Chapter 12

Quantum cascade lasers

The quantum cascade laser is a triumph of solid state engineering. It exists only as a consequence of developments in so-called “band-gap engineering” and unlike a conventional diode laser does not rely on band-to-band recombination. This idea was pioneered by Federico Capasso, then at Bell Labs, who, along with others, started to develop complex quantum devices based on the idea of epitaxy with abrupt interfaces using molecular beams. Several applications of epitaxial growth have already been described in this book, such as double heterostructures, quantum wells, and distributed Bragg reflectors, but of these only the quantum well represents any sort of band gap engineering. In the DBR the electronic properties are not of primary importance beyond achieving the necessary conductivity to allow the mirrors to act as contacts. In the DH laser the width of the active region is still large enough for the material to be considered bulk-like, and the band gap of the active region is identical to the band gap of bulk material. In the quantum well, however, the lowest electron energy level effectively becomes the conduction band, and so the band gap is altered over the bulk material by the zero-point energy of both the electrons and the holes. In a material containing many quantum wells, sometimes with barriers thin enough for the wells to interact with each other, the effective band gap can be made to change throughout the structure simply by altering the well width, even though the well and barrier compositions remain fixed.

The quantum cascade laser uses this type of band gap engineering and consists of a series of quantum wells and barrier which together make up both active regions and injector regions. Usually, but not always, active regions comprise the layers in which the radiative transitions occur among the sub-bands of quantum wells. Injector regions, on the other hand, comprise a series of interacting quantum wells-so-called “superlattices” - which give rise to a band of states, called a “mini band” - such that electron transport perpendicular to the wells can occur. A distinction is made here between mini-bands and sub-bands, which are the discrete states localised in quantum wells whereas mini-bands are the states associated with a superlattice. They arise from a succession of interacting wells, that is to say, wells in which the barriers are sufficiently thin that the wavefunctions are not confined simply to one well but

penetrate through the barriers and extend over the whole of the superlattice. The bands also extend over a limited range in energy and the number of states in a band is equal to the number of wells that make up the superlattice.

The design of the active and injector regions is not straightforward and cannot be reduced to a few simple equations. It will become clear in this chapter that numerical computation is the only tool that the designer has to ensure that a particular choice of material composition and a particular growth sequence will result not only in the right energy levels, but also rates of population and depopulation from the levels that will allow population inversion. Transport through the mini-bands also has to be modelled numerically using Monte Carlo techniques, so what follows in this chapter, unlike the descriptions of other lasers in this book, is not a set of concepts or equations that can be used as design principles, but an exploration of some of the ideas and structures that have so far been reported in the literature. To do anything more requires a level of understanding beyond anything that can be given here.

A quantum cascade laser will contain several injector-active pairs so electrons injected into the system cascade through and are re-used in order to generate multiple photons per electron, as illustrated in figure 12.1 [1]. Consider an electron injected into the upper laser level (ULL) from the right. If the lower laser level (LLL) is empty then by definition population inversion

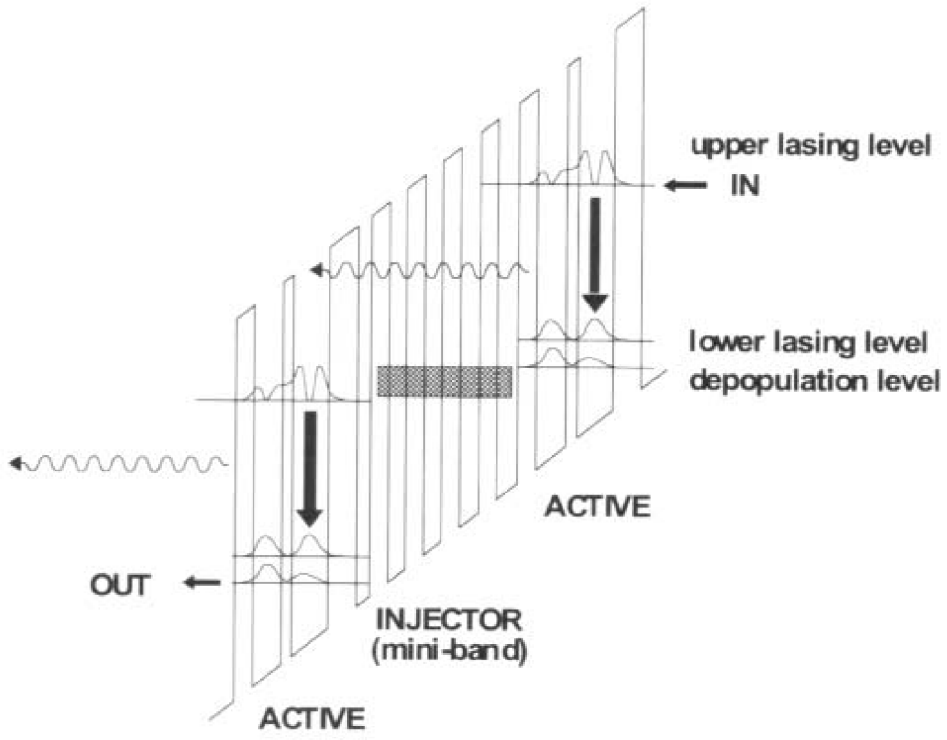

Figure 12.1. A schematic of a quantum cascade laser (after Gmaichl *et al*. [1]) showing two active regions at either end of a superlattice injector. The possibility of more than one photon per electron is clearly indicated, but it should be noted that the photons are emitted in the plane of the layers rather than in the growth direction as indicated above.

between these two levels exists and stimulated emission is possible. If the assumption is made that electrons can only reach the LLL via the ULL, and in an ideal device this will apply, then clearly if the exit from ULL to the depopulation level (DPL) and then to the miniband is slower than the total time from taken to enter the ULL and to make the transition to the LLL, charge will build up within the lower states of the quantum well and population inversion will be lost. For laser action to occur, therefore, it is necessary for the dwell time in the ULL to exceed the relaxation time from the LLL. In an ideal case electrons entering the LLL would be swept out instantaneously, thereby presenting an empty level to the electron in the ULL.

Population inversion in these systems is mathematically very simple. The rate of change of the carrier concentration in the LLL, n_L, can be expressed as a function of the excess concentration in the upper and lower levels, i.e.

$$\frac{dn_L}{dt} = \frac{n_U}{\tau_U} - \frac{n_L}{\tau_L} \tag{1}$$

where $\tau_{U,\,L}$ are the lifetimes of the respective levels. These are global lifetimes that represent all the possible paths out of the state, including relaxation by phonon emission which is, in fact, the dominant relaxation mechanism. Clearly, the presence of multiple pathways will reduce the level lifetime.

As argued, the lower level lifetime has to be shorter than the upper level lifetime.

There is a natural upper limit to the photon energy that can be obtained from a cascade laser. Quite clearly, the energy cannot be greater than the conduction band offset, so even for the most favourable material system, AlGaAs/GaAs, energies will be limited to a few tenths of an electron-volt. However, one of the features of the quantum cascade laser should be apparent by now; the emission wavelength is independent of the band gap of the constituent materials but is dependent instead on the width of the quantum wells in the active region, making this type of device structure very versatile. A very wide range of energies out into the far infra-red are possible. In mid-IR quantum cascade lasers, corresponding to the wavelength range from about 5-20 μm, of which the 8-14 μm atmospheric transmission window is the most important, rapid relaxation from the LLL to the DPL is achieved by making the LLL-DPL energy separation the same as, or very close to, the longitudinal optical phonon energy so that the probability of relaxation by single phonon emission is very high. In GaAs this is ~ 36 meV.

