LASER & PHOTONICS
REVIEWS

Laser Photonics Rev. **6**, No. 1, 24–46 (2012) / **DOI** 10.1002/lpor.201100019

Abstract This paper reviews beam engineering of mid-infrared and terahertz quantum cascade lasers (QCLs), based on two approaches: designer plasmonic structures and deformed microcavities. The plasmonic structures couple laser emission into surface waves and control the laser wavefront in the near-field, thereby greatly increasing beam collimation or introducing new functionalities to QCLs. The plasmonic designs overall preserve laser performance in terms of operating temperature and power output. The deformed microcavity QCLs operate primarily on whispering-gallery modes, which have much higher quality factors than other modes, leading to lower threshold current densities. Cavity deformations are carefully controlled to greatly enhance directionality and output power.

Beam engineering of quantum cascade lasers

Nanfang Yu[1], Qijie Wang[2], and Federico Capasso[1,*]

1. Introduction

There is an intrinsic trade-off between laser threshold and power output in the design of a laser resonator. On one hand, the resonator should confine laser radiation long enough (high quality-factor or Q-factor cavity) so that it experiences sufficient gain; on the other hand, it should provide efficient outcoupling in terms of high optical power and beam quality. Physically, the threshold in semiconductor lasers is proportional to the sum of waveguide losses and power outcoupling losses, and inversely proportional to the gain coefficient and laser mode confinement factor [1]. Laser design hinges on the proper balance between lowering the threshold pump current, which facilitates continuous-wave operation at room temperature, and achieving substantial useable optical power (i. e., with desirable wavelength control such as multi-wavelength operation and broad tunability), and beam characteristics. Concerning the latter, useful features, depending on the applications, include small beam divergence, high beam quality factor [2] (or M^2 factor, which measures how closely the beam cross-section resembles that of an ideal Gaussian beam), beams with certain polarization states (circular, azimuthal, etc), and beams that carry orbital angular momentum [3, 4].

Laser cavities in the shape of disks and cylinders are the most favorable for low threshold due to whispering-gallery modes (WGMs) [5–8]. In WGM lasers photons can only escape via evanescent tunneling due to the finite radius of curvature of the resonator or through scattering at defects on the cavity walls [9]. Lasers based on this principle feature low threshold but also low power output and isotropic emission [9]. One has to employ less rotationally symmetric cavities to construct useful lasers.

The simplest and the most widely used semiconductor laser configuration is a straight ridge waveguide. Its two facets define a Fabry-Pérot cavity and also allow some light to escape. The power out-coupling loss or mirror loss, which is inversely proportional to the ridge length [1], is controllable. Instead of using two discrete facets, first-order Bragg gratings can be incorporated into laser waveguides to provide distributed feedback and to select laser frequency [10]. The drawback of the ridge waveguide design, however, is that the emitted laser beam is typically highly divergent due to diffraction. The effect is especially large for edge-emitting ridge QCLs with emission wavelength typically a few to ten times greater than the thickness of the active core, which leads to divergence angles at least tens of degrees in the direction perpendicular to the substrate. In particular, the emission of double-metal waveguide terahertz (THz) QCLs is essentially uniform over the half space outside the end facet of the waveguide (there are even backward emissions towards the other end of the waveguide). To achieve much smaller beam divergence, one can distribute the light emitting region over a larger area by patterning laser ridges with grating out-couplers. The most popular design is a surface emitting configuration via a second-order grating, which features a small divergence angle inversely proportional to the length of the grating. This strategy has been successfully applied to mid-IR [11–14] and THz [15, 16] ridge QCLs and to surface-emitting ring or disk QCLs [17–19]. A similar approach is based on surface-emitting QCLs with 2D photonic-crystal cavities [20–24]. Such devices operate on photonic band-gap modes (i. e., microcavity-based) or band-edge modes (i. e., distributed-feedback based) and have been shown to achieve controllable far-field emission patterns. Recently, a novel THz QCL based on a third-order grating has been demonstrated to produce highly directional emission in the direction along the laser ridge [25].

A completely different approach for laser resonator design is to break the rotational symmetry of whispering-

---

[1] School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA [2] Present address: School of Electrical and Electronic Engineering and School Physical and Mathematical Sciences, Nanyang Technological University, 50 Nanyang Avenue, 639798 Singapore
[*] Corresponding author: e-mail: Capasso@seas.harvard.edu

gallery mode lasers by using deformed microcavities to improve the power outcoupling efficiency and beam directionality. Successful experimental demonstrations include quadrupolar shaped lasers [26, 27], spiral-shaped lasers [28–30], and rounded triangular lasers [31]. These deformed microcavity lasers, however, operate in the chaotic regime where WGMs are no longer supported, and therefore have degraded Q-factors [32]. Chaotic behavior in these resonators originates from sensitivity to initial conditions of ray trajectories and is characterized by their exponential divergence as the rays propagate in the cavity. Ray dynamics in these non-integrable geometries is described by the nonlinear eikonal equation [33–36]. Circular or cylindrical cavities and ridge waveguides previously discussed, instead, belong to integrable geometries, which have deterministic evolution of ray trajectories. There is yet another type of laser cavity design based on Anderson localization in a random medium [37–39], which however is out of the scope of this paper.

In this paper we present systematic overview of our research aimed at improving QCL power output coupling and directionality, following a brief description of the quantum design of mid-infrared (mid-IR) and terahertz (THz) QCLs. One approach, dubbed wavefront engineering, is to integrate ridge-waveguide QCLs with plasmonic structures that control the far-field emission by tailoring the near-field. These plasmonic structures are less intrusive than the previously-mentioned gratings and photonic-crystal structures since they are separated from the laser waveguide and do not alter the optical mode confinement or introduce additional waveguide losses. To increase power output coupling of THz QCLs, care was taken to reduce index-mismatch between waveguide modes and surface plasmon modes. Mid-IR and THz QCLs with highly directional emission [40–43], controllable polarization [44], multi-beam and multi-wavelength behavior are discussed [45]. Another approach is to change the shape of the laser cavity through suitable deformations. Deformed microcavity lasers and specifically limaçon-shaped QCLs [46, 47] and notched elliptical QCLs [48] are presented. The key aspect of these designs is that the cavity deformation is carefully chosen to preserve WGMs and to generate efficient and directional output without significantly reducing the Q-factor. In the case of the limaçon-shaped QCLs, the deformation is small enough so that they preferentially lase in whispering-gallery-like (WG-like) modes, rather than in a chaotic regime. The notched elliptical QCLs intrinsically support WGMs because an ellipse, like a circle, is an integrable geometry. Their highly unidirectional emission is attained by utilizing the collimation property of ellipses.

2. Band-structure engineering of quantum cascade lasers

First invented at Bell Labs in 1994 [49], the QCL overthrows the operating principle of conventional semiconductor lasers by relying on a radically different process for light emission, which is independent of the bandgap. Instead of using 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 bandgap and h is Planck's constant, QCLs use only one type of charge carriers (electrons) that 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 bandgap by utilizing instead optical transitions between quantized electronic states, are many and far reaching, amounting to a laser with entirely different operating characteristics from laser diodes and superior flexibility and functionality.

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. To achieve this cascading emission of photons, active regions are alternated with doped electron injectors and an appropriate bias voltage is applied. The active-region-injector stages of QCLs give rise to an energy staircase in which photons are emitted at each of the steps. The number of stages typically ranges from 20 to 50 for lasers designed to emit in the mid-IR range, and it is typically a couple of hundred for THz QCLs. This cascade effect is responsible for the very high power that QCLs can attain.

Figure 1 illustrates a typical energy diagram of two stages of a mid-IR QCL, which serves to illustrate the key operating principle. Note that each stage comprises an electron injector and active region. The latter contains three quantized states; the laser transition is defined by the energy difference between states 3 and 2, which is determined primarily by the thickness of the two wider wells. A population inversion between levels 3 and 2 is required for laser action. This translates to a requirement that the electron relaxation time between levels 3 and 2 should be substantially longer than the lifetime of state 2. To achieve that, the lowest level 1 is positioned about one optical phonon energy (~ 34 meV) below level 2, so that electrons in the latter state rapidly scatter by emission of an optical phonon to energy level 1. Because of its resonant nature, this process is very fast, characterized by a relaxation time of the order of 0.1 to 0.2 ps [1]. Electrons in level 3 have instead a substantially longer lifetime because of the much larger energy difference between states 3 and 2 so that the electron-phonon scattering process between the latter states is non-resonant. To achieve lasing, however, one must also suppress the unwanted escape route by tunneling from state 3 to states on the right hand side, which form a broad quasi-continuum. Such escape would reduce the level 3 population. To prevent this occurrence, one designs the injector of the next stage as a superlattice with an electronic density of states such

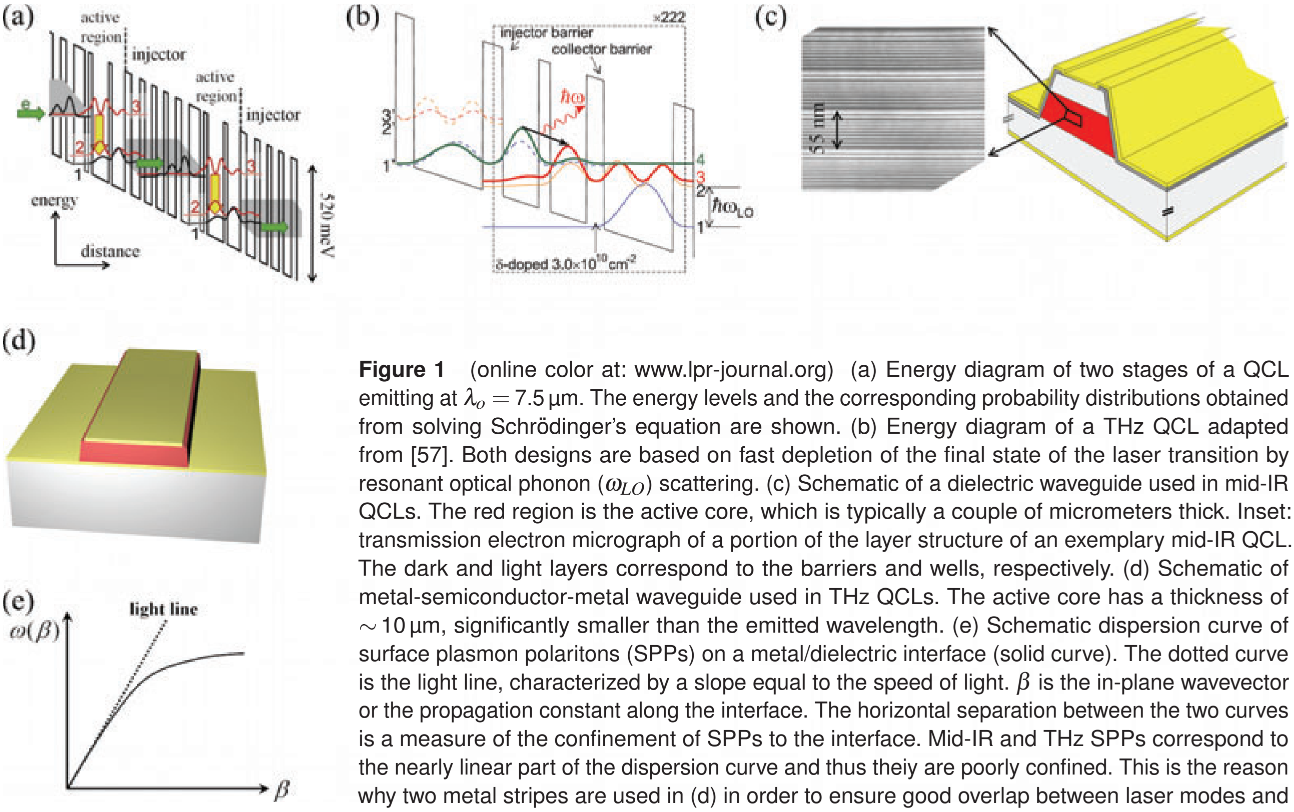


Figure 1 (online color at: www.lpr-journal.org) (a) Energy diagram of two stages of a QCL emitting at $\lambda_o = 7.5\,\mu\text{m}$. The energy levels and the corresponding probability distributions obtained from solving Schrödinger's equation are shown. (b) Energy diagram of a THz QCL adapted from [57]. Both designs are based on fast depletion of the final state of the laser transition by resonant optical phonon (ω_{LO}) scattering. (c) Schematic of a dielectric waveguide used in mid-IR QCLs. The red region is the active core, which is typically a couple of micrometers thick. Inset: transmission electron micrograph of a portion of the layer structure of an exemplary mid-IR QCL. The dark and light layers correspond to the barriers and wells, respectively. (d) Schematic of metal-semiconductor-metal waveguide used in THz QCLs. The active core has a thickness of $\sim 10\,\mu\text{m}$, significantly smaller than the emitted wavelength. (e) Schematic dispersion curve of surface plasmon polaritons (SPPs) on a metal/dielectric interface (solid curve). The dotted curve is the light line, characterized by a slope equal to the speed of light. β is the in-plane wavevector or the propagation constant along the interface. The horizontal separation between the two curves is a measure of the confinement of SPPs to the interface. Mid-IR and THz SPPs correspond to the nearly linear part of the dispersion curve and thus theiy are poorly confined. This is the reason why two metal stripes are used in (d) in order to ensure good overlap between laser modes and the active core.

that at the energy corresponding to E_3 there is no resonant electronic state, but rather a region of low density of states known as a minigap. Notice instead the dense manifold of states (a miniband) facing levels 2 and 1, which favors efficient electron extraction from the active region. Finally, note that electrons are injected into the upper laser level by a process known as resonant tunneling [50], which ensures highly selective injection when the applied voltage is increased to a certain value so that the injector and the upper laser state are aligned.

