We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

5,000
Open access books available

124,000
International authors and editors

140M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the top 1% most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Quantum Confinement in High Electron Mobility Transistors

Shovon Pal, Sascha R. Valentin, Arne Ludwig and Andreas D. Wieck

Abstract

Modulation-doped semiconductor nanostructures exhibit extraordinary electrical and optical properties that are quantum mechanical in nature. The heart of such structures lies in the heterojunction of two epitaxially grown semiconductors with different band gaps. Quantum confinement in this heterojunction is a phenomenon that leads to the quantization of the conduction and the valence band into discrete subbands. The spacing between these quantized bands is a very important parameter that has been perfected over the years into device applications. Most of these devices form low-dimensional charge carriers that potentially allow optical transitions between the subbands in such nanostructures. The transition energy differences between the quantized bands/levels typically lie in the infrared or the terahertz region of the electromagnetic spectrum and can be designed according to the application in demand. Thus, a proper understanding and a suitable external control of such intersubband transitions (ISTs) are not only important aspects of fundamental research but also a necessity for optoelectronic device applications specifically towards closing the terahertz gap.

Keywords: heterojunction, HEMT, terahertz, infrared, intersubband transition

1. Introduction

Low-dimensional semiconductor heterostructures, otherwise known as semiconductor nanostructures, have tremendously revolutionized both the technical and the fundamental aspects of semiconductor industry in terms of device applications. With the ability to grow clean and high-quality samples, device implementations have become a huge success [1–3]. When the dimensions of a region with free carriers (electrons) are reduced as compared to the bulk and approach the deBroglie wavelength, the electronic motion is quantized, thus resulting in carrier confinement that is quantum mechanical in origin. The phenomenon has
been widely used for carrier confinement in one, two and three dimensions that consequently gives rise to nanostructures such as quantum wells, quantum wires and quantum dots, respectively. Due to the quantum confinement, the energy bands (i.e. the conduction and valence bands) are quantized into discrete energy levels/bands and are no longer continuous as in the bulk systems. These quantized energy states are known as subbands for 2D or 1D systems and sublevels for 0D systems. The energetic spacings between these quantized subbands and the sublevels are very important parameters that define the device applications both from an optical and from an electrical point of view.

The intersubband spacings in GaAs-based 2D systems are typically in the order of 10–30 meV [4, 5], as seen in the case of two-dimensional electron gas (2DEGs) with a triangular confinement potential formed across a GaAs/Al_{x}Ga_{1-x}As heterojunction (x being typically 0.3) of a high electron mobility transistor (HEMT) structure. The intersubband transitions (ISTs) typically cover the terahertz (THz) or far-infrared region of the electromagnetic spectrum. However, in the case of a square potential well or in a different material system such as GaN/AlGaN heterojunction, these spacings can be designed to be even in the mid-infrared or near-infrared region. Stacking of quantum wells can further enhance the response of intersubband resonance (ISR), and such designs are the key for various applications like photodetectors or intersubband lasers [6]. One of the very common and sophisticated examples in this regard is the quantum cascade laser [7–9], which is based on the cascade phenomena and intersubband transitions across many layers of quantum wells. Such compact and powerful lasers are used for practical applications in THz spectroscopy [10–13], sensing technology [14, 15], biomedical applications [16, 17] and also in security applications [11, 18]. Structures based on quantum wells have also made significant advancement in the detector technology, for example, quantum well infrared photodetectors [19, 20]. In this chapter, we present a broad overview of the ISTs in a 2DEG formed at the GaAs-Al_{x}Ga_{1-x}As interface of a HEMT structure. We also discuss possible methods to probe the spacing between the subbands and also to tune them significantly by applying an external bias across the sample. Furthermore, we present a fundamental study on the coupling of the ISRs with the 2DEG cyclotron resonance in the presence of tilted magnetic fields. The knowledge of ISTs and the ability of wide electrical tuning of these resonances are then exploited to study the light-matter interaction at THz frequencies in these HEMT structures. The integrated device with 2DEG in a HEMT structure and metamaterials (frequency-selective artificially designed structures) is electrically driven from an uncoupled to a coupled regime of light-matter interaction and then again back to the uncoupled regime. A strong coupling is thus observed when the frequencies of both systems are brought in resonance with each other, manifested as an avoided crossing at that point.

2. High electron mobility transistor design

The low-dimensional charge carriers, trapped in the heterojunction of the HEMT design, form the core of such field-effect transistors. This transistor design also goes by the name of modulation-doped field-effect transistors (MODFET). These designs are used in various high-power [21]
and high-speed [22] electronics, high-resolution imaging [23] and various gas, chemical and biomedical applications [24]. We begin with the design concept of this semiconductor heterostructure along with an overview of its band structure (see Figure 1(a)) that is obtained by solving the Schrödinger-Poisson’s equations self-consistently [25, 26] and adding the band discontinuity at the heterojunctions. A schematic of the layer sequence of the transistor structure is shown in Figure 1(b). On a semi-insulating GaAs substrate/wafer, we start the molecular beam epitaxy (MBE) growth by typically a 50-nm-thin GaAs layer. Then, approximately 10 periods of a GaAs/AlAs short period superlattice (SPS) are grown (not shown in the band diagram). The SPS layers help to smoothen the surface of the bare substrate for the later epitaxial growth and trap eventually surface-segregating unintentional impurities, which have always a tendency to stick at the stoichiometric interfaces of GaAs/AlAs. Moreover, this SPS keeps unwanted charge carriers away, forbidding them to tunnel into the 2DEG layer grown on top. Since the substrate is typically undoped (or semi-insulating), the conduction (or valence) band has no curvature at this point corresponding with Poisson’s equation, which states that the charge density is proportional to the second derivative of the potential with respect to the space coordinate. After the growth of the SPS layer, the first charged layer is the 2DEG that is formed at the heterojunction of the

