INVITED COMMENT

New frontiers in quantum cascade lasers: high performance room temperature terahertz sources

To cite this article: Mikhail A Belkin and Federico Capasso 2015 Phys. Scr. 90 118002

View the article online for updates and enhancements.

Related content

- External cavity terahertz quantum cascade laser sources based on intra-cavity frequency mixing with 1.2–5.9 THz tuning range
  Yifan Jiang, Karun Vijayaraghavan, Seungyong Jung et al.

- Novel InP- and GaSb-based light sources for the near to far infrared
  Sprengel Stephan, Demmerle Frederic and Amann Markus-Christian

- Recent progress in quantum cascade lasers and applications
  Claire Gmachl, Federico Capasso, Deborah L Sivco et al.

Recent citations

- Optical external efficiency of terahertz quantum cascade laser based on erenkov difference frequency generation
  A. Hamadou et al

- Distributed feedback 2.5-terahertz quantum cascade laser with high-power and single-mode emission
  Jiawen Luo et al

- H. L. Hartnagel and V. P. Sirkeli
Invited Comment

New frontiers in quantum cascade lasers: high performance room temperature terahertz sources

Mikhail A Belkin1 and Federico Capasso2

1 Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712, USA
2 Harvard University, School of Engineering and Applied Sciences, 29 Oxford St., Cambridge, MA 02138, USA

E-mail: mbelkin@ece.utexas.edu and capasso@seas.harvard.edu

Received 19 June 2015, revised 25 August 2015
Accepted for publication 11 September 2015
Published 2 October 2015

Abstract

In the last decade quantum cascade lasers (QCLs) have become the most widely used source of mid-infrared radiation, finding large scale applications because of their wide tunability and overall high performance. However far-infrared (terahertz) QCLs have lagged behind in terms of performance and impact due to the inability so far of achieving room temperature operation. Here we review recent research that has led to a new class of QCL light sources that has overcome these limitations leading to room temperature operation in the terahertz spectral range, with nearly 2 mW of optical power and significant tunability, opening up also this region of the spectrum to a wide range of applications.

Keywords: quantum cascade lasers, terahertz, difference frequency generation

(Some figures may appear in colour only in the online journal)

1. Introduction

Last year marked the 20th anniversary of the invention and demonstration of the quantum cascade laser (QCL) at Bell Laboratories[1]. This achievement was the culmination of 20 years of research on bandstructure engineering of semiconductor heterostructures[2] and in particular nanostructures in which quantum effects such as size quantization and resonant tunneling become important[3]. While the design of man-made electronic and optical properties is central to bandstructure engineering, its practical implementation in novel devices was made possible by the epitaxial growth technique of molecular beam epitaxy (MBE) pioneered by Alfred Y Cho. MBE first enabled the growth of ultrathin semiconductor layers with atomic precision down to a few atomic layers as well as unprecedented control of the composition of the material. Later metalorganic vapor phase epitaxy (MOVPE) demonstrated similar capabilities and became the standard growth platform for photonic devices at communication wavelengths. MOVPE and MBE are now central to the production of QCLs.

The QCL avoids the operating principle of conventional semiconductor lasers by relying on a radically different process for light emission, which is independent of the band gap. Instead of using opposite-charge carriers in semiconductors (electrons and holes) at the bottom of their respective conduction and valence bands, which recombine to produce light of frequency \( \nu \approx E_g/h \), where \( E_g \) is the energy band gap and \( h \) is Planck’s constant, QCLs use only one type of charge carriers (electrons), which undergo quantum jumps between energy levels \( E_n \) and \( E_{n-1} \) to create a laser photon of frequency \( (E_n - E_{n-1})/h \). These energy levels do not exist naturally in the constituent materials of the active region, but are artificially created by structuring the active region into quantum wells of nanometric thickness. The motion of electrons perpendicular to the layer interfaces is quantized and...
characterized by energy levels whose difference is determined by the thickness of the wells and by the height of the energy barriers separating them. The implication of this new approach, based on decoupling light emission from the band gap by utilizing instead optical transitions between quantized electronic states in the same energy band (known as intersubband transitions), are many and far reaching, amounting to a laser with entirely different operating characteristics from laser diodes and far superior performance and functionality.

Initially greeted as a scientific curiosity, due to its complex structure and low-temperature operation, the QCL fast established itself as the leading mid-infrared (mid-IR) source of coherent radiation due to the ability to tailor the emission wavelength across the entire mid-IR spectrum, from 3 to beyond 20 μm, its unprecedented tunability and high performance operation at room temperature in the two atmospheric windows with pulsed and continuous-wave (CW) powers up to 100 and 10 W respectively. One should recall that before the advent of QCLs the mid-IR spectrum had a very limited choice of light sources which explains why applications and commercial development in this sector was very limited. Carbon dioxide lasers, while important tools for high power industrial applications such as cutting and welding, have very limited wavelength range and tunability. Diode lasers in the mid-IR have had very limited impact due fundamental difficulties in operating at room temperature, very small tunability and lack of robustness in terms of fabrication and reliability on account of their small bandgap.