From this perspective the LO phonon energy represents a seemingly natural limit on the energy of the laser emission. It follows that if resonant phonon emission can be used to depopulate the LLL then it can also depopulate the ULL. If the emission energy of the laser is lowered by widening the quantum well so that the ULL-LLL energy separation approaches phonon resonance the LLL-DPL separation will naturally move off the resonance, leaving a very short upper lifetime and a relatively longer lower lifetime. The possibility of population inversion is thereby destroyed. It should be emphasised that this is a problem in scattering rates rather than energy levels. Long wavelength luminescence is possible simply by adjusting the well width to achieve the correct energy level separation, and luminescence at 70-80 μm was reported by several groups throughout the late 1990's, but for a laser the additional requirement on the scattering times out of the states is a very severe limitation. If the ULL-LLL transition energy is reduced below the phonon energy then very careful engineering of the LLL-DPL energy level is necessary to ensure the scattering rate will allow population inversion. Other relaxation mechanisms, such as electron-electron scattering, become important at such low subband separations. Even so phonon scattering can still be used with the correct design of laser, but the problem wasn't solved until 2001, since when laser emission wavelengths out to ~ 100 μm have been achieved. The quantum cascade laser is thus a very powerful tool for long wavelength applications in both communication and medicine.

This use of subbands is a significant departure from other diode lasers, but it's not the whole story. It is only the "quantum" but not the "cascade", and this is the second, and perhaps most radical, departure from conventional laser technology, for the electron is re-used in other laser transitions. An electron leaving the DPL enters the miniband states of the next injector region and is transported through the superlattice into the ULL at the other end. Minibands in superlattices are described in detail on page 390. It is sufficient for the present to realise that minibands arise from the interacting states of a series of closely spaced quantum wells. If all the wells and barriers are identical the minibands lie at a common energy throughout the structure, and the bottom of the lowest miniband effectively becomes the conduction band edge within the structure. However, under an applied field this miniband will not be flat, as in figure 12.1, but will in fact break up into a series of discrete states called a Wannier-Stark ladder. In the quantum cascade laser it is necessary to design the injector region so that the electron states of the interacting quantum wells align under the action of an electric field to produce a mini band that is essentially flat. This is achieved by grading the thicknesses of the layers and also the composition within the layers, so that under zero applied bias the band offsets resemble more of a saw-tooth structure than a square well. As the electron is transported through the mini band, its energy relative to the bottom of the well, which is essentially the bulk conduction band edge, increases, and the electron emerges from the mini band at the next ULL. In this way electrons are re-used

in the lasing process, unlike the conventional diode laser where the band-edge recombination removes the electron from the lasing process.

The full extent of the numerical computation necessary to design a quantum cascade laser should be apparent by now. Not only is it necessary to solve the Schrödinger equation to calculate the wavefunctions of the states in the wells, especially if, as in figure 12.1, the states at either end of the laser transition extend over two or three wells, but it is also necessary to simulate the transport and scattering processes. The variation of the energy of the states with applied voltage has to be considered as part of the design, and normally will involve different states of the different wells being brought into resonance at a particular applied field. In addition, of course, the superlattice states also have to be designed to allow transport at the same electric field as the laser transitions occur. The only effective way to do this is to design a structure and to simulate its operation using whatever simplifications and approximations are appropriate. It has been reported, for example, that full details of the transport will not be revealed without a full stochastic Monte Carlo simulation, as details of the electron distributions in the states and the minibands, as well as the effective electron temperature, are all but impossible to determine analytically [2]. Calculations of the electron-electron and electron-phonon scattering suggest, however, that the electrons in the various states reach thermal equilibrium among themselves, but not necessarily with the lattice, so they can be regarded as having the same temperature even if its absolute value is uncertain [3]

Even the full details of the electron-phonon scattering are hard to compute, but the calculations are probably accurate enough for the electron energies involved. Certainly, such calculations are beyond the scope of this book, but the underlying principles can be described.

In summary, quantum cascade lasers differ from conventional diode lasers in several ways:

- the devices are unipolar; that is to say, only one carrier type is involved;
- population inversion is achieved between sub-bands of a quantum well system rather than between electron and hole states;
- the dependence of the lasing wavelength on band gap is therefore broken, and a very wide range of wavelengths can be engineered in well-behaved and well-understood materials systems such as GaAs/AlGaAs;
- the wavelengths extend from the near infra-red to the very far infra-red close to 100 μm;
- the localised and discrete nature of the lasing energy levels means that the system appears to be almost atomic-like, with the exception that of course in-plane dispersion exists
- the in-plane dispersion in each sub-band is of the same sign, unlike bipolar devices, so even if transitions occur between states at in-plane $k \neq 0$ the wavelength is very similar to transitions at $k=0$ (see figure 12.2), and the gain spectrum is correspondingly narrow;
- charge neutrality cannot exist in the same way as in the bipolar device, in which equal concentrations of electrons and holes are injected;
- space charges will exist therefore within the quantum well, which can only be offset by background doping;

In some respects, however, the devices are similar:

- Light emission occurs perpendicular to the direction of current flow.
- Optical confinement is achieved through the use of wave guiding structures.

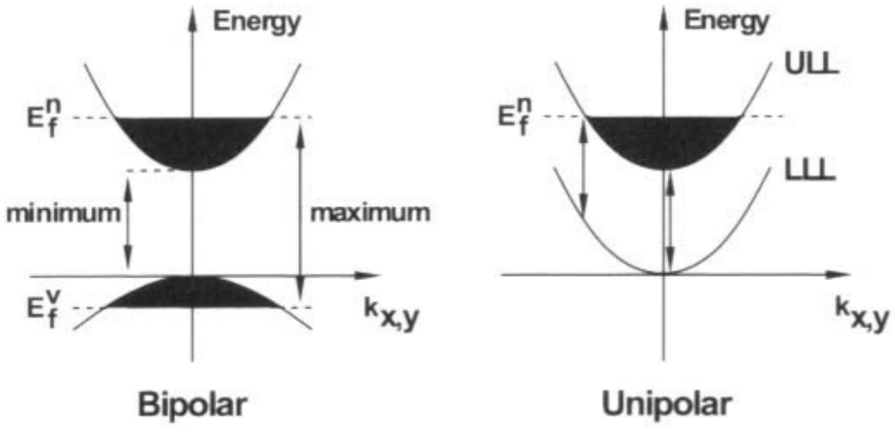

Figure 12.2. The contrast between the gain spectrum of a unipolar quantum cascade structure and a conventional bipolar device. The similar dispersion for the upper and lower laser levels of the uniploar device results in a similar transition energy irrespective of the in-plane wave vector whereas the bipolar device has widely separated maximum and minimum transition energies caused by the seperation of the Fermi levels.

- The cavity reflectors can be formed by cleaving or through the fabrication of distributed Bragg reflectors.

