epsilon d |\hbar\Omega_{21}|^2}{\hbar} \frac{2\gamma}{(\hbar\Delta)^2 + (2\gamma)^2} \left\{ \theta(\Delta)(n_2 - e^{-\beta\hbar|\Delta|} n_1) + \theta(-\Delta)(e^{-\beta\hbar|\Delta|} n_2 - n_1) \right\}$$

(12.2.48)
where $\theta(x)$ is the Heaviside function, with $\theta(x^-) = 0$, $\theta(x^+) = 1$ and $\theta(0) = \frac{1}{2}$, $\beta = 1/kT$, with k the Boltzmann constant, and n_i is the net population of subband i.

The current density is no more driven by the population difference $n_2 - n_1$ but by an effective population term. Let us examine two extreme case: equally populated subbands $n_2 = n_1 = n$, and one empty subband $n_1 = 0$. The first case is shown in Fig. 12.14a, and is the one in which expression 12.2.47 predicts no current, regardless of the detuning. In contrast, the current density that includes the second-order scattering effects exhibits a dispersive shape around the

resonance. A negative current peak occurs when the detuning is negative, which is when the edge of subband 1 is above the edge of subband 2. The current then turns to be positive, after the edge of subband 2 has overcome the edge of subband 1, which is when the detuning is positive. The dispersive shape is the consequence of electron tunneling at a constant energy rather than at a constant wavevector. This case illustrates a superlattice: the current is zero until second-order scattering terms have been taken into account. The case where one subband is empty is shown in Fig. 12.14b. For negative detunings, the current **(p.236)**

between the subbands is exponentially reduced, as only the electrons with a sufficient kinetic energy are able to tunnel to subband 1. In contrast, for positive detunings, the firstand second-order curves overlap perfectly, as all electrons are above the edge of subband 1.

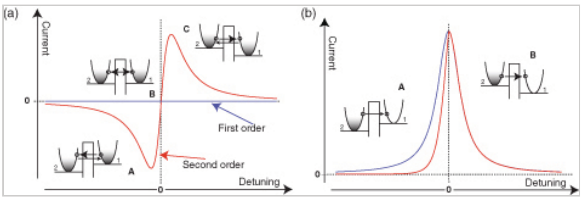


Fig. 12.14. (a) Effects of second-order contributions on tunneling between a pair of equally populated subbands. (A) When the detuning is negative, the subband edge of subband 1 is above the edge of subband 2. As tunneling conserves energy, the current flow from 1 to 2 is greater than the current flow from 2 to 1, yielding a negative net current. (B) When the subbands are aligned the detuning is zero, and both contributions cancel, yielding a zero net current. (C) The detuning is positive and therefore the edge of subband 2 is above the edge of subband 1, yielding a positive net current between subbands. (b) Empty subband 1. (A) The current is reduced, as only a fraction of electrons can tunnel. (B) Models overlap perfectly. Because the mathematical expressions are the same, this figure also depicts the intersubband absorption for various population inversion between the subbands, the detuning Δ being interpreted as the frequency detuning between the transition and incident photon.

The comparison between the current computed using the second-order current of eqn. 12.2.48 and the simplified one from Kazarinov and Suris is shown in Fig. 12.15, applied to a quantum cascade laser structure. In this computation [281] the total current was calculated using all the states of the active region, as described later in this text.

When injecting the current from a filled state to an empty one and as shown in Fig. 12.14, the symmetric Lorenzian lineshape obtained from Kazarinov and Suris's expression for the current tends to *overestimate* the current before a resonance. This effect is clearly apparent in Fig. 12.15, where the predictions using Kazarinov's expression consistently overestimate the experimentally measured current, while the second-order current expression 12.2.48 leads to a much more accurate prediction.

12.2.4 Gain and absorption

The Hamiltonian 12.2.32, although derived for the case of two states coupled by a tunnel barrier, can also be interpreted in a completely different manner. Let us assume a two-level system in the presence of an optical field. In Chapter 4 we saw that the interaction Hamiltonian can be described either by a matrix element of the position (eqn. 4.1.2) or the momentum (eqn. 4.1.5) operators. Using the momentum matrix element given by eqn. 4.1.5, and writing it as a function of the electric field E_ω, we have

(p.237)

$$H_{int} = \frac{-e\mathscr{E}_\omega}{2i\omega m_0} p_{ij}(e^{i\omega t} + e^{-i\omega t})$$

(12.2.49)

Taking advantage of the relation between momentum and dipole matrix element 4.1.7 $p_{ij} = im_0\Omega_{ij}z_{ij}$, we can finally write our interaction Hamiltonian in our basis as:

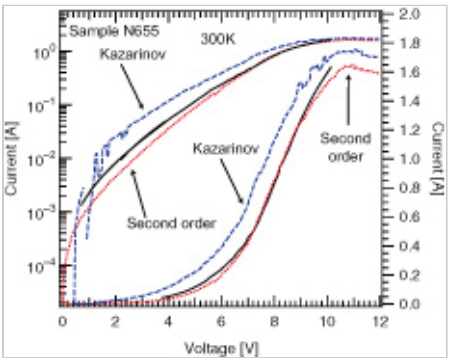


Fig. 12.15. Comparison of the experimental current–voltage curves with the prediction of a density matrix theory (Design described in appendix A.2.4). Kazarinov and Suris's Lorenzian resonant shape is compared to the second-order current expression. Reprinted with permission from [281]. Copyright 2008, APS.

$$H_{int} = \begin{pmatrix} 0 & -\frac{e\mathscr{E}_\omega z_{21}}{\omega}\Omega_{21}e^{-i\omega t} \\ -\frac{e\mathscr{E}_\omega z_{12}}{\omega}\Omega_{12}e^{i\omega t} & 0 \end{pmatrix}$$

(12.2.50)

In the above expression the diagonal terms have been omitted. This is because we are interested in the low-field regime, and therefore will neglect the shift of the levels experienced by the optical field. These terms would be zero in a symmetric potential.

Finding the time-dependence of the density matrix involves solving the equation of motion 12.2.18, assuming that the coherences ρ_{21} and ρ_{12} have an harmonic time dependence $e^{\pm i\omega t}$2. The key result is that the solutions have the same form as the time-independent solutions, being evaluated at an energy $\hbar(\Omega_{21} - \omega)$ instead of $\hbar\Omega_{21}$.

As the Hamiltonian 12.2.50 is, like the one describing the tunneling, diagonal in respect to the in-plane momentum, the above results can also be applied in that case. As a result, we can write the corrections to the coherences in linear response δf_k^{21} and δf_k^{12}, which are related to f_k^{21} and f_k^{12} by

$$\delta f_{k,\omega}^{21} = -\frac{e\mathscr{E}_\omega z_{21}}{\hbar\omega} f_{k,E_{21}-\hbar\omega}^{21} \text{ and } \delta f_{k,\omega}^{12} = \frac{e\mathscr{E}_\omega z_{21}}{\hbar\omega} f_{k,E_{21}+\hbar\omega}^{12}$$

(12.2.51)

which are evaluated at an energy $E_{21} \pm \hbar\omega$ instead of E_{21} due to the time dependence of the ac field. The corrections to the coherences δf_k^{21} will induce a photon-driven ac current, given by

(p.238)

$$\delta j(\omega) \approx e z_{12} \sum_k i(\Omega_{21}\delta f_{k,\omega}^{12} - \Omega_{12}\delta f_{k,\omega}^{21})$$

(12.2.52)
The absorption can be deduced from the ac conductivity, defined as $\sigma(\omega) = \partial(j + \delta j(\omega))/\partial\mathcal{E}_\omega$, using

$$\alpha(\omega) = \frac{\Re(\sigma(\omega))}{\epsilon_0 n_r c}$$

(12.2.53)
Simply replacing the prefactors and by analogy with eqn. 12.2.46 we finally obtain the expression

$$\alpha(\omega) = -\frac{e^2|z_{21}|^2 E_{21}^2}{\epsilon_0 n_r c\hbar^2\omega}\sum_k \frac{\gamma_k^1(f_k^{22} - f_{k_+}^{11}) + \gamma_k^2(f_{k_-}^{22} - f_k^{11})}{(E_{21} - \hbar\omega)^2 + (\gamma_k^1 + \gamma_k^2)^2}$$

(12.2.54)
Note that because of the analogy between the expressions for the current and the absorption, a simplified expression also exists for the absorption lineshape in the case of a classical thermal distribution, similar to the one for the current given by eqn. 12.2.48:

$$\begin{aligned}\alpha(\omega) = -\frac{e^2|z_{21}|^2 E_{21}^2}{\epsilon_0 n_r c\hbar^2\omega}\frac{2\gamma}{(E_{21}-\hbar\omega)^2+(2\gamma)^2}\times \\ \times\{\theta(E_{21} - \hbar\omega)(n_2 - e^{-\beta|E_{21}-\hbar\omega|}n_1)+ \\ +\theta(-(E_{21} - \hbar\omega))(e^{-\beta|E_{21}-\hbar\omega|}n_2 - n_1)\}.\end{aligned}$$

(12.2.55)
As a result of the similarity between the two expressions, the plot of the current as a function of detuning Δ displayed in Fig. 12.14 can be also understood as that of an intersubband absorption where Δ is the frequency detuning between the transition and incident photons.

The most striking feature of the expression for the gain between subbands of eqn. 12.2.54 is that a gain and absorption persist even for equal upperand lowerstate populations. Fig. 12.16 shows the intersubband gain, computed using expression 12.2.54, for increasing values of the lower-state population, the upper one remaining fixed. As the lower population is increased, the gain evolves from a quasi-Lorenzian lineshape at high population inversion to a dispersive lineshape, characteristic of a Bloch oscillator, for equal populations. The gain is still present even when the global subband upperand lower-state populations are equal, because the transitions occur between states with non equivalent in-plane wavevectors.

As will be shown in the following, expression (12.2.54) allows a simple explanation of the intersubband gain mechanism in the presence of strong in-plane scattering. It is instructive to rewrite the differences in populations as

$$\gamma_k^1(f_k^{22} - f_{k+}^{11}) = \gamma_k^1(\underbrace{f_k^{22}(1 - f_{k+}^{11})}_{\text{emission}|2\rangle\to|1\rangle} - \underbrace{f_{k+}^{11}(1 - f_k^{22})}_{\text{absorption}|1\rangle\to|2\rangle})$$

(12.2.56)
which translate directly into the paths depicted in the inset of Fig. 12.17. The two processes above relate the states $|2k\rangle$ and $|1k_+\rangle$ by the emission or absorption of a (non-resonant) photon, $\hbar\omega \neq E_{21}$ for $k \neq k_+$, assisted by relaxation within the lower state via γ_k^1, which

ensures momentum transfer. The second difference in eqn. (12.2.54) is interpreted accordingly, where the relaxation takes place within the upper state.

(p.239)

Bloch oscillator. The above derivation is actually relevant to a relatively old issue in quantum heterostructure research. Ineeded, soon after the original proposal of semiconductor superlattices [2], two apparently quite different schemes to obtain optical gain in such novel systems were put forward. As discussed previously, Kazarinov and Suris pointed out that population inversion and gain can be achieved between electronic subbands in a strongly biased superlattice [7], paving the way to the realization of the quantum cascade laser. On the other hand, based on semi-classical arguments, Ktitorov *et al*. [282] and later Ignatov *et al*. [283] predicted optical gain due to Bloch **(p.240)** oscillations in the miniband transport—despite an apparent missing population inversion.

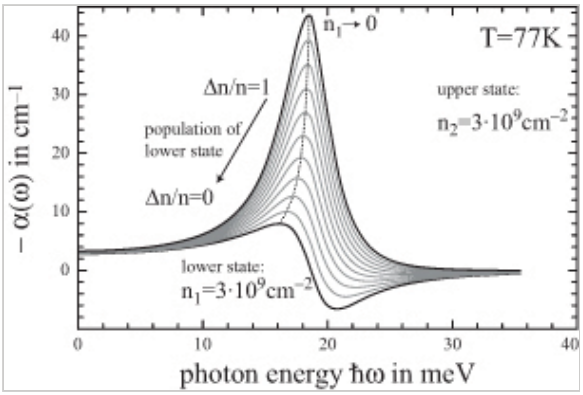


Fig. 12.16. Plot of the intersubband second-order gain from eqn. 12.2.54, computed for various values of the lower-state populations, the upper one remaining fixed. As the lower population is increased, the peak of the gain shifts towards lower energies. A dispersive gain curve is achieved when the populations are equal. Reprinted with permission from [279]. Copyright 2002, APS.

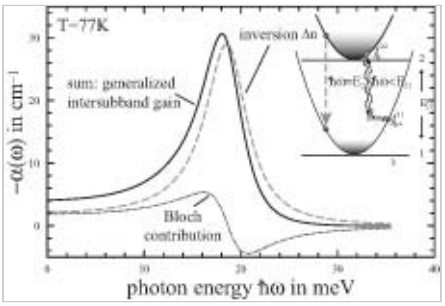


Fig. 12.17. Full line, intersubband gain for a population inversion $\Delta n/n = 0.5$, is the sum of a direct contribution (dashed line) and a Bloch-like contribution (dotted) corresponding to the first and second terms in eqn. 12.2.42. Inset: possible quantum-mechanical paths; for an incident photon with energy $\hbar\omega \neq E_{21}$, absorption or stimulated emission may occur due to a non-resonant absorption or emission into an intermediate state and a subsequent relaxation into the final state.

A key problem hampering the feasibility of the Bloch oscillator, emitting electromagnetic radiation, tunable by the external electric dc field, is the need to stabilize the electric field domains in a biased superlattice at the point of operation. For a long time the relation between both proposals was unclear because the gain in Bloch oscillator was described using a semi-classical model of transport based on the Boltzman equation, while the one in quantum cascade lasers is described using a quantummechanical approach. Interestingly, the result provided by the density matrix model described in eqn. 12.2.54 enables a quantum-mechanical interpretation of the Bloch gain in superlattices, establishing a link

between the intersubband gain originating from a population inversion, with its symmetric spectral shape centered at the transition energy, and the dispersive gain predicted for a periodic superlattice, with its nearly anti-symmetric profile.

Fig. 12.18 shows the prediction of the gain in a biased superlattice for the semiclassical model whose expression is given by Ignatov *et al*. [284], where distinct momentum and energy relaxation times τ_p, τ_e are used:

$$\alpha_{sc}(\omega) = \\ = \frac{e^2 d^2}{\epsilon_0 n_r c} \frac{\Delta}{2\hbar^2} n^{(3)} \frac{I_1(\Delta/2k_BT)}{I_0(\Delta/2k_BT)} \frac{\tau_p}{1+\omega_b^2 \tau_e \tau_p} \Re\left(\frac{1-i\omega\tau_e-\omega_b^2\tau_e\tau_p}{\omega_b^2\tau_e\tau_p+(1-i\omega\tau_e)(1-i\omega\tau_p)}\right),$$

(12.2.57)

where ω_b is the Bloch frequency and Δ the width of the miniband. The ratio of Bessel functions contains the temperature dependence for a non-degenerate electron gas. Fig. 12.18 shows a comparison of the semi-classical results and the quantum-mechanical predictions for the same constant relaxation time, $\tau = \hbar/\gamma =$ 0.2 ps, at **(p.241)** different temperatures *T*. No independent parameters are used. The two approaches agree remarkably well at high temperatures in the semi-classical limit $eF\,d \;\langle\; \Delta$. The narrowing of the Bloch gain profile with lower temperature, compared to the semiclassical curve, reflects an explicit influence of the electron distribution within the subband. This influence is absent in the semi-classical treatment, regardless of the approximation for the distribution function. In the above comparison the full expression of eqn. 12.2.54 was used that includes the poles at negative frequencies

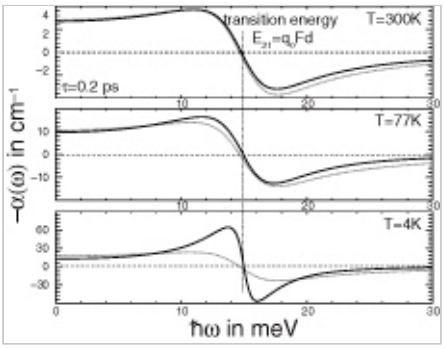


Fig. 12.18. Semi-classical (dotted line) versus quantum-mechanical results (full line) for the absorption in a superlattice for different temperatures *T*. We assume a temperatureindependent scattering time $\tau = 0.2$ ps in the quantum-mechanical model, and set $\tau_{k,\,e} = \tau$ in the semi-classical model. In the semi-classical picture the peak gain scales with the ratio $I_1(\Delta/2k_BT)/I_0(\Delta/2k_BT)$. The quantum-mechanical gain profile exhibits an additional narrowing with lower temperature. Reprinted with permission from [279]. Copyright 2002, APS.

$$\alpha(\omega) = \frac{e^2|z_{21}|^2 E_{21}^2}{\epsilon_0 n_r c \hbar^2 \omega}\Big[\sum_k \left(\frac{\hbar\tau_k^{-1}}{(\epsilon+\hbar\omega)^2+(\hbar\tau_k^{-1})^2} - \frac{\hbar\tau_k^{-1}}{(\epsilon-\hbar\omega)^2+(\hbar\tau_k^{-1})^2}\right)(f_k^{22}-f_k^{11}) + \\ + \sum_k \left(\frac{\gamma_k^2(f_{l-}^{22}-f_k^{22})-\gamma_k^1(f_{l+}^{11}-f_k^{11})}{(\epsilon+\hbar\omega)^2+(\hbar\tau_k^{-1})^2} - \frac{\gamma_k^2(f_{k-}^{22}-f_k^{22})-\gamma_k^1(f_{k+}^{11}-f_k^{11})}{(\epsilon-\hbar\omega)^2+(\hbar\tau_k^{-1})^2}\right)\Big]$$

(12.2.58)

where we set

$$k_\pm = \hbar^{-1}\sqrt{2m^*(\epsilon_k \pm (\epsilon-\hbar\omega))}, l_\pm = \hbar^{-1}\sqrt{2m^*(\epsilon_k \pm (\epsilon+\hbar\omega))} \text{and} \hbar\tau_k^{-1} = \gamma_k^1+\gamma_k^2.$$

The agreement shown in Fig. 12.18 supports the identification of the gain described with the second-order formula as Bloch gain.

Experimental results. To a certain extent, the prediction of the second-order model of gain is already validated by the same expression, derived for the current. In order to test the predictions of this density-matrix model experimentally, it is necessary to measure the spectral profile of the gain in a structure in which the population inversion can be controlled. In [265] the authors compared gain measurements in quantum cascade lasers built either from a single-quantum-well active region, where the extraction lifetime could be varied by changing the barrier thickness, and compared them to measurements carried out in an active region where a resonant optical phonon scattering provided a fast depopulation of the lower state. The gain measurements were performed by recording the electroluminescence in a multi-section cavity structure, where one section was used as a broadband source and the other one as an amplifier, as described in Section 11.5.4.

The measurements, performed at liquid-helium temperature, are compared in Fig. 12.19 with theoretical simulations. The gain spectrum of the single-quantumwell sample (N258) as a function of injected current is displayed in Fig. 12.19a. At the lower current densities (2.1 and 3.3 kA/cm^2) the gain shows a clear dispersive shape, with gain for photon energies below the expected transition energy and loss at higher energies. As the current is increased, a more symmetric shape is recovered, indicating an improvement in the population inversion as the injection is made more resonant. Gain measurements of the control sample N123 are shown in Fig. 12.19c. As expected, the shape of the gain curve is almost symmetric for the whole range of injected current. For a given injection current density, the maximum gain is also larger by a factor of 4. Such a behavior is also predicted by the theoretical model.

The dispersive shape of the gain spectrum, and its depedence on the active region design, are a clear demonstration of the presence of Bloch gain in the low-inversion sample **(p.242)**

.

12.2.5 Resonant tunneling diodes and quantum cascade lasers

The density matrix approach derived above can be used to shed a new light on the relationship between resonant tunneling diodes and quantum cascade lasers [285].

Traditionally, resonant tunneling diode oscillators and quantum cascade lasers have been described as being of a very different nature: the resonant tunneling diode is an electronic device, limited by transit time, using an electronic resonator for

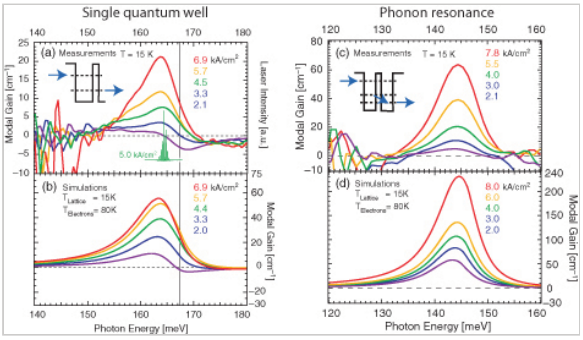


Fig. 12.19. Gain measurements showing evidence for Bloch gain in samples with a low population inversion. Left panels, results for Bloch gain sample (single-quantum well, design described in Appendix A.1.5). (a) Net modal gain measurements and (b) simulations. Measurements achieved for increasing current densities in the amplifier section: 2.1, 3.3, 4.5, 5.7, and 6.9 kA/cm^2. The vertical line represents the transition energy computed from band structure

oscillation. In contrast, the gain in a quantum cascade laser is described as arising from transition between subband states that sustain an oscillation in an optical Fabry–Perot-like cavity.

between subband edges ($k = 0$ states) of the active well. Shown also in (a) is the spectrum of device N258 operated as a laser. Right panels, results for a reference sample (single phonon resonance, design described in Appendix A.1.3). (a) Net modal gain measurements and (b) simulations. Measurements achieved for increasing current densities in the amplifier section: 2.1, 3.0, 4.0, 5.5, and 7.8 kA/cm^2. Adapted from [265].

The only fundamental difference between a quantum cascade laser and a resonant tunneling diode is the existence of an injector in the former that allows active-region periods to be cascaded while maintaining the electrical stability of the structure [20]. The presence of an injector biased in the regime of positive differential resistance, however, automatically induces a loss at low frequencies. The comparison between the schematic band structure of a resonant tunneling diode and that of the simplest quantum cascade laser is presented in Fig 12.20. This specific terahertz quantum cascade laser design is based on the recently demonstrated two-quantum-well structure [194, 193].

(p.243)

As shown in Fig. 12.20, transport and gain in this structure involves only three states per period: the optical transition proceeds across a tunnel barrier from the ground state of a thin quantum well to the first excited state of a wider well, whose transition is resonant with an optical phonon for fast depopulation and whose lower state serves as an injector for the next period.

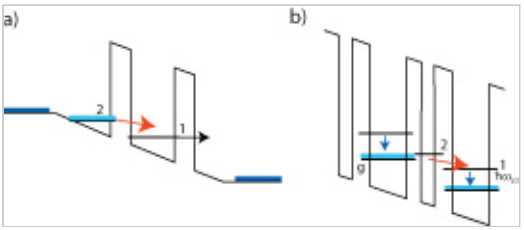


Fig. 12.20. Schematic band structure of (a) a resonant tunneling diode and (b) a quantum cascade laser based on a two-quantum-well active region.

Whilst negative conductance in the resonant tunneling diode is ascribed to the nature of the non-resonant transport between the subband of the injector and the confined state, the gain in the quantum cascade laser is described by a population inversion between states 2 and 1. Both processes can be described, however, by the same density matrix formalism discussed in this chapter. The two subbands 2 and 1, detuned by an energy $\hbar\Delta$, interact through a tunnel barrier with a coupling energy $\hbar\Omega$, while the states suffer a dephasing characterized by a total energy broadening $\hbar\gamma$.

Applying eqn. 12.2.47 to the case of a resonant tunneling diode; assuming no population of the lower state ($n_1 = 0$), the current between the subbands can be written as

$$j(\Delta) = \frac{2e\hbar\Omega^2\gamma}{\Delta^2 + \gamma^2} n_2$$

(12.2.59)

For a device of active area A and length L_{act}, the negative differential conductance is computed by simply differentiating the above equation with respect to Δ, and yields

$$\begin{aligned} G_d &= A\frac{\delta j}{\delta \Delta}\frac{\delta \Delta}{\delta V} \\ &= -A\frac{e^2\Omega^2 4\gamma\Delta}{(\Delta^2+\gamma^2)^2}\frac{d}{L_{act}}n_2 \end{aligned}$$

(12.2.60)

where d is the spatial separation between the upper and lower state, L_{per} is the total length of the RTD, and $\frac{\partial \Delta}{\partial V} = ed/L_{act}$ is the fraction of the electric field dropped across the RTD. We can rewrite the expression as a function of the maximum current at resonance $j^{(max)}$:

$$j^{(max)} = \frac{2e\hbar\Omega^2 n_2}{\gamma}$$

(12.2.61)

so that

$$G_d = -Aj^{(max)}\frac{e\gamma^2 2\Delta}{\hbar(\Delta^2+\gamma^2)^2}\frac{d}{L_{act}}$$

(12.2.62)

(p.244) The above expression has its maximum at approximately $\Delta \approx \gamma/2$:

$$G_d^{(max)} = -Aj^{(max)}\frac{16}{25}\frac{e}{\hbar\gamma}\frac{d}{L_{act}}$$

(12.2.63)

For a resonant tunneling diode, the oscillation condition for a resonant circuit at an angular frequency ω with a quality factor Q is

$$G_d = -\frac{\omega C}{Q}$$

(12.2.64)

so that the maximum current, assuming the capacitance is dominated by the resonant tunneling diode (i.e. $C = \epsilon\epsilon_0 A/L_{act}$):

$$j^{(max)} = \frac{25}{16}\frac{\hbar\epsilon\epsilon_0}{e}\frac{\omega\gamma}{dQ}$$

(12.2.65)

The latter equation, while much too crude to allow an accurate prediction of the resonant current necessary to achieve oscillation, nevertheless describes the key factors entering in the oscillation condition and a correct order of magnitude value ($j^{(max)}$ = 200 kA/cm^2 at 1 THz) for the current densities using typical parameters of [286]. In particular, the current density is expected to increase linearly as a function of operation frequency, as the quality factor Q of electronic resonators tends to remain constant with frequency.

However, in the quantum picture the gain between an upper $n = 2$ state and a lower state $n = 1$ can be described by the scattering-assisted gain formula, written for classical distribution as eqn. 12.2.48. As we are interested in the regime of low frequencies, it is important to keep the non-resonant term in the expression for the gain. This becomes:

$$\alpha(\omega) = \frac{e^2 d^2 |\Omega_{21}|^2}{\epsilon_0 n_r c \omega} \times$$
$$\times \left[\frac{\theta(\Delta-\omega)\left(n_2 - n_1 e^{\frac{-\hbar|(\Delta-\omega)|}{kT}} \right) + \theta(\omega-\Delta)\left(n_2 e^{\frac{-\hbar|(\Delta-\omega)|}{kT}} - n_1 \right)}{(\omega-\Delta)^2+\gamma^2} + \right.$$
$$\left. + \frac{n_2 - n_1 e^{-\frac{\Delta+\hbar\omega}{kT}}}{(\omega+\Delta)^2+\gamma^2} \right]$$

(12.2.66)

The above expression is compared with that derived from the differential conductance, and using the relationship between gain and conductance, eqn. 12.2.53 yields

$$\alpha = \frac{d^2}{\epsilon_0 n c} \frac{4\epsilon^2 \hbar \Omega^2 \gamma \Delta}{(\Delta^2+\gamma^2)^2} \frac{1}{L_{act}} n_2$$

(12.2.67)

in the limit of zero ground-state population. Both equations are compared in Fig. 12.21. As expected, the value of the ac gain given by eqn. 12.2.67 converges towards the value given by the conductance when $\omega \rightarrow 0$.

The resonant response peaking at $\omega \approx \Delta$ in Fig 12.21 is not characteristic of the operation characteristics of a resonant tunneling diode. In these devices the broadening **(p.245)**

is large and the maximum gain is achieved close to $\omega = 0$ for a detuning given by the condition already mentioned, namely $\Delta = \gamma/2$. The positive gain (or equivalently negative differential conductance) exhibited by any inverted subband system extends down to zero frequency, as shown clearly in Fig. 12.21. As a result, the structure is naturally unstable electrically and is impossible to cascade, as it will break into high field domains. To prevent this unfavorable behavior, quantum cascade lasers incorporate injectors that are in effect resonant tunneling diodes biased before their roll-over points, such that the overall differential conductance of the system

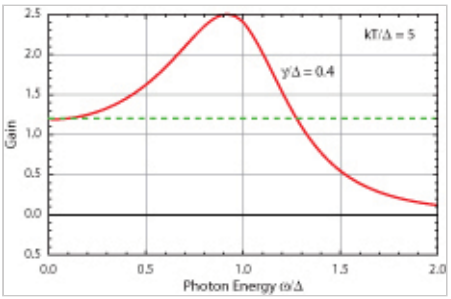


Fig. 12.21. Comparison between the gain computed directly (solid line) and deduced from the computation of the differential conductance (dashed line).

$$G_{tot} = G_{inj}(\omega = 0) + G_{act}(\omega = 0) \geqslant 0$$

(12.2.68)

remains positive. Of course, the desired positive resistance of the injector at low frequency will also extend to higher frequencies and therefore quench the gain. In addition, the continuity of the current requires that

$$J_{inj} = J_{act}$$

(12.2.69)

This in turn will require, in a standard quantum cascade laser, that the population in the injector is greater than the upper-state population $n_g \geqslant n_{up}$, further increasing the absorption.

This effect is clearly apparent in a computation of the absorption in a three-level structure shown in Fig. 12.22. The parameters ($n_g = 2n_2$; $n_1 = 0$; $\Omega_{inj} = \Omega_{act}$) were choosen such that the current continuity condition (eqn. 12.2.69) is met. In this example, reabsorption from the injector strongly decreases the gain.

Injectors in conventional quantum cascade lasers have both the function of selective injection of carriers into the active region and maintaining electrical stability. Distributing the two functionalities may actually solve some of the problems raised by the injector reabsorption. For example, in so-called *scattering-assisted* injectors, the electrons are actually captured by the upper-state. This approach was explored initially in the mid-infrared [287], and is presently the subject of a renewed interest [288]. In this case the electrical stabilization of the structure is made by the shape of the injector. Another intriguing possibility that was recently investigated theoretically [289] is to stabilize the structure using the lower state of the laser transition.