![Figure 1](image_url)

Figure 1. (a) The conduction band diagram of a HEMT structure along the growth direction $z$. The growth starts from the substrate, that is from right to left in the above figure, after the growth of a 50-nm GaAs layer. $\mu_m$ and $\mu_s$ are the quasi-Fermi levels in the metal and semiconductor, respectively ($\mu_m - \mu_s = eV_b$), $V_b$ is the built-in voltage. The Fang-Howard wavefunction of the ground state is also plotted across the heterojunction (which is assumed to be zero in the growth axis). (b) A schematic of the layer sequence.
undoped GaAs and an undoped Al$_{0.33}$Ga$_{0.67}$As spacer layer. Since the 2DEG is essentially electrons and negatively charged, the conduction band curves downwards and reaches the maximum slope at the heterojunction between the GaAs and the Al$_{0.33}$Ga$_{0.67}$As layer, at which point the conduction band ($E_c(z)$) jumps by $\Delta E_c$ due to the band discontinuity. This is followed by the Al$_{0.33}$Ga$_{0.67}$As spacer layer where charge carriers are absent and the slope of the conduction band almost remains constant. In the doped Al$_{0.33}$Ga$_{0.67}$As layer, the positive charges of the donor ions cause the band to bend upwards, thus reversing the slope. Further moving to the GaAs layer, $E_c(z)$ jumps downwards due to the band discontinuity and continues with a constant slope. On top of the GaAs layer, AlAs/GaAs SPS (also known as blocking barrier) is grown to prevent leakage of charge carriers in and out of the 2DEG channel and also to prevent leakage of surface charges into the channel. Finally, the band hits the gate grown on top of the sample with a barrier height equivalent to the Schottky barrier height. Ideally, metals (e.g. Cr or Au) are evaporated on the sample to serve as gates after the completion of the growth. The samples are typically grown by MBE. While a lot of work has been done previously using metallic Schottky gates, nonetheless, these gates suffer from huge drawbacks. These gates limit the forward bias voltage to the turn-on voltage of the Schottky diode. Furthermore, they fail to grow lattice matched on the semiconductor, are poly-crystalline and thus induce potentially a lot of strain on the semiconductor layer below. Moreover, they oxidize over time and thus may become highly ohmic. Due to high reflectivity and certain Drude absorption of their free charge carriers, such gates are also opaque to the incident light, thus limiting their application in optoelectronic devices. Recently, we have introduced epitaxial, complementary-doped, electrostatic and transparent gates that are grown on top of the sample [27–29]. These gates are grown within the UHV conditions of the MBE and thus incorporate a minimum of the unwanted impurities, leading to unprecedented gate perfection, reliability and reproducibility.

These gates circumvent all the abovementioned disadvantages of Schottky gates and are typically composed of a 25-nm-thick bulk carbon-doped GaAs layer (with an acceptor density of $N_A = 3 \times 10^{18}$ cm$^{-3}$) followed by approximately 40 periods of carbon-delta-doped and 0.5-nm carbon-doped GaAs layers with an average acceptor density $N_A = 1 \times 10^{19}$ cm$^{-3}$. In order to solve Poisson’s equation for the evaluation of the band structure, the knowledge of the charge density is necessary. However, it is not possible to calculate the density of charge carriers until the energy bands are known, thus requiring a self-consistent mechanism that is otherwise adopted in the 1D Poisson solver [26].

3. Characteristics of HEMTs

After being introduced in 1980s, these transistors based on high-mobility modulation-doped heterostructures have revolutionized the semiconductor industry in terms of being the most high-performance compound semiconductor FETs.

Due to spatial separation of the electrons from the ionized impurities, the scattering between them is highly reduced as compared to the bulk semiconductors, enhancing the electron mobility especially at low temperatures where the abovementioned scattering mechanism is dominant. The spacer layer further increases the electron-to-donor separation. While the larger
separation reduces the scattering mechanism, as a negative contribution, the carrier concentration is also reduced which reduces the performance. Hence, the spacer thickness should be optimized. Typical values range from 1 to 30 nm. In order to explain how the high mobility of the electrons in HEMT makes them fast transistors, we use the Shockley’s gradual channel approximation model for an FET operation, which states that the rate of change of saturation drain current, $I_{DS,\text{sat}}$, with respect to the change in the gate-source voltage, $V_{GS}$ (also known as the transconductance, $g_{m,\text{sat}}$), scales proportionally to the mobility ($\mu$) and inversely proportionally to the distance between the gate electrode and the electron channel ($d$), both of which are satisfied by the HEMTs:

$$
g_{m,\text{sat}} = \frac{\partial I_{DS,\text{sat}}}{\partial V_{GS}} = \frac{\mu Z}{L_G} \left( V_{GS} - V_{th} \right),
$$

where $Z$ is the total impedance, $L_G$ is the dimension of the gate and $V_{th}$ is the threshold voltage.