QCL based spectroscopy and its applications to chemical sensing, in all its implementations (atmospheric chemistry, trace gas analysis for applications such as pollution monitoring, industrial process control, combustion diagnostics, health care (medical diagnostics such as breath analysis, surgery) have seen an unprecedented growth scientifically, technologically and commercially with about thirty companies active, large and small [4]. Another sector where QCLs are having a major impact is high power (1–10 W CW) applications in areas such as infrared countermeasures to protect aircraft from heat seeking missiles; new emerging applications are material processing for example for 3D printing where one can use wavelength selective chemistry to mold materials.

Recent and earlier reviews of QCLs and their application can be found in [5-9].

In 2002 Alessandro Tredicucci and his team at the University of Pisa demonstrated the first QCL operating in the far-IR also known as the THz spectral region [10]. Until recently, no THz QCL based source operated at room temperature due to physical limitations of standard injection QCLs in that spectral region, as will be clear later in this paper. This situation has now changed through the use of intracavity difference frequency generation in double wavelength mid-IR QCLs which have recently have been reported to emit at room temperature with peak power of nearly 2 mW. Such a power level is adequate for most spectroscopy and sensing applications. A review of such important recent developments is the subject of this paper.

2. Energy diagram and quantum design

This section briefly reviews the design principles of a QCL. The energy diagram of the device at zero bias looks overall like a saw-tooth due to the compositionally graded regions, which consist of a digitally graded alloy of alternating ultrathin semiconductor layers. Under high enough bias such that the applied electric field suppresses the saw-tooth energy barriers, the band diagram is transformed into an energy staircase and electrons are injected efficiently into the excited state of suitably designed quantum wells, emitting a laser photon at each stage. The energy diagram of the QCL drew inspiration from the design of the staircase avalanche photodiode (or solid state photomultiplier), which consists of multiple identical graded gap regions [11]. Under high enough applied bias the sawtooth energy diagram turns into a staircase and photogenerated electrons roll down, gaining enough energy at the steps to create an electron–hole pair by impact ionization at the potential step given by band discontinuity. The number of stages typically ranges from 20 to 35 for lasers designed to emit in the 4–8 μm range, but working lasers can have as few as one or as many as over 100 stages. This cascade effect is responsible for the very high power that QCLs can attain. In QCLs, unlike in a laser diode, an electron remains in the conduction band after emitting a laser photon. The electron can therefore easily be recycled by being injected into an adjacent identical active region, where it emits another photon, and so forth.

The operating principle of a QCL can be understood from figure 1, which illustrates the band diagram of a QCL designed to emit photons of wavelength $\lambda = 7.5 \mu m$. The conduction band slope (units of energy over distance) divided by the electronic charge is the applied electric field. Each QCL stage consists of an active region and of an electron injector. The former contains three energy levels; the laser photon is emitted in the transition between states 3 and 2, which is controlled primarily by the two wider wells thickness. To achieve a large population inversion between states 3 and 2 the lifetime of level 3 is designed to be much longer than that of state 2. To this end state 1 is positioned approximately an optical phonon energy (~34 meV in InGaAs/AlInAs materials) below level 2, which guarantees that electrons in this level resonantly scatter to energy level 1 by emitting an optical phonon, an extremely fast process characterized by a relaxation time of the order of 0.1–0.2 ps. Electrons in state 3 relax much more slowly to level 2 owing to the much larger energy difference, a non-resonant phonon emission process. To ensure lasing however, one must also quench escape by tunneling from the upper laser state to the quasi-continuum, which would reduce its population. This is accomplished by designing the injector downstream as a superlattice with an energy gap (minigap) in correspondence to the upper energy level $E_2$. Notice instead the miniband designed to face $E_2$ and $E_1$ which facilitates electron escape from the active region by preventing their accumulation in those states. Electrons are injected into the upper laser level by resonant tunneling which guarantees efficient injection when the applied bias is high enough to line up the ground
state of the injector with state 3. A transmission electron micrograph of a cross-section of the QCL layered structure is shown in figure 1, where the AlInAs barriers are dark and the GaInAs quantum wells light gray. For wavelengths greater than \( \sim 5 \mu m \), the alloy compositions of the wells and barriers are selected to have the same lattice constant as the substrate (i.e. Al\(_{0.48}\)In\(_{0.52}\)As and In\(_{0.47}\)Ga\(_{0.53}\)As), which leads to a conduction band discontinuity between the barrier and well layers \( \Delta E_c = 0.52 \) eV. For wavelengths considerably shorter than 5 \( \mu m \), the upper laser state is much higher in energy, which promotes thermal escape over the barrier. This process, is a limiting factor for device performance, reducing the maximum operating temperature particularly for CW operation or preventing altogether CW operation at room temperature and severely limiting the maximum achievable optical power. Note that the active region of a QCL can reach temperatures considerably higher than the laser heat sink by tens of degrees, similar to what occurs in laser diodes. By using higher Al content in the barriers and lower In percentage in the quantum wells the conduction band discontinuity can be increased to values in the 0.7 eV–0.8 eV range which suppresses electron thermionic emission above the barriers. CW room temperature QCLs at \( \lambda \leq 5.2 \mu m \), typically employ these strained heterojunctions [12–16]. To further improve temperature performance the separation between level 2 in figure 1 and the injector ground state should be designed to exceed 0.1 eV in order to prevent electrons in the injector to backfill level 2.