(p.246)

The inset of Fig. 12.22 summarizes the concept: an additional injector is placed within the structure, in such a way as to inject it into a state with a very fast lifetime. In this way the population of the injector can be reduced while keeping the same current. As shown by the computation shown by dashed lines in Fig. 12.22, such an approach enables the reduction of the reabsorption at the laser wavelength. However, note that it also (very naturally) decreases the positive conductance of the injector, so that the stability of the structure is also decreased. An even more promising approach would be to insert a "slow and narrow" injector with a long dephasing time and therefore a narrow broadening γ. This could be achieved using transfers to other valley or quantum dot-based injectors.

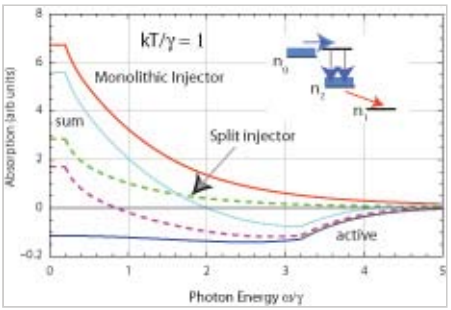


Fig. 12.22. Plot of the absorption contributions from the active region (showing gain, lower solid line) and the injector (upper solid line), as well as the sum of both contributions (solid intermediate line). By comparison, the contribution of an injector that is separated from the active region is shown by the upper dashed line, and the absorption/gain of the total structure by the lower dashed line. Note the stronger gain for the separated injector structure. Inset: schematic drawing of the level scheme for a active region separated from the injector region

12.3 Full density matrix models

The density matrix approach, presented above for only a pair or a few states has been successfully used to interpret and predict the transport and gain in devices for both the terahertz range [183, 198, 290, 291] and the mid-infrared [152, 277]. Generalization to a larger number of states for mid-infrared quantum cascade lasers are more recent [292, 293].

Formally, the Liouville time-evolution eqn. 12.2.19 is now written for all the states of the active region. The choice of a basis of eigenstates on which the transport can be computed is crucial. Indeed the Fermi golden rule will hold only if the energyspacing between the eigenstates ΔE is much larger than the broadening γ of the states themselves. For devices with a relatively thick active-region period comprising a large number of wells, this requirement suggests the use of a basis that spans only a portion of the structure.

An example of such a basis is shown in Fig. 12.23, where the active-region period is divided into two sub-periods: the active wells region that includes the laser doublet and the two depletion states, and the injector region designed to ensure the extraction **(p.247)** and the relaxation of the electrons to the next stage. At the splitting barriers the transport is modelled by sequential resonant tunneling.

It has been shown [72] that the Liouville equation can be rewritten as

$$\mathscr{L}\cdot\rho=(\mathscr{C}+\mathscr{R}+\mathscr{P})\cdot\rho$$

(12.3.70)
where the operators

$$C_{ij,kl}=\mathscr{H}_{ik}\delta_{jl}-\mathscr{H}_{jl}^{*}\delta_{ik}$$

(12.3.71)
are the correction to the Hamiltonian introduced by the coherent coupling between states. It is written as

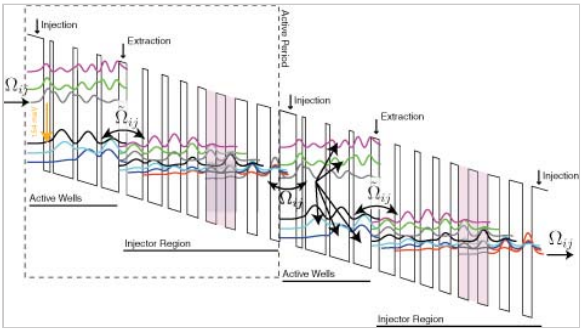


Fig. 12.23. Two active periods shown at the injection resonance (44 kV/cm) of a twophonon resonance active region, shown in Appendix A.2.4. The coupling between the subperiods, achieved by the resonant tunneling (coupling energy matrix Ω for injection, Ω˜ for extraction), is shown through injection/extraction barriers. The rate equations from upper laser state are represented in the active wells of the second period. The grey shaded layers are doped. From [293].

$$\mathscr{C}_{ij,kl}=\hbar\Omega_{ik}\delta_{jl}-\hbar\Omega_{jl}\delta_{ik}\text{with}i\neq k,j\neq l$$
$$\mathscr{C}_{ij,ij}=\hbar\Delta_{ij}$$

where $\hbar\Omega_{ij}$ is the coupling and Δ_{ij} is the detuning between state i and state j. The population decay operator

$$R_{ij,kl}=i\hbar\mathscr{W}_{ik}\delta_{ij}\delta_{kl}$$

(12.3.72)
is written as a function of the scattering rate W_{ik} from state k to state i. The loss of coherence is described by the super-operator

$$\mathscr{P}_{ij,kl}=-i\hbar\tau_{\parallel,ij}^{-1}\delta_{ik}\delta_{jl}$$

(12.3.73)

(p.248) Finally, eqn. 12.3.70 can be summarized as two main equations for the populations $\rho_{ii} = n_i$ and for the coherences ρ_{ij}:

$$\dot{n_i} = \sum_k W_{ik} n_k + \sum_{k \neq i} i\Omega_{ik}(\rho_{ik} - \rho_{ki})$$

(12.3.74)

$$\dot{\rho_{ij}} = i\Omega_{ij}(\sigma_{ij} n_i - \sigma_{ji} n_j) - (i\Delta_{ij} + \tau_{||,ij}^{-1})\rho_{ij}$$

(12.3.75)
These equations can be solved in steady-state $\dot{n}_i = \dot{\rho}_{ij} = 0$, and after some algebraic manipulations the equation of motion can be cast into the form:

$$0 = \sum_k \mathscr{W}_{ik} n_k \text{where} \mathscr{W}_{ik} = W_{ik} + R_{ik}$$

(12.3.76)
W_{ik} equals to the scattering rate from state k to state i, it sums all the elastic and inelastic scattering channels discussed in Chapter 5. The resonant-tunneling contribution to the rate equations reads

$$R_{ik} = \frac{2\Omega_{ik}^2 \tau_{||,ik}}{1 + \Delta_{ik}^2 \tau_{||,ik}^2} \sigma_{ki}$$

(12.3.77)
The effective parameters σ_{ij} account for electrons tunneling at a constant energy rather than at a constant wavevector which effectively implements the second-order scattering in the current calculation. Namely, the parameter σ_{ki} has the classical distribution considered here:

$$\sigma_{ki} = \theta(\Delta_{ki}) + \theta(-\Delta_{ki}) e^{-\beta\hbar|\Delta_{ki}|}$$

(12.3.78)
as in eqn. 12.2.48.

According to eqn. 5.0.3, the dephasing time, which relates the loss of the phase correlation between two states, is computed and has two contributions:

$$\tau_{\|,ij}^{-1} = \frac{1}{2}\left(\tau_i^{-1} + \tau_j^{-1}\right) + \frac{1}{2}\frac{1}{\hbar}\Gamma_{intra}^{IFR}$$

(12.3.79)
The first derives from intersubband transitions, when an electron is scattered from a state of the resonant doublet to another subband. The second Γ_{intra}^{IFR} is the pure phase contribution identified as intrasubband broadening due to interface roughness scattering as expressed in eqn. 5.3.25. We have evaluated Γ_{intra} at the characteristic energy of the electron distribution $k_B T$, where k_B is the Boltzmann constant and T the electronic temperature.

Addition of the optical field. As discussed already in Section 12.1.2, the optical flux has a significant influence on the electronic transport. It is therefore desirable to introduce its effect in the transport using the density matrix formalism. The simplest approach is to introduce it as an additional scattering channel in a rate equation picture. In such an approximation, if we

consider for simplicity two levels *i* and *j* in a sub-period, their equation of motion is affected by the simulated emission and the absorption as

$$\dot{n}_i = (\ldots) - g_{ij}(\hbar\omega_0)S + g_{ji}(\hbar\omega_0)S$$

(12.3.80)

$$\dot{n}_j = (\ldots) + g_{ij}(\hbar\omega_0)S - g_{ji}(\hbar\omega_0)S$$

(12.3.81)

where g_{ij} is the total net gain between level *i* and level *j*. We assume, then, that g_{ij} is a linear functional in the population distributions $f^i(k)$ of the subbands, as this is **(p.249)** the case for the first-order eqn. 4.4.21 and the second-order gain formula eqn. 12.2.54. The gain between upper subband *i* and lower subband *j* can therefore be rewritten using the gain cross sections $g_{ij}^{c,i}$ for states *i* $g_{ij}^{c,j}$ and *j* respectively:

$$g_{ij} = g_{ij}^{c,i} n_i + g_{ij}^{c,j} n_j$$

(12.3.82)

This linear form is particularly suited for rate equations, and we can write

$$\begin{pmatrix} \dot{n}_i \\ \dot{n}_j \end{pmatrix} = (\ldots) + S \begin{pmatrix} -g_{ij}^{c,i} + g_{ji}^{c,i} & g_{ij}^{c,i} - g_{ji}^{c,i} \\ -g_{ij}^{c,j} + g_{ji}^{c,j} & g_{ij}^{c,j} - g_{ji}^{c,j} \end{pmatrix} \cdot \begin{pmatrix} n_i \\ n_j \end{pmatrix}$$

(12.3.83)

If we make the assumption that $g_{ij}^{c,i}, g_{ij}^{c,j}, g_{ji}^{c,i}$ and $g_{ji}^{c,j}$ are computed for an initial population n_i^0 and n_j^0, we can estimate the deviation of the initial populations, due to the photon flux density, by solving these rate equations. In particular, if this system is embeded in a self-consistent routine on the populations, it will converge to the exact solution. The main advantage of this technique is to avoid an inhomogeneous term in the global rate equations. Practically, the gain curve is first computed with populations found with non-radiative processes only. If we decide to let the laser work at the peak gain g_p, given optical losses α for the empty waveguide (i.e. with no active region loss) and for the mirrors, we obtain a threshold condition

$$\Gamma g_p - \alpha = 0$$

(12.3.84)

where Γ is the modal overlap and

$$g_p = \sum_{i>j} g_{ij}$$

(12.3.85)

is the modal gain, summing all the intersubband contributions (intersubband absorption *and* gain). We then match condition 12.3.84 by varying *S*.

The weakness of this rate-equation approach is that it does not include the effect of the coherent coupling directly onto the gain profile, but does it only through the population. In particular, it fails to predict the broadening of the gain curve arising from a too-strong coupling of the upper-state level to the injector through the injection barrier. A more accurate approach is to introduce

the coupling to the light directly into the Liouville operator, as was done for a few levels in the terahertz in [198] and in the mid-infrared in [152].

Comparison with experiment. As an example, we apply the model to a reference two-phonon quantum cascade laser processed in a buried heterostructure [62]. For the simulation we have taken typical parameters for the various scattering mechanisms; the interface roughness is modeled with a value of $\Lambda = 90$ Å for the correlation length of the steps, $\Delta = 1.2$ Å for the step height, and $\kappa = 15$ Å for the correlation length between interfaces as obtained from previous experiments with different $Ga_{0.47}In_{0.53}As$/ $Al_{0.48}In_{0.52}As$ heterostructures [61, 266].

As this model does not solve directly the distribution in *k*-space, two different approaches were compared. In the first one, a very strong intrasubband electron– electron scattering was assumed that brought all the subbands to the same electronic **(p.250)** temperature using eqn. 12.1.14. In the other approach we assumed that the minority of hot carriers would cool fast enough not to heat the electrons inside the injector levels. This result is shown by the simple *k*-resolved rate equation shown in Fig. 12.6 that predicted a negligible electron heating (〈 10 K) for the lower states of the active region at J =3 kA/cm^2.

The two simulated current–voltage curves are compared in Fig. 12.24. A very good agreement is obtained when assuming that the electrons are at the temperature of the lattice. In contrast, the global kinetic balance model is too pessimistic. Note that the excellent agreement between the model and the experiment is achieved in the virtual absence of the fitting parameter: the doping was the nominal one from MBE calibration, and an additional waveguide loss $\alpha_{w,scatt}$ due to scattering from the sidewalls was assumed. The active-region loss is computed implicitly by eqn. 12.3.85, while a Drude model was assumed for the waveguide cladding layers.

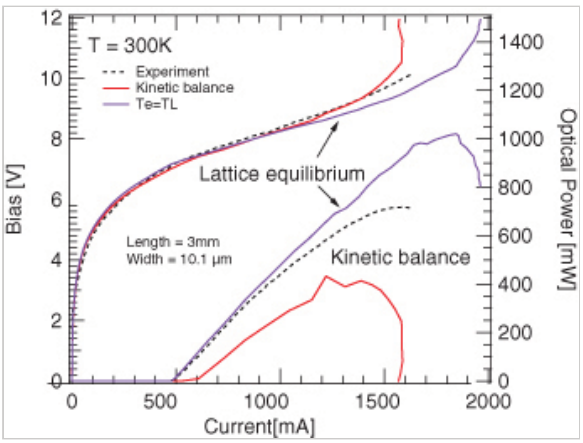


Fig. 12.24. Simulated light–current–voltage curves (full lines) for the kinetic balance model and the electron-lattice thermalization versus experiment (dashed line) at 300 K. The design described in Appendix A.2.4. The losses were measured to 5 cm^{-1}, and the modal overlap is $\Gamma \approx 0.67$. From [293].

The kinetic balance probably overestimates the electronic temperature of most subbands: at the injection resonance (J = 5.2 kA/cm^2) with a lattice temperature of 300 K, the electronic temperature is computed near 600 K. This can be explained by examining the laser transition more closely. In both cases of an electron injected in the upper laser state with no excess of kinetic energy that undergo a non-radiative scattering, optical phonon emission or elastic scattering will end up in the lower subband with a high in-plane momentum. The model will then try to represent these hot electrons with a thermal distribution, causing a strong heating of all the subbands. **(p.251)** The overestimation of the electronic temperature of the injector subbands generates a backfilling in the active wells that account for the strong threshold current increase.

In fact, we can estimate the average time $\bar{\tau}$ spent by an electron in the activewells region. For an applied electric field of 34 kV/cm in the middle of the dynamic range, we compute the sum of the carrier density in the subbands of the active-wells region $n_a \approx 0.26 \cdot 10^{11}$ and the corresponding current density $J \approx 2.6$ kA/cm^2. Using $J = en_a/2\bar{\tau}$ we have $\bar{\tau} \approx 1.6$ ps. During this time the electron can emit approximatively 8 optical phonons [294, 295], using 200 fs as the emission rate of bulk optical phonons. This allows the electron to lose ≈ 256 meV of kinetic energy, largely covering he energy gap formed by the optical transition and the phonon resonances designed between the lower states of the active wells. The majority of the electrons have therefore no excess of kinetic energy when they are extracted to the injector region. In the injector itself the transport is achieved by direct scattering between a dense collection of states that allows efficient thermalization.

We now test the model on a wide range of temperatures. Fig. 12.25a shows the light– current curves at $T = 240$ K and 400 K, showing again a very good agreement for the threshold current density and slope efficiency. The temperature dependence of the threshold current density is shown in Fig. 12.25b, and is fitted by the usual exponential dependence $I = I_0 e^{T/T_0}$. The value extracted from measurements is $T_0 = 174$ K. The simulated value with the kinetic balance model is $T_0 = 111$ K, while the result for equal lattice and electron temperatures is $T_0 = 155$ K, in better agreement with the experiment.

A different situation is, however, observed in strained compensated structures where the electronic temperature was measured clearly above the lattice temperature [244]. The simplification made here no longer holds, as the energy gap between the upper laser state and the injector ground state is much larger.

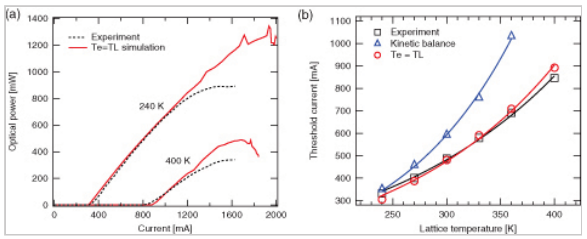


Fig. 12.25. (a) Comparison between theory and experiment of light–current curves for the two extreme temperatures. (b) Threshold current versus lattice temperature. Simulated values are represented with circles for the electron-lattice thermalization, with triangles for the kinetic balance model and with squares for the experiment. The solid lines are fits with an exponential dependence. From [293].

(p.252) The validation of this model enables us to predict quantities that are not accessible directly by experiment, like the injection efficiency. We can readily compute it as $\eta_{inj} = J_{upper}/J_{total}$, where J_{upper} is the current flowing into the upper laser state from the injector region, and J_{total} the net current through the injection barrier. We find a value of $\eta_{inj} \approx 0.86$.

12.4 Monte Carlo

In the Monte Carlo approach, the time evolution of the distribution function $f_{k\alpha}$ is computed using a Boltzman-like equation:

$$\frac{d}{dt} f_{\mathbf{k}\alpha} = \sum_{\mathbf{k}',\alpha'} \sum_{s} \left[P^{s}_{\mathbf{k}\alpha,\mathbf{k}'\alpha'} f_{\mathbf{k}'\alpha'} - P^{s}_{\mathbf{k}'\alpha',\mathbf{k}\alpha} f_{\mathbf{k}\alpha} \right]$$

(12.4.86)

Where $P^{s}_{\mathbf{k}\alpha,\mathbf{k}'\alpha'}$ is the scattering probability for an electron belonging to subband α and in-plane wavector **k** to scatter to state α' and in-plane wavector **k′**. This approach was introduced in quantum cascade lasers by Rossi and Iotti [296]. It can be used to compute the full, *k*-resolved electron distribution. The Boltzman-like eqn. 12.4.86 reflects the balance between probabilities of incoming and outgoing electrons in the eigenstates of one period. In this respect it can be interpreted as a rewriting of a rateequation model 12.1.3 with **k** space resolution, and suffers from the same limitations. The introduction of quantum coherences requires a generalization of eqn. 12.4.86 to a density-matrix formalism, as discussed by Iotti and Rossi [269].

An ensemble Monte Carlo has been used recently to compute the characteristics of quantum cascade lasers operating at cryogenic temperatures and presenting very high wallplug efficiencies [150], and in which, therefore, photon emission is an important transport channel. In Fig. 12.26 the computed light intensity and the wallplug efficiency is compared to the measured values [297]. Due to this ensemble Monte Carlo technique, the authors also demonstrated the effect of the photon transport on the *k*-distribution of the electrons.

12.5 Non-equilibrium Green's function

Non-equilibrium Green's functions allow for a consistent combination of scattering and coherent evolution, and are subject to a growing interest. This approach is based on the seminal work of Kadanoff and Baym [298] and Keldysh [299]. The first application of Non-equilibrium Green's functions to the transport in QCL is due to Wacker [268,300]. The reader is referred to these articles for a thorough presentation of this method, and here we will outline only some of the features of this technique.

In this approach [301], quantum-mechanical coherences are represented by offdiagonal elements of the $\mathbf{G}^{<}(E)$ correlation function, which is related to the density matrix by

$$\rho_{\alpha\mathbf{k},\beta\mathbf{k}} = \frac{1}{2\pi i}\int \mathbf{G}^{<}_{\alpha,\beta,\mathbf{k}}(E)dE$$

(12.5.87)

The theory is formulated with the usual envelope function basis states:

$$\Psi_{\alpha,\mathbf{k}}(\mathbf{r},z) = \frac{1}{\sqrt{S}}e^{i\mathbf{kr}}\chi_{\alpha,\mathbf{k}}(z)$$

(12.5.88)

(p.253)

Here $\chi_{\alpha,\mathbf{k}}(z)$ is the envelope function in the growth direction z. The wavevector **k** and the spatial coordinate **r** are two-dimensional vectors in the plane of each semiconductor layer with normalization area *S*. A numerical challenge is that the correlation function can only be solved self-consistently with the self-energies for each scattering process. This leads to a very numerically intensive

algorithm. As an example, the self-energies and correlations for the interface roughness scattering $\mathbf{\Sigma}^{<,rough}_{\alpha_1,\alpha_1,\mathbf{k}}$ will have to satisfy, in **(p.254)** the Born approximation, the equation

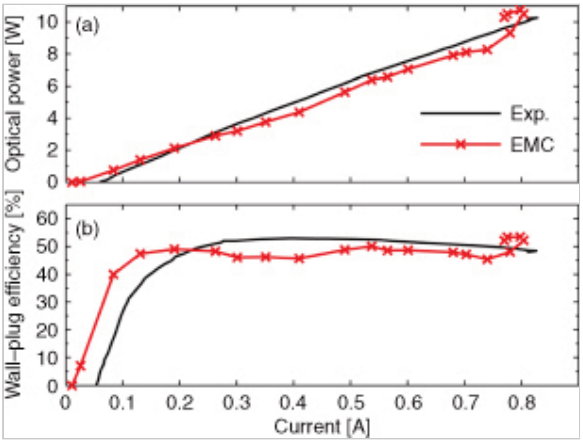


Fig. 12.26. Comparison between the computed and measured light-current (a) and wallplugcurrent (b) characteristic for a high efficiency cryogenic QCL [150]. The computations are performed using an ensemble Monte Carlo technique. Reprinted with permission from [297]. Copyright 2011, AIP.

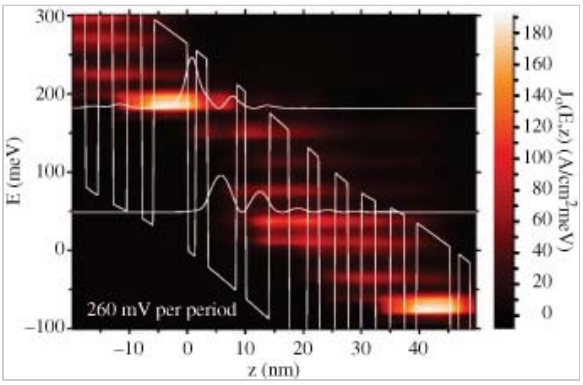


Fig. 12.27. Plot of the spatiallyand energy-resolved current J_0 for a mid-infrared $Al_{0.45}Ga_{0.55}As$ quantum cascade laser. Reprinted with permission from [301]. Copyright 2006, APS.

$$\mathbf{\Sigma}^{<,rough}_{\alpha_1,\alpha_1,\mathbf{k}} = \sum_{\beta,\mathbf{k}'} \langle |V^{rough}_{\alpha_1,\beta}(\mathbf{k}-\mathbf{k}')|^2 \rangle G^{<}_{\beta\beta,\mathbf{k}'}(E)$$

(12.5.89)

and the Keldysh equation

$$G^{<}_{\beta\beta,\mathbf{k}}(E) = \sum_{\beta,\beta'} G^{ret}_{\alpha_1,\beta,\mathbf{k}}(E)\mathbf{\Sigma}^{<}_{\beta,\beta',\mathbf{k}} G^{adv}_{\beta',\alpha_2,\mathbf{k}}(E)$$

(12.5.90)

The knowledge of the correlation function in turn enables us to compute the global quantities such as the energy-resolved current density [301]:

$$J_0(E,z) = \frac{e}{\pi S}\sum_{\alpha,\beta,\mathbf{k}} Re\left[\frac{-\hbar}{m(z)}\chi^*_\alpha(z)\frac{\partial\chi_\beta(z)}{\partial z}G^<_{\beta,\alpha,\mathbf{k}}(E)\right]$$

(12.5.91)

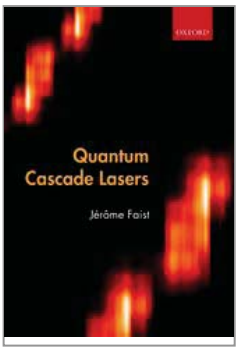
A plot of $J_0(E, z)$, given by eqn. 12.5.91 is shown in Fig. 12.27, computed for a midinfrared $Al_{0.45}Ga_{0.55}As$ quantum cascade laser at 77 K [301]. Note that the current escape from the active region occuring at high kinetic energy for both lower subbands. Such a plot yield precious information on the current path across the structure.

Knowledge of the correlation function also enables the computation of physically relevant quantities such as the total current or the optical gain. In fact, the scatteringassisted or Bloch gain between subbands, derived using the density matrix as shown in Section 12.2.3, was simultaneously derived using the non-equilibrium Green's function approach [300] and shown to be naturally included in it.

Notes:

([1]) position must be handled carefully; a cleaner derivation is provided in[72].

([2]) This result can also be obtained by expanding the Hamiltonian in powers of the field strength and identifying the first order.



Quantum Cascade Lasers
Jérôme Faist

Print publication date: 2013
Print ISBN-13: 9780198528241
Published to Oxford Scholarship Online: May 2013
DOI: 10.1093/acprof:oso/9780198528241.001.0001

Dynamical properties

Jérôme Faist

DOI:10.1093/acprof:oso/9780198528241.003.0013

Abstract and Keywords

This chapter analyzes the dynamical behavior of quantum cascade lasers. It discusses high frequency modulation and multimode instabilities.

Keywords: quantum cascade lasers, high frequency modulation, multimode instabilities, dynamical behavior

13.1 High-frequency modulation

Because of the specific properties of intersubband transitions, dynamical properties of quantum cascade lasers are expected to be very different from those of interband lasers. The first and foremost feature of intersubband transitions is, as shown in Chapter 5, the very fast non-radiative intersubband scattering, proceeding on a timescale of 1 ps or less. As a result, and in strong contrast with interband transitions, the transition is limited by radiative processes only for high optical powers far above laser threshold. A first analysis of the dynamical behavior of quantum cascade lasers can be performed using time-dependent rate equations. Repeating, for convenience eqn.7.4.4, introducing the cavity photon lifetime:

$$\frac{dn_3}{dt} = \frac{J}{e} - \frac{n_3}{\tau_3} - S g_c (n_3 - n_2)$$

(13.1.1)

$$\frac{dn_2}{dt} = \frac{n_3}{\tau_{32}} + S g_c (n_3 - n_2) - \frac{n_2 - n_2^{\text{therm}}}{\tau_2}$$

(13.1.2)

$$\frac{dS}{dt} = \frac{c}{n_{eff}} [(g_c (n_3 - n_2) - \frac{n_{eff}}{c\tau_p}) S + \beta \frac{n_3}{\tau_{sp}}]$$

(13.1.3)

Since these equations do not have, in general, an analytical solution, a small signal analysis is commonly performed, exploring small harmonic variations of the timedependent quantities *J, S, n_i* around a steady-state operating point. One therefore writes

$$S(t) = S^{(0)} + \tilde{S} e^{i\omega t}$$

(13.1.4)

$$J(t) = J^{(0)} + \tilde{J} e^{i\omega t}$$

(13.1.5)

$$n_i(t) = n_i^{(0)} + \tilde{n}_i e^{i\omega t} (i = 3, 2)$$

(13.1.6)

and substitutes in eqn. 7.4.4, keeping only the terms in first order of the time-dependent quantities. The results are best understood when written in terms of the stimulated lifetime

$$\tau_{stim} = \frac{1}{g_c S^{(0)}}$$

(13.1.7)

In the limit of vanishing lower-state lifetime $\tau_2 \to 0$, and negligible backfilling $n_2^{\text{therm}} \to 0$, the modulus squared of the transfer function $h(\omega)$, defined as

(p.256)

$$\frac{\tilde{S}}{\tilde{J}} = h(\omega) \frac{S^{(0)}}{J^{(0)}}$$

(13.1.8)

has the following frequency dependence:

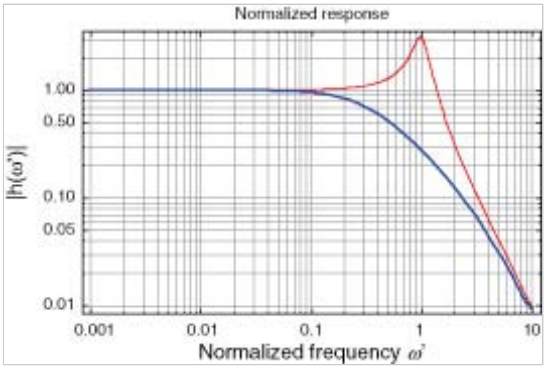


Fig. 13.1. Comparison between the small signal response of a quantum cascade laser (thick solid curve) and a semiconductor laser (thin solid curve) as a function of the normalized angular frequency ω'.

In both cases the stimulated lifetime is ten times the cavity lifetime. In the interband device the electron-hole recombination time is assumed to be a thousand times the photon lifetime.

$$|h(\omega)|^2 = \frac{1}{1+\omega^4\tau_p^2\tau_{stim}^2+\omega^2\tau_{stim}\tau_p\left(\frac{\tau_p}{\tau_{stim}}+2\frac{\tau_p}{\tau_{up}}+\frac{\tau_p\tau_{stim}}{\tau_{up}^2}-2\right)}$$

(13.1.9)

Eqn. 13.1.9 shows that the frequency response of quantum cascade lasers depends on the interplay between the three characteristic times $\tau_{up} = (\tau_{32}^{-1} + \tau_{31}^{-1})^{-1}$, τ_{stim}, and τ_p. The results can be best understood by rewriting the expression as a function of an effective frequency ω' such that

$$\omega'^2 = \omega^2\tau_{stim}\tau_p$$

(13.1.10)

which leads to the expression

$$|h(\omega')|^2 = \frac{1}{1+\omega'^4+\omega'^2\left(\frac{\tau_p}{\tau_{stim}}+2\frac{\tau_p}{\tau_{up}}+\frac{\tau_p\tau_{stim}}{\tau_{up}^2}-2\right)}$$

(13.1.11)

which is a simple second-order response function with a corner frequency defined by $\omega' = 1$. In interband semiconductor lasers, a similar expression can be derived but where the τ_{up} is the electron-hole recombination lifetime, much longer than the photon lifetime. As a result, for low optical powers ($\tau_{stim} \gg \tau_p$) the expression $\frac{\tau_p}{\tau_{stim}}+2\frac{\tau_p}{\tau_{up}}+\frac{\tau_p\tau_{stim}}{\tau_{up}^2}-2\langle 0$ and the response has a resonant response around $\omega' = 1$. Both cases are compared in Fig. 13.1. The maximum frequency response of a quantum cascade laser can be directly estimated by eqn. 13.1.9. For a typical optical cavity loss of $\alpha_{tot} = 10$ cm^{-1}, the photon cavity lifetime is about 10 ps. Assuming zero lower-state **(p.257)**

lifetime, the stimulated lifetime can be simply estimated by how much above the threshold current j_{th} the device is operated:

$$\tau_{stim}^{-1} = \left(\frac{j}{j_{th}} - 1\right)\tau_{up}^{-1}$$

(13.1.12)

In a typical device operated at room temperature, the upper-state lifetime is about $\tau_{up} = 0.6$ ps and the device can be driven at twice the

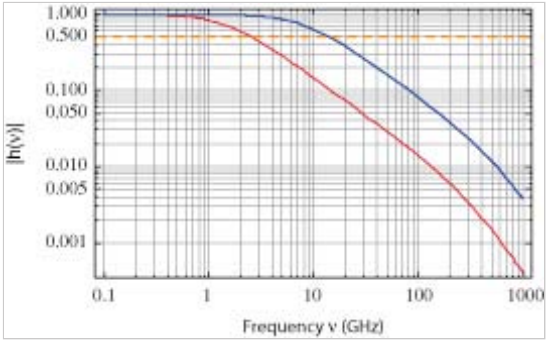


Fig. 13.2. Theoretical computed frequency response of a quantum cascade laser in the limit of no lower-state population. Photon lifetime is assumed to be $\tau_p = 10ps$. Lower curve, laser close to threshold with

threshold without reaching gain saturation, so that τ_{stim} = 0.6 ps.