For mid-IR wavelengths greater than $\sim 5\,\mu\text{m}$, the alloy compositions for the wells and barriers (i. e., $\text{In}_{0.47}\text{Ga}_{0.53}\text{As}$ and $\text{Al}_{0.48}\text{In}_{0.52}\text{As}$) are chosen to be lattice matched to the InP substrate. The quantum well barrier height for these compositions is the conduction band discontinuity between the two alloys: $\Delta E_c = 0.52\,\text{eV}$. For operating wavelengths shorter than $\sim 5\mu\text{m}$, strained AlInAs/GaInAs heterostructures (i. e., higher Al composition for the barrier and lower In content for the quantum wells) are used, so that barrier heights in the 0.7 eV to 0.8 eV range can be achieved, which suppresses electron leakage over the barriers. All reported high performance continuous-wave (CW) room temperature QCLs operating at $\lambda \leqslant 5.2\,\mu\text{m}$ [51–54] used such strained heterojunctions. Another critical design parameter to improve temperature performance is the energy separation between the lower state of the laser transition and the injector ground state. It typically should be in excess of 0.1 eV to minimize thermal backfilling of level 2 by electrons in the injector. In the last few years major improvement in QCL performance (lower threshold and higher power) was achieved by introducing a so-called double phonon resonance design [55]. The active region of this laser has four quantum wells and three energy levels equally separated by the energy of an optical phonon instead of the two levels previously discussed. This active-region design results in a larger population inversion because electrons are more efficiently removed from the lower state of the laser transition and thermal backfilling of the latter is reduced.

THz QCLs were first demonstrated in 2002 [56]. The quantum design of THz QCLs is much more challenging than their mid-IR counterparts primarily because it takes great care to selectively inject electrons into and remove them from, respectively, the upper and lower laser levels, which are closely spaced ($\sim 10\,\text{meV}$). Figure 1b illustrates the bandstructure of a THz QCL of the record operating temperature of 186K [57]. In this design, the laser transition is diagonal (that is, the moduli squared of the wave functions of the upper and lower laser levels have maxima located at different wells and the two wave functions have small spatial overlap). This facilitates selective injection into the upper laser level and reduces nonradiative scattering rate between the two laser levels; both help achieve population inversion. In another design, QCLs based on scattering-assisted injection [58–60] have demonstrated high operating temperature. The key feature of this design is that injection into the upper laser level is by resonant optical phonon scattering instead of by resonant tunneling. This allows one to lower the electron density of the injector, and thus reduces free carrier losses because the electron-phonon scattering rate is faster than the resonant tunneling injection rate. In addition, removing the condition of alignment between the injection level and the upper laser state increases the dynamic range of the laser, making it possible to achieve high optical power [60].

3. Beam engineering of QCLs using designer plasmonic structures

Plasmonics studies a wide range of phenomena involving the interaction between electromagnetic waves and charge density oscillations in metallic structures, generally known as plasmons [61–63]. Plasmonic resonances are characterized by charge oscillations at the plasma frequency and confinement of electromagnetic energy to subwavelength region in the vicinity of the metal. A variety of mechanisms can lead to plasmonic resonances. In the case of optically small metal particles, the resonance is controlled primarily by the dielectric permittivity of the material and also by the shape of the particle. For example, nanometric metal spheres scatter light most efficiently when the real part of the permittivity of the metal, ε_1, is twice that of the surrounding medium, ε_2 [63]. Ancient Greek and medieval artisans used metal inclusion to impart bright colors to glass objects [64]. As the size of the particle increases the plasmon resonance is strongly affected by the phase shift of radiation in accordance with Mie theory [65]. This is the case of optical antennas [66, 67] where plasmonic resonances are primarily due to geometric effects. Antenna resonances occur when the antenna length approximately equals to an odd integer number of the half surface plasmon wavelength [66, 67]. Consider now an extended metal-dielectric interface. Electron oscillations along the metal surface couple to copropagating electromagnetic waves, which have the intensity maximum at the interface and are TM-polarized. These waves, known as surface plasmon polaritons (SPPs), have a dispersion relation with an asymptotic plasma frequency ω_{spp} given by the condition $\text{Real}(\varepsilon_1) = -\text{Real}(\varepsilon_2)$ [61–63]. The dispersion diagram of SPPs on a planar metal-dielectric interface is shown in Fig. 1e. The region of the dispersion curve approaching the asymptote corresponds to surface plasmons localized to a distance from the interface much smaller than the vacuum wavelength and characterized by a strongly enhanced near field caused by the low group velocity charge density oscillations. The portion of the dispersion curve well approximated by the light-line is regarded as "photonic" since at such frequencies SPPs behave very much like photons.

Plasmons have been known for decades, but it was not until two decades ago that researchers started to explore their extremely broad range of applications and physical phenomena made possible by major advances in micro- and nanostructure fabrication and in efficient electromagnetic simulation tools. Today, plasmonics is under active investigation in many research disciplines including near-field optical microscopy [68–70], single molecule spectroscopy [71, 72], chemical and biological sensing [73, 74], subwavelength optics [75–80], metamaterials [81–84], solar cells [85–86], and high-density optical data storage [87, 88].

3.1. 1D plasmonic collimation for mid-IR and THz QCLs

Plasmonic collimation is essentially an antenna array effect [40, 89]. The structure draws inspiration from beaming of optical radiation by metallic aperture-groove structures [78]. The plasmonic collimator consists of an aperture, which is subwavelength in at least one dimension, and a metallic second-order grating patterned on the laser substrate. The aperture is located at the end of a laser waveguide and strongly diffracts light such that a substantial part of the output power is coupled into surface electromagnetic waves on the substrate. The plasmonic grating plays a dual role. First, it progressively and coherently scatters the energy of the surface waves into the far-field, so that the constructive interference between the scattered waves produces a highly collimated laser beam. Second, the interplay between the grating and surface waves improves the confinement of the waves to the substrate, leading to enhanced scattering efficiency. The latter effect is critical for long wavelength surface waves (i. e., beyond near-infrared wavelength), which are more photonic than plasmonic in nature.

Figure 2a shows a scanning electron microscope (SEM) image of a $\lambda_o = 9.9\,\mu\text{m}$ QCL patterned with a 1D plasmonic collimator. The device was fabricated by first sculpting grooves into the cleaved laser substrate using focused ion-beam milling, then conformally coating the laser waveguide facet and the substrate with Al_2O_3 and gold films, and finally opening a slit aperture on the facet. Figure 2b shows the electric-field ($|E|$) distribution around the slit aperture and the first few grating grooves. Most of the near-field intensity is confined within a distance of λ_o from the device substrate, indicating efficient coupling between surface waves and the plasmonic structure. Interference of scattered waves creates a striped pattern in the meso-field. These observations should be compared to the case of a planar interface between gold and air, where mid-IR surface waves in the form of Zenneck waves [90] have loose confinement ($1/e$ decay distance of amplitude in air $\sim 10\lambda_o$). These are also markedly different from the case of a non-optimal grating. For example, poor near-field confinement and less distinguishable interference patterns in the meso-field would occur if the grating period is far from the second-order grating condition, or if the grating groove cavities are well away from resonances. This resonance leads to the maximum scattering efficiency and physically it means that a standing wave builds up along the depth of the grooves or the accumulated phase of near-field during a round trip inside a groove is $\sim 2\pi$.

The measured vertical far-field intensity profile of the 1D collimated mid-IR QCL (Fig. 2a) is plotted in Fig. 2c, indicating that the full-width at half-maximum (FWHM) divergence angle $\Delta\theta$ is about 2.9°. This is similar to that of a side-emitting antenna array, $\Delta\theta \sim \lambda_o/(N\Lambda) \sim 3.2°$, where $N = 20$ is the number of grooves and $\Lambda = 8.9\,\mu\text{m}$ is the grating period. The directivity of the collimated device is ~ 17.7 dB (Fig. 2c black curve), while that of the original unpatterned device is ~ 7.2 dB (Fig. 2c gray curve). The directivity is used here to characterize collimation in the ver-

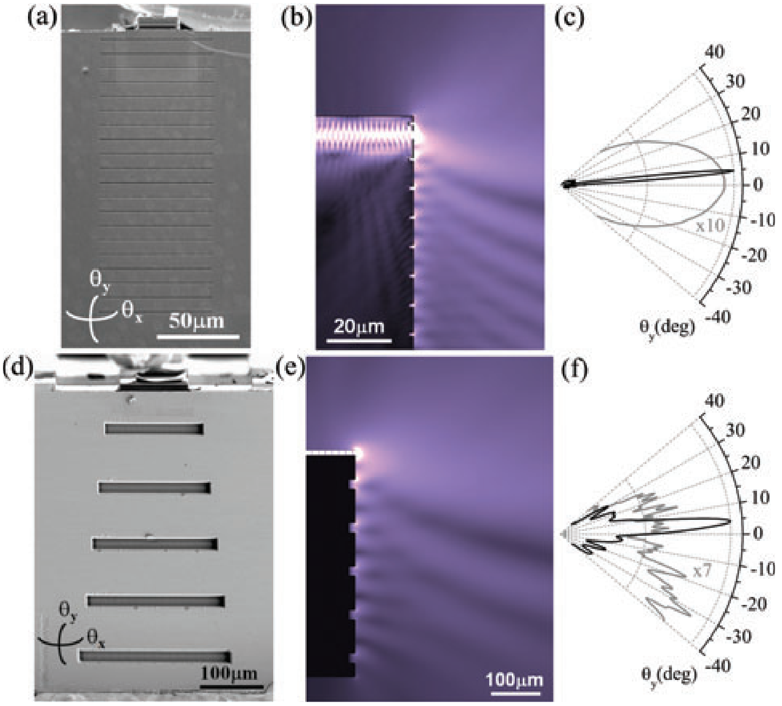


Figure 2 (online color at: www.lpr-journal.org) (a) SEM image of a $\lambda_o = 9.9$ µm QCL with an integrated 1D plasmonic collimator, which is made of gold and has the following geometry: aperture size = 2 µm × 25 µm, grating period = 8.9 µm, distance between aperture and the first grating groove = 7.3 µm, groove width = 0.8 µm, and groove depth = 1.5 µm. (b) Simulated electric-field distribution ($|\mathrm{E}|$) around the slit and the first 7 grating grooves. (c) Measured far-field intensity profiles in the vertical direction of the device in (a) (black curve) and of the original unpatterned device (gray curve). (d) SEM image of a THz QCL with $\lambda_o = 100$ µm and an integrated 1D plasmonic collimator. The collimator is made of gold and has the following geometry: aperture size = 10 µm × 100 µm, grating period = 88 µm, distance between aperture and the first grating groove = 65 µm, groove width = 18 µm, and groove depth = 14 µm. (e) Simulated electric-field distribution ($|\mathrm{E}|$) of the collimated THz QCL. (f) Measured far-field intensity profiles in the vertical direction of the device in (d) (black curve) and of the original unpatterned device (gray curve).

tical direction, and is defined as $D = 10\log_{10}(2\pi I_{\mathrm{peak}}/I_{\mathrm{total}})$, where I_{peak} is the far-field peak intensity and I_{total} is the total intensity under the vertical beam profile. It is advantageous to adjust the scattering strength of the grating grooves by tuning their depth away from the groove cavity resonance to some extent so that the propagation distance of surface waves on the substrate is comparable to the length of the grating, $N\Lambda$. This ensures that the plasmonic grating scatters almost all the energy of the surface waves into the far-field and that the maximum number of scattering grooves are employed to minimize far-field divergence.

A similar design strategy was applied to THz QCLs emitting at $\lambda_o \sim 100$ µm; see Fig. 2d. The fabrication procedure for a THz plasmonic collimator is identical to those for a mid-IR QCL. Fairly wide grating grooves (width $\sim$ 15%-20% of λ_o) were used to enhance the interaction between THz surface waves and the limited number of grooves. The near- and meso-field share similar features as the mid-IR case; see Fig. 2e. The far-field of the collimated THz device has a FWHM divergence angle of about 10° (Fig. 2f). The far-field of the original unpatterned device, on the other hand, is highly divergent due to strong diffraction at the deeply-subwavelength aperture with width $\sim \lambda_o/10$ (Fig. 1d). The large optical background in Fig. 2f is due to the direct emission from the laser aperture, which dominates over the scattering from the grating grooves.

In the 1D collimators for both the mid-IR and THz, the separation between the aperture and the adjacent grating groove, d, is smaller than their respective grating period, Λ. There are two considerations in choosing d<Λ. First, owing to the large spatial frequency components introduced by diffraction at the aperture, the effective wavelength of surface waves in the vicinity of the aperture is smaller than that of surface waves far away from it. The condition d<Λ therefore ensures that the phase difference between the direct emission from the aperture and scattering by the first groove is comparable to the phase difference ($\Delta\varphi \approx 2\pi$) between the waves scattered by adjacent grooves [89]. Second, the efficiency of coupling laser emission into surface waves is a periodic function of d: it reaches a maximum when the reflected surface waves by the grating, and the surface waves that are launched at the aperture and propagating towards the top electrical contact destructively interfere:

$$2dk_{\mathrm{sw}} = (2m+1)\pi \tag{1}$$

where m is an integer and $k_{\mathrm{sw}} \approx k_o$ is the wavevector of surface waves. The equation predicts that efficient coupling occurs when $d = 3\lambda_{\mathrm{sw}}/4$. The values $d = 7.3$ µm $\sim 0.74\lambda_o$, and $d = 65$ µm $\sim 0.65\lambda_o$ were chosen for the mid-IR and the THz QCL, respectively.

3.2. Dipolar model for calculating the far-field of plasmonic gratings

In this section, a dipolar model is presented to calculate the far-field emission pattern of lasers with integrated plasmonic grating structures [91]. The model treats a 2D plasmonic grating as an ensemble of Hertzian electric dipoles distributed along the grating grooves. The dipoles lie in the plane of the substrate and are oriented perpendicular to the grooves; their magnitude and phase correspond to the local electric field of the surface waves. The far-field pattern is calculated by summing the contributions from all the emitting dipoles. This method allows rapid verification and optimization of the plasmonic design without resorting to time-consuming and memory-intensive full-wave simulation techniques.

The most critical issue is to assign the correct magnitude and phase to the Hertzian dipoles, which requires accurate knowledge of surface-wave distribution on the device facet. Several factors complicate this problem.