Not only is the detailed theory of quantum cascade lasers too complex to describe simply, but it is less than 10 years since the first report of a practical quantum cascade laser appeared in the press and the field is still rapidly developing. Much theoretical and experimental work on gain [7-9] and transport [10-12] has already been done in addition to that already referenced. It is left to the interested reader to seek out these works; what follows in this chapter are some of the simplified theoretical considerations by way of explanation of the differences summarised above as well as some of the significant developments that have taken place in recent years.

12.1 Quantum cascade structures

The unipolar nature of the quantum cascade laser offers a degree of flexibility to the device designer that bipolar devices do not. The wavelength range of the bipolar devices is limited by the quality of materials available, and even though low band gap materials exist that will allow laser operation in the 2 μm-5 μm wavelength range (see chapter 10), extending the wavelength beyond this poses real difficulties in terms of epitaxial growth and carrier confinement. The quantum cascade laser overcomes these difficulties by relying not on fundamental material properties such as the band gap but on engineered properties such as the subband spacing. The choice of materials is irrelevant to some extent, provided of course that the conduction band offsets are large enough for effective confinement at the lattice matched composition and the

material growth is well enough controlled to allow abrupt interfaces in well defined layer sequences. Lattice matching is essential because of the total thickness of a typical device. Any sort of strain built up over the layer sequence will result in relaxation unless the strain can be compensated, and whilst such systems have been demonstrated in the laboratory [13] they are far from commercialisation.

Quantum cascade lasers have employed several different schemes for the active regions. The first laser [13] utilised a diagonal transition rather than the vertical transition shown in figure 12.1. In a diagonal scheme the final state of the transition is centred on a different well from the initial state so the electron must move sideways in real space. Such a requirement increases the upper state lifetime, which is essentially dominated by optical phonon scattering. Central to the diagonal transition is the notion of "anti-crossing". The confined states move in energy with the application of an electric field but of course the motion is relative to the point of reference. Some states will move up and others will move down, but when they coincide in energy they will anti-cross, as in figure 12.3. The states are resonant at this point and coupling between the two occurs. The subbands of the active region are designed to anti-cross at the operational electric field of the laser.

Lasers based on inter-subband transitions have been reported at various wavelengths in the near infra-red, at various powers and temperatures of operation. A sense of the progress can be gleaned from Table 12.1, in which some of the experimental devices that have appeared in the literature are

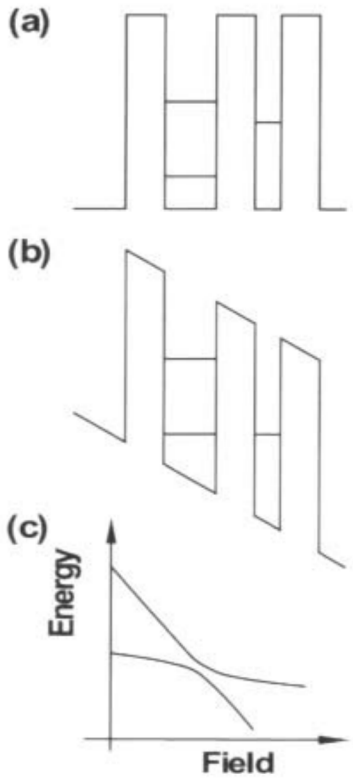

Figure 12.3. Anti-crossing in states of adjacent quantum wells.

Table 12.1. Progess in quantum cascade structures.

λ (μm)	Power (W)	Reflector	T (K)	Material	Reference
4.2	0.008 pulsed	cleaved facet	90	InGaAs/InAlAs	14
4.3	0.004 pulsed	cleaved facet	125	InGaAs/InAlAs	15
4.3	0.03 pulsed	cleaved facet	125	InGaAs/InAlAs	16
5	0.2 pulsed	cleaved facet	320	InGaAs/InAlAs	17
5.3	1.15 pulsed	DFB	420	InGaAs/InAlAs	18
5.9	0.67 pulsed	coated facet	400	InGaAs/InAlAs	19
8.5	0.06 pulsed	DFB	300	InGaAs/InAlAs	20
8.5	0.01pulsed	cleaved facet	305	InAs-GaAs-AlAs	21
10.16	0.23 pulsed	DFB	85	InGaAs/InAlAs	22
	0.08 pulsed		300		

described. Most noticeably, the peak powers and operating temperatures have improved steadily with time, though the extension of the quantum cascade concept to the GaAs material system took some time, the earliest devices all exploiting the InGaAs-InAlAs system. The list of devices and achievements given in table 12.1 is by no means comprehensive and the structures themselves will not be described as they differ mainly in the layer sequence rather than the composition. The distributed feedback reflectors described by Hofstetter *et al*. are chemically wet-etched into the top waveguide layer from the surface down to a depth of 100 nm for the lasers at 5.3 µm [18] (grating period = 825 nm) and 400 nm in the lasers emitting at 10.16 µm [22] (period 1.59 µm). The semiconductor/air interface is thus a vital part of the waveguide structure, which is of course a potential weakness of the device. However, these structures are a long way from commercialisation. Gmachl *et al*. [20] preferred to regrow a thick InP layer over the grating, but the problem with all-semiconductor waveguides, and one neatly avoided by Hofstetter *et al*., is free-carrier absorption, which was shown in chapter 4 to depend on the square of the wavelength. The top part of the waveguide must also serve as the contact layer and some absorption of the laser radiation is inevitable.

In addition to the discrete subbands in quantum wells some lasers have been designed with a superlattice active region so that optical transitions occur from the bottom of one miniband to the top of another. Minibands are less susceptible to imperfections in the growth as there exists a range of energies into which electrons can be injected.A typical such device is a 9 µm $Ga_{0.47}In_{0.53}As/Al_{0.48}In_{0.52}As$ superlattice structure grown by gas source MBE at the Center for Quantum Devices at Northwestern University, Illinois, and described by Razeghi and Slivken [23]. The active region comprises six well and barrier pairs and the injector comprises seven well and barrier pairs as detailed in table 12.2. Several such active/injector stages will be included in a

Table 12.2. Layer sequences used by Razeghi and Slivken [23].

Active		Injector	
Well	Barrier	Well	Barrier
5.6	0.9	2.8	2.5
5.2	0.9	2.8	2.5
4.8	0.9	2.7	2.6
4.5	1.0	2.5	2.6
4.1	1.1	2.4	2.7
3.8	2.5	2.4	2.9
		2.3	3.5

single device, so the necessity to produce 26 layers per stage of an accuracy much greater than 0.1 nm illustrates the demands of the growth technology. Furthermore, in mid-IR ranges the dominant optical interaction in matter is free carrier absorption so the background doping has to be very carefully controlled. If it is too high the absorption losses will be large but if it is too small it will not counteract the space charge induced by high electron injection, which will subsequently distort the internal electric field and disrupt the laser operation.

After growth the devices are patterned into ridge or double channel waveguides by photolithography and chemical or plasma etching. The waveguides are typically 15-25 μm wide (which would normally be considered as a broad area in high power double heterostructures but because of the longer wavelength will still support a single lateral mode) and then cleaved into 3 mm long cavities. At room temperature these lasers are capable of delivering peak powers of 7 W at 3.26 A with a slope efficiency of 4.4 W/A, which equates to 32 emitted photons per injected electron. Operating temperatures up to 470 K have been achieved.