Major improvement in QCL performance (lower threshold and higher power) can be achieved by modifying the energy level diagram of figure 1 through the addition of layers which introduce a double phonon resonance [17]. The active region of this modified structure has four quantum wells and three equally separated energy levels by an optical phonon. This design leads to a larger population inversion due to the more efficient removal of electrons from the lower level of the laser transition.

3. Terahertz QCLs

Owing to meV-level control of energy positions of electronic subbands, QCL technology provides an attractive approach for mass-producible semiconductor lasers also at terahertz frequencies \( (\nu = 1–5 \text{ THz}, \ h\nu \approx 4–20 \text{ meV}) \). The original devices by the group of Alessandro Tredicucci provided laser emission at 4.4 THz with the maximum operating temperature of 50 K [10]. The spectral coverage and the maximum operating temperature of THz QCLs has improved dramatically since then, with the current state-of-the-art device performance summarized in figure 2. The emission frequencies of various THz QCL designs now span 1.2–5 THz range (without the use of strong magnetic fields) and devices with CW output powers in excess of 100 mW [18] and peak powers in excess of 1 W [19] were demonstrated. Furthermore, the emission linewidth of distributed-feedback THz QCLs can be stabilized to a few tens of kHz, making them ideal local oscillators for heterodyne detection [20, 21]. The frequency output of THz QCLs can be continuously-tuned over broad spectral range, with devices tunable over 330 GHz around 4 THz center frequency demonstrated using a moving microelectromechanical plunger to manipulate the lasing transverse mode [22, 23]. Furthermore, THz QCLs show high promise as broadband THz frequency comb sources [24, 25]. Comprehensive reviews of THz QCL design strategies and performance can be found in [26–28].

The best temperature performance of THz QCLs is achieved using double-metal waveguides [29, 30] that are conceptually similar to microstrip waveguides used in high-frequency electronics. These waveguides, depicted in figure 3(a) provide nearly 100% of laser mode confinement in
and other one-dimensional grating structures. Subwavelength modal confinement and highly divergent THz far-field emission from THz QCLs demonstrated to date [33, 34]. While strong THz laser mode confinement in the active region is desired for creating devices with low threshold current and high operating temperature, subwavelength modal confinement also leads to high facet reflectivity [35] and highly divergent THz far-field emission [36] in edge-emitting THz QCLs. A number of approaches have been developed to alleviate this problem by either creating surface-emitting devices or by impedance-matching and collimating THz edge-emission. Surface-extraction of THz radiation in THz QCLs has been achieved by patterning the top metal waveguide cladding to produce ridge-waveguide devices with second-order gratings [37, 38] and other one-dimensional grating structures [39, 40] or to make two-dimensional photonic crystal surface-emitting laser cavities [41, 42]. An excellent summary of surface-extraction schemes employed in THz QCLs is given in [43]. Directional outcoupling from the edge of the double-metal waveguide may be achieved by attaching a silicon microlens [44] or a metallic horn antenna [45] onto the front facet. A monolithic approach to collimate THz output from the edge-emitting double-metal waveguide THz QCLs consists of using surface plasmon gratings fabricated on the device front facet [46]. An example of a THz QCL device with a spoof surface plasmon collimator is shown in figure 3(b) with the experimentally measured THz far field emission profile shown in figure 3(c). Figure 3(d) compares the experimentally measured far-field profiles of the two identical double-metal waveguide THz QCLs with and without the spoof surface plasmon collimator. Recently, by patterning of the waveguide a THz QCL with a high degree of circular polarization has

Figure 2. (a) Summary of the maximum operating temperatures of THz QCLs reported to date. Solid and open symbols refer to CW and pulsed operation, respectively. Red upright triangles are resonant-phonon designs, black squares are scattering-assisted injection design, blue circles refer to bound-to-continuum designs, and green diamonds are THz QCLs based on material system other than GaAs/AlGaAs. See [26, 27, 31] for a description of different THz QCL designs. (b) Timeline for the maximum operating temperatures in pulsed mode achieved by THz QCLs from the initial demonstration in 2002 until mid-2015. The data are taken from [28, 31–33]. Green dashed line at 220 K indicates operating temperatures that may be accessible with compact thermoelectric coolers.

Figure 3. (a) Schematic of a double-metal waveguide for THz QCL. Laser mode in the 10 μm thick active region is shown. Bottom metal cladding is introduced layer via wafer bonding procedure [29, 30]. The ridge width can vary from a few tens of microns to hundreds of microns. (b) Scanning electron microscope image of the fact of a THz QCL with a spoof surface plasmon collimator. The device has a 1.2 mm long, 150 μm wide and 10 μm thick waveguide and lases at approximately 3 THz. (c) Measured 2D far-field THz intensity profiles of the device in (b). (d) The black triangles and black dotted curve are, respectively, measured and simulated laser intensity profiles along φ = 0° for the double-metal waveguide device without the collimator. The blue circles and black solid curve are, respectively, measured and simulated laser intensity profiles along φ = 0° for the device with the collimator. After [46].
been demonstrated for the first time. Such a source could find interesting applications in the spectroscopy of chiral molecules.