$\tau_{stim} = 6ps$, thick upper curve the current is twice threshold and $\tau_{stim} = \tau_{up} = 0.6ps$. The dashed line corresponds to a -3dB response.

The frequency response, computed just above threshold and at twice the threshold, is compared in Fig. 13.2. The 3 dB modulation frequency bandpass extends up to about 11 GHz, with a significant modulation capability extending up to 100 GHz. The frequency bandwidth being proportional to the $1/\sqrt{\tau_{stim}\tau_p}$, an even wider frequency response can be obtained at the cost of larger optical losses and by operating even further above threshold. Indeed, very-high-frequency modulation 3 dB cutoffs above 50 GHz have been predicted for terahertz devices with large optical losses operating high above threshold.

An expression can also be derived in the case of a laser with a finite lower-state lifetime. As shown by its explicit expression, it does not contain n_{therm}:

$$h(\omega) = -\frac{\tau_{32}(\tau_2+\tau_{31})}{(\tau_2-\tau_{32})} \times (\tau_2(\tau_{32}\omega + i) - i\tau_{32})$$
$$\times[\tau_2\{i\tau_{31}\omega^2\tau_p\tau_{\text{stim}} + \omega\tau_{32}\tau_p(1+i\omega\tau_{\text{stim}}) + \tau_{31}\tau_{32}\omega(2+\omega\tau_p(-\omega\tau_{\text{stim}}+2i)) - i\tau_{32}\}$$
$$+\tau_{32}\omega\tau_p\tau_{\text{stim}} + \tau_{31}\{\omega\tau_p\tau_{\text{stim}} + \tau_{32}(\omega\tau_p(1+i\omega\tau_{\text{stim}}) - i)\}]^{-1}$$

(13.1.13)

The previous expression is compared to that derived for vanishing lower-state lifetime in Fig. 13.3 in a case where the lower-state lifetime is close to the upper one (τ_2 = 0.5 ps). Interestingly, the storage of electrons on the lower state is responsible for an enhancement of the high-frequency response at frequencies above 100 GHz.

(p.258)

Of course, the rate equations Eq 7.4.4 can be solved directly by numerical methods. As an example, the response to a double step current function is shown in Fig. 13.4. As expected, the dynamics of the populations, driven by the intersubband times, is much faster than that of the photon population, driven by a combination of the cavity and stimulated lifetimes. The very low value of the spontaneous emission rate, combined with a clamping of the non-saturated gain explains the relatively long turn on time of the laser, here shown to be about 400 ps.

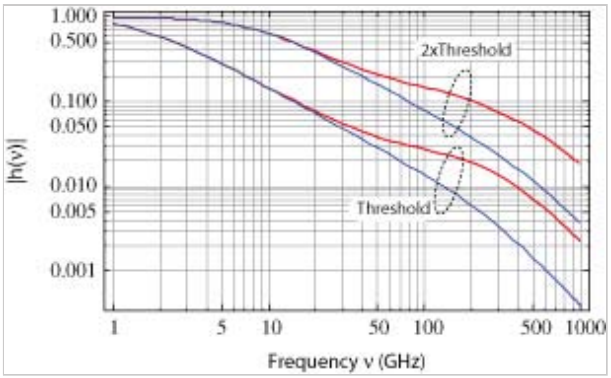


Fig. 13.3. Comparison between the theoretical computed frequency response of a quantum cascade laser in the limit of no lower-state population (thin solid line) and with the lowerstate lifetime relatively close to the upper one, τ_2 = 0.5 ps (thick solid line). Photon lifetime is assumed to be τ_p = 10 ps. Lower curves, laser close to threshold with τ_{stim} = 6 ps.

Upper curves, high optical field situation where $\tau_{stim} = \tau_{up} = 0.6$ ps.

13.1.1 Electrical response

The frequency response discussed in the previous paragraph completely neglects all the electrical aspects of the devices. Assuming one is able to remove all parasitic components, the fundamental limits of the frequency response can be estimated by the RC time product for an active region with a number of periods N_{per} of an area A:

$$RC = \frac{\Delta F N_{per} L_{per}}{\Delta J A} \frac{\epsilon\epsilon_0 A}{N_{per} L_{per}} = \frac{\epsilon\epsilon_0 \Delta F}{\Delta J}.$$

(13.1.14)

The normalized differential resistance $\frac{\Delta F}{\Delta J}$ should in principle be deduced from a model of the transport across the injector. Simulations and experiments show that, in midinfrared quantum cascade lasers, the equivalent of the whole dynamical current range (of $J_{max} = 3 kA/cm^2$ in typical high performance QCLs) is covered over a field range of 20 kV/cm^2. This number can be estimated as the ratio of the typical intersubband broadening parameter $\Gamma = 20$ meV divided by the distance between the injector and upper-state wavefunction (typically 10 nm). Using these numbers, an RC time constant of 7 ps, corresponding to a 3 dB cut-off frequency of $1/(2\pi RC) = 21 GHz$. This number could be increased by a larger J_{max}, obtained by higher active-region doping and a better injection design **(p.259)**

.

13.1.2 High-frequency modulation experiments

The first measurements of the fast optical response of quantum cascade lasers by an electrical drive were limited by parasitic capacitances and response time of the detectors [302]. These early measurements, performed at cryogenic temperatures, established unequivocally the absence of the relaxation resonance peak in the high-frequency response [303]. The use of multi-section devices was also investigated as a mean of increasing the frequency response of these devices [304].

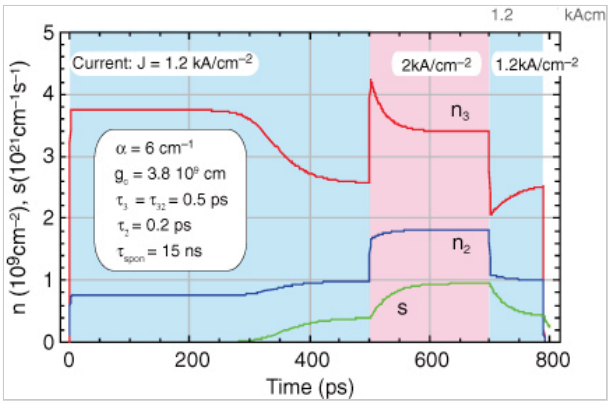


Fig. 13.4. Computed values of the upper and lower sheet electron densities n_3 and n_2, as well as the photon flux s per period for a quantum cascade laser driven by a current with stepwise changes. Grayed regions are the ones where a finite current is driven through the device, with the value indicated directly on the graph. The threshold current of the specific device is 1.02 kA/cm^2.

In the terahertz, the use of metal–metal waveguides is especially well suited for high-frequency modulation experiments, as the waveguide performs well in both microwave and terahertz frequency ranges [305]. Direct modulation has been performed and characterized at 13 GHz [306], and modulation at a frequency of 35 GHz has been used for injection-lock multimode operation [307].

The use of mid-infrared pump–probe experiments enable the study of tQCL ultrafast properties without suffering from the limitations imposed by the electrical circuits. In a series of experiments, the group led by T. Norris has studied the dynamical properties of quantum cascade lasers based on highly diagonal transitions [308], where photon-driven transport is especially efficient, as well as on traditional two-phonon designs [309]. The results achieved on the latter design are summarized in Fig. 13.5. The differential transmission as a function of time delay between pump and probe shows a dip corresponding to the saturation of the gain, followed by a recovery with two characteristic times: the fast one (occurring in a subpicosecond timescale) originates **(p.260)** from the active region dynamics, while the second one is attributed to the transport within the superlattice miniband and the injection process. Note that the latter was neglected in the simplified rate equation approaches shown above. As shown in Fig. 13.5b, both time constants decrease as a function of the increasing applied electric field. As expected, from the above rate-equation analysis, the onset of stimulated emission above threshold (here at I_{th} = 245 mA) will speed up the active region dynamics. Similarly, the transport in the injector miniband will also show a faster response as the field is increased, while being maintained below the onset of negative differential resistance. As shown in Fig. 13.5b, this time can be predicted reasonably well by a superlattice model with an Esaki–Tsu field-velocity relationship [309].

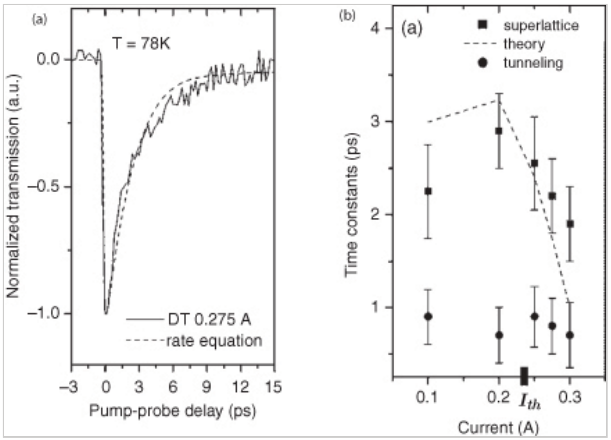


Fig. 13.5. a) Normalized differential transmission signal at 275 mA bias current in a QCL operating at $\lambda \approx 4.6\ \mu$ m. The dashed line is the rate-equation fit. b) Gain recovery time constants at various bias currents. Threshold is indicated as a black tick mark. The filled squares correspond to the superlattice relaxation; the dashed line is the calculated dielectric relaxation in superlattice, taking into account the inhomogeneous effective electric field at each bias current. The filled circles correspond to the resonant tunneling injection from the injector ground state and upper lasing state. Adapted with permission from [309]. Copyright 2008, AIP.

Note that the recovery time found at the largest bias, 2 ps, corresponds to a -3dB modulation limit of 70 GHz, demonstrating experimentally the modulation capabilities of QCLs **(p.261)** .

13.2 Multi-mode instabilities

As shown by the very broad tunability of external cavity devices (for example, in Section 10.3), especially those based on heterogenous cascades, the quantum cascade laser can exhibit very wide-gain bandwidth. It is therefore natural to ask whether such a device can be mode-locked. In fact, such a mode-locking would be extremely interesting for spectroscopie applications; the application of a very short pulse source in the mid-infrared could be numerous, for pump–probe spectroscopy of rotovibrational modes of molecules.

In a traditional mode-locked device, the cavity modes are equally spaced and their respective phases are the same (or have a trivial linear shift) [310]. As a result, in the time domain, the optical output consists of a train of short pulses spaced by the cavity round-trip time. In practice, actually, the requirements are often reversed: a saturable absorber favors (i.e. lowers the average loss for) the appearance of a pulse train, *ensuring therefore both the exact spacing of the modes as well as the correct phase factors*. Similarly, demonstration of mode-locked operation is based on the measurement and characterization, in the spectral and time domains, of the output pulses train.

For this mode-locked operation to be favorable compared to the continuous wave operation, the lifetime of the upper-state must be much longer than the cavity roundtrip time, ensuring that the energy of the inverted system can be stored until it is transferred to the pulse circulating in the cavity. Unfortunately, high-performance quantum cascade lasers operating at room temperature completely fail under that criteria, as the lifetime of the upper-state τ_{up} $\approx$ 0.5 ps is much shorter than the round-trip cavity time of $\tau_{rt} = 2nL/c$ = 42 ps for a 3 mm long cavity device. As a result, the pumping energy cannot be stored in the material during a cavity round-trip to feed a single optical pulse running in the cavity. Such a ratio of lifetimes is very uncommon among laser systems; for a typical solid-state laser the upper-state radiative lifetime is easily 103 longer than the cavity round-trip time. Of course, the upper-state lifetime of a quantum cascade laser can be engineered to much longer values by using a diagonal transition [67]; however, the use of such design parameters usually impact negatively on room-temperature device performances.

One should note, however, that *comb operation*, in which *equally* spaced modes are locked with an “arbitrary” phase shift, is also extremely interesting for its applications in spectroscopy. Actually, a replacement for the Fourier spectrometer has been proposed based on the operation of two such combs with slightly different repetition frequencies [311].

Coherent instabilities had been observed in quantum cascade lasers devices operated in continuous wave. As shown in Fig. 13.6, in relatively high-power devices, the spectrum would switch abruptly from single mode to highly multi-mode operation over a significant total bandwidth ($\rangle$ 50 cm^{-1}) [312]. In addition, photocurrent measurements on a very fast photodetector yielded a strong radio-frequency component at the frequency corresponding

to the round-trip frequency of the chip. The relatively high spectral purity of this beat tone lead the first groups to mistake them for evidence of pulse mode-locking [313] **(p.262)**

.

13.2.1 Origin of the multi-mode instabilities

The origin of these instabilities was discussed in detail in [312]. A key role is played by spatial hole burning. As compared to conventional semiconductor lasers, the quantum cascade laser operates in a different regime of parameters that in particular favors the effect of spatial hole burning in Fabry–Perot cavities. Intuitively, the standing wave created by the lasing mode burns holes in the electron upper-state density, reducing locally the gain, and favoring other longitudinal or transverse modes that exhibit a large field overlap with these high-gain regions. This effect is favored in QCLs because the upper-state is characterized by very low density and a relatively limited electron mobility due to the many interfaces. The in-plane diffusion time of the gain grating is written as

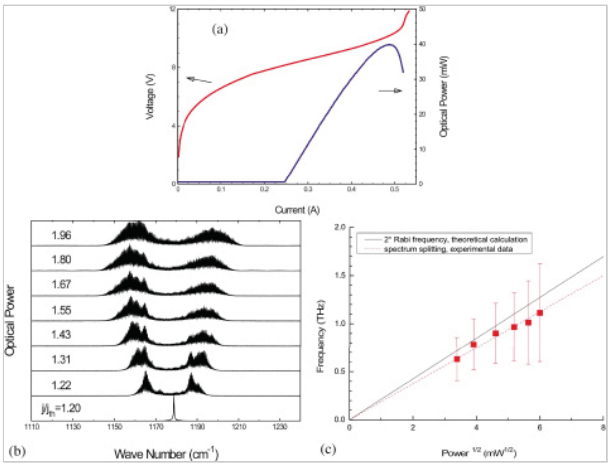


Fig. 13.6. a) V–I and L–I curves and b) optical spectra versus pumping ratio j/j_{th} above threshold obtained in continuous wave at 300 K with a 3 μ m wide ridge laser emitting at 8.38 μ m. c) Spectral splitting and twice the Rabi frequency $\Omega_{Rabi}/2\pi$ versus square root of output power collected from a single laser facet. The dashed line is a least-squares linear fit of the data. Reprinted with permission from [312]. Copyright 2008 by the APS.

$$\tau_{diff} = \frac{1}{k^2 D}$$

(13.2.15)

where D is the diffusion coefficient of the carriers and $k = 2\pi n_{refr}/\lambda$ the wavevector of the gain grating. We can estimate the diffusion constant D using Einstein's relation

(p.263)

$$D = \frac{\mu_e k_B T}{e}$$

(13.2.16)

where k_B is Boltzman's constant and T the temperature. At $T = 300$ K, and for an electron mobility of $3000 cm^2/V\,s$, the diffusion constant is $D = 77$ cm^2/s. Indeed, for a wavelength of $\lambda = 8$ μm the diffusion time $\tau_{diff} = 20$ ps is much longer than the upper-state lifetime τ_{up}

= 0.5 ps. As a result, spatial hole burning effects will be dominant but also very fast, as the lifetime of the gain grating is equal to the upper-state lifetime.

Another characteristic of quantum cascade lasers is their large Rabi frequency of the laser transition $\Omega_{Rabi} = z_{ud}E/\hbar$, due to the high optical powers circulating in the cavity and in the large value of the dipole matrix element z_{ud}. Using the Bloch equations for the upperand lower-state population, a Risken–Nummedal–Graham– Haken (RNGH) instability [314, 315] is predicted, in which a parametric gain appears, peaking at $\pm\Omega_{Rabi}$ off the center laser frequency. However, the assumption of the presence of a saturable absorber is necessary to lower the predicted threshold for this parametric gain from 10 times the threshold current, obtained without saturable absorber, to 1.5 times the threshold which is the typical value observed for the onset of multimode instabilities in quantum cascade lasers. The key evidence presented for the presence of the RNGH instability is the presence, in the laser spectra of some lasers, of two groups of modes whose frequency separation was equal to 2 Ω_{Rabi}. An example of such a situation is shown in Fig. 13.6. The respective strength of the spatial hole burning and of the RNGH instability remains unclear. One problem is that the nature of the saturable absorber effect was never fully clarified. One assumption was that it originated from a Kerr self-lensing that would decrease the losses originating from the interaction of the mode with the sidewalls in ridge lasers. Such an effect would appear mainly in narrow devices processed in the ridge process. However, similar behavior with two group of modes separated by about 40 cm^{-1} was also observed in buried heterostructure lasers [225], where lateral scattering from the waveguide sidewalls should be minimum.

Four wave mixing processes (degenerate and non-degenerate), driven by a nonlinear susceptibility $\chi^{(3)}$, will also favor the proliferation of modes, as they do in high-Q microcavities [316]. In fact, the non-linear interactions mentioned above can also be described by effective $\chi^{(3)}$ interactions.

13.2.2 Autocorrelation experiments

The low pulse energy, as well as the fact that the signal was in the mid-infrared made the typical pulse characterization using autocorrelation onto a non-linear crystal by detection of the second harmonic component unachievable. As shown in the inset in Fig. 13.7a, to characterize the output of the quantum cascade laser structures a different approach was used in which the beam, after passing through a Michelson interferometer, was focused onto a two-photon quantum-well infrared photoconductor [317]. In this detector, the photocurrent I_p^{2ph} is proportional to the fourth power of the beam's electric field $I_p^{2ph} = r^{(2ph)}|E_{inc}(t)|^4$.

Let us define the normalized quantity $\mathcal{J}(\tau)$, equal to the ratio of the photocurrent at the output of the interferometer compared to the sum of the photocurrent from **(p.264)** each arm. Assuming a 50:50 beam-splitter and a stationary process, we therefore have

$$\mathcal{J}(\tau) = \frac{\langle|E(t) + E(t+\tau)|^4\rangle}{\langle|E(t)|^4\rangle + \langle|E(t+\tau)|^4\rangle} = \frac{\langle|E(t) + E(t+\tau)|^4\rangle}{2\langle|E(t)|^4\rangle}$$

(13.2.17)
which can be expanded to

$$\mathcal{J}(\tau) = \frac{2\langle|E(t)|^4\rangle + 4\langle I(t)I(t+\tau)\rangle}{2\langle|E(t)|^4\rangle} + \frac{\langle(E(t)E^*(t+\tau))^2\rangle + \langle(E^*(t)E(t+\tau))^2\rangle + 4\langle I(t)+I(t+\tau))Re(E(t)E^*(t+\tau)\rangle}{2\langle|E(t)|^4\rangle}$$
$$= 1 + 2G_2(\tau) + \frac{\langle Re(G_1(\tau)^2)\rangle + 2\langle(I(t)+I(t+\tau))Re(G_1(\tau))\rangle}{2\langle|E(t)|^4\rangle}$$

(13.2.18)
where

$$G_2(\tau) = \frac{\langle I(t)I(t+\tau)\rangle}{\langle|E(t)|^4\rangle}$$

(13.2.19)
and

$$G_1(\tau) = E(t)E^*(t+\tau)$$

(13.2.20)
As a function of optical delay, the quantities containing $G_1(\tau)$ will average to zero over one optical period, as they all contain terms of the form $e^{i\omega\tau}$ or $e^{2i\omega\tau}$. As a result, a period-averaged measurement of $\mathcal{J}$ will yield $G_2(\tau)$, from where the value of the second-order correlation function $g^{(2)}(\tau)$ can be obtained, using

$$g^{(2)}(\tau) = \frac{\langle I(t)I(t+\tau)\rangle}{\langle I(t)\rangle^2} = \frac{G_2(\tau)}{G_2(\infty)}$$

(13.2.21)
assuming that the intensities are asymptotically uncorrelated with large time separation.

Using eqn. 13.2.18 enables the evaluation of the interferometric autocorrelation for various possible output laser pulses. Let us first consider the case of a pulsed modelocked laser where a single pulse of pulse length δt bounces back and forth in the cavity. In that case, for $\tau = 0$ it is obvious, using eqn. 13.2.17, that $\mathcal{J}(0) = 8$, while for $\tau \rangle\ \delta t$, inspection of eqn. 13.2.18 yields $\mathcal{J}(\tau) = 1$ as $G_2(\tau) = 0$. We then have the well-known result that

$$\mathcal{J}(0)/\mathcal{J}(\tau) = 8:1$$

(13.2.22)
The measurements of the interferometric autocorrelation performed in the devices discussed in detail in [312] are shown in Fig. 13.7, and yield a measured ratio $\mathcal{J}(0)/\mathcal{J}(\tau)$ between 8:2 and 8:3. This ratio unambiguously confirmed that these instabilities were not consistent with single-pulse mode-locking [312]. Nevertheless, the spectrum of the intensity of the same device, taken by focusing the laser output on a fast quantum-well infrared detector and measuring the electrical spectrum of the current using an RF spectrum analyser, is shown in the inset of Fig. 13.7. A beat note of 13 MHz FWHM linewidth is clearly visible at the round-trip frequency of 22 GHz for this 2 mm long device. In [312] a

number of experiments are reported in which the accurate shape of the interferometric autocorrelation line, as well as the width of the beat note, is varying as a function of device design, operation temperature, and ridge width **(p.265)**

.

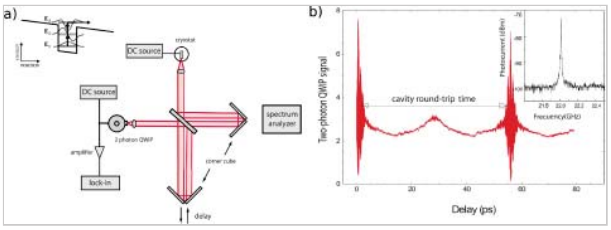


Fig. 13.7. a) The experimental set-up of a two-photon autocorrelation measurement. Inset: conduction-band diagram of the two-photon QWIP showing three equidistant energy levels. b) A second-order interferometric autocorrelation trace of an 8 μm wavelength ridge QC laser. Inset: microwave beat note spectrum of photocurrent generated by a similar laser under the condition of RNGH instability, showing a peak at the round-trip frequency. Reprinted with permission from [312]. Copyright (2008), APS.

13.2.3 Multi-mode with random phases

Given the fact that the interferometric autocorrelation signal is not the one expected from a pulse mode-locking source with its 8:1 ratio, it is interesting to look at a more generic frequency comb model. In the latter we assume that the non-linear coupling introduced by the spatial hole burning effect effectively enhances the width of the gain curve, bringing a set of N equidistant modes, spaced by $\Delta\omega$, above threshold. We further assume that these modes have a (complex) amplitude $E_k \exp(i\phi_k(t))$ where the phases $\phi_k(t)$ are completely random:

$$E(t) = \sum_k E_k \exp(i(\omega_0 + k\Delta\omega)t + \phi_k(t))$$

(13.2.23)

Beat note at the round-trip frequency.The intensity is then

$$I(t) = \langle E^*(t)E(t)\rangle = \sum_k \sum_l E_k E_l \exp(i(l-k)\Delta\omega t)\langle \exp(i(\phi_l(t) - \phi_k(t)))\rangle$$

(13.2.24)

which can be rewritten by grouping the terms having the same values of $m = |l - k|$ as a sum over mode pairs $m = l - k$ as

$$I(t) = \sum_k \sum_m E_k^* E_{k+m} \exp(im\Delta\omega t)\langle \exp(i(\phi_{k+m}(t) - \phi_k(t)))\rangle$$

(13.2.25)

We can restrict the sum to positive values of m by grouping the terms k, m and $k + |m|$, $-|m|$:

$$I(t) = \sum_k \sum_{m\geq 0} E_k E_{k+m} (\exp(im\Delta\omega t)\langle \exp(i(\phi_{k+m}(t) - \phi_k(t)))\rangle + c.c.)$$

(13.2.26)

Considering, for example, the term at the round-trip frequency $\Delta\omega$, we have

(p.266)

$$I^{(\Delta\omega)} = \sum_k E_k E_{k+1} \cos(\phi_{k+1}(t) - \phi_k(t))$$

(13.2.27)

This term has the form of a random walk, and if one assumes for simplicity that all N modes have the same amplitude E, the average intensity at the round-trip frequency, proportional to the current in the photodetector, is just

$$I^{(\Delta\omega)} = \sqrt{N}E^2 = \frac{I_{tot}}{\sqrt{N}}$$

(13.2.28)

As a result, the presence of a peak of the photocurrent at the round-trip frequency arises even for equidistantly spaced modes with random respective phases.

Interferometric autocorrelation.Similarly, the interferometric autocorrelation can also be predicted in this case. Looking first at the intensity:

$$\langle E(t)E^*(t)\rangle = \sum_j \sum_k \langle E_j E_k e^{i(\omega_j - \omega_k)t} e^{i(\phi_j(t) - \phi_k(t))}\rangle$$

(13.2.29)

If the phases are changing with time, the average contains only the terms $j = k$, and we obtain

$$\langle I(t)\rangle = \sum_j E_j^2 = NE^2$$

(13.2.30)

assuming the amplitude of the modes are equal. The intensity correlation is therefore

$$\langle I(t)I(t+\tau)\rangle = \sum_j \sum_k \sum_l \sum_m \langle E_j E_k E_l E_m e^{i(\omega_j - \omega_k)t} e^{i(\omega_l - \omega_m)(t+\tau)} \cdot e^{i(\phi_j(t) - \phi_k(t))} e^{i(\phi_l(t+\tau) - \phi_m(t+\tau))}\rangle$$

(13.2.31)

We agree that the only way to prevent the average of the phases $\langle e^{i(\omega_j - \omega_k)t}\rangle$ from being non-zero is to force the conditions $j = k$, $l = m$, the other being $j = m$, $k = l$. Taking the first possibility,

$$(1) = \sum_j \sum_l \langle |E_j|^2 |E_l|^2\rangle = N^2 |E|^4$$

(13.2.32)

The second possibility is

$$(2) = \sum_j \sum_k \langle E_j^2 E_k^2 e^{i(\omega_j - \omega_k)t} e^{i(\omega_k - \omega_j)(t+\tau)} e^{i(\phi_j(t) - \phi_k(t))} e^{i(\phi_k(t+\tau) - \phi_j(t+\tau))} \rangle$$
$$= \sum_j \sum_k \langle E_j^2 E_k^2 e^{-i(\omega_j - \omega_k)\tau} e^{-i(\Delta\phi_j(t) - \Delta\phi_k(t+\tau))} \rangle$$

(13.2.33)

This term very rapidly vanishes because the sum over all modes of the oscillatory component will tend to zero for large τ. Asa result, we have that

$$\langle I(t)I(t+\tau)\rangle = N^2|E|^4$$

(13.2.34)

(p.267)

Using the same line of argument, we also obtain that for the case $\tau = 0$, the second line in eqn. 13.2.18 also contributes $2NE^4$ and $8NE^4$ respectively. As a result, we obtain for the ratio

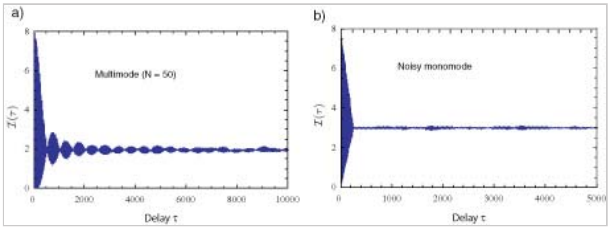


Fig. 13.8. Numerical simulation of the interferometric autocorrelation $\mathcal{J}(\tau)$ for the case of a multi-mode laser with random phases (a) or a single mode but noisy laser (b).

$$\frac{\mathcal{J}(0)}{\mathcal{J}(\tau)} = \frac{(2+4+2+8)N^2E^4}{(2+2)N^2E^4} = \frac{8}{2}$$

(13.2.35)

As expected, the degree of second-order coherence for such a laser is that of an incoherent source $g^{(2)}(0) = G_2(0)/G_2(\infty) = 2$.

Another extreme case can also be investigated: that of a single-mode laser, in which the linewidth is limited by phase jumps. In this case, amplitude fluctuations are cancelled and the quantity $\langle I(t)I(t+\tau)\rangle = \langle I(t)\rangle^2$. We therefore have for the interferometric autocorrelation the ratio $\frac{\mathcal{J}(0)}{\mathcal{J}(\tau)} = 8:3$. As shown in Fig. 13.8, these results are also supported by numerical simulations. This situation also corresponds to that of an FM modulated laser.