First, similar to radio waves propagating on the surface of the earth, mid-IR or THz surface electromagnetic waves are loosely bound to planar metallic surfaces. This is because the decay rate of surface waves normal to the interface in the air, κ, is proportional to the separation between the light line and the surface-plasmon dispersion curve, and at such low frequencies the two curves are very close to each other (corresponding to the region close to the origin in Fig. 1e). Theoretically, the decay rate $\kappa = \sqrt{\beta^2(\omega) - k_o^2(\omega)} \propto \beta(\omega) - k_o(\omega)$, where β is the in-plane wavevector of the surface waves, and k_o is the free-space wavevector. In the mid-IR and THz frequency range, κ is about a hundredth to a tenth of k_o for a planar metal-air interface. In fact, mid-IR and THz waves primarily couple into 3D spherical waves, $\mathrm{E} \propto \exp(ikr)/r$, when they are emitted from a subwavelength hole-aperture defined on a planar metallic surface; similarly, they couple into 3D cylindrical waves, $\mathrm{E} \propto \exp(ik\rho)/\sqrt{\rho}$, when emitted from an infinitely long slit-aperture [92]. This is markedly different from SPPs in the visible, which are essentially 2D waves.

Second, periodic grooves change the dispersion of mid-IR and THz surface waves, such that energy band gaps are opened at multiples of the reciprocal grating vector, $2\pi m/\Lambda$ [93] (that is, $\beta \approx 0$ in the folded dispersion diagram,); see Fig. 3a. By choosing the grating period to be an integer number of the surface-wave wavelength, and therefore operating near these plasmonic band gaps where the surface-wave dispersion deviates from the light line, the surface-wave decay rate κ is significantly enhanced, leading to a considerable improvement in the confinement of the waves. In this way, long-wavelength surface waves can be made essentially two-dimensional; that is, without considering propagation losses (i. e., scattering by the grating grooves and absorption), the field amplitude along the interface would evolve according to $1/\sqrt{r}$ for subwavelength hole-apertures and would be constant for infinitely long slits. For an aperture of arbitrary dimensions, the following formulas can be used to calculate the distribution of 2D surface waves [94]:

$$f(k_z) = \int_{-l/2}^{l/2} \mathrm{E}(z')\exp(-ik_z z')dz' , \tag{2}$$

$$\mathrm{E}(x,z) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(k_z) \times \exp\left[i\left(k_z z + \sqrt{k_{\mathrm{sw}}^2 - k_z^2}\, x\right)\right] dk_z . \tag{3}$$

Equation (2) calculates the spatial frequency components of the source, which is a 1D slit of length l along the z-axis, and $\mathrm{E}(z')$ is the electric-field distribution along the slit. i is the square root of -1. Equation (3) calculates field distributions on the interface (z-x plane), and the integrand stands for a 2D plane wave weighed by the spatial frequency of the source. k_{sw} is the wavevector of the surface waves, and is a function of the wavelength and the geometry of the plasmonic grating (i. e., grating period, groove width and depth). Note that we do not consider the vectorial nature of the surface waves and Eq. (3) only calculates their amplitude. We simply assume that the strength of the Hertzian dipoles is proportional to the amplitude of the surface waves and that the dipoles are orientated perpendicular to the grating grooves.

Third, the model should properly account for the decay of surface waves, which is primarily due to scattering of their energy by the grooves. We choose an exponential factor $\exp[-r/(N\Lambda)]$, assuming that the field amplitude decays to $1/e$ after propagating through N grooves.

The radiated electric field by an ensemble of Hertzian dipoles in the far field can be calculated as [95]

$$\mathrm{E}(\mathbf{r}) = i\omega\mu\frac{\exp(ikr)}{4\pi r}\left(\hat{\theta} f_\theta + \hat{\varphi} f_\varphi\right) \tag{4}$$

where $\hat{\theta}$ and $\hat{\varphi}$ are unit vectors in the spherical coordinate system in the direction of increasing θ, the polar angle, and φ, the azimuthal angle, respectively. f_θ and f_φ are, respectively, the θ and φ components of the vector current momentum, defined as [95]

$$\begin{aligned}\mathbf{f}(\theta,\varphi) &= \iiint d\mathbf{r}'\mathbf{J}(\mathbf{r}')\exp(-i\mathbf{k}\cdot\mathbf{r}') \\ &= \sum_{x',z'} \mathbf{P}(x',z')\exp(-i\mathbf{k}\cdot\mathbf{r}')\end{aligned} \tag{5}$$

which is current density $\mathbf{J}(\mathbf{r}')$ or the dipole ensemble $\sum_{x',z'} \mathbf{P}(x',z')$ on the x-z plane weighted by a phase retardation factor $\exp(-i\mathbf{k}\cdot\mathbf{r}')$.

The measured 2D far-field intensity distributions for the original and the collimated mid-IR QCL are presented in Figs. 3b and c, respectively, demonstrating significant reduction of beam divergence in the vertical direction. The slight curvature of the far-field pattern in Fig. 3c is due to divergence of surface waves as a result of the finite lateral size of the slit aperture. The magnitude and distribution of the Hertzian dipoles correspond to the amplitude and distribution of the surface waves and are illustrated in Fig. 3d. In the dipolar model, the exponentially decaying factor is chosen to be $\exp[-r/(20\Lambda)]$ to represent scattering losses. k_{sw} is taken to be $1.05k_o$ according to 2D simulations. The calculated far field, Fig. 3f, agrees reasonably well with the experimental result, Fig. 3c.

3.3. Multi-beam mid-IR QCLs

In previous designs, the grating period is approximately 10% smaller than the free-space wavelength, yielding an operating point on the lower edge of an energy gap in the dispersion diagram (Fig. 3a). This leads to 1D-collimated beams about 5 degrees from the surface normal towards the top electrical contact for both the mid-IR and THz QCLs;

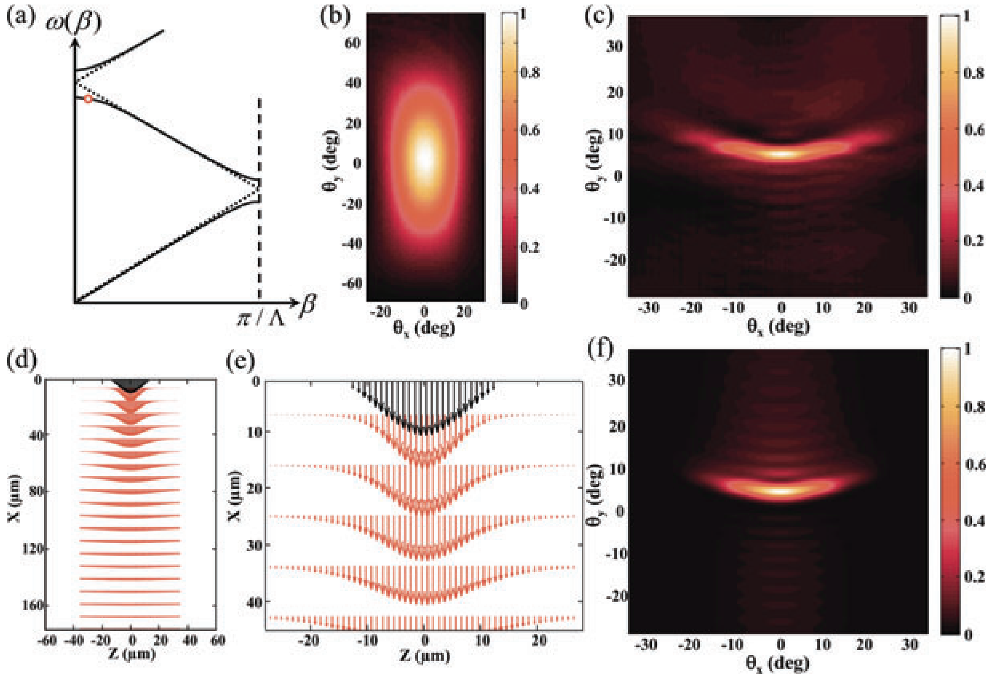


Figure 3 (online color at: www.lpr-journal.org) (a) Schematic dispersion diagram showing that the 1D plasmonic collimator operates on the lower edge of the second energy band gap. The horizontal separation between the operating point (indicated by a red circle) and the light line (dotted curve) is a measure of the confinement of surface waves. For the same grating period, bandgaps in the dispersion diagram open up as the groove depth approaches the cavity resonance condition. (b) and (c) Measured 2D far-field emission patterns before and after patterning the mid-IR 1D plasmonic collimator (Fig. 2a). (d) Dipolar model used to calculate the far-field of 1D collimated QCL in Fig. 2a. Arrows representing Hertzian dipoles are more discernible in the zoom-in plot (e). The length of the arrows represents dipole strength. The dipoles originate from charge oscillations across the two edges of a slit or groove; they are therefore oriented perpendicular to the slit aperture or the grating grooves. Dipoles representing direct emission from the aperture are colored black and those representing scattering from the grooves are colored red. The magnitude of dipoles is proportional to the amplitude of local surface waves, which is calculated based on the 2D model described by Eqs. (2) and (3). For clarity, the magnitude of the black arrows in (d) and (e) is demagnified by a factor of 3. (f) Far-field intensity distribution calculated based on the dipolar model.

see Figs. 2c and f. It is possible to design a larger range of emission angles (e. g., ±25 degrees), provided that the operating point in the dispersion diagram is in the vicinity of the energy gap to maintain sufficient surface-wave confinement and therefore a high scattering efficiency. The emission angle can be estimated as $\arcsin[(2\pi/\Lambda - k_{\mathrm{sw}})/k_o]$. We utilized this concept to split the output of a single-wavelength mid-IR QCL into two 1D-collimated beams by patterning two successive plasmonic gratings with different periodicities on the substrate; see Fig. 4a [45]. The intensity of the dual beams can be independently controlled by tuning the scattering efficiency of grating grooves (i. e., groove width and depth), and by changing the grating length (i. e., number of grooves per grating). For example, to allow the dual beams to have similar peak intensities, more grooves were included in the second grating, which is farther from the aperture and therefore is associated with weaker surface waves.

The measured far field of the dual-beam QCL is shown in Fig. 4d. The grating closer and farther from the aperture diffracts a beam towards ~ 2 and ~ 20 degrees away from the surface normal, respectively. Figures 4e and f are, respectively, 3D full-wave simulation of the far-field, which took about half a day and ~ 10 Gb memory using the Lumerical software [96] in a computer cluster, and analytical calculation based on the dipolar model. The distribution of dipolar emitters in the model is displayed in Figs. 4b and c. The lateral dimension of the second grating was wider than that of the first to account for the divergence of surface waves.

Integrated plasmonic structures can be used to spatially demultiplex emissions from multi-wavelength lasers [45]. To demonstrate this, we used mid-IR QCLs with two stacks of active cores emitting at $\sim 10.5\,\mu\mathrm{m}$ and $\sim 9.3\,\mu\mathrm{m}$. We chose a groove geometry that is the trade-off between the optimal structures for the two wavelengths to provide efficient scattering for both. Specifically, for groove width equal to 0.7 μm, the groove cavity resonances are reached when groove depth h is 1.1 μm and 1.6 μm for $\lambda_o = 9.3\,\mu\mathrm{m}$ and 10.5 μm, respectively; we chose $h = 1.3\,\mu\mathrm{m}$.

Simulation and experimental results for a dual-wavelength device are presented in Fig. 5. A single grating with a periodicity of 8.5 μm was used; it deflects the $\lambda_o = 9.3\,\mu\mathrm{m}$ component away from the surface normal toward the top contact by $\sim 3°$ (Figs. 5b and c) and the $\lambda_o = 10.5\,\mu\mathrm{m}$ component by $\sim 12°$ (Figs. 5d and e). The measured direction and divergence of the two beams are in good agreement with full-wave FDTD simulations; see Figs. 5b–e.

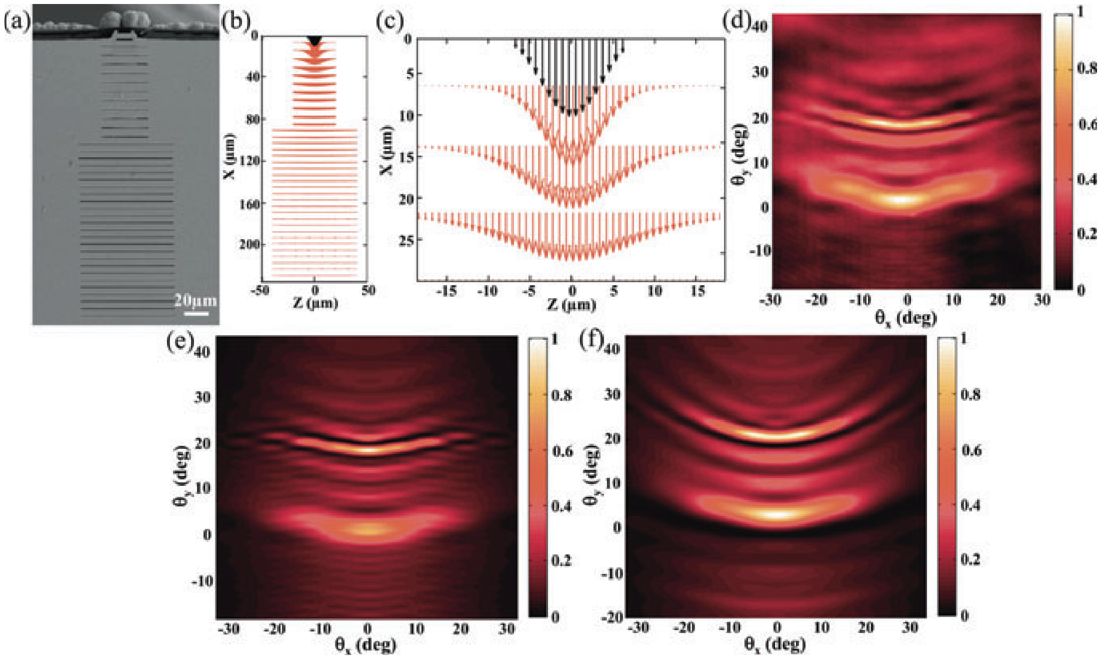


Figure 4 (online color at: www.lpr-journal.org) (a) SEM image of a $\lambda_o = 8.1$ μm QCL integrated with two plasmonic gratings. The grating closer to the aperture contains 11 grooves and has a 7.8-μm periodicity. The other grating contains 25 grooves and has a 6-μm periodicity; it has larger lateral dimensions to account for the lateral spreading of surface waves. All grating grooves have depth and width of 1 μm and 0.6 μm, respectively. (b) Dipolar model used to calculate the far-field intensity distribution of the device. Black and red arrows represent, respectively, the magnitude of dipolar emitters distributed along the aperture and the grating grooves. The magnitude of the black arrows is demagnified by a factor of 5 for clarity. (c) Zoom-in plot of (b). (d) Measured emission profile of the device in (a). (e) Full-wave simulation of the far-field intensity distribution of the device. (f) Calculated far-field intensity distribution of the device using the dipolar model.