One of the key desired characteristics for any semiconductor laser is room-temperature operation. This has not yet been achieved with THz QCLs. The maximum operating temperature demonstrated to date, without using strong magnetic field, is 199.5 K [34] for pulsed operation and 129 K for CW operation [47]. With an applied magnetic field of 20–30 T, THz QCLs can operate to higher temperature, with the maximum operating temperature 225 K demonstrated for 3 THz devices [48]. A summary of the temperature performance of a number of THz QCLs (taken from the literature and operating without a magnetic field) as a function of frequency is shown in figure 2(a) and the timeline of the maximum operating temperature achieved by THz QCLs is shown in figure 2(b). The improvement in temperature performance has been rapid after the initial demonstration of THz QCLs, but it has slowed down considerably in recent years, as shown in figure 2(b). The highest operating temperature is obtained with three-quantum-well resonant depopulation THz QCL designs, originally proposed in [49], and its derivatives [32, 33, 50]. A study [51] of the dependence of the threshold current densities on the heat-sink temperature in THz QCLs based on the three-quantum-well resonant depopulation active region design with nominal emission frequencies of 2.3, 2.7, 3.1, 3.5 and 3.8 THz indicated that the performance of these devices is limited by the interplay of two principal factors: (a) optical phonon scattering of thermal electrons, which dominates at shorter wavelengths, and (b) parasitic current, which dominates at longer wavelengths. These factors result in an optimum frequency (‘sweet spot’) around 3–4 THz for lasers based on this design, visible in figure 2(a). Recently-reported THz QCL designs based on scattering-assisted electron injection [31] and double-phonon depopulation [52] show promise to increase the maximum operating temperatures further, but have not achieved record operating temperatures yet.

All high-performance results are currently achieved with THz QCLs based on GaAs/Al0.15Ga0.85As semiconductor heterostructures grown on GaAs substrates. Alternative materials systems for THz QCLs are being investigated and devices grown on InP substrates based on InGaAs/AlInAs [53], InGaAs/GaAsSb [54], and InGaAs/AlInGaAs [55] heterostructures were experimentally demonstrated with the maximum operating temperatures in the 140–150 K range. Additionally, devices based on GaAs/AlGaAs heterostructures with varying barrier heights have been reported with the maximum operating temperatures in the 180–190 K range [56, 57], close to record temperatures. The first and yet unconfirmed results of THz QCLs based on GaN/AlGaN heterostructure operating at a heat sink temperature of 4 K were reported recently [58]. GaN/AlGaN THz QCLs may potentially achieve operating temperatures above that of GaAs/AlGaAs THz QCLs as large optical phonon energy in GaN/AlGaN (90 meV; in comparison, the optical phonon energy in GaAs is 34 meV) is expected to suppress optical phonon scattering of thermal electrons in the upper laser states, which is believed to be the dominant temperature-degradation mechanism in GaAs/AlGaAs THz QCLs [26, 50, 51].

An alternative approach to THz generation in QCLs is based on intra-cavity THz difference-frequency generation (DFG) in dual-wavelength mid-IR QCLs [60]. These devices, referred to as THz DFG–QCL, have achieved a dramatic progress in recent years. Since THz DFG does not require population across the THz transition and mid-IR QCLs can reliably operate at room temperature, THz DFG–QCLs can operate at room-temperature as well. Figure 4(a) displays the device concept: upon application of bias current device generates two mid-IR pumps at frequencies $\omega_1$ and $\omega_2$ which generate THz output at frequency $\omega_{\text{THz}} = \omega_1 - \omega_2$ via DFG inside of the laser cavity. From the perspective of user operation and fabrication, THz DFG–QCLs are similar to traditional QCLs as both are monolithic electrically-pumped semiconductor sources and have similar requirements for heterostructure growth and fabrication.

DFG is a nonlinear optical process in which two beams at frequencies $\omega_1$ and $\omega_2$ interact in a medium with second-order nonlinear susceptibility, $\chi^{(2)}$, to produce radiation at
frequency $\omega_{THz} = \omega_1 - \omega_2$. The intensity of the wave at frequency $\omega_{THz}$ generated in the QCL waveguide is given by the expression [60, 61]

$$W(\omega_{THz} = \omega_1 - \omega_2) = \frac{\omega_{THz}}{8\pi^2 c^3} n(\omega_1)n(\omega_2)n(\omega_{THz}) \times |\chi^{(2)}|^2 \times \frac{W(\omega_1)W(\omega_2)}{S_{eff}} \times I_{coh}^2,$$

(1)

where $I_{coh} = 1/\left(\left|k_{THz} - (k_1 - k_2)\right|^2 + (\alpha_{THz}/2)^2\right)$ is the coherence length, $W(\omega_1), n(\omega_1)$, and $k_i$ are the power, refractive index, and the wave vector of the beam at frequency $\omega_i$, respectively, $\alpha_{THz}$ stands for the losses at THz DFG frequency, $S_{eff}$ is the effective area of interaction and we assumed that the laser medium is transparent for both mid-IR pumps and neglected the depletion of the pump powers in the DFG process. It follows from equation (1) that, for efficient DFG, one needs to use materials with large $\chi^{(2)}$, input beams of high intensity, and achieve low losses and phase matching, $|k_{THz} - (k_1 - k_2)| \approx 0$. Most of the schemes to generate THz radiation via DFG use focused beams from high-intensity pulsed solid-state lasers (usually $\sim$1 GW cm$^{-2}$, often limited by the damage threshold of the nonlinear crystal) and achieve long coherence length of up to tens of millimeters by either true phase matching or quasi-phase matching in transparent nonlinear crystals [62, 63]. For intracavity DFG in QCLs, the intensities of the pump beams are limited to 1–10 MW cm$^{-2}$ and the coherence length is limited to approximately 100 $\mu$m or smaller values by free carrier absorption at THz frequency as shown in figure 4(b). Under these conditions, efficient THz DFG appears to be impossible, unless the nonlinear susceptibility $|\chi^{(2)}|$ in equation (1) is enhanced well below it typical values of 1–300 pm V$^{-1}$ found in natural nonlinear materials [64].