The HEMTs have a clear advantage of lower access resistance particularly in terms of channel resistance due to the high mobility electrons in the channel in comparison to standard FETs. To summarize, the HEMT design principles allow:

- High carrier concentration of 2DEG in the channel
- High mobility by optimization of the spacer-layer thickness
- Low access resistance by using buried/recessed gates
- Better confinement of carriers in the channel due to high barriers
- Reduced interface and alloy scattering mechanisms, thus enhancing mobility

Typical transistor characteristic curves of GaAs/AlGaAs HEMT structures under dark and after 1 s of illumination with a near-infrared light emitting diode (NIR LED) are shown in Figure 2 (a) and (b), respectively. Let us now briefly discuss the transistor operation. Even under zero bias or for a small positive voltage applied to the gate, an inversion layer is formed at the semiconductor surface, the two-dimensional electron gas. Now, if a small source-drain voltage is applied, a current will flow from the source to the drain through the conducting 2DEG channel. The channel here is highly conducting, that is, offers very small resistance, and the source-drain current ($I_{SD}$) is proportional to the source-drain voltage ($V_{SD}$). This is the linear region. As $V_{SD}$ increases, $I_{SD}$ deviates from the linear relationship since the channel potential reduces the charge near the drain end. This eventually reaches a point at which the inversion charge at the drain end is reduced to nearly zero. This is called the pinch-off point. Beyond this source-drain bias, the drain current remains essentially the same, the pinch-off point starts to move towards the source, but the voltage at this pinch-off point remains the same. Thus, the number of carriers arriving at the pinch-off point from the source and hence the current remains essentially the same. In terms of the conduction band diagram, when a large enough negative bias is applied to the gate, the conduction band is lifted up across the chemical potential (or quasi-Fermi level), thus depleting the channel completely. This results in a zero source-drain current even when the source-drain voltage is increased. As a positive bias is applied, the channel is filled with mobile electrons, and with the increase of the source-drain
voltage, the source-drain current first increases linearly, then non-linearly and finally reaches a saturation value, as described earlier. A schematic of these situations is depicted in Figure 2(c) and (e), while the simplified schematic of the conduction band diagram is shown in Figure 2(d). When the structure is illuminated, it becomes rather difficult to deplete the channel completely out of electrons with the previously applied negative bias, and, moreover, the saturation current also increases with the same gate bias applied before.

### 3.1. Quantum confinement and intersubband transitions in HEMTs

One of the most popular terms in nanoscience is the *quantum confinement* that results from changes in the atomic structure as a consequence of direct influence of ultra-small length scale on the energy band structure [30]. The length scale corresponding to the regime of quantum

---

**Figure 2.** Transistor $I_{SD}$-$V_{SD}$ curves measured at 4.2 K for a range of gate voltage applied from $-0.4$ to $2$ V in steps of $0.2$ V (a) at dark and (b) after illuminating the sample with a near-infrared light emitting diode for 1 s. (c) The schematic of the pinch-off situation. (d) A simplified schematic of the conduction band diagram under two different gate bias: $-0.4$ V (in red for the complete depletion of the channel) and $2$ V (in black when the channel is filled with charge carriers). (e) The schematic of the open channel situation. SI indicates semi-insulating. The complete layer sequence is not shown in the schematic for simplicity.
confinement ranges from 1 to 25 nm for typical IV, III–V or II–VI semiconductors [31]. This leads to the fact that the spatial extent of the electronic wavefunction is comparable to the particle size, making the electrons feel the presence of the particle boundaries and respond to changes in particle size by adjusting their energy. This phenomenon is known as the quantum-size effect. Quantization effects become most important when the particle dimension of a semiconductor is near to and below the bulk semiconductor Bohr exciton radius (in bipolar systems) or the deBroglie wavelength (in unipolar systems), making the properties of the material size-dependent.

In low-dimensional semiconductor nanostructures, the restriction of the electronic motion in one, two and three dimensions leads to the modification of the density of states (DOS) as compared to the bulk states. The electronic DOS is defined as the number of electronic states per unit volume per unit energy, the finiteness of which is a result of the Pauli’s exclusion principle, which states that only two electrons with opposite spins can occupy one volume element in the phase space [32]. The confinement of electronic motion results in the quantization of the conduction and the valence band. With the knowledge of these quantized states, their filling can be explained. The number of occupied subbands depends on the electron density and also on the temperature [33]. In a 2D system, the density of electrons per unit area, \( n_{2D} \), is given by the integration of the product of 2D DOS, \( n(E) \) and the Fermi-Dirac occupation function [34]:

\[
n_{2D} = \int_{-\infty}^{\infty} n(E)f(E, E_F)dE,
\]

where \( E_F \) is the quasi-Fermi energy. The subbands can thus be split as:

\[
n_{2D} = \sum_i n_i,
\]

where \( n_i \) is the number of electrons in the subband with energy \( \varepsilon_i \). The above classical Boltzmann distribution, \( f \), is given by [33]:

\[
f(E, E_F, T) \sim \exp \left( -\frac{E - E_F}{k_B T} \right),
\]

where \( k_B \) is the Boltzmann constant and \( T \) is the temperature in Kelvin. Using the above equation, we obtain:

\[
n_i = \frac{m}{\pi \hbar^2} \int_{E_i}^{\infty} f(E, E_F) dE = \frac{nk_B T}{\pi \hbar^2} \ln \left[ 1 + \exp \left( \frac{E_F - \varepsilon_i}{k_B T} \right) \right].
\]

In the limit of low temperature, where electrons are degenerate, the 2D electron density is given by [33]:

\[
n_{2D} = \sum_i n_i = \frac{m}{\pi \hbar^2} \sum_i (E_F - \varepsilon_i) \Theta(E_F - \varepsilon_i).
\]
Figure 3(a) and (b) show a triangular potential well and a schematic of the filling of the subbands, respectively. Based on the position of the Fermi level, the corresponding subbands are occupied. Under triangular confinement potential (as in the HEMT design), the energy spacing decreases for higher subband energies and finally forms the continuum. When the Fermi energy $E_F$ is higher than $\varepsilon_1$ but less than $\varepsilon_2$, only the ground subband is filled. Similarly, when the Fermi level $E_F$ is above $\varepsilon_2$ but less than $\varepsilon_3$, the lower two subbands are filled with electrons and so on. The position of the quasi-Fermi level can be tuned by changing the band structure externally, that is, by applying either an electric or a magnetic field. With the external field, the conduction band can be raised or lowered with respect to the quasi-Fermi level around the Fermi-pinning point, hence depleting or filling the subbands with electrons.