Efficiency of second harmonic generation. Another procedure that has been used to evaluate the nature of the emission is the efficiency of generation $\beta^{2\omega}$ of the second harmonic [318] emitted in specially designed quantum cascade lasers, as the latter is sensitive to the square of the instantaneous intensity and can be written as proportional to the term

$$\beta^{2\omega} = \frac{\langle |E(t)|^4 \rangle}{\langle |E(t)|^2 \rangle^2}$$

(13.2.36)

We show here that the multi mode operation with random phases also predicts a enhancement of this quantity by a factor of 3. To evaluate the quantity $\langle |E(t)| \rangle^4$ we have to be a little more careful and write the electric field as a real quantity:

$$E(t) = \frac{1}{2} \sum_k E_k \left(e^{i(\omega_k t + \phi_k)} + c.c. \right)$$

(13.2.37)

The evaluation then yields

$$\langle E(t)^4 \rangle = \frac{1}{16} \sum_j \sum_k \sum_l \sum_m \langle E_j E_k^* E_l E_m^* e^{i(\Delta\omega_{jklm} t + \Delta\phi_{jklm})} \rangle$$

(13.2.38)

(p.268)

where the quantities $\Delta\omega_{jklm}$ and $\Delta\phi_{jklm}$ are given by all the combinations of

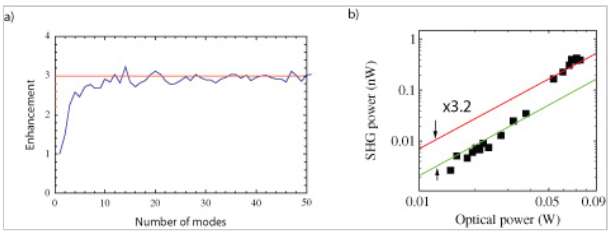


Fig. 13.9. a) Numerical computation of the second harmonic enhancement $\beta^{2\omega}$ in a multimode laser, as a function of the number of modes with random phases. b) Measured second harmonic power below and above the threshold for multimode emission, as a function of laser power in log–log plot. The two fits performed below and above the onset of multimode emission have a slope of 2, and correspond to a ratio of efficiencies of 3.2, close to the value expected for a multimode laser. Reprinted, with permission, from [318]. Copyright 2004 IEEE.

$$\Delta\omega_{jklm} = \pm\omega_j \pm \omega_k \pm \omega_l \pm \omega_m$$

(13.2.39)

and

$$\Delta\phi_{jklm} = \pm\phi_j \pm \phi_k \pm \phi_l \pm \phi_m$$

(13.2.40)

The average is non-zero only when $\Delta\omega_{jklm}$ and $\Delta\phi_{jklm}$ vanish. This requires the conditions ($j = k,\ l = m$), ($j = m,\ k = l$), ($j = m,\ k = l$), each with four possibilities for the signs (two for each pair of indices). For equal amplitudes of the electric fields of each mode, the sum of the $12N^2$ possible non-zero terms yield a total value of

$$\langle E(t)^4 \rangle = \frac{3}{4} N^2 E^4$$

(13.2.41)

Applying the same procedure for the electric field squared yields

$$\langle E(t)^2 \rangle = \frac{1}{4} \sum_j \sum_k \langle E_j E_k^* E_l E_m^* e^{i(\Delta\omega_{jk} t + \Delta\phi_{jk})} \rangle = \frac{1}{2} N E^2$$

(13.2.42)

We therefore obtain for the normalized efficiency of generation

$$\beta^{2\omega} = \frac{\langle |E(t)|^4 \rangle}{\langle |E(t)|^2 \rangle^2} = 3$$

(13.2.43)

This value is well supported by the experimental data of [318], as shown in Fig. 13.9.

In conclusion, instabilities in quantum cascade lasers are operating in a very different regime of parameters compared to the solid-state and semiconductor lasers.

(p.269) The very fast gain recovery stabilizes the device against Q-switching and even modelocking, and on the other side the non-linearities of the active region, originating from either the spatial hole burning or the strong field effects, will try to destabilize the single-mode operation and promote multi-mode operation. The main experimental observations are consistent with a situation where equally spaced modes are interfering, but the respective mode phases not constant or linear with frequency, as required for single pulse mode-locking. In fact, one could go one step further and assume that the very fast recovery time of the gain material will favor a situation where the intensity of the beam is as constant as possible. Actually, recent measurements on broadband quantum cascade lasers have shown that the latter can be made to operate as room temperature frequency combs[1] covering 60-100cm^{-1} and with an intermode beat as narrow as 10Hz. Furthermore, measurements of the relative phases using a beatnote spectroscopy technique demonstrated that these relative phases were very close to the ones of a perfectly FM modulated laser.

13.2.4 Active mode-locking

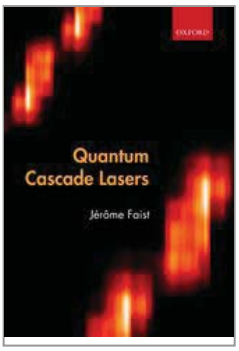
Active mode-locking of a quantum cascade laser has been attempted, usually by modulating a section of the laser at the round-trip frequency and therefore forcing the mode to have equal phases. Mode proliferation has been observed in [319], but true pulsed mode-locking has been achieved only in an experiment where a special active region based on a highly diagonal transition, with a very long upper-state lifetime, was designed for this

purpose. Mode-locking has been measured at cryogenic temperature ($T = 78$ K) over a bandwidth of about $15^{\mathrm{cm}-1}$ at 1585 cm^{-1} [320].

In the terahertz, a bound-to-continuum laser has also been mode-locked by RF injection and phase-locked onto a fiber laser [321]. This enabled the electric field of the terahertz waveform to be sampled coherently by a second femtosecond laser using an asynchronous sampling technique, and this way, imaging the pulse electric field as a function of time. Pulse lengths in the range of 10 ps were achieved, corresponding to the bandwidth of 200 GHz of the up to ten modes locked together. The experiments were carried over at $T = 20$ K, as limited by the QCL gain medium.

Notes:

([1]) A. Hugi, G. Villares, S. Blaser, H. C. Liu, and J. Faist, *Mid-infrared frequency comb based on a quantum cascade laser* Nature, vol. 492, pp. 229233, (2012)



Quantum Cascade Lasers
Jérôme Faist

Print publication date: 2013
Print ISBN-13: 9780198528241
Published to Oxford Scholarship Online: May 2013
DOI: 10.1093/acprof:oso/9780198528241.001.0001

Applications

Jérôme Faist

DOI:10.1093/acprof:oso/9780198528241.003.0014

Abstract and Keywords

This chapter discusses various applications of quantum cascade lasers. These include applications in energy deposition, telecommunications, gas sensing, and broadband spectroscopy.

Keywords: quantum cascade lasers, energy deposition, telecommunications, gas sensing, broadband spectroscopy

The speciffic features of QCLs—operation in a very large-frequency range, high operation temperature, and the capability to operate them with a high modulation frequency—have driven their use in applications. In return, the latter have also conditioned the research in this device. The high-power, high-wallplug-efficiency development of the quantum cascade laser has been driven mostly by countermeasures. The prospect of developing a mid-infrared free space optical link has raised the interest in the high-frequency modulation capability of this device. Finally, spectroscopie applications have stimulated the development of devices over a very wide frequency range with single-mode tunable operation.

14.1 Energy deposition

Many types of man-portable air-defense systems (MANPADS) use an infrared sensor to home in on a heat source on an aircraft. High-power quantum cascade lasers, due to their brightness, small size, and efficient generation of infrared light, are considered to be an ideal source to produce an infrared beam that can be used to disturb the guidance system of

an infrared-guided missile. Another countermeasure scenario using quantum cascade lasers consists in using high-power quantum cascade laser sources in the form of laser pointers to disturb highly sensitive thermal cameras that are used in military equipment such as tanks and helicopters for night vision. Finally, their use as a beacon or target designator has also been considered.

14.2 Telecommunications

One application of quantum cascade lasers is its use as the optical source for a freespace optical link, transmitting telecommunication signals through the atmosphere in a direct line of sight. In contrast to fiber-optical telecommunications, this technique has the advantage that it does not require additional cables to be buried in the ground, leading to very low installation costs. In urban areas where large amounts of fiberoptical connections already exist, fast free-space optical datalinks could be particularly convenient as a means of momentarily increasing the bandwidth between two points. Another interesting application area is to create high-bandwidth access in countries where very little ground infrastructure exists or to connect mobile phone antennae to the fiber backbone.

The choice of the wavelength for such a free-space optical link must combine a number of conflicting requirements:

- Transparency of the atmosphere.

(p.271)

- Availability of high modulation frequency low-cost hardware at the operation wavelength.
- Limitations in optical power caused by eye safety considerations.
- Alignment tolerances and effects of diffraction.
- Effect of solar scintillation.
- Effect of fog and rain.

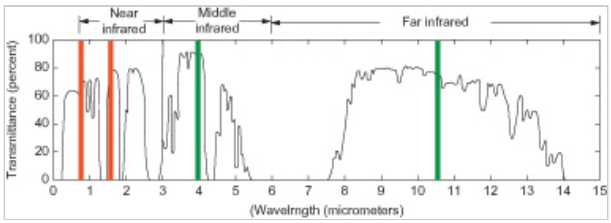


Fig. 14.1. Transmission of the atmosphere and location of potential free space optics transmission windows a 800nm, 1.5 μm, 3–4 μm and 8–10 μm.

Fig. 14.1 shows the transmissivity of the atmosphere at sea level for a clear day over a 7 km pathlength. One can clearly distinguish at least four possible windows for free-space optic communications, at wavelenths of 800 nm, 1.5 μm, 3–4 μm, and 8–10 μm. From the two near-infrared wavelengths, the $\lambda = 1.55$ μm wavelength has the clear advantage of the availability of hardware, potential seamless integration with fiber optics, and relatively eye-safe operation, while the 800 nm offers the possibility of better detectors (Si avalanche photodiodes). However, as shown in Fig. 14.2, these wavelengths are closer to the peak of the solar spectrum, and therefore the link quality may be perturbed by scintillation from

dust or objects in the atmosphere. The shorter wavelength of these near-infrared channels plays both in a positive way, by decreasing the diffraction losses but at the same time negatively by also decreasing the alignment tolerance. In comparison, mid-infrared channels at 3–4 μm and 8–10 μm operate with a lower solar glare, and offer very eye-safe operation, since the light is not even transmitted through the cornea. The longer wavelength offers a good compromise between the size of the optics and the alignment requirements. However, clearly the most attractive feature of mid-infrared is the potential to suffer much less from scattering by fog particles as compared to the visible. Indeed, if one considers the fog as consisting mainly of spherical drops of about 1 μm size, the λ^{-4} dependence of the Rayleigh scattering with wavelength should reduce dramatically the losses in fog, greatly increasing the possible link distance.

It is, however, difficult to back these qualitative considerations by more hard data. Since fog conditions are to a large extent dependent on the location, and temperature, and are difficult to reproduce, the majority of the link evaluations are carried out using numerical models. It is generally agreed that the key advantage of the mid-infrared, as compared to the near-infrared, is to extend a high-availability link distance by a factor of 3 to 10, depending on the nature of the fog, from 200 m to about 1 km [322].

(p.272)

The high-frequency modulation characteristics of the quantum cascade laser discussed in Chapter 13, as well as its high operation power, are two very favorable features for free space optical link applications. Initial experiments have been carried out using first cryogenically cooled [323, 324] and then Peltier-cooled lasers and detectors [325], demonstrating modulation frequencies in the megahertz range. The results from a prototype system developed by the company Maxima are shown in Fig. 14.3. Eye diagrams of a modulated beam are compared directly in the laboratory and after propagation through 810 m of atmosphere. The development of such links using cascade lasers is

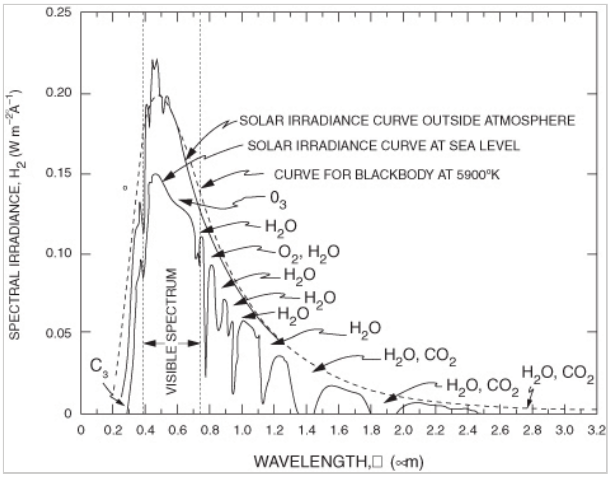
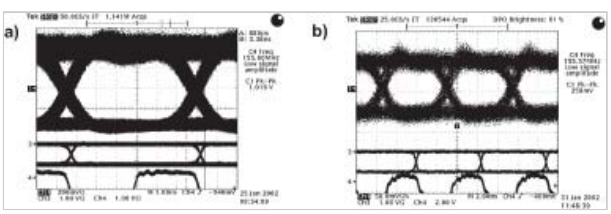


Fig. 14.2. Solar irradiance as a function of wavelength. Solar glare is a major limitation of FSO that significantly decreases with increasing wavelength.



expected to greatly increase with recent improvements in wallplug efficiency, operating temperature, and optical power of the most recent quantum cascade laser sources.

Fig. 14.3. Experimental data transmission using a quantum cascade laser at 9 μm. a) In the laboratory. b) Through an 810 m long link. Data courtesy of J. Plante, Maxima corp.

14.3 Gas-sensing

Quite naturally, the foremost feature of QCL is its ability to operate in a very large wavelength range. As shown in Fig. 14.4, these devices have been operated between wavelengths as short as $\lambda = 2.9$ μm and as long as $\lambda = 250$ μm. The frequency coverage is almost continuous, with the exception of the so-called Restrahlen band, the frequency region between about $\lambda = 28$ μm and $\lambda = 50$ μm where the active material **(p.273)** is optically dense, due to the absorption by the optically active III–V optical phonon modes. The mid-infrared region between approximately $\lambda = 3.3$ μm and $\lambda = 16$ μm is especially attractive, since room-temperature operation is possible and is also achieved in continuous wave over most of this range. In contrast, the maximum operating temperatures achieved by devices operating at frequencies below the Restrahlen band are clearly below the reach of thermoelectric Peltier coolers. This can be seen as the most important limitation to the wide application of terahertz QCL in applications. The ability to operate over a very wide frequency range is especially important for spectrosopy, as the rotovibrational modes of molecules tend to naturally cover a very wide spectrum of frequencies, with fundamental modes covering a wide portion of the mid- and far-infrared spectrum.

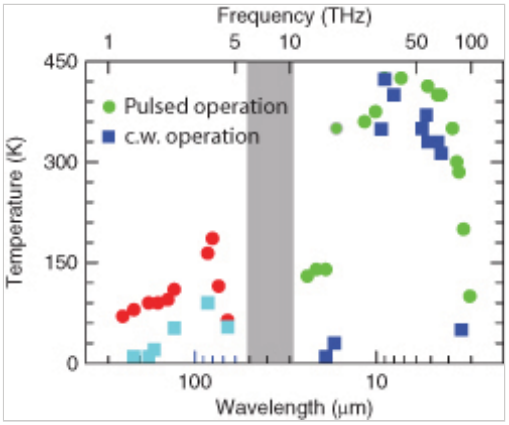


Fig. 14.4. Maximum operating temperatures for continuous (squares) and pulsed (disks) operation, as reported in the literature.

The key advantage of using the fundamental vibration modes instead of the higher harmonics stems from their much larger oscillator strength, yielding, therefore, a much higher detection sensitivity. The plot of a number of these absorption lines, computed using the HITRAN 2000 database, is shown in Fig. 14.5. Some of these compounds can also be detected in the near-infrared using overtone lines. The oscillator strength of these overtone lines is, however, usually much weaker than that of the fundamental mode, the ratio varying depending greatly on the chemical species. Whereas this value is “only” 100 for CH_4, for example, it can reach up to 25,000 for NO.

Due to the much larger oscillator strength in the mid-infrared, a large number of chemicals have been measured with very large sensitivity and accuracy. As a result, the development of quantum cascade laser-based spectroscopy followed the realization of the first distributed feedback quantum cascade lasers in 1997. Using the very first devices operating in pulsed mode at room temperature near $\lambda \approx 8\ \mu$m, N_2O detection and spectroscopy was performed and already yielded sensitivities in the ppm range over very short (10 cm) path lengths [327]. Using liquid nitrogen-cooled devices, high-resolution spectroscopy was initiated using continuous wave devices for direct absorption [328–330] or photoacoustic [331] techniques. As shown in Fig. 10.13, **(p.274)**

the time-resolved spectra measurements of the distributed feedback quantum cascade lasers showed that the device remained single-mode but exhibited rapid chirp during the pulse. This favorable property was exploited by a group in IPM [332] and by Strathclyde university [333] to propose spectrometers based on the intrapulse technique. The advantage of this technique is that the wavelength scanning is performed over a time period where the $1/f$ noise of the set-up is no longer a limitation. The drawback is the necessity of using a very fast detector and electronics to acquire the data.

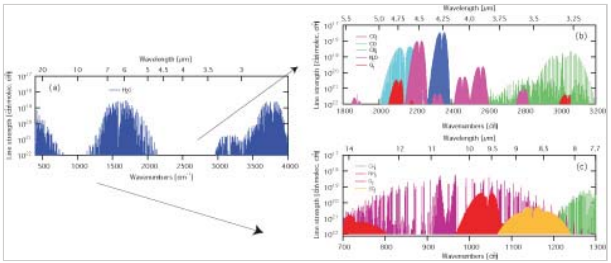


Fig. 14.5. (a) Absorption lines of the H_2O molecule between 400 and 4000 cm^{-1}, showing the two infrared atmospheric windows. The absorption lines of a selection of molecules in the first atmospheric window are shown in (b) and in the second atmospheric window in (c). Data were compiled using the HITRAN04 database [326].

In a competing technique, pioneered for QCLs by Aerodyne and described in more detail in the following paragraph, the QCL is tuned slowly by a DC current while short pulses bring it above threshold. Examples of measurement sensitivities achieved by Aerodyne, using this and similar techniques, are shown in Table 14.1. For all these chemicals, a limit of detection below the ppb is achieved within 1 sec of integration time.

14.3.1 Absorption measurements

Most optical gas detection techniques are based on the measurement of light absorption by a rotovibration molecular mode, using the Beer–Lambert law to extract the gas concentration from the measurement of the light transmitted through a cell. A schematic diagram of such a generic system is shown in Fig. 14.6. In a direct absorption technique, the light transmitted through the cell is measured on the detector. In a photoacoustic technique, the quantum cascade laser is modulated in amplitude, and the acoustic wave generated by the absorbed power in the gas is measured. Because, in general, the absorption due to the gas is small compared to other optical losses of the system, the sensing signal is recorded as the laser is tuned across the absorption line.

Table 14.1 Limit of detection for various gases, using a 76 m path length, for 1 sec and for 100 sec integration time. (Data courtesy of Mark Zahniser, Aerodyne Corp.)

Trace gas	cm^{-1}	ppb 1 s RMS	LoD ppb 100 s
NH_3	967	0.2	0.06
C_2H_4	960	1	0.5
O_3	1050	1.5	0.6
CH_4	1270	1	0.4
N_2O	1270	0.4	0.2
NO_2	1600	0.2	0.3
HONO	1700	0.6	0.3
HNO_3	1723	0.6	0.3
HCHO	1765	0.3	0.15
HCOOH	1765	0.3	0.15
NO	1900	0.6	0.3
OCS	2071	0.06	0.03
CO	2190	0.2	0.1
N_2O	2240	0.2	0.1

Other approaches are also possible: for example, interferometric systems in which the change in the refractive index due to the presence of the gas is measured. From the Beer–Lambert law, the optical power P of the light beam after crossing the gas cell can be written as a function of the initial laser power P_0 as

$$P = P_0 \exp(-\sigma(\nu) n_s L) \tag{14.3.1}$$

where L is the interaction length and $\alpha = \sigma(\nu) n_s$ the

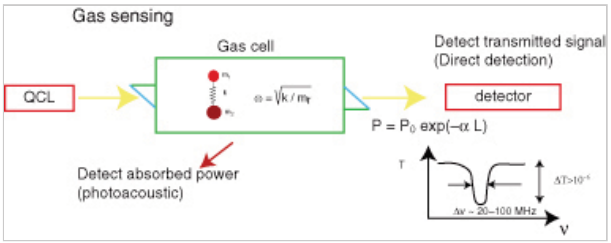


Fig. 14.6. Schematic diagram of a generic gas-sensing system. A quantum cascade laser is tuned across a roto-vibrational line of a molecule at frequency ω. In the direct absorption technique, the transmitted power is measured as a function of laser frequency. In a photoacoustic technique the laser is modulated, and the acoustic wave generated by the absorbed power in the gas is measured.

absorption of the gas, with $\sigma(v)$ the optical cross-section and n_s the gas concentration. The sensitivity of a system where the transmitted fraction of the light is measured therefore depends on

(p.276)

- The accuracy at which a change in transmission $\Delta T_{noise} = (P - P_0)/P_0$ can be recorded.
- The strength of the total absorbed fraction, equal for small absorption to $\alpha(v)L$.

The various systems proposed all try to achieve the lowest possible minimum value of ΔT_{noise} and maximize the value of $\alpha(v)L$. After the selection of a wavelength where the target gas exhibits a strong line with a large value of $\alpha(v)$, the sensitivity will be maximized by a geometry that enables a large path length L. At ambient pressure, all absorption lines are pressure-broadened; as a result the peak absorption is essentially pressure independent, as the increase in the concentration n_s is exactly compensated by the decrease in the optical cross-section $\sigma(v)$ due to the pressure broadening. For this reason it is usually favorable to operate at low pressures where the gas of interest exhibits absorption lines close to their Doppler limit. The narrower lines are less susceptible to interferences from absorption lines from other gases such as H_2O or CO_2. The only drawback is the increased system complexity and cost. As shown later in this text, the interaction length L can be maximized by the use of an open path, a multi-pass cavity, or a cavity ring-down technique.

The sensitivity of a system, measured in ΔT_{noise}, depends on a number of factors, the most important ones being usually:

- The noise of the detector.
- The laser modulation technique.
- The fluctuation in laser power and electronic noise.
- Noise in the optical system.

A rough typical value for the limit in sensitivity is $\Delta T \approx 10^{-5} - 10^{-6}$. The different sensing techniques presented in the literature reffect the various strategies adopted to minimize the value of ΔT_{noise} in the presence of various noise sources. Modulation techniques are used to reduce the effect of $1/f$ noise in the various components of the system.

Detector noise. For an average laser power P_0, and a detector characterized by a noise equivalent power *NEP*, the sensitivity limit of the system caused by the detector noise is $\Delta T_{noise}(\text{in}\sqrt{\text{Hz}})$ can be written as:

$$\Delta T_{noise} = \frac{NEP}{P_0} \tag{14.3.2}$$

for suffciently low incident powers, before the system becomes limited by the shot noise. In the latter case the signal-over-noise for a photoconductive detector is given by

$$\Delta T_{noise} = \sqrt{\frac{4h\nu}{\eta P_0}}$$

(14.3.3)
where η is the quantum efficiency of the detector.

A 200 μm wide liquid nitrogen-cooled mercury–cadmium–telluride detector, with a detectivity of $D^* = 2 \times 10^{10} \mathrm{cm}\sqrt{\mathrm{Hz}}/\mathrm{W}$, will exhibit an $NEP \approx 1$ pW. As a result, a shot noise limit will already be reached for a laser average power of $P_0 \approx 1$ μW, **(p.277)**

with Eqn. 14.3.2 predicting a $\Delta T_{noise} \approx 10^{-6}$. The shot noise limited sensitivity for a power of $P_0 = 1$ mW is $\Delta T_{noise} = 2.5 \times 10^{-8}$. Such low values are usually not reached because of other system limitations. Thermoelectrically cooled detectors are usually preferred to liquid nitrogen-cooled ones for system applications. The latter exhibit approximately 10–50 times lower detectivities $D^* \approx 10^{-9} \mathrm{cm}\sqrt{\mathrm{Hz}}/W$, strongly dependent on the detector wavelength cutoff, raising the required laser power to 10–50 μW to reach shot noise level. One should remember that the required laser powers are the values incident on the detector and do not take into account optical losses in the system. Especially for systems that use multi-pass cells, the overall transmittivity of the system can easily decrease to values in the 1% range.

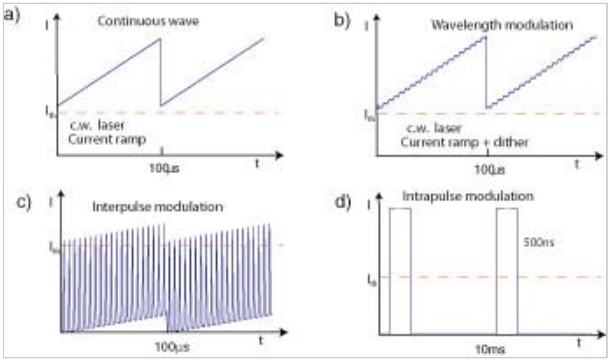


Fig. 14.7. Laser modulation schemes used with quantum cascade lasers. If the laser is operated in a continuous wave, the current can be either simply ramped a), or a small dither can be added for synchronous detection b). For lasers operating in pulsed mode, in the interpulse modulation technique (c) short, 5–10 ns long pulses are added to a continous ramp. In contrast, in the intrapulse technique the laser is turned above threshold and the whole spectrum is acquired during each pulse.

Laser modulation techniques. As already mentioned, the laser must be tuned across the absorption line during a measurement. As discussed in Chapter 10, distributed feedback quantum cascade lasers tune frequency with temperature. When a fast modulation is required, the most efficient way to change the temperature is to vary the drive current of the device. Depending on the drive technique for the laser (pulsed or continuous wave), various modulation techniques have been used. The most common ones are summarized in Fig. 14.7. The choice of driving the laser in continuous wave or in pulse is a combination of various factors. For lasers operated in a continuous wave, the simplest approach is to drive the laser directly with a continuous-wave ramp and to record the transmitted power. In order to reduce the $1/f$ noise of the laser, an additional small modulation can be added to the

drive current, enabling a synchronous detection at the modulation frequency or at its first harmonic. When the laser is operated in pulse mode, two opposite operation philosophies can be used. Either, as shown in Fig. 14.7c, the laser is turned on using very short pulses, designed to maintain a **(p.278)** narrow linewidth. A sub-threshold continuous-wave current ramp tunes the laser between each pulse. Using this technique, one spectrum is acquired during a teth of a milisecond. The advantage of this approach is that it can be used with relatively slow detectors and acquisition boards. In contrast, in the intrapulse modulation technique the laser is turned on with a relatively long pulse. The spectral features of the absorption lines are measured during the pulse by a fast detector and acquisition board. Both approaches have shown their merits, and the choice of one or the other approach depends more on the speciffic application targeted. The interpulse modulation has been successfully used for many applications in environmental monitoring where high sensitivity and selectivity are a primary concern, such as the data shown in Table 14.1. In contrast, the intrapulse modulation has been used for industrial applications such as exhaust gas monitoring or plasma monitoring [334].

Fluctuations in the laser power and electronic noise. A major disadvantage of the direct absorption technique is that the power stability of the laser source must be at least as good as the fractional absorption sensitivity. Assuming that the laser does not introduce additional noise, and assuming a drive current I above the threshold current I_{th}, the required stability in the drive current must satisfy

$$\frac{\Delta I}{I} < \frac{I - I_{th}}{I} \Delta T_{noise}$$

(14.3.4)

For a laser operated 20 mA above a threshold current of 500 mA, aiming at a $\Delta T_{noise} = 10^{-6}$, the current fluctuations must be below $\Delta I = 20$ nA, not very far from the shot noise limit of 0.8 nA. These considerations stress the importance of designing low-noise electronics; an efficient way to reduce the influence of the laser driver noise is by dividing the transmitted intensity by a continuous measurement of the laser power performed in a reference arm of the spectrometer.

Noise in the optical system. The whole optical system is susceptible to changes in its transmission due to thermal drifts, acoustic noise, or mechanical stress. As these noise sources tend to exhibit $1/f$ characteristics, modulation techniques are usually employed along with good construction practice to minimize these noise sources. A notable exception is the fringing effect, or the modulation of the intensity caused by unwanted Fabry–Perot fringes along the path. Most difficult to remove are Fabry– Perot resonances with a free spectral range corresponding to the linewidth of the gas absorption line. In fact, oscillations of the transmission $\Delta T_{fringes}$ equal to

$$\Delta T_{fringes} = 4R$$

(14.3.5)

are obtained in the limit of low reffectivities. As a result, $\Delta T_{fringes} < 10^{-6}$ implies extremely low scattering on the optics, such that effectively $R < 2.5 \times 10^{-7}$.

14.3.2 Open-path measurements

Long pathlengths can be achieved in open-pass systems, and such systems have relatively speciffic advantages and limitations. In particular, a proper background normalization by replacing the gas sample by clean air (”zero air”) cannot be performed **(p.279)**

directly, and the spectra must be performed at ambient pressure, increasing the overlap between neighboring lines. Problems created by the inlet, gas mixing, and tube and cell poisoning are, of course, removed. Atmospheric ozone-level measurements have been performed using quantum cascade lasers at 1044 −1050 cm^{-1} in open-path configurations with good detection levels [335].

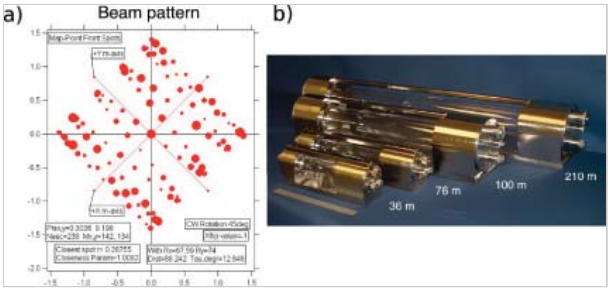


Fig. 14.8. Beam pattern (a) and picture of astigmatic cells (b). A 30 cm long cell with a volume of 0.5 l enables a large number of passes ($N = 238$) with 76 m interaction length. (Data courtesy of Aerodyne Inc.)