12.2 Minibands in superlattices

Minibands can be best understood through the Kronig-Penney model, which was first published in 1931 [24] as an explanation of the band structure of solid state materials in terms of the propagation of an electron inside the periodic potential due to the atoms in the lattice. The form of the model successfully predicts the existence of bands inside solids but the potential is assumed to be square, and therefore not representative of real bulk materials. For this reason it has never really found much use in the calculation of real band structures, although it has not been entirely forgotten [25]. However, with the advent of the

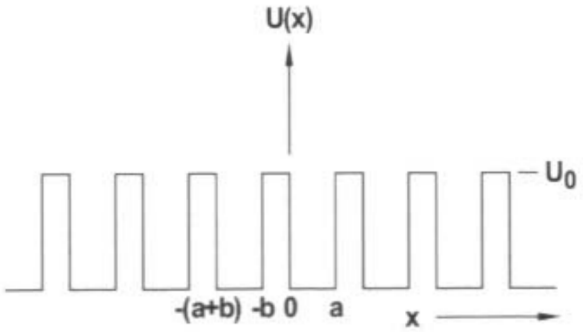

Figure 12.4. The square potential Kronig-Penney model.

superlattice the Kronig-Penney model has found renewed favour because the square potential is nothing less than an idealisation of the abrupt band structure of the heterostructure superlattice. Indeed, Leo Esaki, the founder of modern semiconductor quantum engineering, reports that in making the first super-lattice he was in fact trying to make a practical Kronig-Penney structure [26]. It is not surprising, therefore, that the Kronig-Penney model should apply so well to these structures.

The Kronig-Penney potential is shown in figure 12.4. A potential barrier of width b and height U_0 alternates with a potential well of width a. There are at least three methods of calculating the properties of an electron propagating in such a structure; the boundary matching method, the transfer matrix method, and the Green's function method. The transfer matrix method is a standard formulation for treating wave propagation in periodic structures and relates the wave properties, i.e. amplitude and phase, in one layer to the wave properties in the preceding layer, but it is easy to lose sight of the essential physics in the process. The Green's function method is inherently mathematical and will not be described further, so the boundary matching method will be used here. In essence, this requires the electron wavefunction and its derivative to be continuous across each interface in the structure, thereby matching the wavefunction and its derivative at each boundary.

We start, as before, with Schrödinger's equation.

$$\left[\frac{p^2}{2m_0} + V(r)\right]\psi(r) = E\psi(r) \tag{4}$$

where

$$\frac{p^2}{2m_0} = \frac{-\hbar^2}{2m_0}\nabla^2 \tag{5}$$

and $V_{(r)}$ is the spatially varying potential. All other symbols have their usual meaning. In the Kronig-Penney model as first envisaged the square potential is intended as an approximation to the lattice potential itself, but here the square

potential represents the superlattice conduction band which is super-imposed onto the lattice potential. This is an important difference between the two cases. In the presence of the lattice potential only the wave function is given by the Bloch function (see Appendix III)

$$\psi_{nk} = u_{nk}(r)\,\exp(jkr) \tag{6}$$

but in the superlattice we resolve the potential into its lattice (L) and superlattice (SL) components

$$V(r) = V_L(r) + U_{SL}(r) \tag{7}$$

and write the wave function as

$$\psi(r) = F(r)u_{n0}(r) \tag{8}$$

where u_{n0}(r) is the Bloch function of the n^{th} band at $k = 0$, as explained in appendix III, and $F(r)$ is a slowly varying envelope. Strictly, for this formalism to apply the potential U_{SL}(r) also has to vary slowly, which clearly it does not. It is for this reason that the superlattice is divided into its constituent parts of well and barrier with an effective wavefunction defined in each, which, together with the derivative, are matched at the boundaries. Strictly, it is the envelope function rather than the Bloch function that is matched at the boundary, because the Bloch function varies on the scale of the underlying lattice. Moreover, because the superlattice wells and barriers are lattice matched the Bloch function retains its periodicity across the boundary, and as with the treatment of the confined states in a quantum well, the Bloch function can be discarded. Therefore, the superlattice envelope function is replaced with a wavefunction

$$\psi(z + d) = e^{ikd}\psi(z) \tag{9}$$

which is itself a Bloch function but periodic in the superlattice period $d = a + b$.

The wave functions in both the well (A) and the barrier (B) can be defined. It follows immediately that

$$\psi_A(0) = e^{-ikd}\psi_B(d) \tag{10}$$

so the four boundary conditions of interest can be defined as

$$\psi_A(a) = \psi_B(a) \tag{11}$$

$$\frac{1}{m_A^*}\psi_A'(a) = \frac{1}{m_B^*}\psi_B'(a) \tag{12}$$

$$\psi_A(0) = e^{-ikd}\psi_B(d) \tag{13}$$

and

$$\frac{1}{m_A^*}\psi_A'(0) = \frac{1}{m_B^*}e^{-ikd}\psi_B'(d) \tag{14}$$

The wavefunction can be constructed from a mixture of left and right propagating plane wave functions. Within the well, $0 < z \leq a$

$$\psi_A(z) = C_A e^{ik_A z} + D_A e^{-ik_A z} \tag{15}$$

where C_A and D_A are amplitude coefficients, and within the barrier, $a < z \leq d$

$$\psi_B(z) = C_B e^{ik_B z} + D_B e^{-ik_B z} \tag{16}$$

where the wavevectors are

$$k_A = \frac{1}{\hbar}\sqrt{2m_A^* E} \tag{17}$$

and

$$k_B = \frac{1}{\hbar}\sqrt{2m_B^*(E - U_0)}. \tag{18}$$

For an electron energy $E<U_0$ the wavevector k_B is imaginary, and the transformation can be made

$$k_B = i\mathrm{K}_B \tag{19}$$

where $K_B$is real and corresponds to $E>U_0$. Substituting the well and barrier wavefunctions into the four boundary conditions leads to a set of simultaneous equations for which the determinant must vanish. The solution to this equation for $E<U_0$ can be written [26]

$$\cos(kd) = \cos(k_A a)\cosh(\mathrm{K}_B b) - \frac{1}{2}\left[\mathrm{K} - \frac{1}{\mathrm{K}}\right]\sin(k_A a)\sinh(\mathrm{K}_B b) \tag{20}$$

where

$$\mathrm{K} = \frac{k_A m_B^*}{k_B m_A^*}. \tag{21}$$

A similar expression exists for the case $E>U_0$.