Figure 3(c) shows a schematic of the intersubband transition from the filled ground subband to an empty excited subband. In quantum mechanics, Fermi's golden rule is used to calculate the transition rate (i.e. the probability of a transition to occur per unit time), from one state with a given eigenenergy to another state of higher eigenenergy or to the continuum of energy eigenstates, subjected to some kind of perturbation. According to Fermi's golden rule, this rate of transition, $W_{i\rightarrow f}(\omega)$ from an initial state $|i\rangle$ to a final state $|f\rangle$ under the electric-dipole approximation\(^1\) ($\vec{e} \cdot \vec{p}$), is given by:

$$W_{i\rightarrow f}(\omega) = \frac{2\pi}{\hbar} \left(\frac{eE_0}{m'^*\omega}\right)^2 |\langle f| \vec{e} \cdot \vec{p} |i\rangle|^2 \delta(E_f - E_i - \hbar\omega),$$

(7)

where $E_0$ is the amplitude of the electric field and $e$ and $m'^*$ are the charge and the effective mass of electrons. $E_i$ and $E_f$ are the energies of the initial and the final state. Now, the absorption coefficient, $\alpha_{2D}$, is defined as the ratio of the absorbed electromagnetic energy per unit time and area (considering a 2D system) and the intensity of the incident radiation, summed over all the filled initial and empty final states. In order to ensure that the initial state is filled and the final state is empty, a condition necessary for the transition to occur, we introduce the Fermi factors: $\zeta(E_i)$ for the initial state and $[1 - \zeta(E_f)]$ for the final state. The absorption coefficient is thus given by:

$$\alpha_{2D} = \frac{2e^2}{\hbar cnm'^*\omega^2} \sum_{i,f} |\langle f| \vec{e} \cdot \vec{p} |i\rangle|^2 [\zeta(E_i) - \zeta(E_f)] \delta(E_f - E_i - \hbar\omega),$$

(8)

where $\epsilon_0$ is the absolute permittivity, $c$ is the velocity of light and $n$ is the refractive index of the material. The intersubband absorption takes place within the quantized levels of the conduction or the valence band, schematically shown in Figure 3(c). The total wavefunction can be written as the product of the lattice-periodic Bloch wave (for electrons in a crystal), $'u'$ [35], and

---

1In optical experiments with wavelengths $\lambda$ in the order of micrometers, the width of the quantum well, which practically sets the scale of the electronic wavefunctions, is much shorter than $\lambda$. To a good approximation, the momentum of the photon can be neglected and the electric field can be assumed as constant across the electronic states (with or at $k = 0$). This approximation is also known as the electric-dipole approximation.
a slowly varying envelope function $\psi_0$. According to the Bloch ansatz, the envelope function reduces to the plane wavefunction; hence, the total wavefunction is given by:

$$\psi(r) = \frac{1}{\sqrt{A}} e^{ik \cdot \vec{r}} \cdot q_s(z) \cdot u_\nu(r),$$  \hspace{1cm} (9)$$

where $A$ is the normalization constant, $\nu$ indicates the index for the bands and $s$ represents the subband indices. The complete matrix element in $\langle f | e^{-i\vec{p} \cdot \vec{r}} | i \rangle$ can be split as follows:

$$\langle f | e^{-i\vec{p} \cdot \vec{r}} | i \rangle = e^{-i\vec{p} \cdot \vec{r}} \langle u_\nu | \psi_\nu \rangle \langle \delta_s | \delta_s \rangle + e^{-i\vec{p} \cdot \vec{r}} \langle u_\nu | \psi_\nu \rangle \langle \delta_s | \delta_s \rangle,$$  \hspace{1cm} (10)$$

where the first term indicates the interband transition and the second term is the intersubband transition. The first term has the dipole matrix element of the Bloch functions that explains the selection rule for the interband transition and an overlap integral of the envelope functions. In case of transitions within the subbands of the conduction or the valence band, the first term vanishes and the second term becomes more relevant in the one-band model that consists of an overlap integral of the Bloch function and a dipole matrix element of the envelope function. Further simplification of the matrix elements of the envelope function gives:

$$\langle \delta_{s' \vec{k}'} | \psi_0 | \delta_{s \vec{k}} \rangle = \frac{1}{A} \int d^3r \left[ e^{-i\vec{k} \cdot \vec{r}} q_s^*(z) \{ e_\nu p_\nu + e_\nu p_\nu + e_{-\nu} p_{-\nu} \} e^{-i\vec{k} \cdot \vec{r}} q_s(z) \right].$$  \hspace{1cm} (11)$$

It can be observed that only the third term, $e_{-\nu} p_{-\nu}$ in the curly bracket survives, giving a contribution at a finite frequency. Except for $s' = s'$ and $\vec{k}' = \vec{k}$ (i.e. the initial and the final states are equal), all the other terms vanish, implying the free-carrier absorption at zero frequency when no scattering processes are involved [36]. Hence, only the following matrix element determines the intersubband absorption in the one-band model:
The above equation states that the electric field of the incident radiation must have a component perpendicular to the semiconductor layers or parallel to the growth direction (which is a necessary condition) in order to couple to the ISTs. This is known as the polarization selection rule for the ISTs. In simple words, it states that the electric field vector of the exciting electromagnetic wave or at least a finite component of it must be perpendicular to the 2DEG. Another important quantity in this regard is the oscillator strength \[ f_{s} = \frac{2m^{*} \hbar}{C^{3}} \langle s | p_{z} | s' \rangle \text{ (12)} \]

The above quantity is used to understand and compare the strength of the transitions between initial and final states in different physical systems and obeys the Thomas-Reiche-Kuhn sum rule [37]. It is important to note that for a symmetric quantum well, only parity changing transitions (odd-even or even-odd) are allowed due to the inversion symmetry of the potential well. However, for asymmetric quantum wells, like that of the triangular potential well, the inversion symmetry with respect to the quantum well centre is broken by some means (i.e. internal electric fields or band structure engineering, etc.). This leads to the relaxation of the selection rule, thus allowing transitions between all the subbands.