Quantum well structures can be engineered to provide giant nonlinear response by tailoring intersubband transitions so as to achieve resonant nonlinearity for a specific process [65]. Figure 4(c) provides a schematic of intersubband states tailored to provide giant $\chi^{(2)}$ for THz DFG between $\omega_1$ and $\omega_2$ mid-IR pumps. We note that the electron population is located in the upper state. This point was crucial for operation of THz DFG–QCL technology [60] as such structures provide optical gain, rather than loss for mid-IR pumps. Referring to figure 4(c), the expression for resonant optical nonlinearity for THz DFG becomes [61]

$$\chi^{(2)}(\omega_{THz} = \omega_1 - \omega_2) \approx \Delta N_e \frac{e^3}{\hbar^2 \varepsilon_0} \left| \frac{3_{12} 2_{23} 3_{31}}{\omega_{THz} - \omega_{23} + i\Gamma_{23}} \right| \times \left( \frac{1}{\omega_1 - \omega_{13} + i\Gamma_{13}} + \frac{1}{\omega_2 + \omega_{12} + i\Gamma_{12}} \right),$$

(2)

where $\Delta N_e$ is the population inversion density between state 1 and states 2 and 3 (assuming the population in states 2 and 3 is the same), $\varepsilon_{ij}^\phi, \omega_j$ and $\Gamma_j$ are the dipole matrix element, frequency, and broadening of the transition between states $i$ and $j$.

It is possible to integrate DFG optical nonlinearity with population inversion shown in figure 4(c) into existing high-performance mid-IR QCL designs. In particular, it was shown that a bound-to-continuum [60, 61] and dual upper state [66, 67] active region designs intrinsically possesses strong optical nonlinearity for THz DFG and the QCL bandstructure can be further optimized to enhance $\chi^{(2)}$ for a particular THz DFG process. Figure 5 shows the bandstructures for the bound-to-continuum and dual upper state active regions, along with equivalent energy levels diagram showing triplets of states that provide strong nonlinear optical response.

The values of $\chi^{(2)}$ provided by the QCL active regions for a DFG process may be obtained by applying equation (2) to all the relevant triplets of states. Assuming the mid-IR pump frequencies are sufficiently close to corresponding intersubband transitions, $\omega_1 \approx \omega_{13}$ and $\omega_2 \approx \omega_{12}$, the value of $\chi^{(2)}$ in equation (2) is principally being determined by the values of $\Delta N_e$, transition dipole moments $\varepsilon_{ij}^\phi$, and linewidth broadening factors $\Gamma_j$. The value of $\Delta N_e$ in equation (2) is obtained by computing required population inversion in a QCL to provide modal gain = modal loss in the laser cavity, transition dipole moments $\varepsilon_{ij}^\phi$ and transition energies $\omega_{ij}$ are computed from the QCL bandstructure, and the values of linewidth broadening factors $\Gamma_j$ are typically assumed to be 10–15 meV for the mid-IR transition and 3–5 meV for the THz transition, which are typical values for mid-IR and THz intersubband transitions in InGaAs/AlInAs heterostructures. Assuming further a typical threshold gain of $\sim 15$ cm$^{-1}$ in the active region, one obtains $|\chi^{(2)}| \sim 15000$ pm V$^{-1}$ for THz DFG–QCLs based on bound-to-continuum active region design [68] and $|\chi^{(2)}| \sim 20000$–$25000$ pm V$^{-1}$ for the typical dual upper state active region design [67]. Somewhat larger estimated values of $\chi^{(2)}$ for the dual upper state active region design are due to the constructive addition of $\chi^{(2)}$ from different triplets of states shown in figure 5(b) [67]. The estimated values of $\chi^{(2)}$ are 2–3 orders of magnitude larger than that of typical nonlinear crystals and result in 4–6 orders of magnitude increase in the THz DFG power output in equation (1). We note, however, that no direct measurements of $\chi^{(2)}$ in THz DFG–QCL active regions were performed to date. The values of $\chi^{(2)}$ are very sensitive on the linewidth broadening assumptions and are dependent on the population inversion density $\Delta N_e$ which is dependent on the loss for mid-IR modes in a QCL. Thus, all the numbers given above are theoretical estimates which, however, result in THz DFG power predictions that are in reasonable agreement with experiment [61, 68].

Because of high THz loss in a QCL waveguide, see figure 4(b), only THz radiation generated within $\sim 100$ $\mu$m of the front facet is outcoupled if the THz mode is confined in the same waveguide, which was the case for early THz DFG–QCLs [60, 61, 69]. To enable efficient extraction of THz radiation along the whole length of the QCL waveguide, devices based on Cherenkov phase-matching scheme were proposed and implemented in [70]. Cherenkov phase-
matching scheme is currently used by all state-of-the-art THz DFG–QCL devices [68, 71–74].