In solid state physics texts dealing with the atomic Kronig-Penney model this expression is often simplified by assuming that the barrier thickness tends to zero in the limit as U_0 tends to infinity and also that the effective masses are both unity [27]. This is clearly not appropriate for a superlattice and which the barrier width and height are fixed by the details of the growth, but it is instructive nonetheless to make the assumptions. Then

$$\frac{P}{k_A a}\sin(k_A a) + \cos(k_A a) = F(k_A a) = \cos(ka) \tag{22}$$

where P is a positive finite number. This function is plotted in figure 12.5 for $P = 2\pi$. The straight lines at $F(k_A a) = \pm 1$ represent the limits of the cosine on

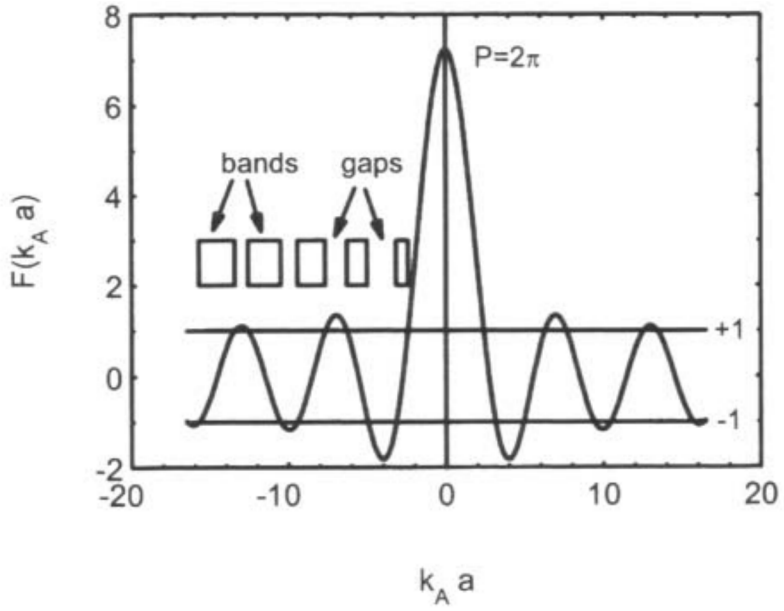

Figure 12.5. Solution of the Kronig-Penney model.

the right hand side. Values of F($k_A a$) lying outside this range are, by definition, invalid, and these define the energy gaps by way of forbidden values of the wavevector. The boxes represent the energy bands spanning the range $|F(k_A a)|<$ 1. At low $k_A a$ the bands are narrow and widely spaced but as $k_A a$ increases the bands become wider and the gaps narrower. Mathematically, the first term on the left hand side dominates for low values of $k_A a$ but this rapidly decays and gives way to the cosine term on the left hand side for large $k_A a$. The function $F(k_A a)$ therefore has no regions of invalidity at large $k_A a$ and the energy gaps disappear as the bands touch. For smaller values of P the bands will merge at lower values of $k_A a$ but there will always be an energy gap at low energies.

The same general results apply to superlattices, but the limitation of a definite barrier width and a finite barrier height means that above barrier states will exist. These will also exhibit band gaps for electron energies just above the barrier but at high electron energies the bands will merge to form a continuum. These states are not of primary interest in the quantum cascade laser, so figure 12.6 shows the confined states of a GaAs/AlGaAs Kronig-Penney superlattice as a function of both well width and barrier width [26]. "Confined" in this context means that the electron energy lies below the barrier height, but of course it is the nature of a superlattice that, unlike a single quantum well with wide barriers, electrons can propagate along the superlattice. For the well-width dependence a fixed barrier width of 5 nm was assumed and similarly for the barrier width dependence a fixed well width of 5 nm was assumed. The conduction band offset, i.e. barrier height, occurs at ~ 190 meV so for a well thickness of zero, corresponding to bulk AlGaAs there are no states in the well. It is easy to imagine the top of the first band continuing up in energy and

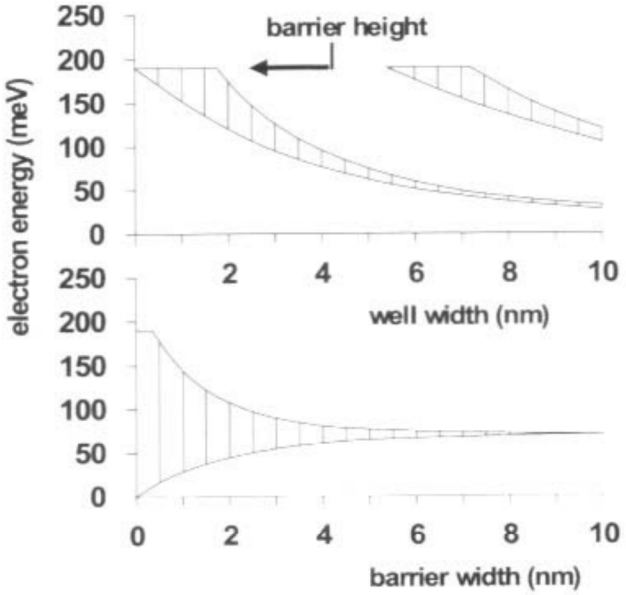

Figure 12.6. Superlattice states in a GaAs/A1GaAs periodic structure. (After Steslicka [26].)

meeting the y-axis at ~ 370 meV, where in fact it coincides with the bottom of the second band. There are therefore no band gaps at zero well-width, which is what would be expected on intuitive grounds.

The band states behave very much like the confined states of an isolated well, with the obvious exception of the width of the band, which is a direct result of the communication with neighbouring wells. Thus, the first band starts to descend into the well immediately the well thickness exceeds zero but doesn't become fully confined until a thickness of 1.7 nm, at which point the top of the band descends into the well. The band energy decreases with well width, as does the width of the band. In effect the electrons spend more time in the well than in the barriers and there is correspondingly a much smaller probability, though still non-zero, of communication with neighbouring wells. Eventually, as the well width increases above 5.3 nm, the second band descends into the well. In this particular configuration no third band is present within the well, but in fact a third and fourth band do exist above the barrier height. These bands exist because even though the wavevector is real at every point within the superlattice the discontinuous change in potential at the interfaces leads to discontinuous changes in the wavevector, and because the system is periodic stop bands will occur for any wavevector satisfying the Bragg condition (see chapter 9). At any given well width the band gap narrows as the electron energy increases and the bands become wider, in accordance with figure 12.5, so eventually a continuum of states forms above the barrier. As in the bulk, the bands can be described with reference to a mini-Brillouin zone corresponding to momenta $\pm m\pi/d$, where d is the superlattice period.

The effect of increasing the barrier width can also be understood by similar intuitive reasoning. At zero barrier width the band structure should

conform to the GaAs band structure, and in figure 12.6 the band extends over the whole energy range of the well. This well is deep enough for one confined state only at a width of 5nm and as the barrier width increases the band becomes progressively narrower until eventually it becomes a discrete state at ~ 70 meV. In effect all communication with neighbouring wells has ceased and the electron is isolated. Above the barrier, bands exist but the bands are in fact widest for narrow barriers. For wide barriers the band widths decrease but so also do the band gaps, because the bands must tend to the continuum of states.

The superlattice structures described above are uniform and therefore the minibands are also uniform, but the electronic structure is modified under the presence of an electric field. In a bulk material the conduction band will simply have the potential associated with the electric field super-imposed upon it, but in a superlattice the effect of the electric field is to separate in energy the individual states of the quantum wells so that they are no longer in resonance and no longer interact. The superlattice breaks up into a series of discrete states at well defined energy intervals, a so-called Wannier-Stark ladder [28]. In a quantum cascade laser the layer thicknesses and compositions are graded so that at zero field states do not align but under the action of an electric field these states are brought into resonance and a miniband is formed. This is somewhat different from the idealised superlattices described above but the principle remains the same. The miniband states arise from interacting states of a series of quantum wells, and the number of states within a mini band corresponds to the number of interacting wells. In addition, the mini band extrema correspond to the mini Brillouin zone boundaries.