4. Intersubband-Landau coupling under tilted magnetic fields

When a magnetic field is applied in a plane perpendicular to the semiconductor surface, the free electrons that carry the electric charge perform an orbital motion in the plane perpendicular to the magnetic field direction. This motion is quantized, and equally spaced levels (called the Landau levels) separated in energy are formed. The Hamiltonian of the quantum mechanical system thus gets decoupled into a magnetic and an electric component, and the energy spectrum consists of a series of Landau ladders for each subband. In the presence of a magnetic field, the Drude conductivity, normalized to \( E_{0} c \), is given by [38–42]:

\[
\sigma(\omega) = \frac{n_{2D} e^{2} \tau}{2m^{*} \epsilon_{0} c} \left[ \frac{1}{1 + i(\omega + \omega_{c}) \tau} + \frac{1}{1 + i(\omega - \omega_{c}) \tau} \right] \text{ (14)}
\]

where \( n_{2D} \) is the electron density, \( \tau \) is the relaxation time constant and \( c \) is the velocity of light in free space. Using Maxwell’s theory, the transmittance, \( t \), can be written as:

\[
t = \frac{2}{1 + \sqrt{\epsilon + \sigma}} \text{ (15)}
\]

The transmission, \( T \), is thus given by [41, 42]:

\[
\langle s' | p_{z} | s' \rangle = \int dz \psi_{p}^{*} (z) p_{z} \psi_{s} (z). \text{ (12)}
\]
where $\sigma_r$ is the real part of the conductivity and $\sigma_i$ is the imaginary part. Using the above equation, the normalized transmission can be written as [41, 42]:

$$
\frac{T(B)}{T(0)} = 1 - \frac{(1 + \omega_p \tau)^2}{2} \left[ \frac{1}{(1 + \omega_p \tau)^2 + (\omega + \omega_c)^2 \tau^2} + \frac{1}{(1 + \omega_p \tau)^2 + (\omega - \omega_c)^2 \tau^2} \right]
$$

where $\omega_p = \frac{\omega_p}{\omega_{pl}(1+\gamma)}$ is the plasma frequency. From the quantum mechanical description of such a system, when electrons in 2DEG are subjected to a space-charge potential, $V(z)$ and a magnetic field $B$ tilted at an angle of $\theta$ with respect to the horizontal direction, the total Hamiltonian of the system is given by [4, 43]:

$$
H = -\frac{\hbar^2}{2m^*} \frac{\partial^2}{\partial x^2} + \frac{e^2B^2}{2m^*} x^2 - \frac{\hbar^2}{2m^*} \frac{\partial^2}{\partial z^2} + V(z) + \frac{e^2B_z^2}{2m^*} x^2 - \frac{e^2B_x B_z}{2m^*} xz.
$$

Using the perturbation theory, one can solve the above Hamiltonian treating $\theta$ as the perturbation parameter. In order to solve the above problem, the product of Airy wavefunctions and Hermite functions ($|n\rangle$) is taken as the basis set. From the first order perturbation theory, there is no correction to the zeroth-order energies, expect for the degenerate situation $E_{i,i} = E_i - E_i = \hbar \omega_{c\perp}$. This is commonly addressed as the full-field coupling regime. In the non-resonant regime, second-order effects are present, and hence using perturbation theory of the second order for non-degenerate levels, the total energy eigenvalues are obtained as [4]:

$$
E_{i,n} = E_i + \hbar \omega_{c\perp} \left( n + \frac{1}{2} \right) + \frac{e^2 B_x (z_{i\perp})_n - (z_{i\parallel})^2}{2m^*} - \frac{e^2 B^2}{2m^*} \sum_{\substack{i,j \neq i \perp j \perp \omega_{c\perp}}} \left( z_{i\perp} \right)^2 \left( 1 - \frac{E_i(2n + 1)}{\hbar \omega_{c\perp}} \right),
$$

where $z_{i\perp}$ are the matrix elements for the ISTs from $i$th subband to $i$th subband. The first two terms represent the linear zeroth-order terms, corresponding to the subband energies and the Landau energies respectively. The third term represents the diamagnetic shift, and the fourth term results from the coupling Hamiltonian. This non-resonant regime is known as the half-field coupling, where the splitting is proportional to $\theta^2$. Thus, larger tilt angles are required for the observation of the splitting in the half-field coupling regime. Transmission measurements
are performed on a HEMT sample using the magnetic field chopping scheme. A fixed number of spectra are taken at a certain magnetic field ($T(B)$), and then the magnetic field is turned off, during which the same number of transmission scans ($T(0)$) are taken. At first, the transmission experiments are performed under no tilt of the magnetic field. A contour plot of the normalized transmission spectra for different fields in the range 3.6–4.2 T is plotted in Figure 4(a). Clearly, the only visible resonance observed is the cyclotron resonance under perpendicular magnetic fields that scales linearly with the field. On introduction of the tilt (approximately 30°), a clear anti-crossing is observed at around 3.9 T (see Figure 4(b)). The apparent observation of the satellite peaks (shown by black arrows in Figure 4(c)) in the presence of the

![Figure 4](image-url)

**Figure 4.** (a) A contour plot of the normalized transmission spectra under perpendicular magnetic fields without any tilt. (b) A contour plot showing the anti-crossing between the cyclotron resonance and the intersubband resonance in the normalized transmission spectra under tilted magnetic fields. This feature appears as the satellite peaks across the cyclotron resonance. (c) Normalized transmission plots for three magnetic fields where satellite peaks (highlighted by black arrows) are seen at 3.9 T.
magnetic field at the anti-crossing point is a manifestation of the subband-Landau coupling and hence the resonance splitting (the splitting between the two satellite peaks across the dominant cyclotron resonance). The spacings of the subbands are twice the value of the splitting at the anti-crossing frequency (according to the half-field coupling regime). This corresponds to $E_{01} = 2hcv_{\perp} = 11.1$ meV where $v_{\perp}$ is the wavenumber corresponding to the cyclotron frequency at the anti-crossing point. All magnetic chopping measurements are performed under zero bias. This phenomenon of avoided crossing is an indirect experimental method to evaluate the intersubband spacing via the so-called magnetic field chopping scheme [44, 45]. A more direct method is by application of the electric fields across the structure and this is discussed in the next section.