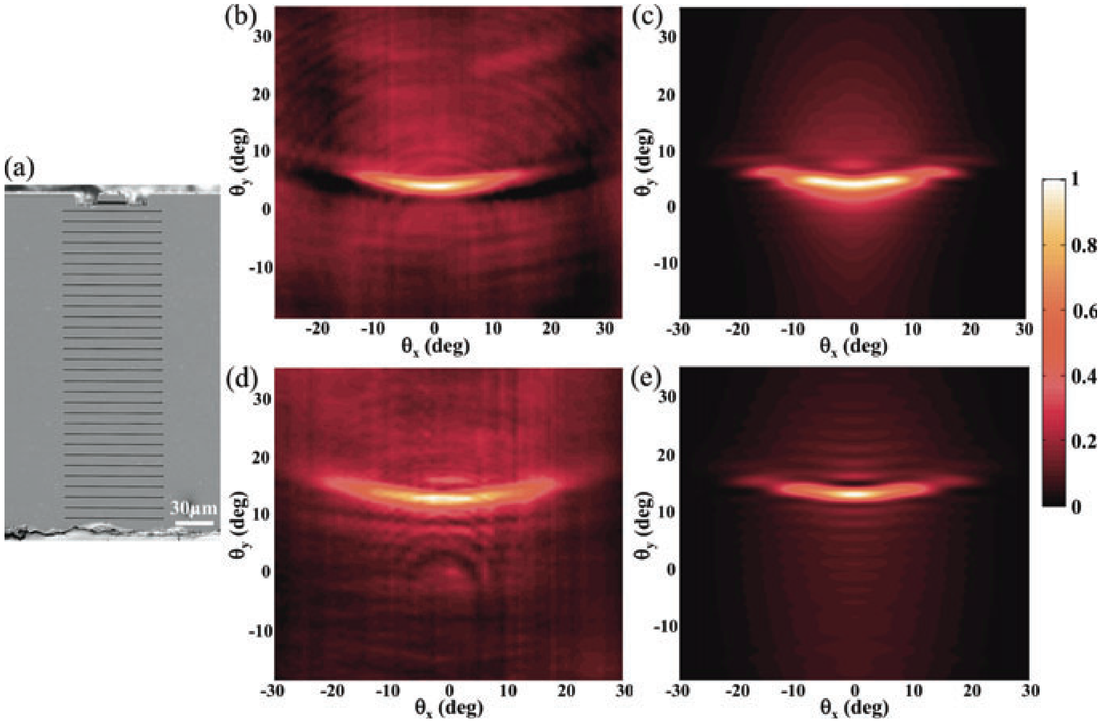


Figure 5 (online color at: www.lpr-journal.org) (a) SEM image of a dual-wavelength QCL patterned with a plasmonic demultiplexer. (b) and (c) Measured and simulated emission patterns of the $\lambda_o = 9.3$ μm component of the laser emission, respectively. (d) and (e) Measured and simulated emission patterns of the $\lambda_o = 10.5$ μm component of the laser emission.

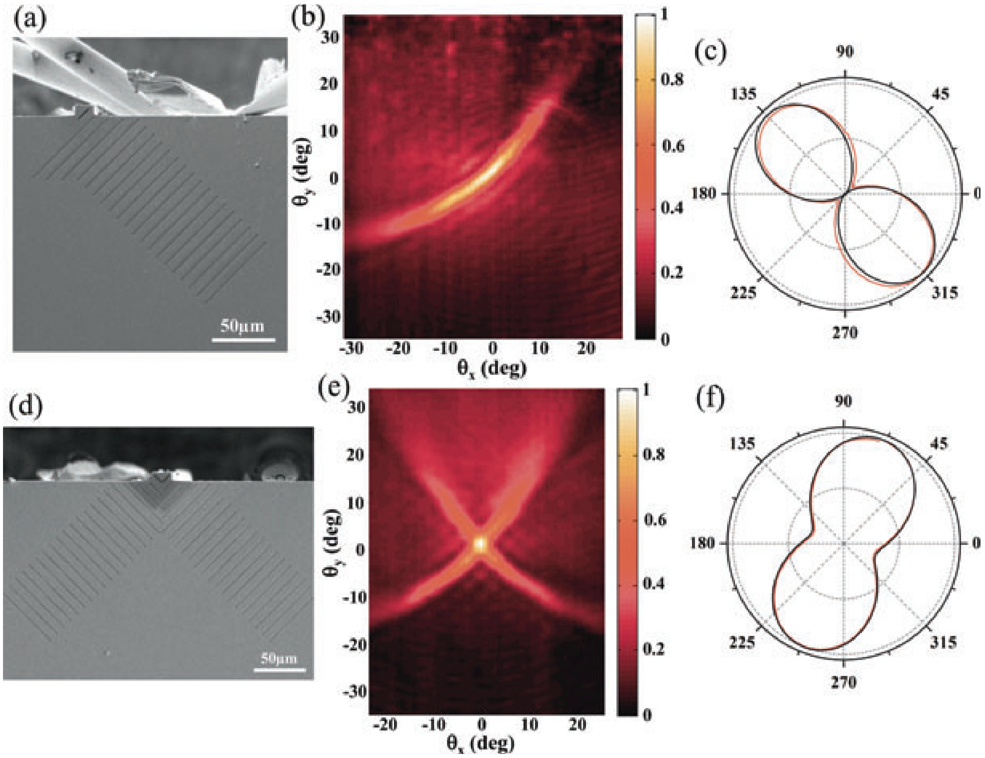


Figure 6 (online color at: www.lpr-journal.org) (a) SEM image of a $\lambda_o = 9.9\,\mu$m QCL integrated with a linear polarizer. The orientation of the slit aperture and the grating is 45° with respect to the vertical direction. (b) Measured emission pattern of the device. (c) Measured device power output as a function of the rotation angle of a wire-grid polarizer, which is placed in front of a mid-IR detector. The red curve is experimental data and the black curve is a calculation assuming a 45° linearly-polarized light beam. (d) SEM image of a $\lambda_o = 9.9\,\mu$m QCL capable of emitting circularly-polarized light. (e) Measured emission pattern of the device in (d). (f) Measured optical power of the central spot in (e) while the wire-grid polarizer was rotated in front of the detector. The red curve is experimental data and the black curve is a fit assuming coherent superposition of a circularly-polarized beam and a linearly-polarized beam.

3.4. QCLs integrated with plasmonic polarizers

Semiconductor lasers are TE or TM polarized, owing to the optical selection rules of the gain medium [10,97]. However, a variety of polarization states will be beneficial for applications (e. g., polarization multiplexing in communications, detection of chiral biological molecules). The demonstration of a circularly-polarized semiconductor laser would be of major significance. Integrated plasmonic structures, without introducing fundamental changes to the quantum design of the active core, may provide a convenient means to control laser polarization by manipulating laser near-field.

The basic plasmonic element we used toward this purpose resembles the 1D plasmonic collimators; see Fig. 6a. The slit aperture, similar to a slot antenna, is illuminated by laser waveguide modes polarized in the vertical direction. Only the component of the laser polarization perpendicular to the long edges of the slit excites antenna modes of the slot (that is, voltage standing waves along the slit), and generates surface waves on the substrate. Interference between waves scattered from the grating grooves creates a beam in the far-field with reduced divergence in the direction perpendicular to the grooves; see Fig. 6b. The polarization of the far-field is along the same direction, as confirmed by experiment (Fig. 6c).

A circularly-polarized beam can be created by coherently combining two linearly-polarized beams with perpendicular polarizations, 90° phase difference, and the same amplitude. We fabricated two 1D collimators on the device substrate with their respective orientations equal to ±45° from the vertical direction; see Fig. 6d. The distance between the aperture and the nearest groove (denoted d_1 and d_2 for the two collimators) introduces a phase shift in the scattered light. The two beams produced by the left and the right gratings will have a 90° phase difference if $|d_1 - d_2| \sim \lambda_{\text{sw}}/4$. Keeping in mind the optimal coupling condition $d \sim 3\lambda_{\text{sw}}/4$ previously discussed, we chose $d_1 \sim 3\lambda_{\text{sw}}/8$ and $d_2 \sim 5\lambda_{\text{sw}}/8$ to make sure the two beams have similar amplitude.

The device shown in Fig. 6d emits a cross-shaped far-field pattern; see Fig. 6e. Figure 6f is the power at the center of the cross measured when a wire-grid polarizer was rotated in front of a detector. Analysis indicated that this part of the laser beam consists of a right circularly-polarized component and a linearly-polarized component. The latter is from direct emission from the two apertures.

Several strategies can be employed to increase the circularly-polarized component in the output. Instead of using straight grating grooves, circular ones will produce 2D-collimated light, as will be shown later, so that the two beams created by the two gratings will have complete spatial overlap. In addition, coupling efficiency into surface waves can be largely improved by including spoof SPP grooves adjacent to the laser aperture. The physics and applications of spoof SPP structures will be discussed later.

3.5. 2D collimated QCLs

For 2D collimation [41,42], it is crucial to let surface waves propagate efficiently in two dimensions on the laser substrate so that the array antenna effect can be utilized in two orthogonal directions. We chose a design comprising a rectangular subwavelength aperture and a circular plasmonic second-order grating; see Fig. 7a. Surface waves propagate preferentially in the vertical direction owing to the TM polarization of QCLs. We relied on diffraction to launch surface waves in the lateral direction. To this end, the lateral width of the aperture w_1 was kept subwavelength, but with the disadvantage of reduced power throughput. Experimentally, as w_1 was increased from $\sim 2\,\mu$m to $\sim 8\,\mu$m (while the vertical aperture size w_2 kept constant at $2\,\mu$m; see Fig. 7c), the power throughput increases by a factor of 3 to $\sim$ 50% of that of the unpatterned device, and the lateral far-field divergence angle only deteriorates slightly from 3.7° to 4.6°; see Figs. 7b and d.

In the dipolar model, the magnitude of dipoles is proportional to the amplitude of local surface waves. The surface waves decay as they propagate away from the laser aperture due to the expansion of wavefront and due to scattering by the plasmonic grating. We chose a factor $\exp[-r/(10\Lambda)]$ to represent the scattering losses. As for the expanding wavefront, we compare two situations by assuming either 2D or 3D surface waves. In the 2D case, the calculated near-field distribution ($|\mathrm{E}|$) according to Eqs. (2) and (3) is shown in Fig. 7e. The field amplitude decreases as $1/\sqrt{r}$ (so that the power in the plane is conserved, that is, $2\pi r|\mathrm{E}|^2$ is constant). In the 3D case (Fig. 7f), the waves emitted from the aperture have a spherical wavefront, therefore the field amplitude decays as $1/r$, which is faster than the 2D case. The calculated far-field emission patterns assuming 2D and 3D surface waves are shown in Figs. 7g and h, respectively. Figure 7g has smaller optical background and agrees better with the experimental result Fig. 7b. It implies that 2D surface waves are a closer approximation to the real situation. This physically makes sense considering that plasmonic gratings improve the surface-wave confinement. We note that in the calculations presented in Figs. 7g and h, the aperture is replaced by a Hertzian dipole to account for direct emission. The strength of the dipole is taken to be 15 times the magnitude of the strongest dipole on the groove adjacent to the aperture, based on 2D simulations of the near-field distribution.

3.6. Designer plasmonics for THz QCLs

The dielectric permittivity (both real and imaginary part) of metals in the THz frequency regime is about two orders of

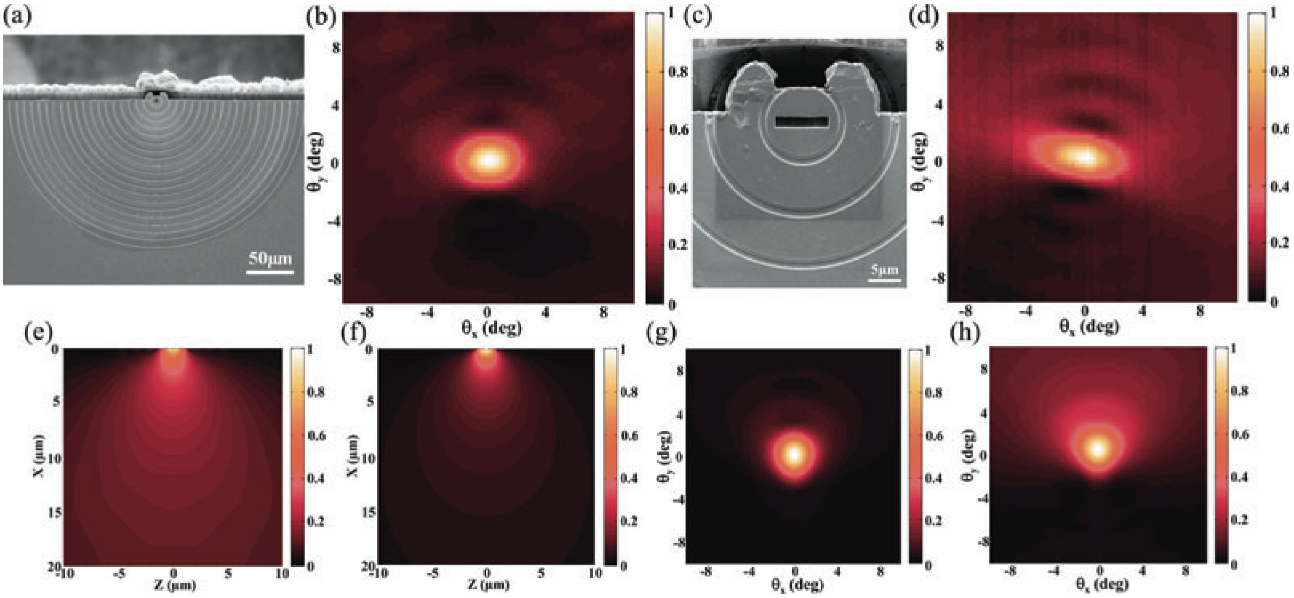


Figure 7 (online color at: www.lpr-journal.org) (a) SEM image of a $\lambda_o = 8.1\,\mu$m QCL integrated with a 2D plasmonic collimator (aperture size 2×2 μm^2). (b) Measured emission pattern of the device in (a). (c) SEM image of the device when the width of the aperture was increased to $\sim 8\,\mu$m. (d) Measured emission pattern of the device in (c). (e) and (f) Calculated electric-field distributions ($|\mathrm{E}|$) generated by a 2-μm-wide slit assuming, respectively, 2D and 3D surface waves. The 2D waves are calculated using Eqs. (2) and (3). The 3D waves are calculated assuming that an array of Hertzian dipoles are placed along the slit and that they emit into the 3D free space. (g) Calculated far-field intensity of the 2D collimated QCL assuming a near-field distribution of (e) in the dipolar model. (h) Calculated far-field intensity of the 2D collimated QCL assuming a near-field distribution of (f) in the dipolar model. Free-space wavelength is $\lambda_o = 8.1\,\mu$m in all calculations (e)–(h).