14.3.3 Multi-pass cavity

Long path lengths are commonly achieved by folding the beam inside a multi pass cavity, or White cell, formed by two spherical mirrors facing each other. Injection of the optical beam at a well-controlled angle enables a beam pattern with N number of bounces before existing the cavity. The use of astigmatic mirrors enables very large values of $N = 200$ to be achieved. Examples of such cavities, along with the beam pattern on the mirror, are shown in Fig. 14.8. A 30 cm long cell with a volume of 0.5 l, enables a large number of passes ($N = 200$) with 76 m interaction length. The number of passes N is limited by both the transmittivity after the N reffection on the mirrors and the necessity to extract the beam after a well-defined number of cell crossings.

A schematic diagram of the experimental set-up used to achieve the high sensitivities realized in Table 14.1 is shown in Fig. 14.9 a) [212]. The quantum cascade laser source is placed in a thermoelectrically cooled enclosure and tuned to the right wavelength range by selecting the appropriate temperature. The laser can be either driven in pulsed mode, in which case short pulses, approximatively 10 ns long at 1 MHz repetition rate, are superimposed on a continuous-wave current ramp that, while keeping the laser just below threshold, tunes it by temperature. Each current ramp lasts about 100 μ s and contains about 100 pulses. After exiting the laser, a small fraction of the beam is used as a reference for line locking and spectrum normalization. The main part of the beam enters the astigmatic multi-pass cell; the $N = 238$ passes yield an effective pathlength of 76 m. The signal from the thermoelectrically cooled **(p.280)**

MCT detector is digitized by a 14 bits data acquisition board. The whole electronics is driven by a computer and controlled

by a proprietary software WINTEL.
The same optical arrangement can be used with the laser driven in a continuous wave with an average power of 3 mW, in which case a small current modulation is added to the laser drive current.

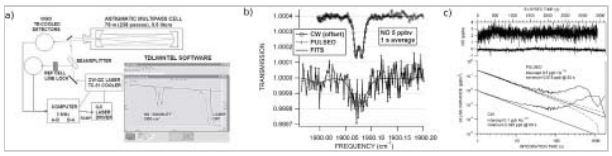


Fig. 14.9. a) Schematic diagram of the optical set-up. b) Comparison of the measurement of the optical doublet of NO at 1900 cm^{-1} measured with a continuous wave (top) and a pulsed QC laser (bottom) for 1 sec integration time. c) Allan variance plot of these measurements with a continuous or pulsed laser system [212].

A comparison of experimental results obtained using these techniques is shown in Fig. 14.9b, where the measurement of the NO doublet at 1900 cm^{-1} is displayed for a concentration of 5 ppb for both continuous wave (top) and pulsed laser condition (bottom) for 1 sec integration time. The much better signal-over-noise of the continous wave measurement is caused first by the larger average power incident on the detector (about 50 μW) and the narrower linewidth of the laser $\Delta\nu = 0.0004$ cm^{-1} or 12 MHz. In contrast, in the pulsed measurement the NO doublet is not resolved, because of the wider laser linewidth $\Delta\nu = 0.006$ cm^{-1}, and the much lower laser power on the detector (2 μW) is responsible for the larger noise.

The plot of the Allan variance of both measurements demonstrates the better sensitivity of the continuous wave system. For 1 sec integration time, the root mean square noise is 0.1 ppbV for the continuous wave system and 0.5 ppbV for the pulsed system. These values decrease to a minimum deviation of 0.025 ppbV and 0.075 ppbV, respectively, for 80 sec integration time. The lowest value, obtained for the continuous wave system, corresponds to a change in transmission of 1.6×10^{-6}, limited by fringing effects. To achieve such high sensitivity a number of mitigating techniques were used, such as dithering one mirror of the cell with a piezo drive and by choosing a beam path in the multi-pass cell that minimizes the fringes with the “wrong” free spectral range.

Using a similar instrument, accurate measurements of the three main isotopologues of CO_2 — $C^{12}O_2^{16}$, $C^{13}O_2^{16}$, and $C^{12}O^{18}O^{16}$ —were measured in the spectral region of 2311 cm^{-1} [336]. By careful referencing and temperature stabilization of the instrument, a standard deviation of 0.026% for the ratio δ^{13} C and 0.029% for the ratio δ^{18} O were achieved. These two values reduced to 0.003% and 0.005%, respectively, after 60 sec integration time. In the field, real-time monitoring of isotopologues using such a laser spectrometer is shown in Fig. 14.10, and compared with reference results from a mass spectrometer, showing the potential of this technique for environmental monitoring. **(p.281)**

The laser spectrometer provides a continuous sampling with very high accuracy.
Isotopologue measurements are

also very interesting for medical applications.

14.3.4 Cavity ring-down spectroscopy

Long interaction lengths can also be achieved using cavity ring-down spectroscopy. In this technique the decay time of a high-finesse Fabry–Perot cavity is measured in the presence (τ) and absence (τ_{empty}) of the gas to be measured. For a cavity length L, the gas absorption α is then obtained as

$$\alpha = \frac{1}{c}\left(\frac{1}{\tau} - \frac{1}{\tau_{empty}}\right) \quad (14.3.6)$$

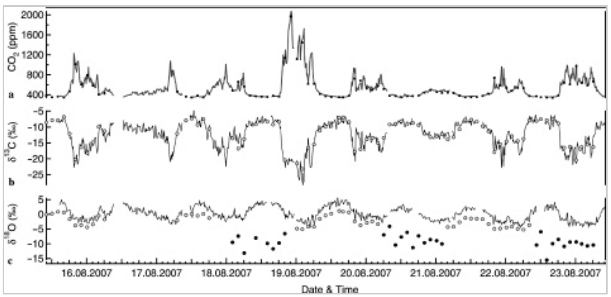


Fig. 14.10. a) Time series of the CO_2 mixing ratio measured by the laser spectrometer (line) and with an mass spectrometer (dots). The corresponding δ ^{13}C and δ ^{18}O values measured by the laser spectrometer are shown in (b) and (c). For comparison, mass spectrometer δ values are also given. The closed symbols for the δ ^{18}O values indicate sampling issues with the metal flasks used for the mass spectrometry measurement [336].

As shown by the above equation, the absorption coefficient is obtained entirely from a time measurement. Very long effective interaction lengths ($\approx 1km$) can be achieved in very compact cavities that, however, require mirrors with very large reffectivities. In a relatively standard approach a continuous-wave laser is abruptly turned off as it is tuned to the cavity resonance. The characteristic decay time of the optical field is then measured by a fast detector, allowing the extraction of τ and τ_{empty}.

A schematic description of an early cavity ring-down experiment [337] is shown in Fig. 14.11. In the latter, the ring-down cavity is formed by two mirrors with R = 99.97% reffectivity separated by 37 cm. The cavity has a free spectral range of 405 MHz, a decay time of τ_{empty} = 3.5 μs, and a linewidth $\Delta\nu_{cavity} = 1/(2\pi\tau_{empty})$ = 45 kHz. As a comparison, a free-running quantum cascade laser is expected to have a linewidth in the megahertz range, whereas a Doppler-broadened line of a typical gas will be in the hundreds-of-megahertz range. As a result, the spectroscopy of a speciffic **(p.282)**

gas absorption line can be performed by setting the relevant wavelength using the current tuning of the laser. Using a piezo control drive, the cavity is tuned to the laser wavelength a large increase in the transmitted signal indicates that the cavity has been filled. Once a trigger level has been reached, the laser is abruptly

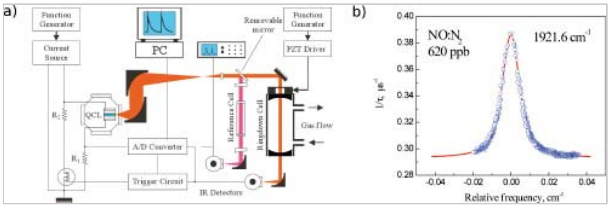


Fig. 14.11. a) Schematic description of a set-up for cavity ring-down spectroscopy using a continuous-wave operation QCL. The device was cooled down to liquid nitrogen temperature for these initial

switched off and the decay time of the cavity light is measured. The measurement of 620 ppb of NO diluted in dry nitrogen is shown in Fig. 14.11b. A sensitivity of 1 ppb has been achieved at 1921.6 cm^{-1}; the equivalent pathlength is $l = c\tau_{empty} \approx 1km$.

experiments b) Ring-down measurement of 620 ppb of NO diluted in N_2. From[337].

14.3.5 Photoacoustic sensing

In the photoacoustic sensing technique the quantum cascade laser, whose power is amplitude modulated, is shined through the gas cell. The pressure wave created by the periodic heating of the gas caused by the absorption is detected by a sensitive microphone. This approach has many advantages compared to direct absorption. The signal is in principle background-free, as the signal picked up by the microphone is proportional to the gas absorption. The system needs an optical detector, and is very modular. To achieve high sensitivity, quantum cascade lasers with higher powers must be used, as the signal-over-noise is directly proportional to the incident laser power. Initial experiments were performed using cryogenically cooled, continuous-wave lasers [331] with about 10 mW power. Later experiments using pulsed, Peltier-cooled devices used even lower powers, in the milliwatt range [210].

The senstivities achieved in these initial experiments were rather limited. Very large increases in the single-mode optical power at room temperature, and availability of devices at the wavelength at which the gases have their largest absorption lines, have greatly improved the recent achieved sensitivities. A summary of recent results [338] achieved by the group at Reims University is shown in Fig. 14.12. It should, however, be noted that even in these results the optical power was never larger than 10 mW.

A very interesting development of photoacoustic spectroscopy is the use of Quartz tuning forks instead of a microphone. The tuning fork acts as a resonant microphone **(p.283)** at a relatively high frequency and high quality factor. The development of quartz enhanced photoacoustic spectroscopy (QEPAS) has been pioneered by the group led by F. Tittel at Rice University, and offers the potential of systems with very low cost and volume while maintaining high sensitivity [339].

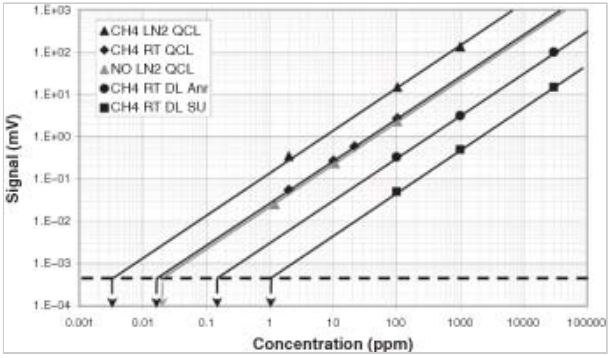


Fig. 14.12. Photoacoustic signal versus gas concentration using various mid-infrared light sources. The signal is clearly proportional to gas concentration as expected in photoacoustic measurements. All the lines are parallel. The detection limits for all the gases and all lasers are

14.4 Broadband spectroscopy

(p.284) The spectroscopy of multi-component gas on the one hand, and of solids or liquids on the other, requires the use of

much wider tuning ranges than those available with a single distributed feedback laser. Multi-component gas detection systems have been developed around multiplexed distributed feedback quantum cascade lasers. A more radical approach is the use of external cavity quantum cascade lasers able, as shown in Chapter 10, to be tuned over a large wavelength range. However, in contrast to distributed feedback quantum cascade lasers, tuning the device without mode jumps is a challenging task. In an instrument developed by G. Wysocki at Rice University, the grating orientation, external cavity length, and laser temperature are tuned together to achieve a mode-hop free tuning of tens of cm^{-1}—a result also achieved by recent commercial instruments [340]. To achieve these results a piezo-activated mode tracking system has been implemented to enable independent control of the EC length and diffraction grating angle. The grating was mounted on a moving platform whose position was controlled by a piezo-acuated linear translation stage and a rotary stage which is equiped with a **(p.285)** motorized coarse angle control and a piezo-actuated (PZT) fine control. The chosen components allowed precise positioning of the EC length and grating angle with a resolution of 〈 0.9 nm and 1

estimated by extending the parallel lines until they cut the noise limit, i.e. 0.4 V at a signal-to-noise ratio of 1. Note the high dynamic range of photoacoustic measurements—up to six orders of magnitude. From [338].

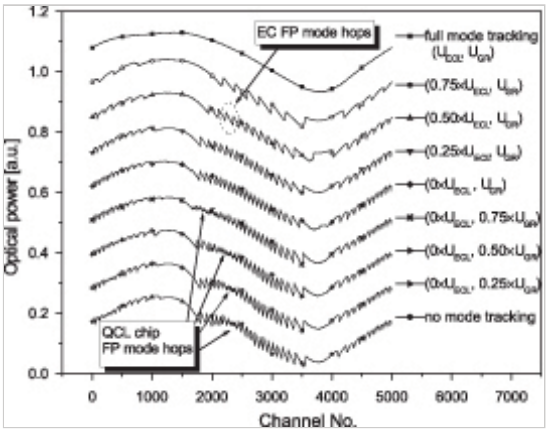


Fig. 14.13. Demonstration of mode-hopfree operation of the EC-QCL using the piezoactuated cavity mode tracking system. From[340]

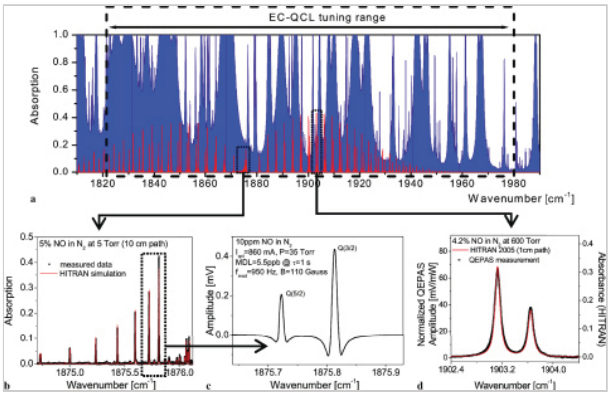


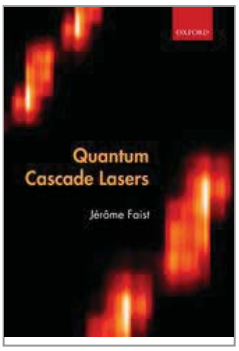
Fig. 14.14. (a) Upper curve, simulation of atmospheric absorption over a 286 m path within the 5.3 μm EC–QCL tuning range. For reference, an absorption spectrum of 1 ppm NO at the same working condition is plotted. The lower panels demonstrate mode-hop-free spectra of NO acquired as (b) a direct absorption spectrum of the NO Q-branch recorded at a scan rate of 10 Hz for 5% NO in N_2 at reduced pressure and with 10 cm optical path length, (c) a Faraday rotation spectrum of NO Q(3/2) and Q(5/2) transitions at ≈ $1875.8cm^{1}$ for 10 ppm

NO in N_2 mixture at 35 Torr and with 42 cm active optical pathlength, and (d) a QEPAS spectrum recorded for 4.2% NO in N_2 at 1903 cm^{-1}, pressure of 600 Torr and with 1 cm optical pathlength (the strongest transition in the fundamental NO R-branch) From [341].

μrad for external cavity length and grating angle, respectively. The total grating angle range provided by a PZT scanner was ±520 μrad. The coarse tuning by the linear motor could be performed within a range of ±6.3° and unidirectional repeatability of 10 μrad, with the actual position measured by a built-in encoder with a resolution of ~ 1.4 μrad/unit. The piezo-actuator controlling the external cavity length had a total travel range of 90 μm, which corresponds to a maximum continuous laser frequency tuning range of ~ 2 cm^{-1} at $\lambda \cong 5$ μm.

Fig. 14.13 illustrates the performance of the mode-tracking system. The output optical power of the EC–QCL as a function of time was monitored for different fractions of the control signals U_{EC} and U_{GR} (for the EC length and for the grating angle respectively) required for full mode tracking. The laser was driven by a ~ 630 mA current and modulated with a sinusoidal waveform. Three particular modes of operation can be observed: no wavelength tracking (control signals U_{EC} and U_{GR} not applied), only grating wavelength tracking (only U_{GR} applied in full), and full wavelength tracking (both U_{EC} and U_{GR} applied in full). Both QCL FP resonator mode-hops and EC FP resonator mode-hops can be observed for partial mode tracking. Increase of the PZT control signals results in progressive separation of the mode-hops, which finally leads to complete laser longitudinal mode tracking.

Examples of spectroscopy achieved using an external cavity quantum cascade laser, implementing these mode-tracking techniques, are shown in Fig. 14.14 around 5.3 μm, comparing various measurement techniques. Fig. 14.14b shows a direct absorption spectrum of the NO Q-branch recorded at a scan rate of 10 Hz for 5% NO in N2 at reduced pressure and with a 10 cm optical path length.



Quantum Cascade Lasers
Jérôme Faist

Print publication date: 2013
Print ISBN-13: 9780198528241
Published to Oxford Scholarship Online: May 2013
DOI: 10.1093/acprof:oso/9780198528241.001.0001

(p.286) Appendix A Designs

(p.286) Appendix A

Designs

A few designs discussed in the text are summarized here. The layer sequence is given starting from the injection barrier and the thickness indicated in nanometers. Barrier material is denoted in bold face, and quantum wells are in Roman. The doped layers are underlined.

A.1 First-generation designs

All these devices were grown using $Ga_{0.47}$ $In_{0.53}$ As for the wells and $Al_{0.48}$ $In_{0.52}$ As for the barrier lattices matched to the InP substrates.

A.1.1 First QCL at $\lambda \approx 4.3$ μm

This design was based on a diagonal transition with one-phonon extraction. The sequence was **4.5**, 9, **3.5**, 3.5, **3.0**, 2.8, **3.0**, 1.8, **1.2**, 1.6, **1.4**, 1.3, **1.7**, 1.1, **1.9**, 0.9, **2.1**, 0.7, **2.3**, 0.6. The Si doping level was $N_d = 1.5 \times 10^{17}$ cm^{-3}.

A.1.2 Vertical transition at 4.6 μm

This design was based on a two-quantum-well active region, with a vertical transition and a single-phonon extraction. The layer sequence was **6.8**, 4.8, **2.8**, 3.9, **2.7**, 2.2, **2.2**, 2.1, **2.1**, 2.0, **2.0**, 2.0, **2.0**, 1.8, **1.8**, 1.7, **2.0**, 1.6, **2.2**, 1.6, **2.4**, 1.4. The Si doping level was $N_d = 3 \times 10^{17}$ cm^{-3}.

A.1.3 Vertical two-quantum-well active region at 8.55 μm

This design, also based on a vertical transition and a two-quantum-well active region, was similar to the previous one but adapted for longer wavelengths. The layer sequence was **4.5**, 8.0, **1.0**, 5.7, **2.4**, 4.4, **1.4**, 3.6, **1.2**, 3.6, **1.2**, 3.4, **1.0**, 3.4. The Si doping level was $N_d = 3 \times 10^{17}$ cm^{-3}.

A.1.4 Super-diagonal transition at λ ≈ 10.5 μm

The particularity of this design was a very diagonal transition, where the upper-state of the laser transition was also the ground state of the active region. The layer sequence was **2.5**, 7.4, **2.5**, 5.5, **2.5**, 4.6, **2.4**, 4.2. The Si doping level was 9.2×10^{17} cm^{-3}.

A.1.5 Single-quantum-well active region: Bloch gain at 7.5 μm

The gain arises despite a very low population inversion in a single-quantum well. The ground state of the well was "pushed" up due to an $Al_{0.48}$ $In_{0.52}$ As spike in the center **(p.287)** of the quantum well. The layer sequence was **4.8**, 3.6, **0.2**, 3.6, **3.5**, 5.1, **1.1**, 5.0, **1.2**, 4.5, **1.3**, 3.5, **1.5**, 3.4, **1.6**, 4.4, **1.8**, 3.2, **2.1**, 3.0, **2.5**, 3.0, **2.9**, 2.9. The Si doping level was $N_d = 3 \times 10^{17}$ cm^{-3}.

A.1.6 Three-quantum-well (λ ≈ 10.3 μm) diagonal

This three-quantum-well active region design uses a diagonal transition and a single phonon extraction. The layer sequence was **4.2**, 3.1, **0.9**, 6.4, **1.0**, 6.0, **2.8**, 3.9, **1.0**, 3.8, **1.2**, 3.7, **1.5**, 3.9, **1.7**, 4.0. The Si doping level was $N_d \approx 2.5 \times 10^{17}$ cm^{-3}.

A.1.7 Three-quantum-well (λ ≈ 5.3 μm) vertical

In this three-quantum-well active region, the transition was vertical. The layer sequence was **5.0**, 0.9, **1.5**, 4.7, **2.2**, 4.0, **3.0**, 2.3, **2.3**, 2.2, **2.2**, 2.0, **2.0**, 2.0, **2.3**, 1.9, **2.8**, 1.9. The Si doping level was $N_d = 2 \times 10^{17}$ cm^{-3}.

A.2 Bound-to-continuum and two-phonon designs

The high performance was achieved using these design concepts with better population inversion and high-temperature characteristics. In these designs, $Ga_{0.47}$ $In_{0.53}$ As and $Al_{0.48}$ $In_{0.52}$ As were used for the well and barrier materials.

A.2.1 Two-phonon resonance at 9 μm

In this design, extraction from the lower state was achieved by placing a cascade of two-phonon resonances. The layer sequence of the structure was **4.0**, 1.9, **0.7**, 5.8, **0.9**, 5.7, **0.9**, 5.0, **2.2**, 3.4, **1.4**, 3.3, **1.3**, 3.2, **1.5**, 3.1, **1.9**, 3.0, **2.3**, 2.9, **2.5**, 2.9. The Si doping level was $N_d = 4 \times 10^{17}$ cm^{-3}.

A.2.2 Bound-to-continuum at 9 μm

In this design, the transition occurs between an isolated state and the lower miniband. The layer sequence was **4.0**, 2.0, **0.7**, 6.0, **0.9**, 5.9, **1.0**, 5.2, **1.3**, 3.8, **1.4**, 3.5, **1.5**, 3.3, **1.6**, 3.1, **1.9**, 3.1, **2.3**, 3.0, **2.4**, 2.9. The Si doping level was $N_d = 3 \times 10^{17}$ cm^{-3}.

A.2.3 Broad bound-to-continuum at 9 μm

This design was bound-to-continuum design specially designed for broad gain operation. The layer sequence was **3.9**, 2.2, **0.8**, 6, **0.9**, 5.9, **1**, 5.2, **1.3**, 4.3, **1.4**, 3.8, **1.5**, 3.6, **1.6**, 3.4, **1.9**, 3.3, **2.3**, 3.2, **2.5**, 3.2, **2.9**, 3.1. The SI doping level was $N_d = 2.3 \times 10^{17}$ cm^{-3}.

A.2.4 Two-phonon at 8 μm

This was another version of the two-phonon resonance design. The layer sequence was **4.3**, 1.7, **0.9**, 5.4, **1.1**, 5.3, **1.2**, 4.7, **2.2**, 4.3, **1.5**, 3.8, **1.6**, 3.4, **1.8**, 3.0, **2.1**, 2.8, **2.5**, 2.7, **3.2**, 2.7, **3.6**, 2.5. The SI doping level was $N_d = 1.5 \times 10^{17}$ cm^{-3}.

(p.288) #### A.2.5 Bound-to-continuum at 16 μm

This was a long-wavelength version of the bound-to-continuum design. The layer sequence of one period of structure was **3.3**, 3.2, **0.5**, 6.5, **0.6**, 6.6, **0.7**, 6.3, **0.8**, 5.8, **1.0**, 4.6, **1.2**, 4.4, 1.4, 4.4, 1.7, 4.2, **2.0**, 4.1, **2.2**, 4.0. The Si doping level was $N_d = 3 \times 10^{17}$ cm^{-3}

A.3 Strain-compensated designs

In these designs, strain-compensated combinations of (In,Ga,Al)As were used. The composition of the well and barrier materials were indicated.

A.3.1 Short-wavelength (λ ≈ 3.6 μm) three-quantum wells

This was a three-quantum-well active region based on a vertical transition with a single optical phonon extraction. The barriers were in $In_{0.4}$ $Al_{0.6}$ As, the quantum wells in $In_{0.7}$ $Ga_{0.3}$ As, and the layer sequence is **4.5**, 0.5, **1.2**, 3.5, **2.3**, 3.0, **2.8**, 2.0, **1.8**, 1.8, **1.8**, 1.9, **1.8**, 1.5, **2.0**, 1.5, **2.3**, 1.4, **2.5**, 1.3, **3.0**, 1.3, **3.4**, 1.2, **3.6**, 1.1. The SI doping level was $N_d = 2.5 \times 10^{17}$ cm^{-3}.

A.3.2 Three-quantum-well strain-compensated at 4.6 μm

This was a three-quantum-well active region based on a diagonal transition with a single optical phonon extraction [153]. The well material was $Ga_{0.365}$ $In_{0.635}$ As and the barriers were $Al_{0.665}$ $In_{0.335}$ As. The layer sequence was **3.5**, 1.3, **1.4**, 4.1, **1.7**, 3.7, **2.5**, 2.8, **1.5**, 2.6, **1.6**, 2.4, **1.7**, 2.2, **1.9**, 2.1, **2.1**, 2.0, **2.3**, 1.8, **2.4**, 1.8. The Si doping level was $N_d = 2 \cdot 10^{17}$ cm^{-3}.

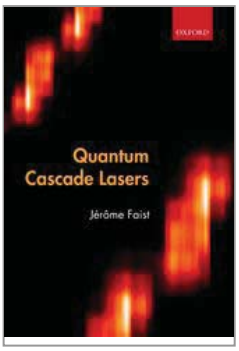
A.3.3 Short-wavelength (λ = 3.3 μm) Sb-free

This design was based on a bound-to-continuum transition. The quantum wells were made in $In_{0.74}$ $Ga_{0.26}$ As, whereas the composite barriers were fabricated using latticematched $Al_{0.48}$ $In_{0.52}$ As (italic) and AlAs (bold). The layer sequence was **0.5**, *0.7*, **1.0**, 1.2, **1.0**, 3.4, **1.0**, 3.1, **0.9**, 2.7, **0.9**, 2.3, ***0.2*****, 0.8**, 2.1, *0.3*, **0.7**, 2.0, *0.4*, **0.8**, 1.9, *0.4*, **0.8**, 1.8, *0.5*, **0.8**, 1.7, *0.7*, **0.9**, 1.2, *1.0*, **1.0**, 1.2, *1.3*, **1.0**, 1.1, *1.3*. The Si doping level was $N_d = 6 \times 10^{17}$ cm^{-3}.

A.3.4 Strain-compensated at 4.8 μ m

In this bound-to-continuum design, both wells *and* barrierswere composite. The $Ga_{0.391}$ $In_{0.609}$ As layer were in Roman, the $Al_{0.546}$ $In_{0.454}$ As in italic, the AlAs and the InAs layers in Roman bold. Composite barriers and wells are further emphasized by parentheses and brackets. The layer sequence was (*46*, 10, *14*), [20, **2**, 21],(*7*, **2**, *7*), [19, **2**, 18], (*8*, **2**, *7*), [13, **2**, 2, **2**, 13], (*9*, **2**, *2*, **2**, *2*), [11, **2**, 2, **2**, 10], (*7*, **2**, *2*, **2**, *8*), [12, **2**, 12], (*10*, **2**, *10*), 24, *23*,

22, 25, 21, *30*, 20, *33*, 19, *37*, 18. The Si doping level was adjusted such that the sheet carrier density per period was $n_s = 9 \times 10^{10}$ cm^{-2}.



Quantum Cascade Lasers
Jérôme Faist

Print publication date: 2013
Print ISBN-13: 9780198528241
Published to Oxford Scholarship Online: May 2013
DOI: 10.1093/acprof:oso/9780198528241.001.0001

(p.289) References

Bibliography references:

[1] W. P. McCray, "MBE deserves a place in the history books," *Nat Nanotechnol*., vol. 2, no. 5, pp. 259–261, 2007.

[2] L. Esaki and R. Tsu, "Superlattice and negative differential conductivity in semiconductors," *IBM J. Res. Develop*., vol. 14, no. 61, pp. 61–65, 1970.

[3] R. Dingle, W. Wiegmann, and C. Henry, "Quantum states of confined carriers in very thin $Al_xGa1_{-x}As$-GaAs-$Al_xGa1_{-x}As$ heterostructures," *Phys. Rev. Lett*., vol. 33, pp. 827–830, 1974.

[4] C. Henry, *The origin of quantum wells and the quantum well laser*. Academic Press, P.S. Zory ed., Feb. 1993.

[5] T. Ando, A. Fowler, and F. Stern, "Electronic properties of two-dimensional systems," *Reviews of Modern Physics*, vol. 54, no. 2, pp. 437–672, 1982.

[6] L. West and S. Eglash, "First observation of an extremely large-dipole infrared transition within the conduction band of a GaAs quantum well," *Applied Physics Letters*, vol. 46, no. 12, pp. 1156–1158, 1985.

[7] R. Kazarinov and R. Suris, "Possibility of the amplification of electromagnetic waves in a semiconductor with a superlattice," *Sov. Phys. Semicond*., vol. 5, no. 4, pp. 707–709, 1971.

[8] R. Kazarinov and R. Suris, “Electric and electromagnetic properties of semiconductors with a superlattice,” *Sov. Phys. Semicond*., vol. 6, no. 1, pp. 120–131, 1972.

[9] F. Capasso, K. Mohammed, and A. Cho, “Resonant tunneling through double barriers, perpendicular quantum transport phenomena in superlattices, and their device applications,” *IEEE J. Quantum Electron*., vol. 22, no. 9, pp. 1853–1869, 1986.

[10] H. Liu, “A novel superlattice infrared source,” *J. Appl. Phys*., vol. 63, no. 8, pp. 2856–2858, 1988.

[11] A. Kastalsky, V. Goldman, and J. Abeles, “Possibility of infrared laser in a resonant tunneling structure,” *Applied Physics Letters*, vol. 59, no. 21, pp. 2636–2638, 1991.

[12] A. Kastalsky and A. Efros, “Magnetic field-induced suppression of acoustic phonon emission in a superlattice,” *J. Appl. Phys*., vol. 69, no. 2, pp. 841–845, 1991.

[13] Q. Hu and S. Feng, “Feasibility of far-infrared lasers using μ ltiple semiconductor quantum wells,” *Applied Physics Letters*, vol. 59, no. 23, pp. 2923–2925, 1991.