Cherenkov emission occurs when the phase-velocity of the nonlinear polarization wave in a thin slab of nonlinear optical material is faster than the phase-velocity of the generated radiation in the medium surrounding the slab [75]. The Cherenkov emission angle $\theta_c$ is given as

$$\theta_c = \cos^{-1} \left( \frac{k_{nl}}{k_{THz}} \right),$$

where $k_{nl}$ is the propagation constant of the nonlinear polarization wave and $k_{THz}$ is the propagation constant of the terahertz wave in the substrate. We can write an expression for the nonlinear polarization wave at $\omega_{THz} = \omega_1 - \omega_2$ in the slab waveguide approximation as:

$$P^{(2)}_z(x, z) = \varepsilon_0 \chi^{(2)}(z) E^{(1)}_z(z) E_z^{(2)}(z) e^{i(\omega_{THz} - (\beta_1 - \beta_2)x)},$$

where the $z$-direction is normal to the QCL layers and the $x$-direction is along the waveguide, $\beta_1$ and $\beta_2$ are the propagation constants for mid-IR pump modes, $E^{(1)}_z(z)$ and $E_z^{(2)}(z)$ are $z$-components of $E$-field of the mid-IR pump modes, and $\chi^{(2)}(z)$ is the giant intersubband optical nonlinearity for DFG in the QCL active region defined in equation (2). The propagation constant of the nonlinear polarization wave is thus given as $\beta_1 - \beta_2$. Since the two mid-IR pump frequencies are close, $\omega_1 \approx \omega_2$, one can write [76]

$$|\beta_1 - \beta_2| \approx \frac{n_g \omega_{THz}}{c},$$

where $n_g = n_{eff} + \omega_1 \frac{\partial n_{eff}}{\partial \omega} \bigg|_{\omega_1}$ is the group effective refractive index for mid-IR pumps. In order to produce Cherenkov DFG emission into the substrate, the substrate refractive index at $\omega_{THz}$ must be larger than $n_g$ and this condition is satisfied throughout the 1–5 THz spectral range for InP/GaInAs/AlInAs QCLs grown on semi-insulating (SI) InP. In particular, the value of $n_g$ is typically in the range 3.3–3.4 for mid-IR pumps [70, 77] and the refractive index of InP in 1–5 THz range is 3.5–3.7 [59]. The value of Cherenkov angle in typical devices [68, 70] is approximately 20°. Since SI InP substrate is non-conducting, lateral current extraction is implemented in all Cherenkov THz DFG–QCL.

In the slab-waveguide approximation, Cherenkov THz emission may be modeled analytically as a leaky slab-waveguide mode produced by the polarization source described in equation (4) following the formalism described in [78] (generalized for the multi-layer waveguide structure of DFG–QCLs). In a physical picture, two Cherenkov waves are generated by the nonlinear polarization wave: one propagates towards the top contact and the other one towards the bottom substrate as shown in figure 6(a). These two waves are partially reflected by various waveguide layers and may pass through the active region multiple times and interfere with

Figure 5. (a) Bandstructure (top panel) and energy level schematic (bottom panel) of a bound-to-continuum active region design for THz DFG–QCL. State 4 is the upper laser state, while states 3, 2, and 1 represent the lower laser state manifold. Relevant triplets of states that produce strong nonlinearity for THz DFG are shown in the bottom panel. (b) Bandstructure (top panel) and energy level schematic (bottom panel) of a dual upper state active region design for THz DFG–QCL. States 5 and 4 are the upper laser state, while states 3, 2, and 1 represent the lower laser state manifold. Relevant triplets of states that produce strong nonlinearity for THz DFG are shown in the bottom panel.
each other before finally exiting to the substrate. The calculated square magnitude of the H-field for the TM-polarized Cherenkov wave ($\langle |H|^2 \rangle$) for the devices reported in [68] are shown in figure 6(b).

To avoid total internal reflection of the Cherenkov wave at InP/air interface, the front facet of the substrate has to be polished as shown in figure 6(a). Experimentally, 30° substrate polishing results in the Cherenkov THz wave outcoupling in forward direction at 3.5 THz for typical THz DFG–QCLs. The angle varies for other THz frequencies due to InP refractive index dispersion [71].

We note that the ridge width in THz DFG–QCLs reported to date is typically in the range of 15–25 μm, which is smaller than THz wavelength of 60–300 μm (corresponds to 1–5 THz frequencies). As a result, THz radiation is emitted from ridge-waveguide QCLs in a cone of angles into the substrate, instead of a plane wave in a slab-waveguide model, as shown schematically in figure 6(a). Initial evidence of Cherenkov emission cone in THz DFG–QCLs was reported [73]. It is estimated [79] that less than 20% of the total THz power in the forward-going Cherenkov is outcoupled into the free space in the current devices with flat-polished substrates (as shown in figure 6(c)); in additional nearly all THz power emitted into back-ward going Cherenkov wave is lost. In total, it is estimated [79] that less than 10% of the total THz power produced in THz DFG–QCLs is outcoupled to free space in the state-of-the-art THz DFG–QCLs.

To enable narrow-line THz emission from THz DFG–QCLs, distributed feedback (DFB) gratings for mid-IR pumps are fabricated. Three approaches are currently used. In the first approach two grating sections are etched along the QCL waveguide for selecting the two mid-IR pump frequencies [80] as shown schematically in figure 7(a). In the second approach, a single DFB grating that contains two Fourier components to select the two mid-IR frequencies is etched [69]. Finally, sampled gratings may also be used to enable dual-single-mode mid-IR pump selection [81].