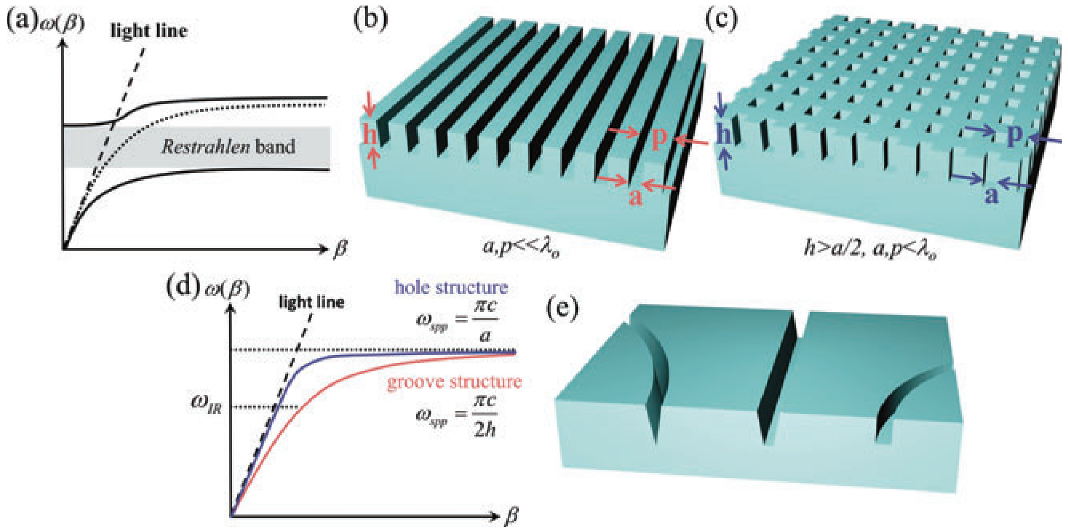


Figure 8 (online color at: www.lpr-journal.org) (a) The *Reststrahlen* band in compound semiconductors divides the dispersion curve of surface waves into two branches. The asymptote of the lower branch, determined by the lower edge of the band, can be much smaller than that of the original dispersion curve. For example, the *Reststrahlen* band for GaAs and InSb is ∼8.1–8.8 THz and ∼5.4–5.7 THz, respectively [109]. This places the asymptote to be around 8.1 and 5.4 THz for the two materials. As a result, the dispersion curves of planar semiconductor-air interfaces are well separated from the light lines at THz frequencies. (b) A spoof SPP structure consists of an array of grooves with a periodicity much smaller than the free-space wavelength λ_o. (c) Another spoof SPP structure consists of a square lattice of holes. The lattice constant in both directions should be much smaller than λ_o. (d) Schematic dispersion diagrams of the two spoof SPP structures in (b) and (c). Their asymptotic frequencies are determined by the geometry of the surface pattern as well as the materials properties of the substrate. The value of ω_{spp} indicated in the figure is for a perfectly conducting substrate. The dispersion curve of the hole structure in (c) is typically steeper compared to that of the groove structure in (b). (e) Schematic of trench structures supporting channel polaritons.

magnitude larger than that in the mid-IR [98]. As a result the confinement of surface waves becomes poorer: the 1/e decay distance of the magnitude of electric field normal to the metal-air interface in the air increases from $\sim 10\lambda_o$ for the mid-infrared to $\sim 100\lambda_o$ for the far-infrared. In correspondence, the skin depth (i. e., 1/e decay distance of the magnitude of electric field inside the metal) decreases from $\sim 1/100\lambda_o$ for the mid-infrared to $\sim 1/1000\lambda_o$ for the far-infrared. That is, essentially all the energy of the light-metal coupled system is carried by electromagnetic waves in the free space while little is associated with oscillating charges in the metal. THz surface waves are difficult to manipulate precisely because it is hard to control the quasi-3D Zenneck waves in the air.

Several strategies can be used to tackle the problem. One method is to use new plasmonic media; of particular interest are conductive oxides (such as indium tin oxide and aluminum zinc oxide) [99, 100] and doped semiconductors [43, 56, 101–103]. These materials have much smaller carrier densities, greatly reducing the asymptotic surface plasmon frequency, ω_{spp}, which is proportional to the square root of carrier densities. For example, GaAs n-doped to $\sim 10^{18}\ \text{cm}^{-3}$ leads to ω_{spp} of a few tens of terahertz. The new plasmonic media have their own peculiarities. Oxide semiconductors are non-stoichiometric and their optical properties can be largely tuned by using different growth methods and annealing conditions [99, 100]. Compound semiconductors have a *Reststrahlen* band, which splits the surface-wave dispersion curve into two branches and effectively reduces ω_{spp} [104–106]; see Fig. 8a. Within the *Reststrahlen* band, electromagnetic energy couples to optical phonons (that is, collective vibrational motions of positively and negatively charged ions against each other in the crystal). In the expression for permittivity, a second term should be added to the Drude term to account for this additional channel of energy dissipation [104–109]:

$$\varepsilon(\omega)=\varepsilon_\infty\left(1-\frac{\omega_p}{\omega(\omega+i\gamma)}\right)+\frac{\omega_{LO}^2-\omega_{TO}^2}{\omega_{TO}^2-\omega^2-i\Gamma\omega} \qquad (6)$$

where ω_{LO} and ω_{TO} are, respectively, the frequencies of longitudinal and transverse optical phonons. ε_∞ is the high-frequency dielectric constant and ε_p is the plasma frequency. γ and Γ are electron and phonon damping constants, respectively.

Another approach to get better control of THz surface waves is to pattern surface structures, which could consist of hole or groove indentations with subwavelength period or trench-like waveguides with subwavelength cross-section. This approach works by introducing localized modes that couple with surface waves. The localized modes are primarily due to the geometry of the textures rather than the material properties of the substrate; that is, the modes exist even for perfectly conducting substrates. In Fig. 8b, when surface waves propagate perpendicular to the grooves, their energy constantly feeds the standing waves inside the groove cavities. As a result, the group velocity of the propagation greatly slows down, leading to a plateau in the dispersion

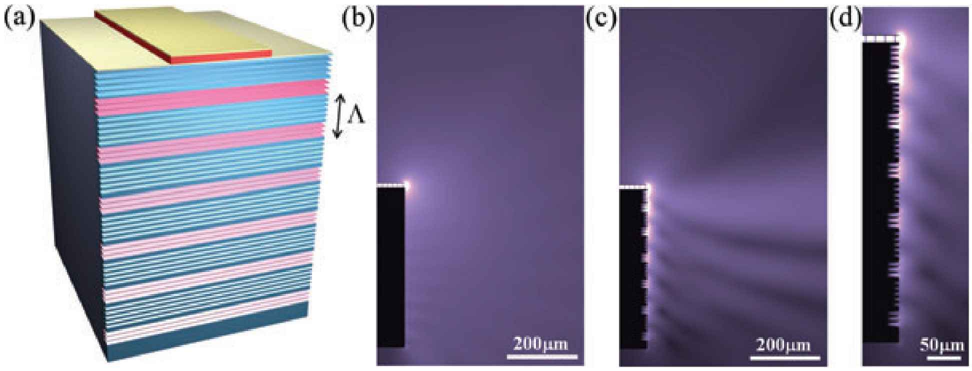


Figure 9 (online color at: www.lpr-journal.org) (a) Schematic of a THz QCL integrated with a metasurface collimator. Spoof SPP grooves are defined directly in the semiconductor substrate of the device. Color indicates deep (pink) and shallow (blue) grooves, which modulate the propagation constant of surface waves and thereby create a second-order grating of period Λ for the surface waves. The grating coherently scatters the energy of surface waves to produce a collimated beam. The metasurface structure increases device power output by enhancing the coupling of laser waveguide modes into surface-waves. (b) and (c) Simulated electric-field distributions ($|\mathrm{E}|$) of the original unpatterned device and the collimated device, respectively. (d) Zoom-in plot of (c).

curve (Fig. 8d). In the case of a 2D square lattice of deep holes, Fig. 8c, the plateau originates from coupling of surface waves to waveguide modes of the holes. The asymptotic frequency is therefore equal to the cut-off frequency of the waveguide. In contrast, there is no cut-off frequency in the previous case: the waves inside the grooves are TEM-like modes. This explains the difference in steepness of the dispersion curve between groove and hole arrays.

The designs in Figs. 8b and c are known as spoof SPP structures because they support well confined surface waves that mimic SPPs in the visible and near-infrared. Spoof SPP structures were introduced into the optics community thanks to Pendry, Martín-Moreno, and García-Vidal's seminal papers [79, 80]. It is interesting to note that similar structures have been studied by radio-wave and millimeter-wave engineers half a century ago [110, 111] and the concept has been used to design corrugated surface-wave antennas [112, 113].

Care should be taken in implementing spoof SPP structures. In the case of an array of grooves (Fig. 8b), the array period needs to be much smaller than the free-space wavelength (preferably $p<\lambda_o/10$), to eliminate diffraction or reflection of surface waves. In addition, the grooves should be sufficiently deep (at least $h>\lambda_o/10$) so that the operating point in the dispersion curve is close to the asymptote ($\omega_{\text{spp}} = \frac{\pi c}{2h}$ or equivalently $h=\lambda_o/4$) to exploit well confined spoof SPPs. The two constraints combined imply fabrication of grooves with considerable aspect ratio. In the case of a 2D hole array, one should avoid approaching the asymptote ($a=\lambda_o/2$), because in that limit one is no longer in the effective medium regime and light strongly diffracts.

As mentioned previously, another type of surface texture is a trench sculpted in a metallic surface (Fig. 8e). In the simplest sense, a trench mimics a coplanar waveguide and it supports TEM-like modes called "channel polaritons" [114–119] that travel along its extension. The modes have larger-than-unity effective mode index and their energy tends to concentrate at the opening of the trench. For trenches with triangular or trapezoidal cross-sections, as the taper angle of the two walls increases, light power is gradually "squeezed" out of the trench up to a point where channel polaritons are no longer sustained. For trenches with a rectangular cross-section, its depth should be sufficiently large to support well-localized channel polaritons.

We utilized several concepts discussed in the preceding paragraphs to create a "metasurface" collimator for THz QCLs [43]; see Fig. 9a. The structure is so named because it is essentially 2D and has tailorable optical properties for controlling surface waves. The collimator was sculpted directly in the heavily doped GaAs laser substrate; no metal coatings were used in order to take advantage of a much lower carrier concentration of the semiconductor and for ease of fabrication. The metasurface pattern has lateral dimensions a few times the waveguide width. When channel polaritons are excited adjacent to the laser aperture, they propagate along the grooves to spread light in the lateral direction, effectively enlarging the emitting aperture. The vertical cross-section of the metasurface structure, as can be seen in Figs. 9c and d, shows subwavelength-spaced grooves with different depths. The groove depth is arranged in such as way to alter the propagation constant of surface waves periodically, thereby effectively creating a plasmonic second-order grating. Keep in mind that spoof SPP structures with deeper grooves correspond to an operating point closer to the asymptote of the dispersion curve and a larger propagation constant.

Compared to the conventional second-order grating previously discussed (Fig. 2), the current design has two advantages. First, the confinement of surface waves is further improved: most of the near-field power is localized within a distance equivalent to ~ 20–$30\%\ \lambda_o$ from the device substrate. Second, the metasurface structure helps extract more power from the laser waveguide and couples a larger percentage of the output power into surface waves. The second effect helps alleviate the problem of poor power outcoupling of THz QCLs with double-metal waveguides. In the original unpatterned THz QCLs (Fig. 9b), the power reflectivity at the laser aperture is $\sim 90\%$ due to a mismatch in propagation constant between the laser waveguide modes, $n_{\text{wg}}k_o$ where $n_{\text{wg}} \sim 3.5$, and the free space modes, k_o. The

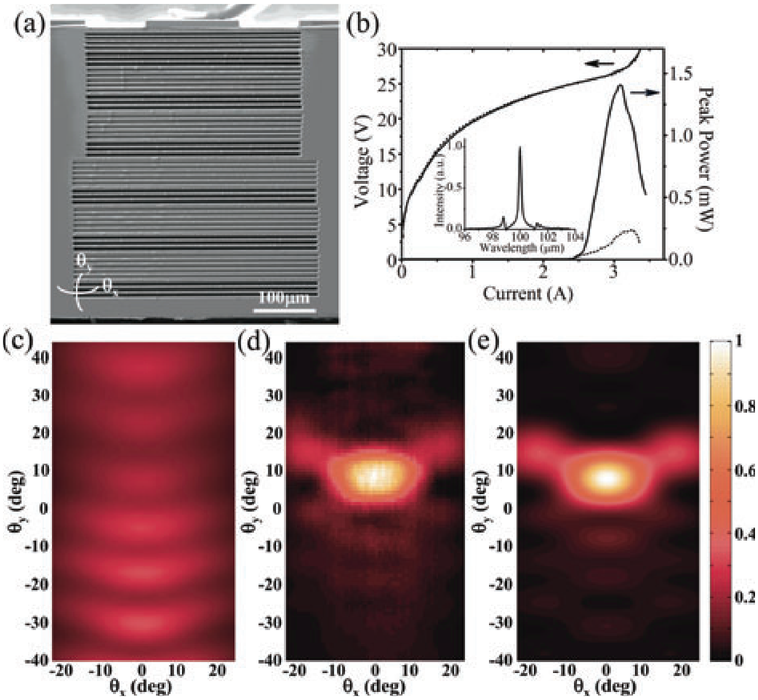


Figure 10 (online color at: www.lpr-journal.org) (a) SEM image of a THz QCL integrated with a metasurface collimator of the design in Fig. 9. The laser emission wavelength is 100 μm and it has a 1.2-mm-long, 150-μm-wide, and 10-μm-thick waveguide. (b) L-I-V characteristics of the device. The dotted, and solid curves are for the original unpatterned device and the device with the metasurface collimator, respectively. Inset: emission spectrum of collimated device at a pump current of 3 A. (c) Simulated far-field intensity distribution of the original unpatterned device. (d) and (e) Measured and simulated far-field intensity distributions of the collimated device.

spoof SPP grooves adjacent to the laser aperture support surface waves with increased propagation constant ($>k_o$), representing a more efficient channel for power outcoupling. In our design, Fig. 9d, the grooves in the close vicinity of the laser aperture have a depth of 12 μm, giving rise to a propagation constant of $\sim 1.25k_o$. As a result, 25% more power is extracted from the laser cavity, and ∼ 50% of all the emitted power is coupled into surface waves.