12.5 Miniband cascade lasers

There are several advantages if the transitions take place between minibands rather than discrete subbands, and several lasers emitting over a wide range of wavelengths have been demonstrated. The wavelength is determined principally by the energy separation between the minibands

Population inversion is easier to achieve in a superlattice active region provided the electron temperature is low enough for the bottom miniband to remain largely empty and provided the width of the miniband is larger than the optical phonon energy. If the electron temperature is too high then all the states in the miniband will be full and scattering between states will not occur. However, in an empty miniband the intra-miniband scattering times are usually much shorter than inter-miniband scattering times so that depopulation of the lower laser level is readily achieved. Similarly the doping level must be such that the Fermi level lies well below the top of the lower miniband. However, there is an additional requirement on the doping level. It should be optimised to ensure that the electric field does not penetrate significantly into the

superlattice so that the field is dropped almost entirely across the injector region.

These ideas were first demonstrated at $\lambda = 8$ μm in an eight period superlattice with 1 nm thick barriers of $Al_{0.48}In_{0.52}As$ and 4.3 nm thick $Ga_{0.47}In_{0.53}As$ wells [31]. The fact of eight periods means that each mini-band comprises 8 states with an electroluminescence spectrum 30 meV wide compared with 10 meV for similar quantum cascade structures. This increases the threshold current somewhat despite the high oscillator strength of approximately 60, and an intra-miniband relaxation time of ~0.1 ps compared with the optical mode phonon scattering time of ~ 10 ps from the bottom of the second miniband to the top of the first miniband. An intrinsic superlattice has also been proposed [32], the main difference being that the dopants are placed in the injection and relaxation regions so that the dopants are separated from the extrinsic electrons. The electric field generated by this mechanism exists primarily in the active regions and is then cancelled by the applied field to give a similar profile as shown in figure 12.12.

As well as inter-miniband transitions, structures can be designed so that transitions occur between a bound state and a miniband [33, 34] (figure 12.13). The essential idea is that a chirped superlattice gives rise to minibands under the influence of an electric field. These bands are never truly flat, but extend reasonably uniformly over a limited region of space until they break up to form the next miniband. In this scheme there is no separation of the active and the injection regions, as occurs in other devices, as the lower laser level and the injector are one and the same miniband. However, a single narrow well placed at strategic points within the superlattice gives rise to a discrete state which acts as the upper laser level. The upper miniband is effectively redundant. One of the potential disadvantages of this design is the low oscillator strength, as understood from the sum rule, and the relatively large linewidth. Transitions from the upper laser level to various points within the miniband are possible,

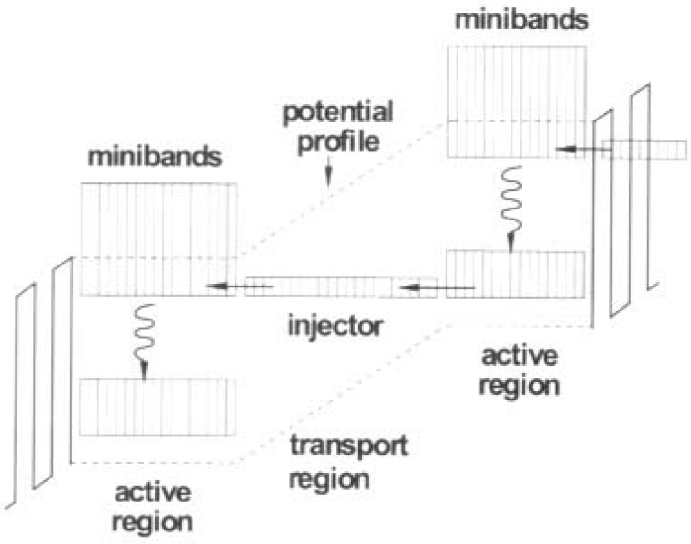

Figure 12.12. A schematic of a superlattice cascade laser.

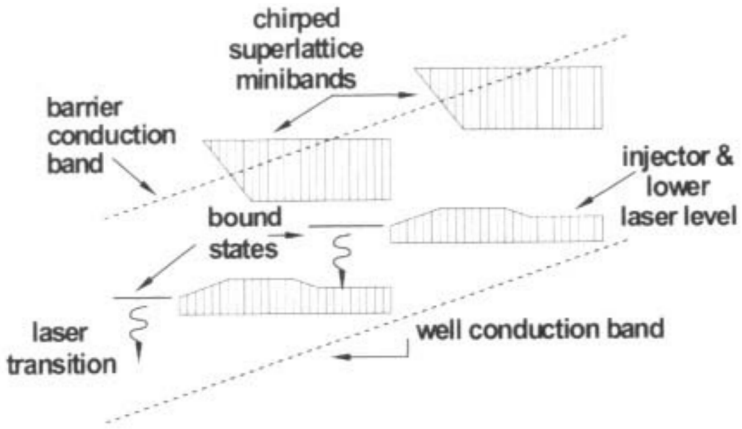

Figure 12.13. A bound-continuum cascade laser. (After Faist *et al*. [33].)

leading to a linewidth of some 20 meV or so [34]. However, this is offset by the realisation that room temperature operation will in any case cause a thermal broadening of this magnitude, and the advantage of the fast depopulation rate of the lower laser level makes this an attractive scheme for high temperature operation.

12.7 Waveguides in quantum cascade structures

Waveguiding structures in quantum cascade lasers present a serious technological challenge to the laser manufacturer. In waveguides based on the conventional refractive index difference the optical penetration into the cladding layers is proportional to the wavelength, which can be very long in the case of THz devices. If the cladding layer is not thick enough to contain the entire optical field some irreversible leakage, i.e. loss, out of the guide will occur. Of course, if the core of the guide is itself optically thick then for the fundamental mode the optical field at the core-cladding interface is small compared with the field at the centre of the guide so the loss from this mode will be small. However, power will also propagate in the higher order modes and these will be more lossy. If conventional waveguides are to be built then structures several times thicker than the wavelength need to be grown. Even in the GaAs/AlGaAs system difficulties occur for thicknesses greater than ~1.5 μm due to residual strain in the AlGaAs, so alternative approaches are necessary.

Sirtori *et al*. [45] proposed the use of heavily doped GaAs cladding layers for mid-IR wavelengths in the range 5-20 μm, and demonstrated the principle on a laser emitting at 8.92 μm. Heavily doped GaAs ($n \approx 5 \times 10^{18}$ cm^{-3}) has a plasma frequency around 11 μm, and around this wavelength the real part of the refractive index decreases. Heavily doped GaAs cladding layers will therefore provide a large refractive index difference at the wavelength of interest. The active region itself was quite small and consisted of only 36 periods, totalling just over 1.5 μm thickness, but was embedded in ≈ 3.5 μm of moderately doped GaAs ($n \approx 4 \times 10^{16}$ cm^{-3}) on either side adjacent to which was the highly doped layer. GaAs rather than AlGaAs was used for the simple reason given above that the thickness of AlGaAs is limited but GaAs on the other hand can be grown much thicker. The structure is illustrated in figure 12.16, along with the refractive index profile. The highly doped GaAs layers

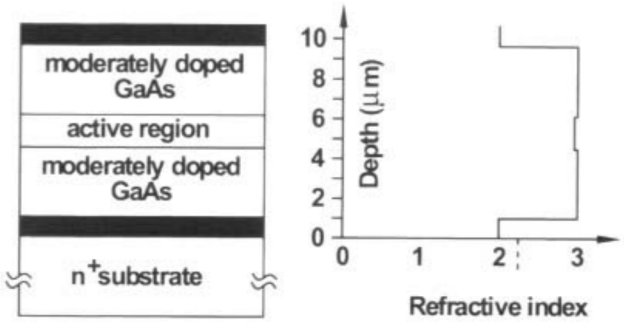

Figure 12.16. A waveguide structure based on carrier induced refractive index changes in the infra red. The heavily doped layers are shown in black.