[14] B. Lax, *Quantum Electronics, A Symposium*. New York: Columbia University, 1960.

[15] A. Andronov, “Proposed optical-phonon-mediated population inversion and stimulated FIR emission in superlattices,” *Semicond. Sci. Technol*., vol. 7, pp. B629–B632, 1992.

[16] E. Gornik and D. Tsui, “Voltage-tunable far-infrared emission from Si inversion layers,” *Phys. Rev. Lett*., vol. 37, no. 21, pp. 1425–1428, 1976.

[17] M. Helm, P. England, E. Colas, F. DeRosa, and S. Allen, “Intersubband emission from semiconductor superlattices excited by sequential resonant tunneling,” *Phys. Rev. Lett*., vol. 63, no. 1, pp. 74–77, 1989.

[18] H. Schneider and H. Liu, *Quantum Well Infrared Detectors, Physics and Applications*. Springer Series in Optical Science, 1st ed., 2007.

[19] B. Levine, K. Choi, C. Bethea, J. Walker, and R. Malik, “New 10 μ m infrared detector using intersubband absorptioon in resonant tunneling GaAlAs superlattices,” *Applied Physics Letters*, vol. 50, no. 16, pp. 1092–1094, 1987.

[20] J. Faist, F. Capasso, D. Sivco, C. Sirtori, A. Hutchinson, and A. Cho, “Quantum cascade laser,” *Science*, vol. 264, pp. 553–556, 1994.

[21] J. Faist, C. Gmachl, F. Capasso, C. Sirtori, D. Sivco, J. Baillargeon, and A. Cho, “Distributed feedback quantum cascade lasers,” *Applied Physics Letters*, vol. 70, no. 20, pp. 2670–2672, 1997.

[22] M. Beck, D. Hofstetter, T. Aellen, J. Faist, U. Oesterle, M. Ilegems, E. Gini, and H. Melchior, “Continuous wave operation of a mid-infrared semiconductor laser at room temperature,” *Science*, vol. 295, no. 5553, pp. 301–305, 2002.

[23] R. Köhler, A. Tredicucci, F. Beltram, H. Beere, E. Linfield, A. Davies, D. Ritchie, R. Iotti, and F. Rossi, "Terahertz semiconductor-heterostructure laser," *Nature*, vol. 417, pp. 156–159, 2002.

(p.290) [24] R. Yang, "Infrared laser based on intersubband transitions in quantum wells," *Superlatt. Microstruct.*, vol. 17, no. 1, pp. 77–77, 1995.

[25] C. L. Canedy, W. W. Bewley, M. Kim, C. S. Kim, J. A. Nolde, D. C. Larrabee, J. R. Lindle, I. Vurgaftman, and J. R. Meyer, "High-temperature interband cascade lasers emitting at λ=3.6- 4.3 μ m," *Applied Physics Letters*, vol. 90, no. 18, p. 181120, 2007.

[26] M. Tacke, "New developments and applications of tunable ir lead salt lasers," *Infrared Phys. Technol.*, vol. 36, pp. 447–463, 1995.

[27] J. Roberts, R. Green, L. Wilson, E. Zibik, D. Revin, J. Cockburn, and R. Airey, "Quantum cascade lasers grown by metalorganic vapor phase epitaxy," *Applied Physics Letters*, vol. 82, no. 24, pp. 4221–4223, 2003.

[28] L. Diehl, D. Bour, S. Corzine, J. Zhu, G. Höfler, M. Lonèar, M. Troccoli, and F. Capasso, "High-power quantum cascade lasers grown by low-pressure metal organic vapor-phase epitaxy operating in continuous wave above 400K," *Applied Physics Letters*, vol. 88, no. 20, p. 201115, 2006.

[29] D. Bowen and B. K. Tanner, *High Resolution X-Ray Diffractometry and Topography*. CRC Press, 1st ed., Feb. 1998.

[30] V. Holy, U. Pietsch, and T. Baumbach, *High-Resolution X-Ray Scattering from Thin Films and Multilayers*. Springer-Verlag Telos, Mar. 1999.

[31] A. Bismuto, M. Beck, and J. Faist, "High power Sb-free quantum cascade laser emitting at 3.3μ m above 350 K," *Applied Physics Letters*, vol. 98, no. 19, p. 191104, 2011.

[32] J. Faist, F. Capasso, C. Sirtori, D. Sivco, J. Baillargeon, A. Hutchinson, S. Chu, and A. Cho, "High power mid-infrared (λ greater than or similar to 5 μ m) quantum cascade lasers operating above room temperature," *Applied Physics Letters*, vol. 68, no. 26, pp. 3680–3682, 1996.

[33] G. Dehlinger, L. Diehl, U. Gennser, H. Sigg, J. Faist, K. Ensslin, D. Grutzmacher, and E. Muller, "Intersubband electroluminescence from silicon-based quantum cascade structures," *Science*, vol. 290, no. 5500, pp. 2277–+, 2000.

[34] S. Tsujino, A. Borak, C. Falub, T. Fromherz, L. Diehl, H. Sigg, and D. Grutzmacher, "Intravalence-band mixing in strain-compensated SiGe quantum wells," *Physical Review B*, vol. 72, no. 15, pp. –, 2005.

[35] G. Binnig and H. Rohrer, "Scanning tunneling microscopy," *Ibm Journal of Research and Development*, vol. 44, no. 1.2, pp. 279–293, 2000.

[36] P. Offermans, P. Koenraad, J. Wolter, M. Beck, T. Aellen, and J. Faist, “Digital alloy interface grading of an InAlAs/InGaAs quantum cascade laser structure studied by cross-sectional scanning tunneling microscopy,” *Applied Physics Letters*, vol. 83, no. 20, pp. 4131–4133, 2003.

[37] E. Muller, J. Panitz, and S. McLane, “The AtomProbe Field Ion Microscope,” *Review of Scientific Instruments*, vol. 39, no. 1, pp. 83–86, 1968.

[38] T. Kelly, D. Larson, K. Thompson, R. Alvis, J. Bunton, J. Olson, and B. Gorman, “Atom probe tomography of electronic materials,” *Annu. Rev. Mater. Res*., vol. 37, pp. 681–727, 2007.

[39] H. Beere, J. Fowler, J. Alton, E. Linfield, D. Ritchie, R. Kohler, A. Tredicucci, G. Scalari, L. Ajili, J. Faist, and S. Barbieri, “MBE growth of terahertz quantum cascade lasers,” *Journal of Crystal Growth*, vol. 278, no. 1-4, pp. 756–764, 2005.

[40] T. Aellen, M. Beck, N. Hoyler, M. Giovannini, J. Faist, and E. Gini, “Doping in quantum cascade lasers. I. InAlAs-InGaAs/InP midinfrared devices,” *Journal of Applied Physics*, vol. 100, no. 4, p. 043101, 2006.

[41] C. Gmachl, F. Capasso, J. Faist, A. Hutchinson, A. Tredicucci, D. Sivco, J. Baillargeon, S. Chu, and A. Cho, “Continuous-wave and high-power pulsed operation of index-coupled distributed feedback quantum cascade laser at λ approximate to 8.5 μ m,” *Applied Physics Letters*, vol. 72, no. 12, pp. 1430–1432, 1998.

[42] J. Faist, A. Muller, M. Beck, D. Hofstetter, S. Blaser, U. Oesterle, and M. Ilegems, “A quantum cascade laser based on an n-i-p-i superlattice,” *IEEE Photonics Technology Letters*, vol. 12, no. 3, pp. 263–265, 2000.

[43] C. Gmachl, H. Hwang, R. Paiella, D. Sivco, J. Baillargeon, F. Capasso, and A. Cho, “Quantum cascade lasers with low-loss chalcogenide lateral waveguides,” *IEEE Photonics Technology Letters*, vol. 13, no. 3, pp. 182–184, 2001.

[44] A. Bismuto, “Mid-infrared quantum cascade lasers: active medium andwaveguide engineering,” *Diss ETH 19924*, pp. 1–198, 2011.

[45] A. Wittmann, “High-performance quantum cascade laser sources for spectroscopic applications,” *DISS. ETH Nr. 18363*, pp. 1–188, 2009.

[46] P. Klein, J. Furneaux, and R. Henry, “Time-dependent photoluminescence of InP: Fe,” *Physical Review B*, vol. 29, no. 4, p. 1947, 1984.

(p.291) [47] Picture courtesy of B. Hinkov and Y. Bonetti.

[48] J. Chelikovsky and M. Cohen, “Nonlocal pseudopotential calculations for the eletronic structure of eleven diamond and zinc-blende semiconductors,” *Physical Review B*, vol. 14, no. 2, pp. 556– 582, 1976.

[49] G. Bastard, *Wave Mechanics Applied to Semiconductor Heterostructures*. Les Ulis, France: Les éditions de physique, 1988.

[50] E. Kane, “Band Structure of Indium Antimonide,” *Journal of Physics and Chemistry of Solids*, vol. 1, no. 4, pp. 249–261, 1957.

[51] C. Sirtori, F. Capasso, J. Faist, and S. Scandolo, “Nonparabolicity and a sum rule associated with bound-to-bound and bound-to-continuum intersubband transitions in quantum wells,” *Phys. Rev. B*, vol. 50, no. 12, pp. 8663–8674, 1994.

[52] L. Diehl, “Development of Si/SiGe quantum cascade emitters,” *PhD Thesis, University of Neuchatel*, pp. 1–143, 2003.

[53] H. Liu, M. Buchanan, and Z. Wasilewski, “How good is the polarization selection rule for intersubband transitions?,” *Applied Physics Letters*, vol. 72, no. 14, pp. 1682–1684, 1998.

[54] C. Sirtori, F. Capasso, D. Sivco, and A. Cho, “Giant, triply resonant, third-order nonlinear susceptibility chi-3-omega(3) in coupled quantum wells,” *Physical Review Letters*, vol. 68, no. 7, pp. 1010–1013, 1992.

[55] F. Capasso, C. Sirtori, J. Faist, D. Sivco, S. Chu, and A. Cho, “Observation of an electronic bound state above a potential well,” *Nature*, vol. 358, no. 6387, pp. 565–567, 1992.

[56] R. Yang, J. Xu, and M. Sweeny, “Selection rules of intersubband transitions in conduction-band quantum wells,” *Physical Review B*, vol. 50, no. 11, pp. 7474–7482, 1994.

[57] J. Faist, C. Sirtori, F. Capasso, L. Pfeiffer, and K. West, “Phonon limited intersubband lifetimes and linewidths in a two-dimensional electron gas,” *Applied Physics Letters*, vol. 64, no. 7, pp. 872–874, 1994.

[58] S. Luin, V. Pellegrini, F. Beltram, X. Marcadet, and C. Sirtori, “Interplay between disorder and intersubband collective excitations in the two-dimensional electron gas,” *Physical Review B*, vol. 64, no. 4, p. 041306, 2001.

[59] K. Campman, H. Schmidt, A. Imamoglu, and A. Gossard, “Interface roughness and alloy-disorder scattering contributions to intersubband transition linewidths,” *Applied Physics Letters*, vol. 69, no. 17, pp. 2554–2556, 1996.

[60] T. Unuma, M. Yoshita, T. Noda, H. Sakaki, and H. Akiyama, “Intersubband absorption linewidth in GaAs quantum wells due to scattering by interface roughness, phonons, alloy disorder, and impurities,” *J. Appl. Phys*., vol. 93, no. 3, pp. 1586–1597, 2003.

[61] S. Tsujino, A. Borak, E. Muller, M. Scheinert, C. Falub, H. Sigg, D. Grutzmacher, M. Giovannini, and J. Faist, “Interface-roughness-induced broadening of intersubband

electroluminescence in p-SiGe and n-GaInAs/AlInAs quantum-cascade structures," *Applied Physics Letters*, vol. 86, no. 6, p. 062113, 2005.

[62] A. Wittmann, Y. Bonetti, J. Faist, E. Gini, and M. Giovannini, "Intersubband linewidths in quantum cascade laser designs," *Applied Physics Letters*, vol. 93, no. 14, p. 141103, 2008.

[63] R. Warburton, C. Gauer, A. Wixforth, J. Kotthaus, B. Brar, and H. Kroemer, "Intersubband resonances in InAs/AlSb quantum wells: Selection rules, matrix elements, and the depolarization field," *Physical Review B*, vol. 53, no. 12, pp. 7903–7910, 1996.

[64] R. Warburton, K. Weilhammer, J. Kotthaus, M. Thomas, and H. Kroemer, "Influence of collective effects on the linewidth of intersubband resonance," *Physical Review Letters*, vol. 80, no. 10, pp. 2185–2188, 1998.

[65] A. Harwit and J. Harris, "Observation of Stark Shifts in Quantum-Well Intersubband Transitions," *Applied Physics Letters*, vol. 50, no. 11, pp. 685–687, 1987.

[66] Y. Mii, R. Karunasiri, K. Wang, M. Chen, and P. Yuh, "Large stark shifts of the local to global state intersubband transitions in step quantum-wells," *Applied Physics Letters*, vol. 56, no. 20, pp. 1986–1988, 1990.

[67] J. Faist, F. Capasso, C. Sirtori, D. Sivco, A. Hutchinson, and A. Cho, "Laser action by tuning the oscillator strength," *Nature*, vol. 387, no. 6635, pp. 777–782, 1997.

[68] A. Muller, M. Beck, J. Faist, U. Oesterle, and M. Ilegems, "Electrically tunable, room-temperature quantum-cascade lasers," *Applied Physics Letters*, vol. 75, no. 11, pp. 1509–1511, 1999.

[69] Y. Duboz, B. Rosencher, and N. Laurent, "Monolithic spectrometer based on intersubband transitions in quantum wells," *IEEE Journal of Quantum Electronics*, vol. 30, no. 7, pp. 1507– 1510, 1994.

(p.292) [70] C. Sirtori, F. Capasso, D. Sivco, A. Hutchinson, and A. Cho, "Resonant stark tuning of second-order susceptibility in coupled quantum wells," *Applied Physics Letters*, vol. 60, no. 2, pp. 151– 153, 1992.

[71] T. Ando, "Broadening of Inter-Subband Transitions in Image-Potential-Induced Surface-States Outside Liquid-Helium," *Journal of the Physical Society of Japan*, vol. 44, no. 3, pp. 765–773, 1978.

[72] R. Terazzi, "Transport in quantum cascade lasers," *Diss ETH No. 20036*, pp. 1–337, 2011.

[73] P. Price, "Two-dimensional electron transport in semiconductor layers i: phonon scattering," *Ann. Phys. (San Diego)*, vol. 133, no. 2, pp. 217–239, 1981.

[74] J. Faist, F. Capasso, C. Sirtori, D. Sivco, A. Hutchinson, S. Chu, and A. Cho, "Measurement of the intersubband scattering rate in semiconductor quantum wells by excited state differential absorption spectroscopy," *Applied Physics Letters*, vol. 63, no. 10, pp. 1354–1356, 1993.

[75] S. Lutgen, R. Kaindl, M. Woerner, T. Elsaesser, A. Hase, H. Kunzel, M. Gulia, D. Meglio, and P. Lugli, "Nonequilibrium dynamics in a quasi-two-dimensional electron plasma after ultrafast intersubband excitation," *Physical Review Letters*, vol. 77, no. 17, pp. 3657–3660, 1996.

[76] S. Lutgen, R. Kaindl, M. Woerner, T. Elsaesser, A. A.Hase, and H. Kunzel, "Nonlinear intersubband absorption of a hot quasi-two-dimensional electron plasma studied by femtosecond infrared spectroscopy," *Phys. Rev. B*, vol. 54, no. 24, pp. R17343–R17346, 1996.

[77] R. Ferreira and G. Bastard, "Evaluation of some scattering times for electrons in unbiased and biased single-and μ ltiple-quantum-well structures," *Phys. Rev. B*, vol. 40, no. 2, pp. 1074–1086, 1989.

[78] J. Faist, F. Capasso, C. Sirtori, D. Sivco, A. Hutchinson, S. Chu, and A. Cho, "Narrowing of the intersubband electroluminescent spectrum in coupled-quantum-well heterostructures, "*Applied Physics Letters*, vol. 65, no. 1, pp. 94–96, 1994.

[79] G. Scamarcio, F. Capasso, C. Sirtori, J. Faist, A. Hutchinson, D. Sivco, and A. Cho, "High-power infrared (8-micrometer wavelength) superlattice lasers," *Science*, vol. 276, no. 5313, pp. 773– 776, 1997.

[80] G. Bastard, "Energy levels and alloy scattering in InPIn (Ga) As heterojunctions," *Applied Physics Letters*, vol. 43, no. 6, pp. 591–593, 1983.

[81] T. Ando, "Self-Consistent Results for a GaAs/AlGaAs Heterojunction. II. Low Temperature Mobility," *Journal of the Physical Society of Japan*, vol. 51, no. 12, pp. 3900–3907, 1982.

[82] A. Bismuto, R. Terazzi, M. Beck, and J. Faist, "Influence of the growth temperature on the performances of strain-balanced quantum cascade lasers," *Applied Physics Letters*, vol. 98, no. 9, p. 091105, 2011.

[83] S. Borenstain and J. Katz, "Intersubband auger recombination and population inversion in quantum-well subbands," *Phys. Rev. B*, vol. 39, no. 15, pp. 10852–10857, 1989.

[84] J. Smet, C. Fonstad, and Q. Hu, "Intrawell and interwell intersubband transitions in μ ltiple quantum wells for far-infrared sources," *J. Appl. Phys*., vol. 79, no. 12, pp. 9305–9320, 1996.

[85] B. S. Williams, "Terahertz quantum cascade lasers," *Ph.D Thesis, MIT*, p. 310, 2003.

[86] P. Hyldgaard and J. Wilkins, “Electron-electron scattering in far-infrared quantum cascade lasers,” *Phys. Rev. B*, vol. 53, no. 11, pp. 6889–6892, 1996.

[87] D. Oberli, D. Wake, M. Klein, J. Klem, T. Henderson, and H. Morkoc, “Time-resolved raman scattering in GaAs quantum wells,” *Phys. Rev. Lett*., vol. 59, no. 6, pp. 696–699, 1987.

[88] M. Tatham, J. Ryan, and C. Foxon, “Time-resolved raman measurements of intersubband relaxation in GaAs quantum wells,” *Phys. Rev. Lett*., vol. 63, no. 15, pp. 1637–1640, 1989.

[89] A. Seilmeier, H. Hubner, G. Abstreiter, G. Weimann, and W. Schlapp, “Intersubband relaxation in GaAs-Al_xGa1-_xAs quantum well structures observed directly by an infrared bleaching technique,” *Phys. Rev. Lett*., vol. 59, no. 12, pp. 1345–1348, 1987.

[90] F. Julien, J. Lourtioz, N. Herschkorn, D. Delacourt, J. Pocholle, M. Papuchon, R. Planel, and G. Leroux, “Optical saturation of intersubband absorption in GaAs-AlxGa1-xAs quantum wells,” *Applied Physics Letters*, vol. 53, no. 2, pp. 116–118, 1988.

[91] F. Julien, J. Lourtioz, N. Herschkorn, D. Delacourt, J. Pochoclle, M. Papuchon, R. Planel, and G. Leroux, “Optical saturation of intersubband absorption in GaAs-AlxGa1-xAs quantum wells (vol. 53, p. 116, 1988),” *Applied Physics Letters*, vol. 62, no. 18, pp. 2289–2289, 1993.

[92] H. Grahn, H. Schneider, W. Ruhle, K. Von Klitzing, and K. Ploog, “Nonthermal occupation of higher subbands in seconductor superlattices via sequential resonant tunneling,” *Physical Review Letters*, vol. 64, no. 20, pp. 2426–2429, 1990.

[93] T. Asano, M. Tamura, S. Yoshizawa, and S. Noda, “Pump–probe measurement of ultrafast all-optical modulation based on intersubband transition in n-doped quantum wells,” *Applied Physics Letters*, vol. 77, no. 1, pp. 19–21, 2000.

(p.293) [94] J. Levenson, G. Dolique, J. Oudar, and I. Abram, “Time resolved inter-subband relaxation in GaAs/GaAlAs μ ltiple quantum wells,” *Solid-state electronics*, vol. 32, no. 12, pp. 1869–1873, 1989.

[95] B. Murdin, G. Knippels, A. Van der Meer, C. Pidgeon, C. Langerak, M. Helm, W. Heiss, K. Unterrainer, E. Gornik, K. Geerinck, N. Hovenier, and W. Weckebach, “Excite-probe determination of the intersubband lifetime in wide GaAs/AlGaAs quantum-wells using a far-infrared free electron laser,” *Semiconductor Science and Technology*, vol. 9, no. 8, pp. 1554–1557, 1994.

[96] M. Hartig, J. Ganière, P. Selbmann, B. Deveaud, and L. Rota, “Density dependence of carriercarrier-induced intersubband scattering in GaAs/Al_xGa1-_xAs quantum wells,” *Phys. Rev. B*, vol. 60, no. 3, pp. 1500–1503, 1999.

[97] M. Rochat, J. Faist, M. Beck, and U. Oesterle, “Electrically pumped Terahertz quantum well sources,” *Physica E-Low-Dimensional Systems & Nanostructures*, vol. 7, no. 1-2, pp. 44–47, 2000.

[98] J. Faist, F. Capasso, L. Pfeiffer, and K. West, “Phonon limited intersubband lifetimes and linewidths in a two-dimensional electron gas,” *Applied Physics Letters*, vol. 64, no. 7, pp. 872– 874, 1994.

[99] J. Heyman, K. Unterrainer, K. Craig, B. Galdrikian, M. Sherwin, K. Campman, P. Hopkins, and A. Gossard, “Temperature and intensity dependence of intersubband relaxation rates from photovoltage and absorption,” *Physical Review Letters*, vol. 74, no. 14, pp. 2682–2685, 1995.

[100] J. Heyman, K. Unterrainer, K. Craig, J. Williams, M. Sherwin, K. Campman, P. Hopkins, A. Gossard, B. Murdin, and C. Langerak, “Far-infrared pump–probe measurements of the intersubband lifetime in an AlGaAs/GaAs coupled-quantum well,” *Applied Physics Letters*, vol. 68, no. 21, pp. 3019–3021, 1996.

[101] M. Hartig, S. Haacke, P. Selbmann, B. Deveaud, R. Taylor, and L. Rota, “Efficient intersubband scattering via carrier-carrier interaction in quantum wells,” *Phys. Rev. Lett*., vol. 80, no. 9, pp. 1940–1943, 1998.

[102] J. Faist, F. Capasso, C. Sirtori, D. Sivco, A. Hutchinson, S. Chu, and A. Cho, “Mid-infrared field-tunable intersubband electroluminescence at room temperature by photon-assisted tunneling in coupled-quantum wells,” *Applied Physics Letters*, vol. 64, no. 9, pp. 1144–1146, 1994.

[103] E. Benveniste, A. Vasanelli, A. Delteil, J. Devenson, R. Teissier, A. Baranov, A. M. Andrews, G. Strasser, I. Sagnes, and C. Sirtori, “Influence of the material parameters on quantum cascade devices,” *Applied Physics Letters*, vol. 93, no. 13, p. 131108, 2008.

[104] L. Landau and E. Lifchitz, *Mécanique quantique, théorie non relativiste*. Moscou: Editions Mir, 1967.

[105] A. Vasanelli, A. Leuliet, C. Sirtori, A. Wade, G. Fedorov, D. Smirnov, G. Bastard, B. Vinter, M. Giovannini, and J. Faist, “Role of elastic scattering mechanisms in GaInAs/AlInAs quantum cascade lasers,” *Applied Physics Letters*, vol. 89, no. 17, p. 172120, 2006.

[106] S. Blaser, M. Rochat, M. Beck, J. Faist, and U. Oesterle, “Far-infrared emission and stark-cyclotron resonances in a quantum cascade structure based on photon-assisted tunneling transition,” *Phys. Rev. B*, vol. 61, no. 12, pp. 8369–8374, 2000.

[107] J. Ulrich, R. Zobl, K. Unterrainer, G. Strasser, and E. Gornik, “Magnetic-field-enhanced quantum-cascade emission,” *Applied Physics Letters*, vol. 76, no. 1, pp. 19–21, 2000.

[108] C. Becker, C. Sirtori, O. Drachenko, V. Rylkov, D. Smirnov, and J. Leotin, “GaAs quantum box cascade lasers,” *Applied Physics Letters*, vol. 81, no. 16, pp. 2941–2943, 2002.

[109] C. Becker, C. Sirtori, H. Page, A. Robertson, V. Ortiz, and X. Marcadet, “Influence of confined phonon modes on the thermal behavior of AlAs/GaAs quantum cascade structures,” *Physical Review B*, vol. 65, no. 8, p. 4, 2002.

[110] A. Leuliet, A. Vasanelli, A. Wade, G. Fedorov, D. Smirnov, G. Bastard, and C. Sirtori, “Electron scattering spectroscopy by a high magnetic field in quantum cascade lasers,” *Physical Review B*, vol. 73, no. 8, p. 9, 2006.

[111] H. E. Tureci, L. Ge, S. Rotter, and A. D. Stone, “Strong interactions in μ ltimode random lasers,” *Science*, vol. 320, no. 5876, pp. 643–646, 2008.

[112] C. Sirtori, C. Gmachl, F. Capasso, J. Faist, D. Sivco, A. Hutchinson, and A. Cho, “Long-wavelength (λ = 8-11.5 μ m) semiconductor lasers with waveguides based on surface plasmons,” *Opt. Lett*., vol. 23, no. 17, pp. 1366–1368, 1998.

[113] J. Faist, C. Sirtori, F. Capasso, D. Sivco, J. Baillargeon, A. Hutchinson, and A. Cho, “High-power long-wavelength (λ = 11.5 μ m) quantum cascade lasers operating above room temperature,” *IEEE Photon. Technol. Lett*., vol. 10, no. 8, pp. 1100–1102, 1998.

[114] R. Colombelli, K. Srinivasan, M. Troccoli, O. Painter, C. Gmachl, D. Tennant, A. Sergent, D. Sivco, A. Cho, and F. Capasso, “Quantum cascade surface-emitting photonic crystal laser,” *Science*, vol. 302, pp. 1374–1377, 2003.

(p.294) [115] A. Bousseksou, Y. Chassagneux, J. R. Coudevylle, R. Colombelli, C. Sirtori, G. Patriarche, G. Beaudoin, and I. Sagnes, “Surface-plasmon distributed-feedback quantum cascade lasers operating pulsed, room temperature,” *Applied Physics Letters*, vol. 95, no. 9, p. 091105, 2009.

[116] C. Sirtori, P. Kruck, S. Barbieri, H. Page, J. Nagle, M. Beck, J. Faist, and U. Oesterle, “Low-loss al-free waveguides for unipolar semiconductor lasers,” *Applied Physics Letters*, vol. 75, no. 25, pp. 3911–3913, 1999.

[117] J. Devenson, D. Barate, O. Cathabard, R. Teissier, and A. N. Baranov, “Very short wavelength (λ=3.1-3.3 μ m) quantum cascade lasers,” *Applied Physics Letters*, vol. 89, no. 19, p. 191115, 2006.

[118] C. Gmachl, A. Tredicucci, F. Capasso, A. Hutchinson, D. Sivco, J. Baillargeon, and A. Cho, “High-power gimel approximate to 8 μ m quantum cascade lasers with near optimum performance,” *Applied Physics Letters*, vol. 72, no. 24, pp. 3130–3132, 1998.

[119] C. Faugeras, S. Forget, E. Boer-Duchemin, H. Page, J. Bengloan, O. Parillaud, M. Calligaro, C. Sirtori, M. Giovannini, and K. Faist, “High-power room temperature emission

quantum cascade lasers at a λ=9 μ m," *IEEE Journal of Quantum Electronics*, vol. 41, no. 12, pp. 1430– 1438, 2005.

[120] H. Page, S. Dhillon, M. Calligaro, C. Becker, V. Ortiz, and C. Sirtori, "Improved CW operation of GaAs-Based QC lasers: T-max=150 K," *IEEE Journal of Quantum Electronics*, vol. 40, no. 6, pp. 665–672, 2004.

[121] A. Bismuto, T. Gresch, A. Bächle, and J. m. Faist, "Large cavity quantum cascade lasers with InP interstacks," *Applied Physics Letters*, vol. 93, no. 23, p. 231104, 2008.

[122] C. Gmachl, A. Sergent, A. Tredicucci, F. Capasso, A. Hutchinson, D. Sivco, J. Baillargeon, S. Chu, and A. Cho, "Improved CW operation of quantum cascade lasers with epitaxial-side heat-sinking," *IEEE Photonics Technology Letters*, vol. 11, no. 11, pp. 1369–1371, 1999.

[123] A. Lops, V. Spagnolo, and G. Scamarcio, "Thermal modeling of GaInAs/AlInAs quantum cascade lasers," *Journal of Applied Physics*, vol. 100, p. 043109, 2006.

[124] K. Choi, B. Levine, R. Malik, J. Walker, and C. Bethea, "Periodic negative conductance by sequential resonant tunneling through an expanding high-field superlattice domain," *Phys. Rev. B*, vol. 35, no. 8, pp. 4172–4175, 1987.

[125] J. Faist, F. Capasso, C. Sirtori, D. Sivco, A. Hutchinson, and A. Cho, "Laser action by tuning the oscillator strength," *Nature*, vol. 387, pp. 777–782, 1997.

[126] J. Faist, F. Capasso, C. Sirtori, D. Sivco, A. Hutchinson, and A. Cho, "Vertical transition quantum cascade laser with bragg confined excited state," *Applied Physics Letters*, vol. 66, no. 5, pp. 538–540, 1995.

[127] A. Tredicucci, F. Capasso, C. Gmachl, D. Sivco, A. Hutchinson, and A. Cho, "High performance interminiband quantum cascade lasers with graded superlattices," *Applied Physics Letters*, vol. 73, no. 15, pp. 2101–2103, 1998.

[128] M. Troccoli, G. Scamarcio, V. Spagnolo, A. Tredicucci, C. Gmachl, F. Capasso, D. Sivco, A. Cho, and M. Striccoli, "Electronic distribution in superlattice quantum cascade lasers," *Applied Physics Letters*, vol. 77, no. 8, pp. 1088–1090, 2000.