Figure 10a shows an SEM image of the facet of a fabricated device. The collimator occupies a small footprint with dimensions $\sim 4\lambda_o \times 4.5\lambda_o$ just below the laser waveguide. The central beam has vertical and lateral FWHM divergence angles of $\sim 12°$ and $\sim 16°$, respectively, in good agreement with the 3D full-wave simulation results; compare Figs. 10d and e. The optical background has average intensity below 10% of the central lobe peak intensity, owing to a more balanced direct emission from the aperture and scattered light from the facet. The measured beam directivity is increased from ∼ 8 dB for the original unpatterned laser (Fig. 9c), to ∼ 19 dB for the device with the collimator (Fig. 10d). The directivity of a 2D far-field distribution is defined as $D = 10\log_{10}(4\pi I_{\text{peak}}/I_{\text{total}})$, where I_{peak} is the far-field peak intensity and I_{total} is the total intensity of the 2D far-field. The collected power of the collimated device increases by a factor of ∼ 6 compared to the original unpatterned device; see Fig. 10b. The maximum operating temperature of the patterned device is 135 K, the same as that of the original device. The lasing threshold is primarily

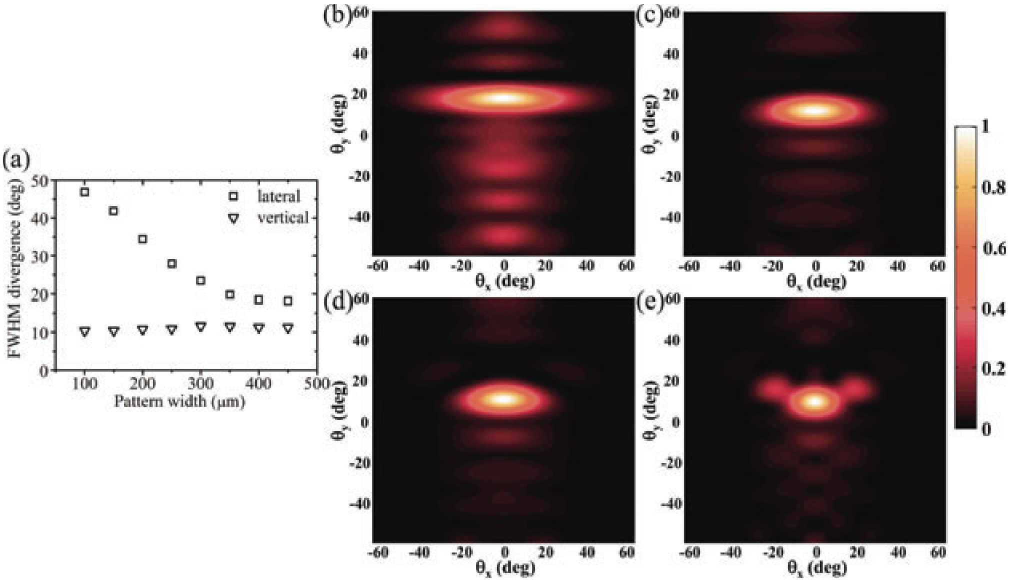


Figure 11 (online color at: www.lpr-journal.org) (a) Lateral and vertical far-field divergence versus lateral size of the THz metasurface collimator. (b)–(e) Simulated far-field intensity distributions corresponding to metasurface patterns with different lateral dimensions 100, 200, 250, and 400 μm, respectively. The metasurface patterns have the same cross-section as in Fig. 9d.

determined by waveguide loss (estimated to be ∼ 30 times larger than mirror loss in our devices) and had little change after defining the collimator; see Fig. 10b.

The effect of channel polaritons is demonstrated by conducting a series of 3D full-wave simulations in which one changes the lateral dimension of the metasurface pattern; see Figs. 11b–e. A summary of the results, Fig. 11a, shows that lateral far-field divergence, $\theta_{||}$, is initially inversely proportional to the pattern width, W, and that the reduction in $\theta_{||}$ saturates for large W. The vertical far-field divergence keeps constant. Note that reduction of lateral beam divergence originates from a propagation effect. A delay in phase occurs as channel polaritons travel sideways. This results in a curved lateral wavefront and ultimately limits the achievable reduction in lateral divergence.

4. Deformed microcavity QCLs

4.1. Limaçon-shaped microcavity

Optical microcavities can be designed to take advantage of total internal reflection, which results in resonators supporting whispering-gallery modes with a high Q-factor. One of the crucial problems of these devices for practical applications such as designing microcavity lasers, however, is that their emission is nondirectional due to their radial symmetry, in addition to their inefficient power output coupling.

Recently, a limaçon-shaped microcavity has been proposed to achieve directional emission and a high Q-factor [120]. The limaçon profile is defined as $\rho(\theta) = \rho_0(1+\delta\cos\theta)$, where δ is the deformation factor and ρ_0 is the radius when $\theta = \pi/2$; see Fig. 12a. The limaçon geometry was first formally studied by Étienne Pascal in the 17th century. Limaçon is a diminutive of limace, which in French means slug, a mollusk that lacks a shell.

The performance of the limaçon-shaped microcavity laser can be explained by studying the phase diagram of the cavity, i. e., Poincaré surface of section; see Fig. 12b. In simple words, this diagram describes the distribution of rays inside and outside of the cavity. The x-axis of the diagram represents the normalized arc-length and the y-axis represents the sine of the incidence angle χ of a ray trajectory in the cavity (χ is taken to be positive in the second and fourth quadrants and negative in the first and third quadrants; see Fig. 12a). The simulations were carried out by launching a large number of rays along the cavity boundary with whispering gallery-like mode initial conditions ($0.8 < |\sin\chi| < 1$). The rays propagate following the laws of specular reflection and Snell's law of refraction, and each event that one ray intersects the resonator boundary is represented by a point in the phase diagram. Note that since the radius of curvature at any point along the resonator boundary is at least ten times larger than the wavelength in the material, the use of specular reflection and Snell's law is valid. Two red lines in the phase diagram mark the border between regions corresponding to total internal reflection and the "leakage" region, which means that after a certain amount of reflections the rays eventually escape from the cavity. In the leakage region, the ray intensity is weighted by the Fresnel coefficients and is shown in gray scale in Fig. 12b; it represents the so-called unstable manifold that determines the far-field emission pattern.

With the help of phase diagrams, it is predicted that for limaçon-shaped microcavities with deformation near 0.4 and material refractive index ∼ 3.2, high directionality and large Q-factor can be achieved simultaneously. In the phase diagram, the ray trajectories in the leakage region are concentrated in a few spots (Fig. 12b), implying that light can only escape from the cavity at certain locations along the boundary. The resulting far-field profile possesses pronounced emissions around $\theta = 0°$ and $\theta = \pm 135°$ directions; see Fig. 12c. The main emission lobes A, B, C, and D in Fig. 12c can be related to the escape regions A, B, C, and D, respectively, in Fig. 12b. We also carried out wave simulations by solving Maxwell equations numerically using the finite-element method; detailed information on the simulation technique can be found in [121].

In experiment, QCLs were used as a model system because of the absence of undesirable surface recombination (as in the case of diode lasers) and because of low scattering losses in the mid-infrared wavelength range. In addition, the intrinsic TM polarization of QCLs prohibits emissions normal to the plane of the laser cavities (far-field emission can be regarded as collective contributions from dipolar emitters distributed on the cavity sidewalls and aligned perpendicular to the plane of the laser cavities; these dipoles emit preferentially in the in-plane direction). We fabricated devices with different sizes $\rho_0 = 50$, 80, and 110 μm and different deformations with δ ranging from 0.20 to 0.80. Photolithography was used to define the contour of laser cavities, and dry etching was used to define the sidewalls of the cavities;

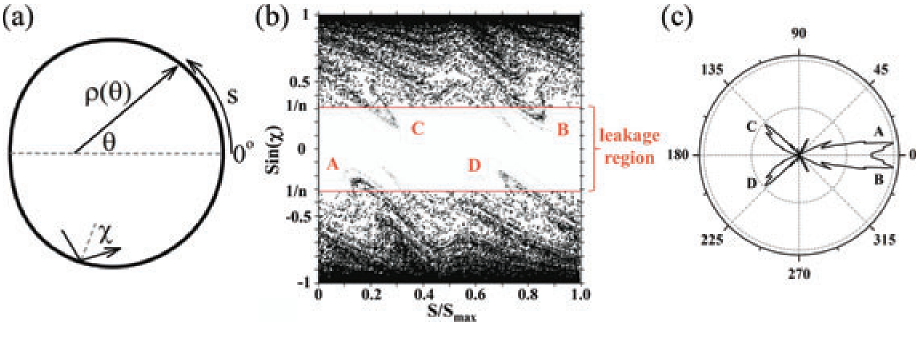


Figure 12 (online color at: www.lpr-journal.org) (a) Schematic of a limaçon microcavity with deformation $\delta = 0.4$. (b) Phase diagram (Poincaré surface of section) used to describe distribution of rays inside and outside of the cavity. TM-polarization (electric field perpendicular to the plane of the microcavity) is assumed. (c) Ray simulation of far-field intensity profile of the cavity in polar coordinates.

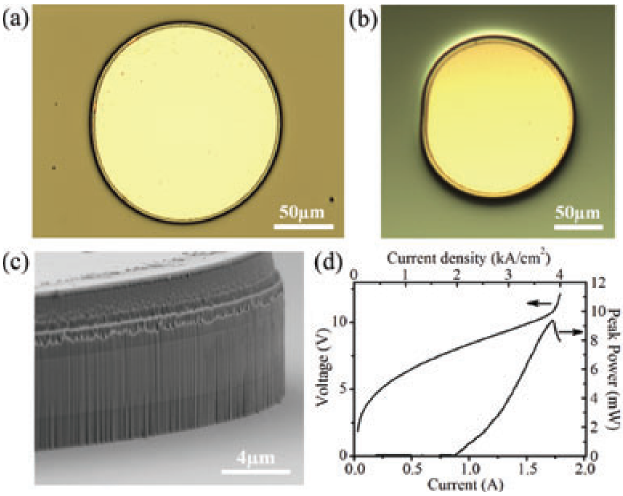


Figure 13 (online color at: www.lpr-journal.org) (a) Optical microscope image of a limaçon microcavity laser with deformation $\delta = 0.4$ and radius $\rho_0 = 80\,\mu$m. (b) Another device with $\delta = 0.55$ and radius $\rho_0 = 80\,\mu$m. The main lobe of emission for both devices points towards right. (c) SEM image of a device showing a smooth sidewall. (d) L–I–V characteristics of a limaçon microcavity laser with $\delta = 0.4$ and radius $\rho_0 = 110\,\mu$m measured at pulsed mode and at room temperature.

see Figs. 13a–c. Devices were electrically driven through top metallization Ti/Au (10 nm/200 nm). Figure 13d shows the light power and voltage versus current characteristics of a representative device with $\delta = 0.4$ and $\rho_0 = 110\,\mu$m measured at 300 K. A peak output power close to 10 mW was obtained; this represents the main lobe of the far-field along $\theta = 0$, which contains more than 50% of the total output power. The threshold current density of the microcavity laser, J_{th}, is about 2 kA cm^{-2}, smaller than that of ridge QCLs (waveguide length 2.5 mm, ridge width 14 µm) fabricated from the same wafer, 2.6 kA cm^{-2}. In the microcavity laser, only a small fraction of the pumped volume along the periphery of the cavity participates in the generation of photons, its output power is therefore lower than that of the ridge laser, which is about ~ 50 mW from a single facet. Lasers with smaller size usually have smaller threshold current densities. This is probably because the percentage of the pump current that passes through the central regions of these devices without generating photons is smaller compared to devices with larger size. For example, J_{th} is about 1.8 kA cm^{-2} for a microcavity laser with $\delta = 0.4$ and $\rho_0 = 80\,\mu$m. The device has a peak output power of about 4 mW.

The Q factor of microcavity lasers can be calculated using [122]

$$Q = 2\pi n_{\mathrm{eff}}/(\lambda_o J_{\mathrm{th}} g\Gamma) \tag{7}$$

where λ_o is the vacuum wavelength, J_{th} the threshold current density, g the gain coefficient, and Γ the laser mode confinement factor. $g\Gamma$ in the equation should be similar to that of ridge lasers and can be estimated from the dependence of the threshold current density on the waveguide length of ridge lasers:

$$J_{\mathrm{th,ridge}} = (\alpha_m + \alpha_{w,\mathrm{ridge}})/(g\Gamma) \tag{8}$$

where $\alpha_m = -\log_e R/L$ (R is the power reflectivity at laser facets) is mirror loss and α_w is waveguide loss. Similarly we have

$$J_{\mathrm{th,cavity}} = (\alpha_{\mathrm{out}} + \alpha_{w,\mathrm{cavity}})/(g\Gamma) \tag{9}$$

for microcavity lasers, where α_{out} is the out-coupling loss.

We calculated based on Eq. (8) that $g\Gamma = 8.0\times10^{-3}$ cmA^{-1}, and the Q-factor is estimated to be ~ 1200 for a microcavity laser with $\delta = 0.4$ using Eq. (7). It is larger than that of circular or deformed microcavity lasers emitting at similar wavelengths, Q $\sim$ a few hundred [122, 123]. We also calculated based on Eq. (8) that $\alpha_{w,\mathrm{ridge}} = 15.6\,\mathrm{cm}^{-1}$. The waveguide loss of microcavity lasers is probably larger, $\alpha_{w,cavity} \geqslant 15.6\,\mathrm{cm}^{-1}$, because of a stronger interaction between laser modes and sidewall roughness in microcavity lasers compared to that in conventional ridge waveguides; therefore, according to Eq. (9) we estimated that $\alpha_{\mathrm{out}} \leqslant 0.4\,\mathrm{cm}^{-1}$. Thus, the optical losses are primarily determined by waveguide losses associated with free-carrier absorption. The measured Q-factor in our devices is much smaller than the value obtained in simulations, because free-carrier absorption was not included in simulations.