(black) have a refractive index of 2 compared with 3 for the moderately doped GaAs, and this is sufficient to confine the optical field fully within this structure. The active region has a slightly lower refractive index because of the presence of AlGaAs barriers, but the effect is small and doesn't significantly alter the confinement properties.

Such high doping levels inevitably leads to absorption at long wavelengths, and in this structure the extinction coefficient, i.e. the imaginary part of the refractive index, is estimated to be $k = 0.1$ falling abruptly to 10^{-4} in the moderately doped regions. By way of comparison, the absorption depth at this wavelength for $k = 0.1$ is about 7 μm, but that strictly applies to a normally incident plane wave. Nonetheless, the cavity length will be several hundreds of micrometres long and the fact of such a high extinction coefficient will lead to considerable loss, calculated by the authors of this structure to be 1740 cm^{-1}, which is approximately 90% of the total waveguide loss despite an optical overlap in the heavily doped regions of 0.008. Losses in heavily doped cladding layers such as this therefore represent the most significant source of loss in the cavity, but are in fact an unavoidable consequence of this type of waveguiding structure.

Waveguides for longer wavelengths utilise a particular property of thin metal films that at long wavelengths where the real part of the dielectric constant is large and negative such films, when bounded by a dielectric with a real and positive dielectric constant, will support an electromagnetic mode that will propagate over large distances [46, 47]. In fact, the condition on the dielectric constants is not so strict and modes will propagate over a wide variety of wavelengths. Under the conditions described above the electromagnetic mode that propagates has a magnetic vector parallel to the interface and an electric vector perpendicular to the plane of the film, which is consistent with the polarisation properties of the emitted radiation. Therefore a metal contact placed over the active region serves two purposes; to act as a contact and to act as waveguide. This waveguide will act as such over a wide range of frequencies, but of course if the imaginary part of the dielectric constant of either the metal or the semiconductor is non-zero some attenuation of the propagating beam will occur.

The fact of a perpendicular electric vector requires a surface charge density on the metal. Induced by the electric field and oscillating with it, this charge density corresponds to the collective oscillation known as a plasmon, hence the guide is often called a surface plasmon guide. The propagation depth into the semiconductor is quite long so the overlap from a mode supported by a single metal layer is quite small. Penetration of the optical field into the heavily doped substrate will lead to some attenuation but unless the substrate is doped so heavily that it becomes almost metallic it will not significantly alter the mode profile. Encapsulating the active region between two metal layers increases the overlap almost to unity, so whilst the attenuation of the mode

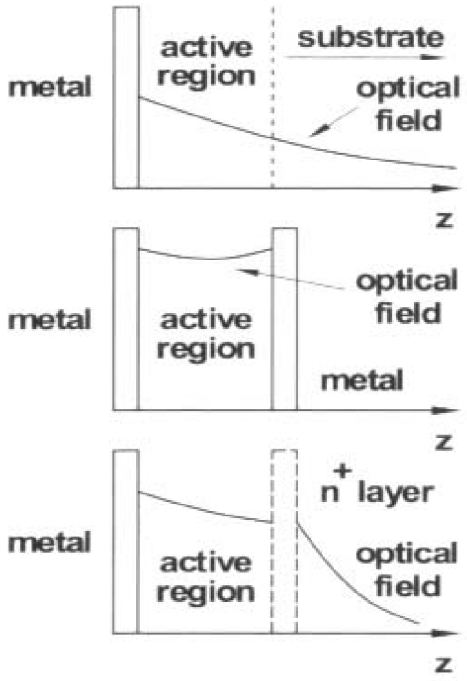

Figure 12.17. Surface plasmon wave guide structures using a single metal layer, two metal layers, or a metal layer in conjunction with a doped semiconductor.

might increase, the figure of merit α/Γ may well decrease. However, in order to achieve this sort of structure it is necessary to bond the wafer to a substrate after depositing the top metal contact in order to provide some mechanical support while the substrate is etched off and a metal layer deposited on to what was the bottom of the device. Such structures have been built and demonstrated at $\lambda = 100\ \mu m$ [48]. An alternative to a double metal wave guide is a heavy doped semiconducting or highly conducting intermetallic layer beneath the active region. Intermetallics such as NiAl compounds can be grown epitaxially on GaAs [49] and so have the advantage that the substrate does not need to be etched. Heavily doped semiconducting layers have the same advantage but suffer from the disadvantage that the mode field can leak out in the case of a thin layer and reduce the overlap. The optical fields from the different configurations are illustrated schematically in figure 12.17.

12.8 Summary

Quantum cascade lasers represent a radical departure from conventional diode lasers in which recombination of electrons and holes across the band gap results in photon emission. Necessarily the electron is removed from the active region by this process so that the maximum possible internal efficiency is one photon per electron-hole pair. In the cascade laser population inversion is achieved between electron subbands in a multiple quantum well structure, which means that the electron is available for further stimulated emission in another set of subbands. Quantum cascade lasers therefore contain more than one active region separated by injector regions designed for the sole purpose of transporting the electron from one active region to another.

The design of these systems is a complicated exercise in numerical computation involving solution of Schrödinger's equation, Fermi's Golden Rule for electron-phonon and electron-electron scattering, and in some instances Monte Carlo formulations of transport. For these reasons no simple design rules exist, but some principles of operation can be discerned. These are:

- At least three levels are needed in an active region.
- The lowest level serves the dual purpose of depopulating the lowest laser level and feeding the electron into the next injector region.
- The lifetime of the lower laser level has to be shorter than the lifetime of the upper laser level in order to achieve population inversion.
- Depopulation by optical mode phonon scattering is one of the most effective scattering mechanisms from the lower laser level, so the energy separation between the lowest levels should correspond to the phonon energy.
- Emission wavelengths can be extended out to the very far infra-red ($\approx 100\ \mu m$) by careful design of the active region.

- Superlattices are usually used for the injector regions but they can also be used for the active regions in both bound-to-continuum and continuum-to-continuum designs.
- Using a superlattice as the lower laser level has the advantage that inter-miniband scattering times are extremely fast and population inversion is easier to achieve.
- Waveguide structures are more difficult to fabricate than in conventional edge-emitting devices but metal waveguides based on surface plasmons are common.

It should be borne in mind that whilst a very wide diversity of designs and operating wavelengths have been demonstrated, the quantum cascade laser is still a relatively recent invention and the technology is still very much under development. Many of the ideas described in this chapter may well find their way into commercial devices, but this is by no means clear at present.