Typical pulsed-mode room-temperature device performance of THz DFG–QCLs is shown in figures 7(b)–(d). Depending on the mid-IR pump spacing, THz emission in these devices can be varied in the entire 1–6 THz range and beyond, limited only by the materials losses and inevitable reduction of THz DFG efficiency which falls down with $\omega_{THz}^2$ as shown in equation (1). The data presented in figure 7 is taken from [68]. The active region of these devices was made up of two stacks of QCL stages designed for emission at $\lambda_1 = 8.2$ μm and $\lambda_2 = 9.2$ μm, based on the bound-to-continuum active region design shown in figure 5(a). The waveguide structure is shown in figure 6(b). The THz wall plug efficiency (WPE) of a 4 THz device was approximately $0.9 \times 10^{-6}$ [68]. In comparison, the highest WPE achieved by THz QCLs is approximately $5 \times 10^{-3}$ at 10 K [43].

High extraction efficiency of THz radiation and relaxed phase-matching conditions in Cherenkov THz DFG–QCLs devices enable broad spectral tunability of THz output in these lasers. An example of a broadly-tunable external cavity (EC) THz DFG–QCL system [68, 71] is presented in figure 8. To enable THz DFG emission tuning in our device, we configure EC system to continuously tune one mid-IR pump frequency, while having the second mid-IR pump frequency fixed. We fix one mid-IR pump frequency by a DFB grating etched into the QCL waveguide and provide tuning of the second mid-IR pump frequency by an external grating in a Littrow-type EC configuration as shown schematically in figure 8(a). Due to spatial hole burning, gain competition between the two mid-IR pumps is reduced and simultaneous two-color emission can be achieved experimentally [68, 71].

Figure 8(b) shows a photograph of the EC system. Experimental results obtained with a 2.3 mm long ridge waveguide THz DFG–QCL that contains a 1.6 mm long DFB grating section and back facet anti-reflection coating are shown in figures 8(c) and (d). The device structure was
identical to that described in figure 6 and [68]. The DFB grating was designed to fix $\omega_1$ pump at 963 cm$^{-1}$. The substrate front facet of the device was polished at a 30° angle to avoid total internal reflection of the Cherenkov THz emission and allow for THz extraction in a forward direction. Figure 8(b) shows the emission spectra and power output of the two mid-IR pumps at different EC grating positions taken at a current density of 8.0 kA cm$^{-2}$. Also shown is the mid-IR power for $\omega_1$ (circles) and $\omega_2$ (triangles) pumps as a function of $\omega_2$ pump wavenumber. The dashed line shows the transmission spectrum of mid-IR long-pass filter used for power measurements. (c) THz emission spectra of the EC THz DFG–QCL system in (a)–(c) taken at a current density of 8.0 kA cm$^{-2}$. Also plotted are the THz peak power (blue squares and right axis) and mid-IR-to-THz conversion efficiency (red circles and left axis) as a function of THz frequency.
continuously from 1009 cm$^{-1}$ to 1182 cm$^{-1}$, resulting in the available difference-frequency-tuning from 1.4 THz to 5.9 THz. Due to gain competition, the mid-IR emission switches from dual-frequency ($\omega_1$ and $\omega_2$) to single-frequency ($\omega_2$) output as the EC grating tunes pump frequency $\omega_2$ towards DFB frequency $\omega_1$ [71].

The THz emission spectra and peak power for different EC grating positions taken at a current density of 8.0 kA cm$^{-2}$ through the device are displayed in figure 8(d). Also shown is mid-IR-to-THz nonlinear conversion efficiency at a current density of 8.0 kA cm$^{-2}$, defined as the ratio of the measured THz peak power to the product of the two mid-IR pump powers. THz peak power of 90 μW and mid-IR-to-THz conversion efficiency of nearly 250 μW W$^{-2}$ were observed at 3.8 THz. The conversion efficiency in figure 8(d) peaks in the 3.7–4.5 THz range of frequencies and falls off at both high- and low-frequency ends of the tuning curves, which is consistent with results obtained with other THz DFG–QCL devices [81, 82]. On the high-frequency end, efficiency of THz generation is limited by the onset of high optical losses in InGaAs/AlInAs/InP materials due to tails of optical-phonon absorption bands (Reststrahlen band). On the low-frequency end, THz generation efficiency is principally limited by high free carrier absorption in the QCL waveguide and the $\omega_\text{THz}$ dependence of the DFG efficiency, see equation (1). Additional factors include spectral dependence of intersubband optical nonlinearity and residual absorption in SI InP substrate through which THz radiation is extracted, see [71]. Besides EC systems, a variety of monolithic THz DFG–QCL sources have been demonstrated [82–84] based on thermal tuning of mid-IR pumps (using direct thermal tuning or Vernier tuning mechanism) and/or QCL array concepts [85].