[129] P. Kruck, H. Page, C. Sirtori, S. Barbieri, M. Stellmacher, and J. Nagle, "Improved temperature performance of $Al_{0.33}Ga_{0.67}As$/GaAs quantum-cascade lasers with emission wavelength at $\lambda \sim 11\mu$ m," *Applied Physics Letters*, vol. 76, no. 23, pp. 3340–3342, 2000.

[130] J. Faist, M. Beck, T. Aellen, and E. Gini, "Quantum-cascade lasers based on a bound-tocontinuum transition," *Applied Physics Letters*, vol. 78, no. 2, pp. 147–149, 2001.

[131] M. A. Belkin, F. Capasso, F. Xie, A. Belyanin, M. Fischer, A. Wittmann, and J. Faist, "Room temperature terahertz quantum cascade laser source based on intracavity difference-frequency generation," *Applied Physics Letters*, vol. 92, no. 20, p. 201101, 2008.

[132] C. Gmachl, F. Capasso, A. Tredicucci, D. Sivco, R. Köhler, A. Hutchinson, and A. Cho, “Dependence of the device performance on the number of stages in quantum-cascade lasers,” *IEEE J. Select. Topics Quantum Electron*., vol. 5, no. 3, pp. 808–816, 1999.

[133] J. Faist, M. Beck, T. Aellen, and E. Gini, “Quantum cascade lasers based on a bound-tocontinuum transition,” *Applied Physics Letters*, vol. 78, no. 2, pp. 147–149, 2001.

[134] T. Unuma, T. Takahashi, T. Noda, M. Yoshita, H. Sakaki, M. Baba, and H. Akiyama, “Effects of interface roughness and phonon scattering on intersubband absorption linewidth in a GaAs quantum well,” *Applied Physics Letters*, vol. 78, no. 22, pp. 3448–3450, 2001.

[135] J. Yu, S. Darvish, A. Evans, J. Nguyen, S. Slivken, and M. Razeghi, “Room-temperature continuous-wave operation of quantum-cascade lasers at $\lambda \sim 4\mu$ m,” *Applied Physics Letters*, vol. 88, pp. 041111–1–041111–3, 2006.

(p.295) [136] J. Yu, A. Evans, S. Slivken, S. Darvish, and M. Razeghi, “Short wavelength ($\lambda \sim 4.3\mu$ m) high-performance continuous-wave quantum-cascade lasers,” *IEEE Photon. Technol. Lett*., vol. 17, no. 06, pp. 1154–1156, 2005.

[137] A. Evans, J. Yu, S. Slivken, and M. Razeghi, “Continuous-wave operation of $\lambda \sim 4.8\mu$ m quantum-cascade lasers at room temperature,” *Applied Physics Letters*, vol. 85, no. 12, pp. 2166–2168, 2004.

[138] A. Evans, J. Nguyen, S. Slivken, J. Yu, S. Darvish, and M. Razeghi, “Quantum-cascade lasers operating in continuous-wave mode above 90° c at $\lambda \sim 5.25\mu$ m,” *Applied Physics Letters*, vol. 88, pp. 051105–1–051105–3, 2006.

[139] L. Diehl, D. Bour, S. Corzine, J. Zhu, G. Höfler, M. Loncar, M. Troccoli, and F. Capasso, “High-temperature continuous wave operation of strain-balanced quantum cascade lasers grown by metal organic vapor-phase epitaxy,” *Applied Physics Letters*, vol. 89, pp. 081101–1–081101–3, 2006.

[140] A. Evans, J. Yu, J. David, L. Doris, K. Mi, S. Slivken, and M. Razeghi, “High-temperature, high-power, continuous-wave operation of buried heterostructure quantum-cascade lasers,” *Applied Physics Letters*, vol. 84, no. 03, pp. 314–316, 2004.

[141] R. Maulini, A. Lyakh, and A. Tsekoun, “λ 7.1 μ m quantum cascade lasers with 19% wall-plug efficiency at room temperature,” *Optics Express*, vol. 19, no. 18, p. 17203, 2011.

[142] Y. Bai, N. Bandyopadhyay, S. Tsao, S. Slivken, and M. Razeghi, “Room temperature quantum cascade lasers with 27% wall plug efficiency,” *Applied Physics Letters*, vol. 98, no. 18, p. 181102, 2011.

[143] M. Rochat, D. Hofstetter, M. Beck, and J. Faist, “Long-wavelength ($\lambda \sim 16\mu$ m), room-temperature, single-frequency quantum-cascade lasers based on a bound-to-continuum transition,” *Applied Physics Letters*, vol. 79, no. 26, pp. 4271–4273, 2001.

[144] T. Gresch, M. Giovannini, N. Hoyler, and J. Faist, “Quantum cascade lasers with large optical waveguides,” *IEEE Photon. Technol. Lett.*, vol. 18, no. 3, pp. 544–547, 2006.

[145] J. Nguyen, J. Yu, A. Evans, S. Slivken, and M. Razeghi, “Optical coatings by ion-beam sputtering deposition for long-wave infrared quantum cascade lasers,” *Applied Physics Letters*, vol. 89, pp. 111113–1–111113–3, 2006.

[146] J. Faist, A. Tredicucci, F. Capasso, C. Sirtori, D. Sivco, J. Baillargeon, A. Hutchinson, and A. Cho, “High-power continuous-wave quantum cascade lasers,” *IEEE Journal of Quantum Electronics*, vol. 34, no. 2, pp. 336–343, 1998.

[147] A. Lyakh, R. Maulini, A. Tsekoun, R. Go, C. Pflügl, L. Diehl, Q. J. Wang, F. Capasso, and C. Kumar N Patel, “3 W continuous-wave room temperature single-facet emission from quantum cascade lasers based on nonresonant extraction design approach,” *Applied Physics Letters*, vol. 95, no. 14, p. 141113, 2009.

[148] S. Katz, A. Vizbaras, G. Boehm, and M.-C. Amann, “High-performance injectorless quantum cascade lasers emitting below 6 μ m,” *Applied Physics Letters*, vol. 94, no. 15, p. 151106, 2009.

[149] K. J. Franz, P. Q. Liu, J. J. J. Raftery, M. D. Escarra, A. J. Hoffman, S. S. Howard, Y. Yao, Y. Dikmelik, X. Wang, J.-Y. Fan, J. B. Khurgin, and C. Gmachl, “Short Injector Quantum Cascade Lasers,” *IEEE Journal of Quantum Electronics*, vol. 46, no. 5, pp. 591–600, 2010.

[150] P. Q. Liu, A. J. Hoffman, M. D. Escarra, K. J. Franz, J. B. Khurgin, Y. Dikmelik, X. Wang, J.- Y. Fan, and C. F. Gmachl, “Highly power-efficient quantum cascade lasers,” *Nature Photonics*, vol. 4, no. 2, pp. 95–98, 2010.

[151] Y. Bai, S. Slivken, S. Kuboya, S. R. Darvish, and M. Razeghi, “Quantum cascade lasers that emit more light than heat,” *Nature Photonics*, vol. 4, no. 2, pp. 99–102, 2010.

[152] J. B. Khurgin, Y. Dikmelik, P. Q. Liu, A. J. Hoffman, M. D. Escarra, K. J. Franz, and C. F. Gmachl, “Role of interface roughness in the transport and lasing characteristics of quantum-cascade lasers,” *Applied Physics Letters*, vol. 94, no. 9, p. 091101, 2009.

[153] Q. Yang, R. Loesch, W. Bronner, S. Hugger, F. Fuchs, R. Aidam, and J. Wagner, “High-peakpower strain-compensated GaInAs/AlInAs quantum cascade lasers (λ similar to 4.6 μ m) based on a slightly diagonal active region design,” *Applied Physics Letters*, vol. 93, no. 25, p. 251110, 2008.

[154] Y. Bai, N. Bandyopadhyay, S. Tsao, E. Selcuk, S. Slivken, and M. Razeghi, “Highly temperature insensitive quantum cascade lasers,” *Applied Physics Letters*, vol. 97, no. 25, p. 251104, 2010.

[155] J. Faist, F. Capasso, D. Sivco, A. Hutchinson, C. Sirtori, S. Chu, and A. Cho, “Quantum cascade laser: temperature dependence of the performance characteristics and

high $T 0$ operation," *Applied Physics Letters*, vol. 65, no. 23, pp. 2901–2903, 1994.

[156] H. Page, C. Becker, A. Robertson, G. Glastre, V. Ortiz, and C. Sirtori, "300 K operation of a GaAs-based quantum-cascade laser at λ approximate to 9 μ m," *Applied Physics Letters*, vol. 78, no. 22, pp. 3529–3531, 2001.

(p.296) [157] H. Page, A. Robertson, C. Sirtori, C. Becker, G. Glastre, and J. Nagle, "Demonstration of (λ = 11.5-μ m) GaAs-based quantum cascade laser operating on a Peltier cooled element," *Photonics Technology Letters, IEEE*, vol. 13, no. 6, pp. 556–558, 2001.

[158] A. Wittmann, M. Giovannini, J. Faist, L. Hvozdara, S. Blaser, D. Hofstetter, and E. Gini, "Room temperature, continuous wave operation of distributed feedback quantum cascade lasers with widely spaced operation frequencies," *Applied Physics Letters*, vol. 89, pp. 141116–1– 141116–3, 2006.

[159] J. Faist, F. Capasso, D. Sivco, A. Hutchinson, S. Chu, and A. Cho, "Short wavelength ($\lambda \sim$3.4 μ m) quantum cascade laser based on strained compensated InGaAs/alinas," *Applied Physics Letters*, vol. 72, no. 6, pp. 680–682, 1998.

[160] M. P. Semtsiv, S. Dressler, and W. T. Masselink, "Short-Wavelength (λ 3.6 μ m)In0.73Ga0.27As-AlAs-InP Quantum-Cascade Laser," *IEEE Journal of Quantum Electronics*, vol. 43, no. 1, pp. 42–46, 2007.

[161] J. P. Commin, D. G. Revin, S. Y. Zhang, A. B. Krysa, K. Kennedy, and J. W. Cockburn, "High peak power λ 3.3 and 3.5μ m InGaAs/AlAs(Sb) quantum cascade lasers operating up to 400 K," *Applied Physics Letters*, vol. 97, no. 3, p. 031108, 2010.

[162] I. P. Marko, A. R. Adams, S. J. Sweeney, R. Teissier, A. N. Baranov, and S. Tomić, "Evidence of carrier leakage into the L-valley in InAs-based quantum cascade lasers under high hydrostatic pressure," *physica status solidi (b)*, vol. 246, no. 3, pp. 512–515, 2009.

[163] S. R. Jin, C. N. Ahmad, S. J. Sweeney, A. R. Adams, B. N. Murdin, H. Page, X. Marcadet, C. Sirtori, and S. Tomic, "Spectroscopy of GaAs/AlGaAs quantum-cascade lasers using hydrostatic pressure," *Applied Physics Letters*, vol. 89, no. 22, p. 221105, 2006.

[164] C. Van de Walle, "Band linups and deformation potentials in the model-solid theory," *Physical Review B*, vol. 39, no. 3, pp. 1871–1883, 1989.

[165] I. Vurgaftman, J. Meyer, and L. Ram-Mohan, "Band parameters for III-V compound semiconductors and their alloys," *Journal of Applied Physics*, vol. 89, no. 11, pp. 5815–5875, 2001.

[166] M. Sugawara, N. Okazaki, T. Fuji, and S. Yamazaki, "Conduction-band and valence-band structures in strained In1-xGaxAs/InP quantum-wells on (001) InP substrates," *Physical Review B*, vol. 48, no. 11, pp. 8102–8118, 1993.

[167] D. G. Revin, L. R. Wilson, E. A. Zibik, R. P. Green, J. W. Cockburn, M. J. Steer, R. J. Airey, and M. Hopkinson, "InGaAsAlAsSb quantum cascade lasers," *Applied Physics Letters*, vol. 85, no. 18, p. 3992, 2004.

[168] D. G. Revin, J. W. Cockburn, M. J. Steer, R. J. Airey, M. Hopkinson, A. B. Krysa, L. R. Wilson, and S. Menzel, "InGaAs/AlAsSb/InP quantum cascade lasers operating at wavelengths close to 3 μ m," *Applied Physics Letters*, vol. 90, no. 2, p. 021108, 2007.

[169] M. P. Semtsiv, M. Wienold, S. Dressler, and W. T. Masselink, "Short-wavelength (λ approximate to 3.3μ m) InP-based strain-compensated quantum-cascade laser," *Applied Physics Letters*, vol. 89, no. 21, p. 211124, 2006.

[170] M. P. Semtsiv, M. Wienold, S. Dressler, and W. T. Masselink, "Short-wavelength (λ approximate to 3.05 μ m) InP-based strain-compensated quantum-cascade laser," *Applied Physics Letters*, vol. 90, no. 5, p. 051111, 2007.

[171] B. S. Williams, "Terahertz quantum-cascade lasers," *Nature Photonics*, vol. 1, no. 9, pp. 517– 525, 2007.

[172] G. Scalari, C. Walther, M. Fischer, R. Terazzi, H. Beere, D. Ritchie, and J. Faist, "THz and sub-THz quantum cascade lasers," *Laser& Photonics Reviews*, vol. 3, no. 1-2, pp. 45–66, 2009.

[173] M. Rochat, L. Ajili, H. Willenberg, J. Faist, H. Beere, G. Davies, E. Linfield, and D. Ritchie, "Low-threshold terahertz quantum-cascade lasers," *Applied Physics Letters*, vol. 81, no. 8, pp. 1381–1383, 2002.

[174] S. Kohen, B. Williams, and Q. Hu, "Electromagnetic modeling of terahertz quantum cascade laser waveguides and resonators," *J. Appl. Phys*., vol. 97, pp. 053106–1–053106–9, 2005.

[175] L. Ajili, G. Scalari, J. Faist, H. Beere, E. Linfield, D. Ritchie, and G. Davies, "High power quantum-cascade lasers operating at $A \sim 87$ and 130 μ m," *Applied Physics Letters*, vol. 85, no. 18, pp. 3986–3988, 2004.

[176] K. Unterrainer, R. Colombelli, C. Gmachl, F. Capasso, H. Hwang, A. Sergent, D. Sivco, and A. Cho, "Quantum cascade lasers with double metal-semiconductor waveguide resonators," *Applied Physics Letters*, vol. 80, no. 17, pp. 3060–3062, 2002.

[177] M. A. Belkin, M. J. Fan, S. Hormoz, F. Capasso, S. Khanna, M. Lachab, A. Davies, and E. Lindfield, "Terahertz quantum cascade lasers with copper metal-metal waveguides operating up to 178 k," *Optics Express*, vol. 16, no. 05, pp. 3242–3248, 2008.

(p.297) [178] J. Gao, J. Hovenier, Z. Yang, J. Baselmans, A. Baryshev, M. Hajenius, T. Klapwijk, A. Adam, T. Klaassen, B. Williams, S. Kumar, Q. Hu, and J. Reno, "Terahertz heterodyne receiver based on a quantum cascade laser and a superconducting bolometer," *Applied Physics Letters*, vol. 86, pp. 244104–1–244104–3, 2005.

[179] A. Lee, Q. Qin, S. Kumar, B. S. Williams, Q. Hu, and J. L. Reno, "High-power and high-temperature thz quantum-cascade lasers based on lens-coupled metalmetal waveguides," *Opt. Lett*., vol. 32, pp. 2840–2842, 2007.

[180] M. Amanti, M. Fischer, C. Walther, G. Scalari, and J. Faist, "Horn antennas for terahertz quantum cascade lasers," *IEE Elect. Lett*., vol. 43, no. 10, pp. 573–574, 2007.

[181] G. Scalari, L. Ajili, J. Faist, H. Beere, E. Linfield, D. Ritchie, and G. Davies, "Far-infrared ($\lambda \sim 87\mu$ m) bound-to-continuum quantum-cascade lasers operating up to 90 k," *Applied Physics Letters*, vol. 82, no. 19, pp. 3165–3167, 2003.

[182] S. Barbieri, J. Alton, H. Beere, J. Fowler, E. Linfield, and D. Ritchie, "2.9 thz quantum cascade lasers operating up 70 k in continuous wave," *Applied Physics Letters*, vol. 85, no. 10, pp. 1674– 1676, 2004.

[183] G. Scalari, R. Terazzi, M. Giovannini, N. Hoyler, and J. Faist, "Population inversion by resonant tunneling in quantum wells," *Applied Physics Letters*, vol. 91, no. 3, p. 032103, 2007.

[184] B. Williams, H. Callebaut, S. Kumar, Q. Hu, and J. Reno, "3.4-thz quantum cascade laser based on longitudinal-optical-phonon scattering for depopulation," *Applied Physics Letters*, vol. 82, no. 7, pp. 1015–1017, 2003.

[185] B. Williams, S. Kumar, Q. Qin, Q. Hu, and J. Reno, "Terahertz quantum cascade lasers with double-resonant-phonon depopulation," *Applied Physics Letters*, vol. 88, pp. 261101–1–261101– 3, 2006.

[186] H. Luo, S. R. Laframboise, Z. R. Wasilewski, G. C. Aers, H. C. Liu, and J. C. Cao, "Terahertz quantum-cascade lasers based on a three-well active module," *Applied Physics Letters*, vol. 90, no. 4, p. 041112, 2007.

[187] S. Kumar, Q. Hu, and J. L. Reno, "186 K operation of terahertz quantum-cascade lasers based on a diagonal design," *Applied Physics Letters*, vol. 94, no. 13, p. 131105, 2009.

[188] C. Walther, M. Fischer, G. Scalari, R. Terazzi, N. Hoyler, and J. Faist, "Quantum cascade lasers operating from 1.2 to 1.6 thz," *Applied Physics Letters*, vol. 91, pp. 131122–1–131122–3, 2007.

[189] G. Scalari, N. Hoyler, M. Giovannini, and J. Faist, "Terahertz bound-to-continuum quanutmcascade lasers based on optical-phonon scattering extraction," *Applied Physics Letters*, vol. 86, pp. 181101–1–181101–3, 2005.

[190] M. I. Amanti, G. Scalari, R. Terazzi, M. Fischer, M. Beck, J. Faist, A. Rudra, P. Gallo, and E. Kapon, "Bound-to-continuum terahertz quantum cascade laser with a single-quantum-well phonon extraction/injection stage," *New Journal of Physics*, vol. 11, p. 125022, 2009.

[191] S. Fathololoumi, E. Dupont, C. Chan, Z. Wasilewski, S. Laframboise, D. Ban, A. Mátyás, C. Jirauschek, Q. Hu, and H. Liu, "Terahertz quantum cascade lasers operating up to 200 K with optimized oscillator strength and improved injection tunneling," *Optics Express*, vol. 20, no. 4, pp. 3866–3876, 2012.

[192] G. Scalari, M. Amanti, R. Terazzi, M. Beck, and J. Faist, "Two-well quantum cascade laser emitting from 2.7 to 4.1 THz," in *Proceedings of the Tenth International Conference on Intersubband Transitions in Quantum Wells* (http://www.itqw2009.com/index.php., ed.), Montreal, Canada, 2009.

[193] S. Kumar, C. W. I. Chan, Q. Hu, and J. L. Reno, "Two-well terahertz quantum-cascade laser with direct intrawell-phonon depopulation," *Applied Physics Letters*, vol. 95, no. 14, p. 141110, 2009.

[194] G. Scalari, M. Amanti, C. Walther, R. Terazzi, M. Beck, and J. Faist, "Broadband THz lasing from a photon-phonon quantum cascade structure," *Optics Express*, vol. 18, no. 8, pp. 8043– 8052, 2010.

[195] F. Peeters, A. Matulis, M. Helm, T. Fromherz, and W. Hilber, "Oscillator strength and sum rule for intersubband transitions in a superlattice," *Physical Review B*, vol. 48, no. 16, pp. 12008– 12015, 1993.

[196] A. Wacker, G. Bastard, F. Carosella, R. Ferreira, and E. Dupont, "Unraveling of free-carrier absorption for terahertz radiation in heterostructures," *Physical Review B*, vol. 84, no. 20, p. 205319, 2011.

[197] C. Walther, G. Scalari, J. Faist, H. Beere, and D. Ritchie, "Low frequency terahertz quantum cascade laser operating from 1.6 to 1.8 THz," *Applied Physics Letters*, vol. 89, no. 23, p. 231121, 2006.

(p.298) [198] E. Dupont, S. Fathololoumi, and H. Liu, "Simplified density-matrix model applied to three-well terahertz quantum cascade lasers," *Physical Review B*, vol. 81, no. 20, p. 205311, 2010.

[199] M. I. Amanti, M. Fischer, G. Scalari, M. Beck, and J. Faist, "Low-divergence single-mode terahertz quantum cascade laser," *Nature Photonics*, vol. 3, no. 10, pp. 586–590, 2009.

[200] H. Kogelnik and C. Shank, "Coupled-wave theory of distributed feedback lasers," *J. Appl. Phys*., vol. 43, no. 05, pp. 2327–2335, 1972.

[201] N. Finger, W. Schrenk, and E. Gornik, "Analysis of TM-Polarized DFB laser structures with metal surface gratings," *IEEE Journal of Quantum Electronics*, vol. 36, no. 7, pp. 780–786, 2000.

[202] W. Schrenk, N. Finger, S. Gianordoli, L. Hvozdara, G. Strasser, and E. Gornik, "GaAs/AlGaAs distributed feedback quantum cascade lasers," *Applied Physics Letters*, vol.

76, no. 3, pp. 253– 255, 2000.

[203] M. Carras, M. Garcia, X. Marcadet, O. Parillaud, A. De Rossi, and S. Bansropun, "Top grating index-coupled distributed feedback quantum cascade lasers," *Applied Physics Letters*, vol. 93, no. 1, p. 011109, 2008.

[204] D. Hofstetter, J. Faist, M. Beck, A. Muller, and U. Oesterle, "Demonstration of high-performance 10.16 μ m quantum cascade distributed feedback lasers fabricated without epitaxial regrowth," *Applied Physics Letters*, vol. 75, no. 5, pp. 665–667, 1999.

[205] D. Hofstetter, M. Beck, T. Aellen, and J. Faist, "High-temperature operation of distributed feedback quantum-cascade lasers at 5.3 μ m," *Applied Physics Letters*, vol. 78, no. 4, pp. 396– 398, 2001.

[206] A. Wittmann, Y. Bonetti, M. Fischer, J. Faist, S. Blaser, and E. Gini, "Distributed-Feedback Quantum-Cascade Lasers at 9 μ m Operating in Continuous Wave Up to 423 K," *IEEE Photonics Technology Letters*, vol. 21, no. 9-12, pp. 814–816, 2009.

[207] D. Hofstetter, J. Faist, M. Beck, A. Muller, and U. Oesterle, "Demonstration of high-performance 10.16 μ m quantum cascade distributed feedback lasers fabricated without epitaxial regrowth," *Applied Physics Letters*, vol. 75, no. 5, pp. 665–667, 1999.

[208] C. Gmachl, A. Straub, R. Colombelli, F. Capasso, D. Sivco, A. Sergent, and A. Cho, "Single-mode, tunable distributed-feedback and μ ltiple-wavelength quantum cascade lasers," *IEEE J. Quantum Electron.*, vol. 38, no. 6, pp. 569–581, 2002.

[209] S. Blaser, D. Yarekha, L. Hvozdara, Y. Bonetti, A. Muller, M. Giovannini, and J. Faist, "Room-temperature, continuous-wave, single-mode quantum-cascade lasers at $A \sim$ 5.4 μ m," *Applied Physics Letters*, vol. 86, pp. 41109–1–41109–3, 2005.

[210] D. Hofstetter, M. Beck, J. Faist, M. Nägele, and M. Sigrist, "Photoacoustic spectroscopy with quantum cascade distributed-feedback lasers," *Opt. Lett.*, vol. 26, no. 12, pp. 887–889, 2001.

[211] D. Nelson, J. McManus, S. Herndon, J. Shorter, M. Zahniser, S. Blaser, L. Hvozdara, A. Muller, M. Giovannini, and J. Faist, "Characterization of a near-room-temperature, continuous-wave quantum cascade laser for long-term, unattended monitoring of nitric oxide in the atmosphere," *Opt. Lett.*, vol. 31, no. 13, pp. 2012–2014, 2006.

[212] J. B. McManus, D. D. Nelson, S. C. Herndon, J. H. Shorter, M. S. Zahniser, S. Blaser, L. Hvozdara, A. Muller, M. Giovannini, and J. Faist, "Comparison of cw and pulsed operation with a TE-cooled quantum cascade infrared laser for detection of nitric oxide at 1900 cm(-1)," *Applied Physics B: Lasers and Optics*, vol. 85, no. 2-3, pp. 235–241, 2006.

[213] A. Schawlow and C. Townes, "Infrared and Optical Masers," *Physical Review*, vol. 112, no. 6, pp. 1940–1949, 1958.

[214] C. Henry, “Theory of the linewdith of semiconductor lasers,” *IEEE Journal of Quantum Electronics*, vol. 18, no. 2, pp. 259–264, 1982.

[215] T. Aellen, R. Maulini, R. Terazzi, N. Hoyler, M. Giovannini, J. Faist, S. Blaser, and L. Hvozdara, “Direct measurement of the linewidth enhancement factor by optical heterodyning of an amplitude-modulated quantum cascade laser,” *Applied Physics Letters*, vol. 89, no. 9, p. 091121, 2006.

[216] M. Yamanishi, T. Edamura, K. Fujita, N. Akikusa, and H. Kan, “Theory of the intrinsic linewidth of quantum-cascade lasers: Hidden reason for the narrow linewidth and line-broadening by thermal photons,” *IEEE Journal of Quantum Electronics*, vol. 44, no. 1-2, pp. 12–29, 2008.

[217] S. Bartalini, S. Borri, P. Cancio, A. Castrillo, I. Galli, G. Giusfredi, D. Mazzotti, L. Gianfrani, and P. De Natale, “Observing the Intrinsic Linewidth of a Quantum-Cascade Laser: Beyond the Schawlow-Townes Limit,” *Physical Review Letters*, vol. 104, no. 8, p. 083904, 2010.

[218] M. Zhu and J. Hall, “Stabilization of the optical phase/frequency of a laser system: application to a commercial dye laser with an external stabilizer,” *Journal of The Optical Society of America B-Optical Physics*, vol. 10, no. 5, pp. 802–816, 1993.

(p.299) [219] M. G. Littman and H. J. Metcalf, “Spectrally narrow pulsed dye laser without beam expander (et),” *Appl. Opt*., vol. 17, pp. 2224–2230, 1978.

[220] B. Saleh and M. Teich, *Fundamentals of Photonics*. Wiley series in pure and applied optics, 2nd ed., Feb. 1991.

[221] G. Luo, C. Peng, H. Le, S. Pei, W. Hwang, B. Ishaug, J. Um, J. Baillargeon, and C. Lin, “Grating-tuned external-cavity quantum-cascade semiconductor lasers,” *Applied Physics Letters*, vol. 78, no. 19, pp. 2834–2836, 2001.

[222] R. Maulini, M. Beck, J. Faist, and E. Gini, “Broadband tuning of external cavity bound-tocontinuum quantum-cascade lasers,” *Applied Physics Letters*, vol. 84, no. 10, pp. 1659–1661, 2004.

[223] R. Maulini, D. Yarekha, J. Bulliard, M. Giovannini, J. Faist, and E. Gini, “Continuous-wave operation of a broadly tunable thermoelectrically cooled external cavity quantum-cascade laser,” *Opt. Lett*., vol. 30, no. 19, pp. 2584–2586, 2005.

[224] A. Mohan, A. Wittmann, A. Hugi, S. Blaser, M. Giovannini, and J. Faist, “Room-temperature continuous-wave operation of an external-cavity quantum cascade laser,” *Opt. Lett*, vol. 32, pp. 2792–2794, 2007.

[225] A. Wittmann, T. Gresch, E. Gini, L. Hvozdara, N. Hoyler, M. Giovannim, and J. Faist, “High-performance bound-to-continuum quantum-cascade lasers for broad-gain applications,” *IEEE Journal of Quantum Electronics*, vol. 44, no. 1-2, pp. 36–40, 2008.

[226] R. Maulini, I. Dunayevskiy, A. Lyakh, A. Tsekoun, C. K. N. Patel, L. Diehl, C. Pfluegl, and F. Capasso, "Widely tunable high-power external cavity quantum cascade laser operating in continuous-wave at room temperature," *Electronics Letters*, vol. 45, no. 2, pp. 107–U30, 2009.

[227] C. F. Lin, Y. S. Su, and B. R. Wu, "External-cavity semiconductor laser tunable from 1.3 to 1.54 μ m for optical communication," *IEEE Photon. Technol. Lett.*, vol. 14, p. 3, 2003.

[228] C. Gmachl, D. Sivco, J. Baillargeon, A. Hutchinson, F. Capasso, and A. Cho, "Quantum cascade lasers with a heterogeneous cascade: Two-wavelength operation," *Applied Physics Letters*, vol. 79, no. 5, pp. 572–574, 2001.

[229] C. Gmachl, D. Sivco, R. Colombelli, F. Capasso, and A. Cho, "Ultra-broadband semiconductor laser," *Nature*, vol. 415, no. 6874, pp. 883–887, 2002.

[230] R. Maulini, A. Mohan, M. Giovannini, J. Faist, and E. Gini, "External cavity quantum-cascade laser tunable from 8.2 to 10.4 μ m using a gain element with a heterogeneous cascade," *Applied Physics Letters*, vol. 88, no. 20, p. 201113, 2006.

[231] A. Wittmann, A. Hugi, E. Gini, N. Hoyler, and J. Faist, "Heterogeneous High-Performance Quantum-Cascade Laser Sources for Broad-Band Tuning," *IEEE Journal of Quantum Electronics*, vol. 44, no. 11-12, pp. 1083–1088, 2008.

[232] A. Hugi, R. Terazzi, Y. Bonetti, A. Wittmann, M. Fischer, M. Beck, J. Faist, and E. Gini, "External cavity quantum cascade laser tunable from 7.6 to 11.4μ m," *Applied Physics Letters*, vol. 95, no. 6, p. 061103, 2009.