Several cavity modes coexist and were observed in experiments. Figure 14a–f show the intensity distribution of three TM modes calculated for a structure with $\delta = 0.40$ and $\rho_0 = 80\,\mu$m. The modes in Figs. 14a and c are WG-like modes with the highest Q factor ($Q > 10^7$ theoretically), and can be excited at the lowest pump current. It is reasonable to assume that they correspond to the two sets of spectra observed in Fig. 14g labeled by blue and red arrows. Note that the average mode spacing of each set is approximately 6.0 cm^{-1}, which agrees well with the calculated value 6.2 cm^{-1}for WGMs, given by $1/Ln_{\mathrm{eff}}$, where L is the perimeter of the cavity and n_{eff} is the effective mode index. At a higher pump current, several additional unequally spaced modes appeared, indicated by gray arrows in Fig. 14g. They correspond to lower Q-factor non-WGMs of the type shown in Fig. 14e. We observed essentially the same spectra from all far-field lobes. Note that although intensity distributions within a laser cavity are markedly different for different modes, they all show similar far-field patterns (Fig. 14h). This is called "universal far-field behavior" also observed in other deformed microcavities [124]. The reason for this phenomenon is that the far-field pattern is primarily determined by the structure of the unstable manifolds in the leaky region, which are determined by the geometry of the deformed microcavity, regardless of the different spatial distributions of cavity modes.

4.2. Notched elliptical microcavity QCLs

Unlike the limaçon-shaped microcavities, the notched elliptical microcavity lasers support regular WGMs just as circular microcavities. Our design is based on the collimation property of ellipses [125], which states that if the eccentricity of the ellipse ε is equal to $n_{\mathrm{out}}/n_{\mathrm{in}}$, light rays emitted from a point source located on one of the foci of the

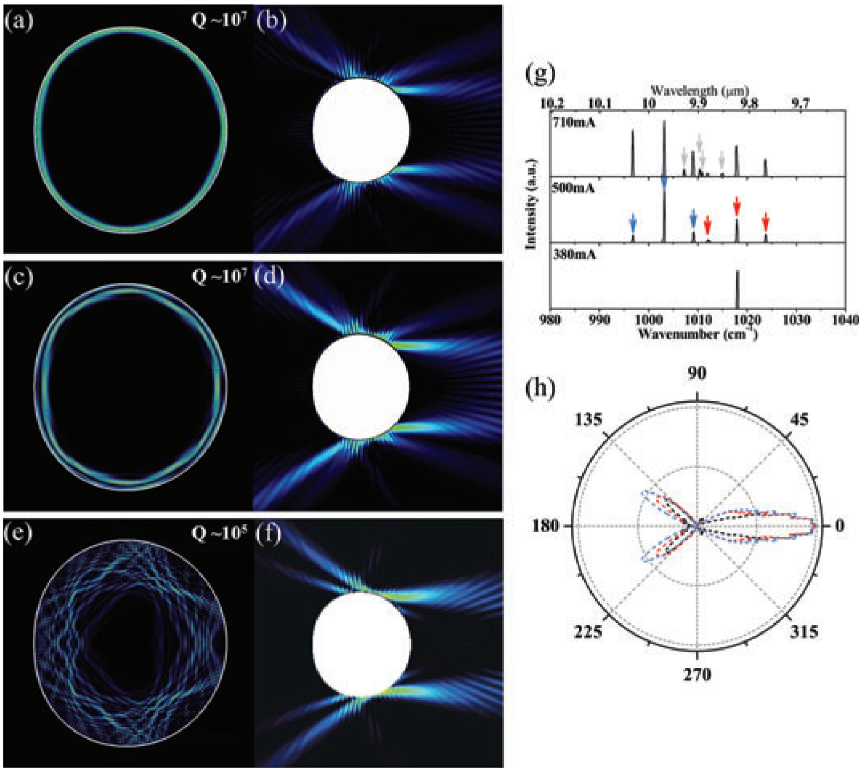


Figure 14 (online color at: www.lpr-journal.org) Wave simulation of intensity distribution of WG-like modes with the first radial order and the highest Q-factor (a), the second radial order with comparable Q-factor (c), and a non-WGM with much smaller Q-factor (e). The limaçon microcavity has deformation $\delta = 0.4$ and radius $\rho_0 = 80\,\mu\text{m}$. (b), (d), and (f) are intensity distributions of the external meso-field corresponding to (a), (c), and (e), respectively. (g) Emission spectra of a limaçon microcavity QCL with $\rho_0 = 80\,\mu\text{m}$ and $\delta = 0.4$ at different pump currents. The threshold current is around 380 mA. At a pump current of 500 mA, two sets of WG-like modes, indicated by blue and red arrows, occurred and they correspond to the first- and second-order radial modes illustrated in (a) and (c). At a higher pump current of 710 mA, several additional non-WGMs, like the one shown in (e), appeared. These modes are labelled by gray arrows. (h) The three modes in (a), (c), and (e) have similar far-field emission profiles. Red, blue, and black curves correspond to modes in (a), (c), and (e), respectively.

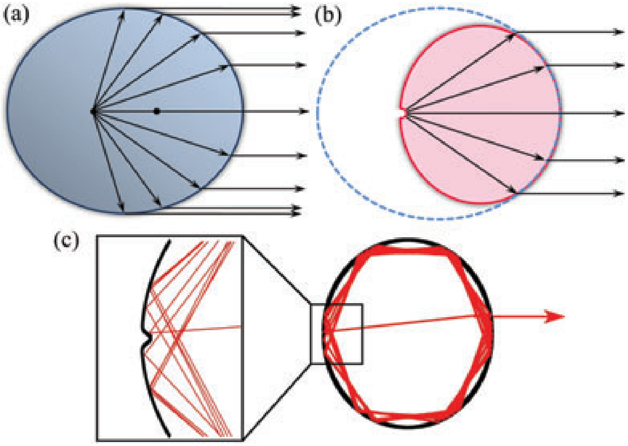


Figure 15 (online color at: www.lpr-journal.org) (a) Schematic demonstrating the collimation effect of an elliptical disk. The eccentricity of the ellipse should be equal to the ratio between the refractive index of the medium outside the ellipse and that of the medium inside. The two dots represent the foci of the ellipse and the arrows represent light rays. (b) Schematic showing the method to construct a notched elliptical resonator. The notch on the left boundary of the resonator coincides with the left focus of the virtual ellipse (dashed curve). The right boundaries of the two ellipses match over the largest possible range. (c) Evolution of a light ray inside the notched elliptical resonator based on ray simulations.

ellipse will be refracted by the opposite half-side boundary of the ellipse into parallel rays (Fig. 15a). Here n_{out} and n_{in} are, respectively, the refractive indices of the media outside and inside the ellipse.

Let the dashed curve in Fig. 15b represent a virtual ellipse that is exactly the same as the ellipse in Fig. 15a. We place a notched ellipse (pink in Fig. 15b) inside the virtual ellipse such that: first, the notch is located on the left focus of the virtual ellipse, and second, the right half-boundary of the notched ellipse approximates that of the virtual ellipse within the largest possible angular range. As a result, the WGMs circulating along the circumference of the notched ellipse will be scattered by the notch and the scattered light, represented by arrows in Fig. 15b, will be refracted to form a collimated beam propagating towards right.

Figure 15c shows ray simulations of WGM dynamics. A single ray is started at a certain position on the resonator boundary such that it experiences total internal reflection. It is then specularly reflected and circulates around the elliptical resonator many times, forming a WG-like mode, until it eventually hits the notch. The ray then gets scattered to the opposite side, refracted at the right boundary, and exits the cavity in a direction almost parallel to the horizontal axis. A magnification of the dynamics near the notch is shown in Fig. 15c.

An analytical expression can be derived for the optimal eccentricity of the elliptical cavity. As mentioned previously, the eccentricity of the virtual ellipse is determined by the effective mode index n_{eff}:

$$\varepsilon_{\text{virtual}} = \sqrt{a'^2 - b'^2}/a' = 1/n_{\text{eff}} \qquad (10)$$

where a' and b' are its semimajor and semiminor axes, respectively. For optimal match of boundaries, we require the two ellipses have the same radius of curvature at point P (see Fig. 16a):

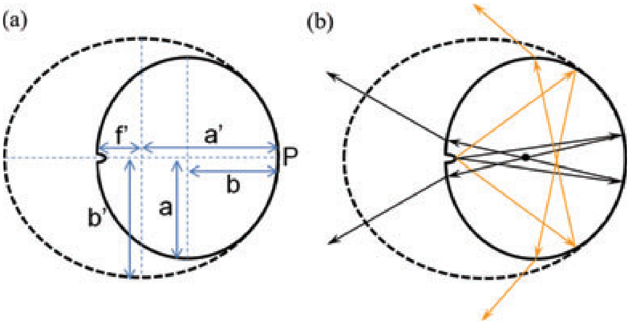


Figure 16 (online color at: www.lpr-journal.org) (a) Geometries used to calculate the optimal eccentricity of the notched elliptical resonator. The semimajor and semiminor axes of the notched resonator (black solid curve) and the virtual ellipse (black dashed curve) are, respectively, a, b and a', b'. The distance from one focus to the center in the virtual ellipse is f'. The two ellipses match best in their right half-side when the radius of curvature is the same at their rightmost boundaries. If the refractive index of the notched resonator is larger than two, the major axes of the two ellipses should be perpendicular to each other (as in (a)); whereas if the refractive index is smaller than two, their major axes should overlap. (b) Schematic showing that minor emission occurs when the scattered light from the notch is reflected and impinges on cavity boundaries with an incident angle smaller than the critical angle for total internal reflection. Emission from the backside of the resonator (black arrows) comes from scattering of WGMs at the notch with small angles with respect to the horizontal axis; emission from the upper and lower sides of the resonator (orange arrows) originates from large-angle scattering of WGMs at the notch.

$$b'^2/a' = a^2/b \tag{11}$$

where a and b are, respectively, the semimajor and semiminor axes of the elliptical resonator.

In addition, referring to Fig. 16a, we have

$$f' + a' = 2b \tag{12}$$

where f' is the distance between a focal point and the center in the virtual ellipse. Based on the above three equations, it can be readily shown that the ratio between semimajor and semiminor axes of the elliptical resonator is

$$a/b = \sqrt{2 - 2/n_{\mathrm{eff}}} \tag{13}$$

which corresponding to an eccentricity of

$$\varepsilon = \sqrt{(n_{\mathrm{eff}} - 2)/(2n_{\mathrm{eff}} - 2)}\,. \tag{14}$$

Therefore, the optimal ratio of a/b is ~ 1.2 for mid-IR QCLs with $n_{\mathrm{eff}} = 3.2$. Note that the above derivations are valid for $n_{\mathrm{eff}} \geqslant 2$. If $n_{\mathrm{eff}} < 2$, the major axis of the elliptical resonator should be aligned with that of the virtual ellipse (instead of perpendicular to it in Fig. 16a), and one can show in a similar way that

$$a/b = 1/\sqrt{2 - 2/n_{\mathrm{eff}}} \tag{15}$$

or

$$\varepsilon = \sqrt{2/n_{\mathrm{eff}} - 1}\,. \tag{16}$$

Figure 17a shows, in a linear scale, simulated intensity distribution of the lowest-radial-order WGM of a notched elliptical microcavity QCL with $\lambda_o = 10\,\mu\mathrm{m}$ and $n_{\mathrm{eff}} \sim 3.2$. The laser cavity is optimized to minimize far-field beam divergence. This mode has two maxima in the radial direction (Fig. 17a), corresponding to a radial order n=2, which is different from limaçon-shaped microcavity lasers discussed previously, where the lowest-radial-order WGMs have n=1 (Fig. 14a). Furthermore, only the outermost part of the mode in Fig. 17a interacts with the notch, resulting in a high Q-factor. Figure 17b is the same intensity distribution plotted in a logarithmic scale, showing clearly the scattered light by the notch inside the cavity.

Figure 17c is a plot of the light intensity distribution outside the laser cavity ina linear scale. The weak emissions from the vicinity of the notch and from the high curvature ends of the elliptical resonator (where the major axis intersects the ellipse) are due to reflected light from the right half-boundary of the ellipse; see Fig. 16b. Note that because of the high refractive index of the laser material, the critical angle for total internal reflection is small ($\sim 18°$); basically only scattered light waves by the notch or their reflection can

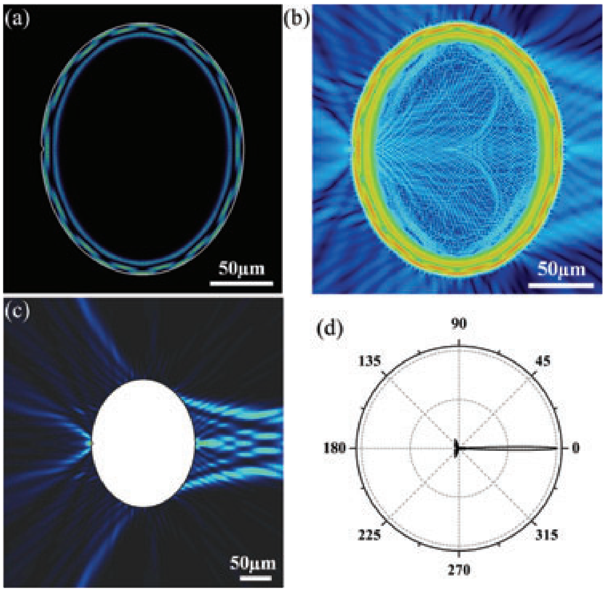


Figure 17 (online color at: www.lpr-journal.org) (a) and (b) Wave simulations of intensity distribution of the lowest-radial-order WGM in a $\lambda_o = 10\,\mu\mathrm{m}$ QCL with semiminor axis X = 80 μm and semimajor axis Y = 96 μm, in linear and logarithmic scales, respectively. The Q-factor of this mode is 590,000. The notch, located on the left boundary of the ellipse, has an opening of $o = 3\,\mu\mathrm{m}$ and depth of $d = 2\,\mu\mathrm{m}$. In the simulations, the equation used to describe the boundary of the notched elliptical resonator is $x = \left\{X - d\exp\left[-(\varphi - \pi)^2/\delta^2\right]\right\}\cos\varphi$ and $y = Y\sin\varphi$ in polar coordinates, where φ is the polar angle and the origin is at the center of the ellipse. δ determines the opening of the notch. The effective mode index is $n_{\mathrm{eff}} \sim 3.2$. (c) External in-plane intensity distribution of the mode shown in (a). The majority of the output power is collimated toward the right-hand-side direction. (d) Simulated far-field intensity distribution of the mode showing an in-plane FWHM divergence angle $\sim 5°$.