Since their first demonstration in 2007, the power output of THz DFG–QCLs has experienced exponential growth, as shown in figure 9. Recent advances include demonstration of room-temperature 3.5 THz DFG–QCLs with THz peak power output of 1.9 mW [74] and the THz WPE of $0.7 \times 10^{-3}$ and room-temperature CW devices with over $3 \mu$W of THz power output at 3.6 THz [73, 74] by the Razeghi group. High performance was achieved by optimizing the waveguide layers thickness and doping, processing lasers as buried-heterostructure devices, and employing flip-chip bonding onto diamond substrates for efficient current extraction through side-contacts and thermal management [73, 74]. The active region of these devices was based on the single phonon resonance depopulation design with thinned extraction barrier which is conceptually similar to the bound-to-continuum active region presented in figure 5(a). New active region design for THz DFG QCLs based on dual upper state design (figure 5(b)) with 1.5–2 times higher optical nonlinearity for THz DFG, compared to bound-to-continuum active region, was also reported recently [67].

It is estimated that less than 20% of generated THz power in a Cherenkov wave in current devices being outcoupled to free space [79]. Thus, similar to THz QCLs [43], significant improvements in THz power output in THz DFG–QCLs are expected to be produced by optimizing THz outcoupling efficiency using various photonic structures, such as properly-optimized second-order gratings. Overall, Manley–Rowe relations may be viewed as the fundamental limit on WPE for THz DFG systems. From this perspective, given the current record values for room-temperature, pulsed-operation WPE of mid-IR QCLs of 27% at $\lambda \approx 5 \mu$m [86] and 6% at $\lambda \approx 10 \mu$m [87], one obtains the Manley–Rowe-limited WPE value of $\sim 1\%$ for 4 THz generation in THz DFG–QCLs at room temperature. Given that the best WPE of THz DFG–QCLs reported to date is $\sim 0.7 \times 10^{-3}\%$ [74], we expect that the WPE of THz DFG–QCLs may be improved by up to three orders of magnitude and devices with room-temperature THz peak power output in excess of 100 mW may be produced through waveguide structure, active region, and packaging optimization.

THz DFG–QCLs may also be an ideal source for THz frequency comb generation with the spectral bandwidth comparable to or larger than that of THz QCLs [88, 89], and room-temperature operation. Room-temperature frequency-comb generation in mid-IR QCL has been reported recently [90] and the intra-cavity down conversion of such combs in THz DFG–QCLs appears to be a promising route towards achieving multi-octave-spanning THz frequency comb sources.

Acknowledgments

The authors would like to acknowledge Alfred Cho and Deborah Sivco of Bell Laboratories; Feng Xie and Alexey Belyanin of Texas A&M University; Douglas Oakley, Chris Vineis, and George Turner of MIT Lincoln Laboratory; Milan Fischer, Andreas Wittmann, and Jerome Faist of ETH Zurich; Augustinas Vizbaras, Frederic Demmerle, Christian Grasse, Gerhard Boehm, and Markus-Christian Amann of the Walter Schottky Institute of the Technical University of Munich; Xiajun Wang and Mariano Troccoli of AdTech Optics; Kazuue Fujita, Masahiro Hitaka, Akio Ito, Tadataka Edamura, and Masamichi Yamanishi of Hamamatsu Photonics, and Seunyong Jung, Karun Vijayaraghavan, Yifan Jiang, Aiting Jiang, Ming Jang, Robert Adams, and Karthik Choutagunta, of the University of Texas at Austin for their

![Figure 9](image-url)
contributions to the development of THz DFG–QCL technology discussed in this review.

The results discussed in this review were produced with the financial support from the Air Force Office of Scientific Research, National Science Foundation, Defense Advanced Research Projects Agency, and Norman Hackerman Advanced Research Program of the State of Texas. The preparation of this review paper was supported in part by the National Science Foundation grants number ECCS-1150449 (CAREER) and ECCS-1408511.

References

[31] Kumar S, Chan C W I, Hu Q and Reno J L 2012 A 1.8 THz quantum cascade laser operating significantly above the temperature of Nat. Phys. 7 166–71


Adams R W, Vijayaraghavan K, Wang Q J, Fan J, Capasso F, Khanna S P, Davies A G, Linfield E H and Belkin M A 2010 GaAs/Al\textsubscript{0.15}Ga\textsubscript{0.85}As terahertz quantum cascade lasers with double-phonon resonant depopulation operating up to 172 K Appl. Phys. Lett. 97 131111


Ohtani K, Beck M, Scalari G and Faij J 2013 Terahertz quantum cascade lasers based on quaternary AlInGaAs barriers Appl. Phys. Lett. 103 041103

Fischer M, Scalari G, Walther C and Faij J 2009 Terahertz quantum cascade lasers based on In\textsubscript{0.53}Ga\textsubscript{0.47}As/In\textsubscript{0.52}Al\textsubscript{0.48}As Appl. Phys. Lett. 95 9382–7


Lu Q Y, Bandyopadhyay N, Slivken S, Bai Y and Razeghi M 2013 Room temperature terahertz quantum cascade laser sources with 215 μW output power through epilayer-down mounting Appl. Phys. Lett. 103 011101


Askaryan G A 1962 Cerenkov radiation and transition radiation from electromagnetic waves Sov. Phys.—JETP 15 943


Belkin M A (unpublished)

Straub A et al 2002 Simultaneously at two wavelengths (5.0 and 7.5 μm) singlemode and tunable quantum cascade distributed feedback laser Electron. Lett. 38 565–7


monolithic room-temperature terahertz quantum cascade laser sources Nat. Commun. 5 4267