5. Tuning and probing of intersubband transitions electrically

Intersubband transitions are the most fundamental optical transitions that can be excited in low-dimensional semiconductor nanostructures. The observation of ISRs is a result of the fact that the component of the incident infrared electric field perpendicular to the semiconductor layers or parallel to the growth direction selectively couples, thus exciting the electrons from the lower occupied subband to the higher empty subband. By applying a voltage across the structure, it is possible to deplete and selectively populate the subbands. Thus, a more direct scheme of transmission measurement is proposed to study the intersubband spacing in such semiconductor nanostructures (viz. HEMT) even in the absence of an external magnetic field. The change in the transmission, $T(V_g,n)/T(V_r,0)$, due to the 2D space-charge layer for a normally incident light polarized in the perpendicular direction (say $z$) is given by [38]:

$$\frac{T(V_g,n)}{T(V_r,0)} = 1 - \frac{2\text{Re}(\sigma_{zz})}{\sqrt{\epsilon_0/\mu_0} \left(1 + \sqrt{\epsilon_{SE}/\epsilon_0}\right) + \sigma_g},$$

(20)

where $\epsilon_{SE}$ is the dynamic dielectric function of the substrate, $\sigma_g$ is the conductivity of the gate and $\sigma_{zz}$ is the conductivity tensor element of the 2D layer. $\epsilon_0$ and $\mu_0$ are the absolute permittivity and permeability of the free space, respectively. The Drude model very well describes the dynamic conductivity response of the quasi-free charge carriers in the 2D space-charge layer [41, 42]:

$$\sigma_{\text{Drude}}(\omega) = \frac{e^2 n_{2D} \tau}{m^*} \frac{1}{1 + i\omega \tau},$$

(21)

where $\tau$ is the scattering time, $n_{2D}$ is the carrier density and $m^*$ is the effective mass. However, for the ISTs, observed under normally incident light, with the sample tilted at an angle, the conductivity can be described by replacing $\sigma_{zz}$ by $\sigma_{\text{ISR}}$ [41, 42]:

$$\sigma_{\text{ISR}}(\omega) \propto \frac{|\epsilon_{\text{ISR}}|^2}{\omega^2 - \omega_{\text{ISR}}^2 + i\omega / \tau},$$

(22)
where $|\langle z_{j_0} \rangle|^2$ is the matrix element for the IST from the $i$th state to the $i_0$th state. The ISR frequency can however differ from the observed one due to resonance screening (depolarization shift) arising from the many body effects [46]. In the density-chopping scheme, a certain number of scans are taken at the reference voltage (a voltage much below the threshold voltage), when the 2DEG is completely depleted of charge carriers. Then, the gate voltage is slowly increased to a value when the subbands start populating (this can be well seen from the capacitance-voltage spectrum in Figure 5(a) where the change in capacitance is measured during the broadband absorption onset upon changing the gate voltage and modulating it with the LockIn technique). The same number of scans is taken at this higher gate voltage. The voltages are then changed alternatively, and the respective scans are co-added and averaged over long measurement times. The long measurement time ensures that any drift arising from the complicated experimental setup can be averaged out to zero and a good signal-to-noise ratio is obtained. Figure 5(b) shows the density-chopped transmission spectra of a HEMT.
sample. With the increase of the gate voltage, the conduction band is pulled below the quasi-
Fermi level. This results in the steepening of the triangular potential well, resulting in an
increase of the intersubband spacing. Thus, by applying a bias on the gate, the intersubband
resonance can be significantly tuned over a wide frequency (1–3 THz) or energy (4–12 meV)
range, as can be seen in Figure 5(c).

6. An access to the interior of HEMT via artificial structures

Artificial structures such as metamaterials are engineered in the sub-wavelength sizes for
certain desired properties. They are designed in assemblies of multiple individual elements
called unit cells. These structures possess unique properties such as negative [47] or zero [48
refractive indices, magnetism at optical frequencies [49], etc. The special properties are not
inherent to the materials but the design of the structures and the way electromagnetic field
interacts with them. They can also be treated as planar cavities with certain resonance frequen-
cies. When electromagnetic radiation with a certain polarization is incident on these structures,
the electric or the magnetic field couples to the cavity and exhibits a resonance that is known as
the cavity resonance or resonance frequency of the metamaterials. In transmission measure-
ments, this appears as a dip at that particular resonance frequency. An array of interconnected
double split-ring resonators (see Figure 6(a–d)) is adopted for the metamaterial design, whose
dimensions and the characteristic frequency response are first simulated by the standard finite
difference time domain solver (like CST microwave studio). For simplicity and small computa-
tion time, only one unit cell, as shown in Figure 6(e), is used for the simulation with a
periodic boundary condition in the planar directions. Moreover, these meta-atoms are placed
far apart from each other to avoid any influence of inter-meta-atom interactions. For the right
coupling of the electromagnetic radiation, the electric fields are confined in the two narrow
capacitor arms of the double split-ring resonator (see Figure 6(e)). Moreover, the fringing field
effect ensures that there is a strong electric field component along the growth direction that
extends over a few 100 nm [29]. This component of the electric field couples with the HEMT to
excite the ISRs in accordance with the polarization selection rule as discussed before.