12.9 References

[1] Gmachl C, Capasso F, Sivco D L and Cho A Y 2001 *Reports on Progress in Physics* **64** 1533–1601
[2] Harrison P and Kelsall R W 1998 *Solid State Electronics* **42** 1449–1451
[3] Harrison P 1999 Appl. Phys. Lett. 75 2800-2830
[4] Hyldgaard P and Wilkins J W 1996 *Physical Review* **B53** 6889–6892
[5] Tripathi P and Ridley B K 2003 *J. Phys.: Condem. Matter* **15** 1057–1069
[6] Kinsler P, Harrison P and Kelsall R W 1998 *Phys. Rev*. **B58** 4771–4778
[7] Gorfinkel V B, Luryi S and Gelmont B 1996 *IEEE J. Quant. Electr*. **32** 1995–2003
[8] Yang Q K and Li A Z 2000 *J. Phys.: Condens. Matter* **12** 1907–1914
[9] Suchalkin S, Bruno J, Tober R, Westerfield D, Kisin M and Belenky G 2003 *Appl. Phys. Lett*. **83** 1500–1502
[10] Helm M 1995 *Semicond. Sci. Technol*. **10** 557–575
[11] Iotti R C and Rossi F 2001 *Phys. Rev. Lett*. **8**7 art. no 146603
[12] Ohtsuka T, Schrottke L, Key R, Kostial H and Grahn H T 2003 *j. Appl. Phys*. **94** 2192–2198
[13] Faist J, Capasso F, Sivco D L, Hutchinson A L, Chu S-N G and Cho A Y 1998 *Appl. Phys. Lett*. **72** 680–682
[14] Faist J, Capasso F, Sivco D L, Hutchinson A L and Cho A Y 1994 *Science* **264** 553–556
[15] Faist J, Capasso F, Sivco D L, Hutchinson A L, Sirtori C, Chu S-N G and Cho AY 1994 *Appl. Phys. Lett*. **65** 2901–2903
[16] Faist J, Capasso F, Sivco D L, Sirtori C, Hutchinson A L and Cho A Y 1994 *Electronics Letters* **30** 865–868
[17] Faist J, Capasso F, Sirtori C, Sivco D L, Baillargeon J N, Hutchinson A L, Chu S-N G and Cho A Y 1996 *Appl. Phys. Lett* **68** 3680–3682
[18] Hofstetter D, Beck M, Allen T and Faist J 2001 *Appl. Phys. Lett*. **78** 396–398

[19] Yu J S, Slivken S, Evans A, David J and Razeghi M 2003 *Appl. Phys. Lett*. **82** 3397–3399
[20] Gmachl C, Capasso F, Faist J, Hutchinson A L, Tredicucci A, Sivco D L, Baillargeon J N, Chu S-N G and Cho A Y 1998 *Appl. Phys. Lett*. **72** 1430–1432
[21] Carder D A, Wilson L R, Green R P, Cockburn J W, Hopkinson M, Steer M J, Airey R and Hill G 2003 *Appl. Phys*. Lett. **82**? 3409–3411
[22] Hofstetter D, Faist J, Beck M, Müiller A and Oesterle U 2000 *Physica* **E7** 25–28
[23] Razhegi M and Slivken S 2003 *Physica Status Solidi* (a) **195** 144–150
[24] Kronig R de L and Penney W G 1931 *Proc. Roy. Soc. (London)* **A130** 499
[25] Eldib A M, Hassan H F and Mohamed M A 1987 *J. Phys. C: Solid State Phys*. **20** 3011–3019
[26] Steslicka M, Kucharczyk R, Akjouj A, Djafari-Rouhani B, Dobrzynski L and Davison S G 2002 *Surface Science Reports* **47** 93–196
[27] Kittel C 1976 *Introduction to Solid State Physics* 5th Edition (New York: John Wiley & Sons) p 192
[28] Manenti M, Compagnone F, Di Carlo A, Lugli P, Scamarcio G and Rizzi F 2003 *Appl. Phys. Lett*. **82** 4029–4031
[29] West L C and Eglash S J 1985 *Appl. Phys. Lett*. **46** 1156–1158
[30] Gelmont B, Gorfinkel V and Luryi S 1996 *Appl. Phys. Lett*. **68** 2171–2173
[31] Scamarcio G, Capasso F, Sirtori C, Faist J, Hutchinson A L, Sivco D L and Cho AY 1997 *Science* **276** 773–776
[32] Tredicucci A, Capasso F, Gmachl C, Sivco D L, Hutchinson A L, Cho A Y, Faist J and Scamarcio G 1998 *Appl. Phys. Lett*. **72** 2388–2390
[33] Faist J, Beck M, Aellen T and Gini E 2001 *Appl. Phys. Lett*. **78** 147–149
[34] Rochat M, Hofstetter D, Beck M and Faist J 2001 *Appl. Phys. Lett*. **79** 4271–4273
[35] Kelsall R W, Kinsler P and Harrison P 2000 *Physica E* **7** 48–51
[36] Xin Z J and Rutt H N 1997 *Semicond. Sci. Technol*. **12** 1129–1134
[37] Lyubomirsky I and Hu Q 1998 *Appl. Phys. Lett*. **73** 300–302
[38] Kohler R, lotti R C, Tredicucci A and Rossi F 2001 *Appl. Phys. Lett*. **79** 3920–3922
[39] Callebaut H, Kumar S, Williams B S, Hu Q and Reno J L 2003 *Appl. Phys. Lett*. **83** 207–209
[40] Xin Z J and Rutt H N 1997 *Semicond. Sci. Technol*. **12** 1129–1134
[41] Harrison P 1997 *Semicond. Sci. Technol*. **12** 1487–1490
[42] Scalari G, Ajili L, Faist J, Beere H, Linfield E H, Ritchie D A and Davis G 2003 *Appl. Phys. Lett*. **82** 3165–3167
[43] Kohler R, Tredicucci A, Beltram F, Beere H, Linfield E H, Davis G, Ritchie D A, Dhillon S H and Sirtori C 2003 *Appl. Phys. Lett*.**82** 1518–1520
[44] Rochat M, Ajili L, Willenberg H, Faist J, Beere H, Davies G, Linfield E and Ritchie D 2002 *Appl. Phys. Lett*.**81** 1381–1383
[45] Sirtori C, Kruck P, Barbieri S, Page H, Nagle J, Beck M, Faist J and Oesterle U 1999 *Appl. Phys. Lett*. **75** 3911–3913
[46] Burke J J, Stegeman G I and Tamir T 1986 *Phys. Rev*. B. **33** 5186–5201
[47] Yang F, Sambles J R and Bradberry G W 1991 *Phys. Rev*. B. **44** 5855–5872
[48] Williams B S, Kumar S, Callebaut H, Hu Q and Reno J L 2003 *Appl. Phys. Lett*. **83** 2124–2126

[49] Indjin D, Ikonić Z, Harrison P and Kelsall R W 2003 *J. Appl. Phys*. **94** 3249–3252

Problems

1. Assuming $m_e/m^* \approx 15$ for GaAs, show, by calculating the oscillator strength, that the strongest transitions between the subband of a quantum well are those occurring between states adjacent in energy, i.e. 1→2, 2→3, 3→4, etc. to confirm the requirement that the parity flips in inter-sub band transitions.

2. Using the idealised infinite square well quantum well estimate the thickness of two adjacent wells that from the active region of a laser emitting at $\lambda = 50$ μm. Assume a phonon energy of 36.0 meV similar to bulk GaAs. As an exercise, you might try calculating the equivalent quantities in a real quantum well formed from GaAs/AlAs just to illustrate how complicated is the procedure for designing cascade lasers.

3. Estimate the change in wave vector required for an electron to relax from the top of a miniband to the bottom in a GaAs superlattice with a period of 10 nm. Compare this with the kinetic energy of an electron in bulk GaAs at a similar wavevector.