[233] B. G. Lee, H. A. Zhang, C. Pfluegl, L. Diehl, M. A. Belkin, M. Fischer, A. Wittmann, J. Faist, and F. Capasso, "Broadband Distributed-Feedback Quantum Cascade Laser Array Operating From 8.0 to 9.8 μ m," *IEEE Photon. Technol. Lett.*, vol. 21, no. 13, pp. 914–916, 2009.

[234] B. G. Lee, M. A. Belkin, R. Audet, J. MacArthur, L. Diehl, C. Pflugl, and F. Capasso, "Widely tunable single-mode quantum cascade laser source for mid-infrared spectroscopy," *Applied Physics Letters*, vol. 91, no. 23, p. 231101, 2007.

[235] E. Mujagic, C. Schwarzer, W. Schrenk, J. Chen, C. Gmachl, and G. Strasser, "Ring-cavity surface-emitting lasers as a building block for tunable and coherent quantum cascade laser arrays," *Semiconductor Science and Technology*, vol. 26, no. 1, p. 014019, 2011.

[236] L. Mahler and A. Tredicucci, "Photonic engineering of surface-emitting terahertz quantum cascade lasers," *Laser& Photonics Reviews*, vol. 5, no. 5, pp. 647–658, 2011.

[237] E. Mujagic, C. Schwarzer, Y. Yao, J. Chen, C. Gmachl, and G. Strasser, "Two-dimensional broadband distributed-feedback quantum cascade laser arrays," *Applied Physics Letters*, vol. 98, no. 14, p. 141101, 2011.

[238] E. Mujagic, M. Nobile, H. Detz, W. Schrenk, J. Chen, C. Gmachl, and G. Strasser, “Ring cavity induced threshold reduction in single-mode surface emitting quantum cascade lasers,” *Applied Physics Letters*, vol. 96, no. 3, p. 031111, 2010.

[239] http://www.agilent.com.

[240] http://www.avtech.com.

[241] http://www.alpeslasers.ch.

[242] V. Spagnolo, M. Troccoli, G. Scamarcio, C. Gmachl, F. Capasso, A. Tredicucci, A. Sergent, A. Hutchinson, D. Sivco, and A. Cho, “Temperature profile of GaInAs/AlInAs/InP quantum cascade-laser facets measured by microprobe photoluminescence,” *Applied Physics Letters*, vol. 78, no. 15, pp. 2095–2097, 2001.

(p.300) [243] V. Spagnolo, G. Scamarcio, H. Page, and C. Sirtori, “Simultaneous measurement of the electronic and lattice temperatures in GaAs/Al0.45Ga0.55As quantum-cascade lasers: Influence on the optical performance,” *Applied Physics Letters*, vol. 84, no. 18, pp. 3690–3692, 2004.

[244] M. S. Vitiello, T. Gresch, A. Lops, V. Spagnolo, G. Scamarcio, N. Hoyler, M. Giovannini, and J. Faist, “Influence of InAs, AlAs delta layers on the optical, electronic, and thermal characteristics of strain-compensated GaInAs/AlInAs quantum-cascade lasers,” *Applied Physics Letters*, vol. 91, no. 16, p. 161111, 2007.

[245] V. Spagnolo, G. Scamarcio, M. Troccoli, F. Capasso, C. Gmachl, A. Sergent, A. Hutchinson, D. Sivco, and A. Cho, “Nonequilibrium optical phonon generation by steady-state electron transport in quantum-cascade lasers,” *Applied Physics Letters*, vol. 80, no. 23, pp. 4303–4305, 2002.

[246] Y. Varshni, “Temperature dependence of the enegy gap in semiconductors,” *Physica*, vol. 34, no. 1, pp. 149–154, 1967.

[247] C. Pflugl, M. Litzenberger, W. Schrenk, D. Pogany, E. Gornik, and G. Strasser, “Interferometric study of thermal dynamics in GaAs-based quantum-cascade lasers,” *Applied Physics Letters*, vol. 82, no. 11, pp. 1664–1666, 2003.

[248] T. Gresch, “Gain and waveguide engineering in mid-infrared quantum cascade lasers,” *Diss. ETH No. 18732*, vol. 18732, pp. 1–308, 2009.

[249] M. S. Vitiello, G. Scamarcio, V. Spagnolo, S. S. Dhillon, and C. Sirtori, “Terahertz quantum cascade lasers with large wall-plug efficiency,” *Applied Physics Letters*, vol. 90, no. 19, p. 191115, 2007.

[250] M. Vitiello, G. Scamarcio, V. Spagnolo, B. Williams, S. Kumar, Q. Hu, and J. Reno, “Measurement of subband electronic temperatures and population inversion in THz quantum-cascade lasers,” *Applied Physics Letters*, vol. 86, no. 11, p. 111115, 2005.

[251] M. S. Vitiello, G. Scamarcio, J. Faist, G. Scalari, C. Walther, H. E. Beere, and D. A. Ritchie, “Probing quantum efficiency by laser-induced hot-electron cooling,” *Applied Physics Letters*, vol. 94, no. 2, p. 021115, 2009.

[252] D. Marcuse, “Mode conversion caused by surface imperfections of a dielectric slab waveguide,” *Bell System Technical Journal*, vol. 48, no. 10, pp. 3187–+, 1969.

[253] F. Toor, D. L. Sivco, H. E. Liu, and C. F. Gmachl, “Effect of waveguide sidewall roughness on the threshold current density and slope efficiency of quantum cascade lasers,” *Applied Physics Letters*, vol. 93, no. 3, p. 031104, 2008.

[254] B. Hakki and T. Paoli, “Gain spectra in GaAs double-heterotructure injection lasers,” *Journal of Applied Physics*, vol. 46, no. 3, pp. 1299–1306, 1975.

[255] J. Faist, F. Capasso, C. Sirtori, D. Sivco, A. Hutchinson, and A. Cho, “Continuous-wave operation of a vertical transition quantum cascade laser above T=80K,” *Applied Physics Letters*, vol. 67, no. 21, pp. 3057–3059, 1995.

[256] J. Faist, F. Capasso, C. Sirtori, D. Sivco, A. Hutchinson, S. Chu, and A. Cho, “Continuous wave operation of quantum cascade lasers based on vertical transitions at λ=4.6 μ m,” *Superlattices and Microstructures*, vol. 19, no. 4, pp. 337–345, 1996.

[257] D. Hofstetter and J. Faist, “Measurement of semiconductor laser gain and dispersion curves utilizing Fourier transforms of the emission spectra,” *IEEE Photonics Technology Letters*, vol. 11, no. 11, pp. 1372–1374, 1999.

[258] D. Revin, L. Wilson, D. Carder, J. Cockburn, M. Steer, M. Hopkinson, R. Airey, M. Garcia, C. Sirtori, Y. Rouillard, D. Barate, and A. Vicet, “Measurements of optical losses in mid-infrared semiconductor lasers using Fabry-Perot transmission oscillations,” *Journal of Applied Physics*, vol. 95, no. 12, pp. 7584–7587, 2004.

[259] E. Benveniste, S. Laurent, A. Vasanelli, C. Manquest, C. Sirtori, F. Teulon, M. Carras, and X. Marcadet, “Measurement of gain and losses of a midinfrared quantum cascade laser by wavelength chirping spectroscopy,” *Applied Physics Letters*, vol. 94, no. 8, p. 081110, 2009.

[260] D. G. Revin, L. R. Wilson, J. W. Cockburn, A. B. Krysa, J. S. Roberts, and R. J. Airey, “Intersubband spectroscopy of quantum cascade lasers under operating conditions,” *Applied Physics Letters*, vol. 88, no. 13, p. 131105, 2006.

[261] W. Parz, T. Mueller, J. Darmo, M. Austerer, G. Strasser, L. Wilson, J. Cockburn, A. Krysa, J. Roberts, and K. Unterrainer, “Intersubband gain-induced dispersion,” *Optics Letters*, vol. 34, no. 2, pp. 208–210, 2009.

[262] J. Kroell, S. S. Dhillon, X. Marcadet, M. Calligaro, C. Sirtori, and K. Unterrainer, “Phase-resolved measurements of stimulated emission in a laser,” *Nature*, vol. 449, no. 7163, pp. 698–U4, 2007.

(p.301) [263] S. Barbieri, C. Sirtori, H. Page, M. Beck, J. Faist, and J. Nagle, “Gain measurements on GaAs-based quantum cascade lasers using a two-section cavity technique,” *IEEE Journal of Quantum Electronics*, vol. 36, no. 6, pp. 736–741, 2000.

[264] M. Rochat, M. Beck, J. Faist, and U. Oesterle, “Measurement of far-infrared waveguide loss using a μ ltisection single-pass technique,” *Applied Physics Letters*, vol. 78, no. 14, pp. 1967– 1969, 2001.

[265] R. Terazzi, T. Gresch, M. Giovannini, N. Hoyler, N. Sekine, and J. Faist, “Bloch gain in quantum cascade lasers,” *Nature Physics*, vol. 3, no. 5, pp. 329–333, 2007.

[266] T. Gresch, R. Terazzi, J. Faist, and M. Giovannini, “Bloch gain in quantum cascade lasers at high temperature,” *Applied Physics Letters*, vol. 94, no. 3, p. 031102, 2009.

[267] T. Gresch, J. Faist, and M. Giovannini, “Gain measurements in strain-compensated quantum cascade laser,” *Applied Physics Letters*, vol. 94, no. 16, p. 161114, 2009.

[268] A. Wacker, “Semiconductor superlattices: a model system for nonlinear transport,” *Physics Reports: Review Section of Physics Letters*, vol. 357, no. 1, pp. 1–111, 2002.

[269] R. Iotti and F. Rossi, “Nature of charge transport in quantum-cascade lasers,” *Physical Review Letters*, vol. 87, no. 14, p. 146603, 2001.

[270] F. Rossi and T. Kuhn, “Theory of ultrafast phenomena in photoexcited semiconductors,” *Reviews of Modern Physics*, vol. 74, no. 3, p. 895, 2002.

[271] R. Iotti, E. Ciancio, and F. Rossi, “Quantum transport theory for semiconductor nanostructures: A density-matrix formulation,” *Physical Review B*, vol. 72, no. 12, p. 125347, 2005.

[272] D. Indjin, P. Harrison, R. Kelsall, and Z. Ikonic, “Influence of leakage current on temperature performance of GaAs/AlGaAs quantum cascade lasers,” *Applied Physics Letters*, vol. 81, no. 3, pp. 400–402, 2002.

[273] S. Blaser, L. Diehl, M. Beck, J. Faist, U. Oesterle, J. Xu, S. Barbieri, and F. Beltram, “Characterization and modeling of quantum cascade lasers based on a photon-assisted tunneling transition,” *IEEE Journal of Quantum Electronics*, vol. 37, no. 3, pp. 448–455, 2001.

[274] J. Faist, D. Hofstetter, M. Beck, T. Aellen, M. Rochat, and S. Blaser, “Bound-to-continuum and two-phonon resonance quantum-cascade lasers for high duty cycle, high-temperature operation,” *IEEE Journal of Quantum Electronics*, vol. 38, no. 6, pp. 533–546, 2002.

[275] A. Muller, M. Beck, J. Faist, U. Oesterle, and M. Ilegems, “Electrically tunable, room-temperature quantum-cascade lasers,” *Applied Physics Letters*, vol. 75, no. 11, pp. 1509–1511, 1999.

[276] P. Harrison, D. Indjin, and R. Kelsall, “Electron temperature and mechanisms of hot carrier generation in quantum cascade lasers,” *Journal of Applied Physics*, vol. 92, p. 6921, 2002.

[277] C. Sirtori, F. Capasso, J. Faist, A. Hutchinson, D. Sivco, and A. Cho, “Resonant tunneling in quantum cascade lasers,” *IEEE Journal of Quantum Electronics*, vol. 34, no. 9, pp. 1722–1729, 1998.

[278] S. Barbieri, C. Sirtori, H. Page, M. Stellmacher, and J. Nagle, “Design strategies for GaAs-based unipolar lasers: Optimum injector-active region coupling via resonant tunneling,” *Applied Physics Letters*, vol. 78, no. 3, pp. 282–284, 2001.

[279] H. Willenberg, G. Dohler, and J. Faist, “Intersubband gain in a Bloch oscillator and quantum cascade laser,” *Physical Review B*, vol. 67, no. 8, pp. –, 2003.

[280] W. Kohn and J. Luttinger, “Quantum theory of electrical transport phenomena,” *Physical Review*, vol. 108, no. 3, p. 590, 1957.

[281] R. Terazzi, T. Gresch, A. Wittmann, and J. Faist, “Sequential resonant tunneling in quantum cascade lasers,” *Physical Review B*, vol. 78, no. 15, p. 4, 2008.

[282] S. Ktitorov, G. Simin, and V. Sindalovskii, “Bragg reflections and the high-frequency conductivity of an electronic solid-state plasma,” *Fiz. tverd. Tela.*, vol. 13, pp. 2230–2233, 1971.

[283] A. Ignatov and Y. Romanov, “Nonlinear electromagnetic properties of semiconductors with a superlattice,” *Phys. Stat. Sol. B*, vol. 73, pp. 327–333, 1976.

[284] A. Ignatov and Y. Romanov, “Nonlinear Electromagnetic Properties of semiconductors with a superlattice,” *Physica Status Solidi B-Basic Research*, vol. 73, no. 1, pp. 327–333, 1976.

[285] J. Faist and G. Scalari, “Unified description of resonant tunnelling diodes and terahertz quantum cascade lasers,” *Electronics Letters*, vol. 46, no. 26, p. S46, 2010.

[286] M. Asada, S. Suzuki, and N. Kishimoto, “Resonant Tunneling diodes for sub-terahertz and terahertz oscillators,” *Jpn. J. Appl. Phys.*, vol. 47, no. 6, pp. 4375–4384, 2008.

[287] G. Scamarcio, M. Troccoli, F. Capasso, A. Hutchinson, D. Sivco, and A. Cho, “High peak power (2.2W) superlattice quantum cascade laser,” *Electronics Letters*, vol. 37, no. 5, pp. 295–296, 2001.

(p.302) [288] K. Fujita, S. Furuta, A. Sugiyama, T. Ochiai, T. Edamura, N. Akikusa, M. Yamanishi, and H. Kan, “High-Performance λ similar to 8.6 μ m Quantum Cascade Lasers With Single Phonon-Continuum Depopulation Structures,” *IEEE Journal of Quantum Electronics*, vol. 46, no. 5, pp. 683–688, 2010.

[289] A. Wacker, “Extraction-controlled quantum cascade lasers,” *Applied Physics Letters*, vol. 97, no. 8, p. 081105, 2010.

[290] H. Callebaut and Q. Hu, “Importance of coherence for electron transport in terahertz quantum cascade lasers,” *Journal of Applied Physics*, vol. 98, no. 10, p. 104505, 2005.

[291] S. Kumar and Q. Hu, “Coherence of resonant-tunneling transport in terahertz quantum-cascade lasers,” *Physical Review B*, vol. 80, no. 24, p. 245316, 2009.

[292] A. Gordon and D. Majer, “Coherent transport in semiconductor heterostructures: A phenomenological approach,” *Physical Review B*, vol. 80, no. 19, p. 195317, 2009.

[293] R. Terazzi and J. Faist, “A density matrix model of transport and radiation in quantum cascade lasers,” *New Journal of Physics*, vol. 12, p. 033045, 2010.

[294] J. Faist, F. Capasso, C. Sirtori, D. Sivco, A. Hutchinson, M. Hybertsen, and A. Cho, “Quantum cascade lasers without intersubband population inversion,” *Phys. Rev. Lett*., vol. 76, no. 3, pp. 411–414, 1996.

[295] J. Faist, D. Hofstetter, M. Beck, T. Aellen, M. Rochat, and S. Blaser, “Bound-to-continuum and two-phonon resonance quantum cascade lasers for high duty cycle, high temperature operation,” *IEEE J. Quantum Electron*., vol. 38, no. 6, pp. 533–546, 2002.

[296] R. C. Iotti and F. Rossi, “Carrier thermalization versus phonon-assisted relaxation in quantum-cascade lasers: A Monte Carlo approach,” *Applied Physics Letters*, vol. 78, no. 19, pp. 2902– 2904, 2001.

[297] A. Mátyás, P. Lugli, and C. Jirauschek, “Photon-induced carrier transport in high efficiency midinfrared quantum cascade lasers,” *Journal of Applied Physics*, vol. 110, no. 1, p. 013108, 2011.

[298] L. Kadanoff and G. Baym, *Quantum Statistical Mechanics*. Benjamin, New York, Mar. 1962.

[299] L. Keldysh, “Diagram technique for nonequilibrium processes,” *Soviet Physics Jetp-Ussr*, vol. 20, no. 4, p. 1018, 1965.

[300] S.-C. Lee and A. Wacker, “Nonequilibrium Green’s function theory for transport and gain properties of quantum cascade structures,” *Physical Review B*, vol. 66, no. 24, p. 18, 2002.

[301] S.-C. Lee, F. Banit, M. Woerner, and A. Wacker, “Quantum mechanical wavepacket transport in quantum cascade laser structures,” *Physical Review B*, vol. 73, no. 24, p. 6, 2006.

[302] R. Paiella, F. Capasso, C. Gmachl, C. Bethea, D. Sivco, J. Baillargeon, A. Hutchinson, and A. Cho, “High-speed operation of gain-switched midinfrared quantum cascade lasers,” *Applied Physics Letters*, vol. 75, no. 17, pp. 2536–2538, 1999.

[303] R. Paiella, R. Martini, F. Capasso, C. Gmachl, H. Hwang, D. Sivco, J. Baillargeon, A. Cho, E. Whittaker, and H. Liu, "High-frequency modulation without the relaxation oscillation resonance in quantum cascade lasers," *Applied Physics Letters*, vol. 79, no. 16, pp. 2526–2528, 2001.

[304] D. Hofstetter, M. Beck, T. Aellen, and S. Blaser, "High-frequency modulation of a quantum-cascade laser using a monolithically integrated intracavity modulator," *IEEE Photonics Technology Letters*, vol. 15, no. 8, pp. 1044–1046, 2003.

[305] W. Maineult, L. Ding, P. Gellie, P. Filloux, C. Sirtori, S. Barbieri, T. Akalin, J. Lampin, I. Sagnes, and H. Beere, "Microwave modulation of terahertz quantum cascade lasers: a transmission-line approach," *Applied Physics Letters*, vol. 96, p. 021108, 2010.

[306] S. Barbieri, W. Maineult, S. S. Dhillon, C. Sirtori, J. Alton, N. Breuil, H. E. Beere, and D. A. Ritchie, "13 GHz direct modulation of terahertz quantum cascade lasers," *Applied Physics Letters*, vol. 91, no. 14, p. 143510, 2007.

[307] P. Gellie, S. Barbieri, J.-F. Lampin, P. Filloux, C. Manquest, C. Sirtori, I. Sagnes, S. P. Khanna, E. H. Linfield, A. G. Davies, H. Beere, and D. Ritchie, "Injection-locking of terahertz quantum cascade lasers up to 35GHz using RF amplitude modulation," *Optics Express*, vol. 18, no. 20, pp. 20799–20816, 2010.

[308] H. Choi, L. Diehl, Z.-K. Wu, M. Giovannini, J. Faist, F. Capasso, and T. B. Norris, "Gain recovery dynamics and photon-driven transport in quantum cascade lasers," *Physical Review Letters*, vol. 100, no. 16, p. 167401, 2008.

[309] H. Choi, T. B. Norris, T. Gresch, M. Giovannini, J. Faist, L. Diehl, and F. Capasso, "Femtosecond dynamics of resonant tunneling and superlattice relaxation in quantum cascade lasers," *Applied Physics Letters*, vol. 92, no. 12, p. 122114, 2008.

(p.303) [310] H. Haus, "Mode-locking of lasers," *IEEE Journal of Selected Topics In Quantum Electronics*, vol. 6, no. 6, pp. 1173–1185, 2000.

[311] F. Keilmann, C. Gohle, and R. Holzwarth, "Time-domain mid-infrared frequency-comb spectrometer," *Optics Letters*, vol. 29, no. 13, pp. 1542–1544, 2004.

[312] A. Gordon, C. Y. Wang, L. Diehl, F. X. Kaertner, A. Belyanin, D. Bour, S. Corzine, G. Hoefler, H. C. Liu, H. Schneider, T. Maier, M. Troccoli, J. Faist, and F. Capasso, "Multimode regimes in quantum cascade lasers: From coherent instabilities to spatial hole burning," *Physical Review A*, vol. 77, no. 5, p. 053804, 2008.

[313] R. Paiella, F. Capasso, C. Gmachl, D. Sivco, J. Baillargeon, A. Hutchinson, A. Cho, and H. Liu, "Self-mode-locking of quantum cascade lasers with giant ultrafast optical nonlinearities," *Science*, vol. 290, no. 5497, pp. 1739–1742, 2000.

[314] H. Risken and K. Nummedal, "Self pulsing in lasers," *Journal of Applied Physics*, vol. 39, no. 10, pp. 4662–&, 1968.

[315] R. Graham and H. Haken, “Quantum Theory of Light Propagation in a Fluctuating Laser-Active Medium,” *Zeitschrift Fur Physik*, vol. 213, no. 5, pp. 420–450, 1968.

[316] P. Del’haye, A. Schliesser, O. Arcizet, T. Wilken, R. Holzwarth, and T. J. Kippenberg, “Optical frequency comb generation from a monolithic microresonator,” *Nature*, vol. 450, no. 7173, pp. 1214–1217, 2007.

[317] K. Mogi, K. Naganuma, and H. Yamada, “A novel real-time chirp measurement method for ultrashort optical pulses,” *Japanese Journal of Applied Physics Part 1: Regular Papers Short Notes & Review Papers*, vol. 27, no. 11, pp. 2078–2081, 1988.

[318] A. Soibel, F. Capasso, C. Gmachl, M. Peabody, A. Sergent, R. Paiella, D. Sivco, A. Cho, and H. Liu, “Stability of pulse emission and enhancement of intracavity second-harmonic generation in self-mode-locked quantum cascade lasers,” *IEEE Journal of Quantum Electronics*, vol. 40, no. 3, pp. 197–204, 2004.

[319] A. Soibel, F. Capasso, C. Gmachl, M. Peabody, A. Sergent, R. Paiella, H. Hwang, D. Sivco, A. Cho, H. Liu, C. Jirauschek, and F. Kartner, “Active mode locking of broadband quantum cascade lasers,” *IEEE Journal of Quantum Electronics*, vol. 40, no. 7, pp. 844–851, 2004.

[320] C. Y. Wang, L. Kuznetsova, V. M. Gkortsas, L. Diehl, F. X. Kaertner, M. A. Belkin, A. Belyanin, X. Li, D. Ham, H. Schneider, P. Grant, C. Y. Song, S. Haffouz, Z. R. Wasilewski, H. C. Liu, and F. Capasso, “Mode-locked pulses from mid-infrared Quantum Cascade Lasers,” *Optics Express*, vol. 17, no. 15, pp. 12929–12943, 2009.

[321] S. Barbieri, M. Ravaro, P. Gellie, G. Santarelli, C. Manquest, C. Sirtori, S. P. Khanna, E. H. Linfield, and A. G. Davies, “Coherent sampling of active mode-locked terahertz quantum cascade lasers and frequency synthesis,” *Nature Photonics*, vol. 5, no. 5, pp. 306–313, 2011.

[322] M. Achour, “Free-space optics wavelength selection: 10 m versus shorter wavelengths [Invited],” *Journal of Optical Networking*, vol. 2, no. 6, p. 127, 2003.

[323] R. Martini, C. Gmachl, J. Falciglia, F. Curti, C. Bethea, F. Capasso, E. Whittaker, R. Paiella, A. Tredicucci, A. Hutchinson, D. Sivco, and A. Cho, “High-speed modulation and free-space optical audio/video transmission using quantum cascade lasers,” *IEE Elect. Lett.*, vol. 37, no. 3, pp. 102–103, 2001.

[324] R. Martini, C. Bethea, F. Capasso, C. Gmachl, R. Paiella, E. Whittaker, H. Hwang, D. Sivco, J. Baillargeon, and A. Cho, “Free-space optical transmission of μ ltimedia satellite data streams using mid-infrared quantum cascade lasers,” *IEE Elect. Lett.*, vol. 38, no. 4, pp. 181–183, 2002.

[325] S. Blaser, D. Hofstetter, M. Beck, and J. Faist, “Free-space optical data link using peltier-cooled quantum cascade laser,” *IEE Elect. Lett.*, vol. 37, no. 12, pp. 778–780, 2001.

[326] L. S. Rothman, D. Jacquemart, A. Barbe, D. Chris Benner, M. Birk, L. R. Brown, M. R. Carleer, C. Chackerian Jr., K. Chance, L. H. Coudert, V. Dana, V. M. Devi, J. M. Flaud, R. R. Gamache, A. Goldman, J. M. Hartmann, K. W. Jucks, A. G. Maki, J. Y. Mandin, S. T. Massie, J. Orphal, A. Perrin, C. P. Rinsland, M. A. H. Smith, J. Tennyson, R. N. Tolchenov, R. A. Toth, J. Vander Auwera, P. Varanasi, and G. Wagner, "The HITRAN 2004 molecular spectroscopic database," *Journal of Quantitative Spectroscopy and Radiative Transfer*, vol. 96, no. 2, pp. 139–204, 2005.

[327] K. Namjou, S. Cai, E. Whittaker, J. Faist, C. Gmachl, F. Capasso, D. Sivco, and A. Cho, "Sensitive absorption spectroscopy with a room-temperature distributed-feedback quantum-cascade laser," *Opt. Lett*., vol. 23, no. 3, pp. 219–221, 1998.

[328] S. Sharpe, J. Kelly, J. Hartmann, C. Gmachl, F. Capasso, D. Sivco, J. Baillargeon, and A. Cho, "High-resolution (Doppler-limited) spectroscopy using quantum-cascade distributed-feedback lasers," *Opt. Lett*., vol. 23, no. 17, pp. 1396–1398, 1998.

(p.304) [329] A. Kosterev, R. Curl, F. Tittel, C. Gmachl, F. Capasso, D. Sivco, J. Baillargeon, A. Hutchinson, and A. Cho, "Methane concentration and isotopic composition measurements with a mid-infrared quantum-cascade laser," *Optics Letters*, vol. 24, no. 23, pp. 1762–1764, 1999.

[330] A. Kosterev, R. Curl, F. Tittel, C. Gmachl, F. Capasso, D. Sivco, J. Baillargeon, A. Hutchinson, and A. Cho, "Effective utilization of quantum-cascade distributed-feedback lasers in absorption spectroscopy," *Applied Optics*, vol. 39, no. 24, pp. 4425–4430, 2000.

[331] B. Paldus, T. Spence, R. Zare, J. Oomens, F. Harren, D. Parker, C. Gmachl, F. Cappasso, D. Sivco, J. Baillargeon, A. Hutchinson, and A. Cho, "Photoacoustic spectroscopy using quantum-cascade lasers," *Optics Letters*, vol. 24, no. 3, pp. 178–180, 1999.

[332] T. Beyer, M. Braun, and A. Lambrecht, "Fast gas spectroscopy using pulsed quantum cascade lasers," *Journal of Applied Physics*, vol. 93, no. 6, pp. 3158–3160, 2003.

[333] E. Normand, M. McCulloch, G. Duxbury, and N. Langford, "Fast, real-time spectrometer based on a pulsed quantum-cascade laser," *Optics Letters*, vol. 28, no. 1, pp. 16–18, 2003.

[334] J. Roepcke, G. Lombardi, A. Rousseau, and P. B. Davies, "Application of mid-infrared tuneable diode laser absorption spectroscopy to plasma diagnostics: a review," *Plasma Sources Science & Technology*, vol. 15, no. 4, pp. S148–S168, 2006.

[335] R. Jimenez, M. Taslakov, V. Simeonov, B. Calpini, F. Jeanneret, D. Hofstetter, M. Beck, J. Faist, and H. Van den Bergh, "Ozone detection by differential absorption spectroscopy at ambient pressure with a 9.6 μ m pulsed quantum-cascade laser," *Applied Physics B: Lasers and Optics*, vol. 78, no. 2, pp. 249–256, 2004.

[336] B. Tuzson, J. Mohn, M. J. Zeeman, R. A. Werner, W. Eugster, M. S. Zahniser, D. D. Nelson, J. B. McManus, and L. Emmenegger, “High precision and continuous field measurements of delta C-13 and delta O-18 in carbon dioxide with a cryogen-free QCLAS,” *Applied Physics B: Lasers and Optics*, vol. 92, no. 3, pp. 451–458, 2008.

[337] A. Kosterev, A. Malinovsky, F. Tittel, C. Gmachl, F. Capasso, D. Sivco, J. Baillargeon, A. Hutchinson, and A. Cho, “Cavity ringdown spectroscopic detection of nitric oxide with a continuous-wave quantum-cascade laser,” *Applied Optics*, vol. 40, no. 30, pp. 5522–5529, 2001.

[338] V. Zeninari, A. Grossel, L. Joly, T. Decarpenterie, B. Grouiez, B. Bonno, and B. Parvitte, “Photoacoustic spectroscopy for trace gas detection with cryogenic and room-temperature continuous-wave quantum cascade lasers,” *Central European Journal of Physics*, vol. 8, no. 2, pp. 194–201, 2010.

[339] L. Dong, A. Kosterev, D. Thomazy, and F. Tittel, “Qepas spectrophones: design, optimization, and performance,” *Applied Physics B: Lasers and Optics*, vol. 100, pp. 627–635, 2010.

[340] G. Wysocki, R. Curl, F. Tittel, R. Maulini, J. Bulliard, and J. Faist, “Widely tunable mode-hop free external cavity quantum cascade laser for high resolution spectroscopic applications,” *Applied Physics B: Lasers and Optics*, vol. 81, no. 6, pp. 769–777, 2005.

[341] G. Wysocki, R. Lewicki, R. F. Curl, F. K. Tittel, L. Diehl, F. Capasso, M. Troccoli, G. Hofler, D. Bour, S. Corzine, R. Maulini, M. Giovannini, and J. Faist, “Widely tunable mode-hop free external cavity quantum cascade lasers for high resolution spectroscopy and chemical sensing,” *Applied Physics B: Lasers and Optics*, vol. 92, no. 3, pp. 305–311, 2008.