Two transmission minima (or dips) are observed—one at 1.2 THz and the other at 2.4 THz (see
Figure 6(f)). The experimental characterization of the metamaterial array is performed by a
standard THz time-domain spectroscopy at room temperature, where the thermal energy, $k_B T$
(= 25 meV), is higher than the subband spacings. This thermal occupation of higher subbands
consequently prevents us from observing the ISR (ground-to-first excited state) in the 2DEG
layer. Hence, the response from the sample is purely due to metamaterials. A Ti:Sa laser with
an 80 fs pulse duration (a centre wavelength of 800 nm) and a repetition rate of 80 MHz is used
to generate the THz radiation by exciting an inter-digitated photoconductive antenna [9]
processed on a GaAs substrate. A fixed DC bias is applied on the antenna. The THz generation
is obtained under the transmission geometry of the antenna. Four 90° off-axis parabolic
mirrors are used for the collection and collimation of the THz beam. The detection is based on
free space electro-optic sampling [8, 29] of the THz electric field by using a birefringent, 2-mm-
thick ZnTe crystal. As compared to the simulation, the transmission is normalized with respect
to the orientation of the metamaterial. The electric field component of the THz source is in the plane parallel to the optical table. Hence, when the metamaterials are oriented at 0° (solid-black arrow in Figure 6(f)), the incident infrared radiation couples into the structures. When the metamaterials are oriented at 90° (black-dashed arrow in Figure 6(f)), the field does not couple. By normalizing the transmitted spectrum of the metamaterial at 0° with respect to the one at 90°, two transmission dips are obtained—one at 1.2 THz and the other at 2.4 THz as shown in Figure 6(f).

Once the sample is cooled down to liquid helium temperatures, at first, the characterization of the voltage range is performed over which the density-chopping measurements are to be taken. The change in capacitance with the gate voltage is measured by capacitance-voltage spectroscopy, mentioned before, as shown in Figure 7(a). A typical charging spectrum of
2DEG has a capacitance close to zero in the beginning and then as the gate voltage is increased, the conduction band is pulled below the quasi-Fermi level and subsequently the 2DEG subbands are filled with electrons. The filling is observed as a steep increase in the capacitance. The region of interest is the steep slope, where increasing the gate voltage increases the 2DEG ISRs. This is due to the fact that with more positive gate voltage, the slope of the triangular potential confinement steepens and hence increases the subband spacings, thus shifting the ISRs to higher energies. This phenomenon is well known as the quantum-confined Stark effect. It is also necessary to completely ionize the donor-exchange centres (DX centres). As more DX centres are ionized, less forward bias is required to charge the 2DEG subbands with electrons. With longer illumination, all the DX centres are successively ionized, leading to the shift of the charging slope in the capacitance-voltage spectra towards more negative biases. The spectrum shown in Figure 7(a) is obtained after 3 h of continuous illumination. The shaded region in the charging spectra, shown in Figure 7(a), indicates the region where the density-chopped infrared transmission measurements are performed. The density-chopping scheme is similar to that explained before, where the change in transmission, \( T(V_{g,n})/T(V_{r,0}) \), is recorded. According to the charging spectrum, the reference voltage \( V_r \) is chosen much below the threshold voltage (\(-0.2 \text{ V}\)), that is at \(-2 \text{ V}\), and the gate voltage \( V_g \) is varied from \(-0.25\) to \(0 \text{ V}\) in steps of \(0.02\) mA. The spectra are recorded alternatively between the gate voltage and the reference voltage and successively co-added and averaged over long measurement times. A contour plot of all the normalized transmission spectra is shown in Figure 7(b). It is observed that at low temperatures, the cavity resonance slightly shifted to a higher frequency as a result of the lower losses in the cavity in comparison to the room temperature measurements. At \( V_g = -0.14 \text{ V}\), a clear splitting of the ISR can be observed when the ISR (ground-to-first excited state, \( E_{01} \)) of the 2DEG crosses the resonance of the metamaterials at 2.4 THz. The width of the splitting is found to be 0.47 THz. By applying bias on the gated and networked-metamaterial layer, a significant modulation of the intensity of cavity resonance is observed [50, 51]. Thus, in our chopping scheme, the metamaterial resonance at 2.4 THz disappears for biases between \(-0.16 \text{ V}\) and \(-0.04 \text{ V}\), where new states appear, known as the upper and the lower polaritons. Polaritons are defined as the quasi particles that result from intense light-matter interactions. The separation between them determines how strong the interaction is and is given by the light-matter coupling constant, \( \Omega \), which is given by [42]:

\[
\Omega = \frac{f_{01}}{2m} \sqrt{\frac{n_{2D}}{2\epsilon_0 \hbar \omega_{01}} V_{\text{mode}}}
\]

(23)

where \( f_{01} \) is the oscillator strength of the ground-to-first excited state ISR with an energy of \( \hbar \omega_{01} \). \( V_{\text{mode}} \) is the mode volume of the microcavity, given by (to a very good approximation) [39]:

\[V_{\text{mode}} = \frac{2 \pi r^2}{n_{2D}}\]

They are isolated donor atoms, which can be occupied by the electrons in connection with a large lattice relaxation also known as a deep-donor complex (DX center) in Al\(_x\)Ga\(_{1-x}\)As.
where $E$ is the electric field. The coupling strength depends on three important parameters. First, it is proportional to $1/\sqrt{V_{\text{mode}}}$, implying that the cavity mode volume should be small for higher coupling strength. Second, the higher the transition energy, the smaller the coupling strength ($1/\sqrt{\hbar\omega_{ij}}$). And finally, the coupling strength scales as $\sqrt{n_{\text{fw}}}$, which is a characteristic feature of the fermionic systems. The higher the carrier density, the greater is the coupling. The voltage tuning of our device is based on the quantum-confined Stark effect. The dependence of the coupling strength, $\Omega/2$, on the number of quantum wells as shown Gabbay et al. [52] can be written as:

$$\Omega/2 = \frac{N}{2} \sqrt{1 - e^{-2a_{\text{avg}}\Delta Z_{\text{QW}}}}} \frac{1 - e^{-2a_{\text{avg}}\Delta Z}}{1 - e^{-2a_{\text{avg}}\Delta Z}}$$

where $a_{\text{avg}}$ is the average absorption coefficient, $N$ is a constant, which is proportional to the average light-matter interaction and $\Delta Z$ is the distance between the QWs. For a single QW, as in the present investigation, the coupling strength is proportional to the value of $N/2$. In the theoretical studies, Gabbay et al. found the value of $N$ to be 1, which implies that for a single QW, the coupling strength is 0.5. The splitting in our experiments is found to be 0.47 THz, which agrees well with the theoretical value. It is well known that if the splitting is significantly above the sum of the full width at half maximum of both the ISR and the metamaterial resonance, then the coupling can be assigned to be in the strong coupling regime. Thus, an
ultra-strong light-matter interaction regime is achieved by employing a single triangular quantum well in a HEMT heterostructure with a normalized coupling ratio of 0.19.

7. Conclusion

In conclusion, we have reviewed the quantum mechanical phenomenon that governs various electrical and optical properties in the low-dimensional semiconductor nanostructures such as a HEMT. We have demonstrated how one could electrically, or in combination with magnetic fields, probe and tune the intersubband transitions in the heterojunction of a HEMT structure. Such structures primarily have a triangular confinement potential. In the presence of a magnetic field, each subband is further split into a series of Landau levels or cyclotron orbits. Upon optical excitation with an infrared source, the intersubband resonances couple to the cyclotron resonance under tilted magnetic fields. This leads to the appearance of satellite peaks at the anti-crossing point. From the values of splitting at the anti-crossing points, the spacing between the corresponding subbands can be evaluated. Experiments performed in the absence of magnetic fields demonstrated that it is also possible to directly measure and tune these spacings via density-chopped infrared transmission spectroscopy. The subband spacings are measured directly and found to be in the far-infrared region (wide electrically tunable from 6 to 12 meV) of the electromagnetic spectrum. New epitaxial, complementary-doped, semitransparent electrostatic gates that have better optical transmission are introduced [29]. The integrated device with a 2DEG in a high electron mobility transistor structure and artificial metamaterials forms a strongly coupled system that can be electrically driven from an uncoupled to a coupled and again back to the uncoupled regime. In the strongly coupled regime, a periodic exchange of energy between the two systems is observed as a splitting of 0.47 THz at the point of avoided crossing. This is a very high-energy separation, considering the fact that only one quantum well is employed and thus the achievement of a strong coupling regime can be safely claimed. The tuning mechanism is attributed to the quantum-confined Stark effect. This device architecture is particularly interesting in designing devices like modulators and detectors specifically in the THz regime. The integrated device has the high-speed dynamic characteristics of the HEMT design and the appropriate frequency-controlling ability of the metamaterials. From the design perspective of the metamaterials, they can be made particularly for the THz regime with appropriate dimensions (like the one used in this chapter). Upon excitation with a broadband source, this layer selects the desired frequency for which it is designed, and under the application of an external electrical field across the structure, the transmission of this frequency can be controlled and also modulated. This control dynamics can be very fast, simply owing to the fast dynamics of the HEMT design [52, 53]. Furthermore, this design can also be used to detect THz frequencies. Various other 2D materials (like graphene [54–56] or black phosphorous [57, 58]) are also used these days in the transistor configuration for developing THz detectors, simply utilizing the fast dynamics of the transistor design. These novel devices have thus helped to reduce the long-debated THz gap in the electromagnetic spectrum, where there is a severe lack of fast electronic devices.
Author details

Shovon Pal¹⁻³*, Sascha R. Valentín¹, Arne Ludwig¹ and Andreas D. Wieck¹

*Address all correspondence to: shovon.pal@ruhr-uni-bochum.de

1 Applied Solid State Physics, Ruhr University Bochum, Bochum, Germany
2 Terahertz Spectroscopy and Terahertz Technology, Ruhr University Bochum, Bochum, Germany
3 Department of Materials, ETH Zurich, Vladimir-Prelog-Weg, Zurich, Switzerland

References


Kumar S. Recent progress in terahertz quantum cascade lasers. IEEE Journal of Selected Topics in Quantum Electronics. 2011;17:38

Corrigan P, Martini R, Whittaker EA, Bethea C. Quantum cascade lasers and the Kruse model in free space optical communication. Optics Express. 2009;17:4355


Pickwell E, Wallace VP. Biomedical applications of terahertz technology. Journal of Physics D. 2006;39:R301


Ren F, Pearnton SJ. Semiconductor Device-Based Sensors for Gas, Chemical, and Bio Applications. CRC Press; Cleveland, OH; Boca Raton, FL; 2011.


[31] Sattler KD. Handbook of Nanophysics: Principles and Methods. CRC Press; Cleveland, OH; Boca Raton, FL; 2010


[38] Batke E, Heitmann D. Rapid-scan Fourier transform spectroscopy of 2-D space charge layers in semiconductors. Infrared Physics. 1984;24:189


[42] Pal S. Terahertz transmission spectroscopy of charge carriers under surface electric fields. Ruhr-Universität Bochum; 2015.