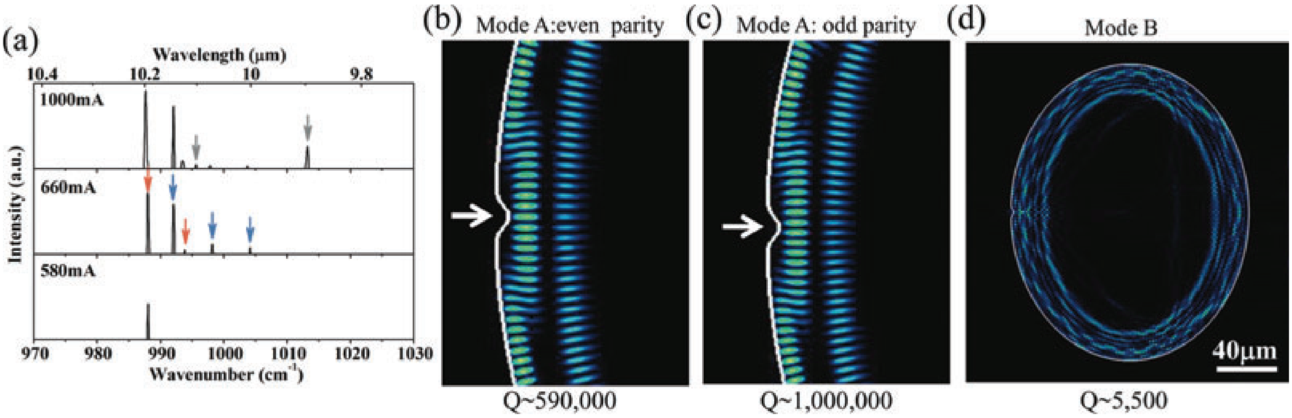


Figure 18 (online color at: www.lpr-journal.org) (a) Emission spectra of a a notched elliptical microcavity QCL at different pump currents. The device has semiminor axis 80 μm, semimajor axis 96 μm, notch opening 3 μm, and notch depth 2 μm. The threshold current is 520 mA. Two sets of WGMs, indicated by blue and red arrows, occur at a pump current of 660 mA. At a higher pump current of 1 A, several additional modes, indicated by gray arrows, appear and they correspond to lower Q-factor WGMs modes. (b) and (c) Details of the intensity distribution of the lowest-radial-order WGMs with even and odd parities, respectively. Note their difference indicated by the arrows. The two degenerate modes likely contribute to the two sets of WGMs in (a). (d) A higher-radial-order WGM of the cavity with a Q-factor of ~5,500.

escape the elliptical resonator. Figure 17d is the calculated far-field intensity profile. The majority of the output power is collimated along the horizontal direction toward the right-hand-side direction, leading to a unidirectional emission. In Fig. 17, the ratio of semiminor axis to the wavelength in the laser material is relatively high (~ 26); simulations show that the smallest notched elliptical resonators with directional emission for TM light have a ratio of ~ 3.

The lowest-radial-order WQM consists of two degenerate states with different parities: the mode with even (odd) parity has an antinode (a node) adjacent to the notch; see Figs. 18b and c. Without accounting for free-carrier absorption, the Q-factors for the modes with odd and even parities are estimated to be ~ 1,000,000 and ~ 590,000, respectively. Owing to unavoidable material absorption losses, both odd and even parity modes have similar Q-factors on the order of 1,000 and start to lase at similar pump currents (Fig. 18a). However, due to a weaker interaction with the notch, the odd parity mode has a smaller power out-coupling compared to the even parity mode. Figure 18c shows a higher-radial-order WGM with a lower Q-factor of 5,500. This mode produces a directional emission with divergence comparable to that of the lowest-radial-order WGMs.

Devices with different dimensions and notch sizes were fabricated and tested at room temperature. Figure 19a shows an SEM image of a representative device emitting at $\lambda_o = 10\,\mu\text{m}$ and the zoom-in view of the notch. Figure 19b presents L-I-V characteristics of the device measured in pulsed mode operation with a duty cycle of 1%. A peak output power of ~ 5 mW, a threshold current density ~ 2 kA/cm^2, and a slope efficiency of ~ 11 mW/A were obtained. The measured Q-factor of devices with optimal geometry is ~ 1,260, which was deduced from threshold current density and the gain coefficient measurements and is comparable to that of limaçon-shaped microcavity lasers. Figure 19c shows that the FWHM widths of the main lobe are ~ 6° and ~ 19° in the lateral and vertical directions, respectively, representing a reduction by a factor of 5 and 2, respectively, compared to those of limaçon-shaped microcavity lasers. The lateral divergence angle is approaching the diffraction limit, which is estimated to be ~ 4.1° (free-space wavelength, ~ 10 μm, divided by the effective emission aperture, ~ 140 μm, according to Fig. 17c). The emission of the notched laser is unidirectional (Fig. 19d): the suppression ratio in peak power between the main and side lobes in the far-field is ~ 10 dB. This compares favorably with other microcavity resonators, which typically have multiple emission lobes of comparable intensity or small suppression ratios. The laser cavity design is robust for different emission wavelengths. Figure 19e shows unidirectional emission from a device with a shorter wavelength $\lambda_o = 7.8\,\mu\text{m}$. The far-field profiles are essentially the same at different pump currents and are insensitive to the shape of the notch and variations of the notch sizes (from 2 μm to 4 μm, a deviation well within fabrication uncertainties). The divergence angle of notched elliptical QCLs seems to be diffraction limited by the dimensions of the laser cavities. Figure 20shows that the FWHM divergence angles of the main lobes are 8.3°, 6.7°, and 4.8°, respectively, for devices with semiminor axis 50, 80, and 110 μm. These devices have the same deformation ($Y/X = 1.2$), notch sizes, and emission wavelength.

Simulations demonstrate that the far-field of TE polarized modes (i. e., in-plane electric field) supported by notched elliptical resonators is also highly unidirectional. This indicates that the design is applicable to near-infrared and visible diode lasers based on interband electronic transitions, where free carrier absorption is negligible and optical losses are primarily due to scattering at cavity sidewall roughness and non-radiative recombination on the free surface of the devices [122]. The scheme of notched resonators is promising for realizing microcavity diode lasers with low pump current threshold, and highly unidirectional and efficient power out-coupling compared to conventional microcavity lasers.

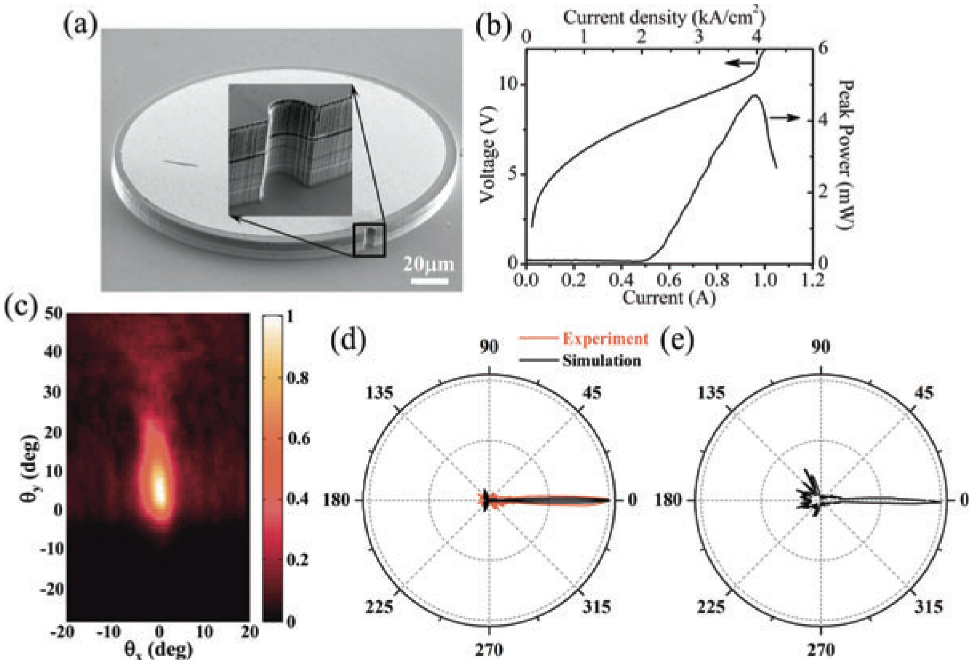


Figure 19 (online color at: www.lpr-journal.org) (a) SEM image of a notched elliptical microcavity laser with semiminor axis $X = 80$ μm, semimajor axis $Y = 96$ μm, notch dimensions $o = 3$ μm, and $d = 2$ μm, and emission wavelength $\lambda_o = 10$ μm. Inset: zoom-in view of the device showing the notch and the smooth sidewalls of the laser cavity. (b) Room temperature L-I-V characteristics of the device. (c) Measured 2D far-field emission profile of the device at a pump current of 720 mA. (d) Experimental and simulated in-plane far-field intensity profiles of the device in polar coordinates. The small discrepancy is most likely due to the surface roughness that adds optical background and broadens the main lobe of the emission. (e) Measured far-field intensity profile of a $\lambda_o = 7.8$ μm notched elliptical microcavity QCL with X = 80 μm, Y = 96 μm, $o = 3$ μm, and $d = 3$ μm.

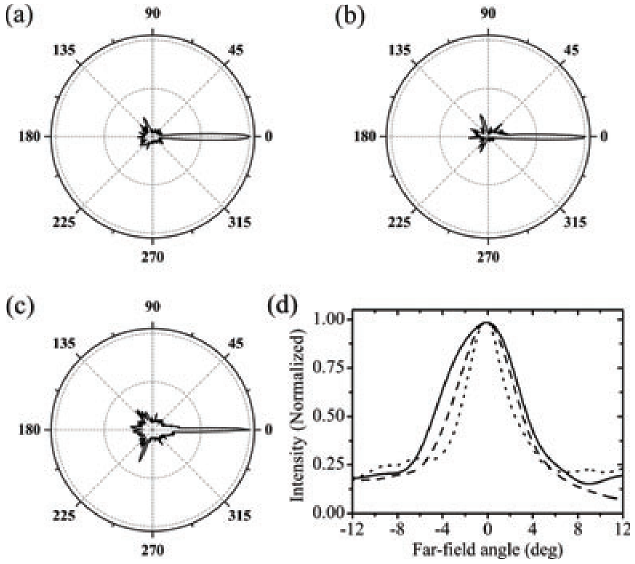


Figure 20 (a)–(c) Measured far-field emission profiles of three notched elliptical QCLs with optimal deformation (ratio between semimajor and semiminor axes equal to 1.2) but different sizes. Semiminor axis is 50, 80, and 110 μm for (a), (b), and (c), respectively. Laser wavelengths are all ∼ 10 μm. (d) Main lobes of the far-field. The solid, dashed, and dotted curves correspond to (a), (b), and (c), respectively.

5. Conclusions and perspectives

We have demonstrated QCLs with new or improved beam characteristics. This is achieved, in the first approach, by monolithically integrating properly designed plasmonic structures on the cleaved device substrates below the laser waveguide, and in the second approach, by processing devices into deformed microcavity lasers. The plasmonic structures create desirable far-field radiation by coupling laser emission into surface waves propagating on the substrates and by coherently scattering the energy of the waves into the free space. We have shown that plasmonic second-order gratings improve confinement of infrared surface waves and function as antenna arrays. Metasurface plasmonic structures provide more flexibility in tailoring the dispersion of surface waves, enabling us to engineer a two-dimensional space for controlling the propagation constant and confinement of surface waves. We have discussed two types of deformed microcavity lasers. The limaçon microcavities are intrinsically chaotic because the geometry is non-integrable. However, we limit the degree of deformation so that the microcavities support whispering-gallery like modes, which have a high quality factor and dominate over chaotic modes. Limaçon microcavity QCLs with optimal deformation produce a main lobe with FWHM divergence angle $\sim 30°$ and two smaller side lobes. Notched elliptical microcavity QCLs are superior in producing highly unidirectional emission with small in-plane divergence ($\sim 5°$). The laser resonators

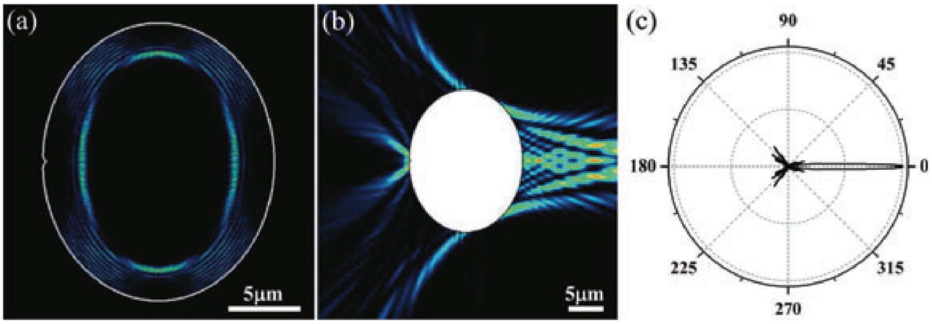


Figure 21 (online color at: www.lpr-journal.org) (a) Intensity distribution of a TE mode with the highest Q factor. The emission wavelength is 1 μm. The device has semiminor and senimajor axis of 8 μm, and 9.6 μm, respectively, and a notch with $o = 0.3$ μm, and $d = 0.2$ μm. The Q-factor of the mode is ∼270,000. (b) External in-plane intensity distribution of the mode. (c) Simulated far-field intensity profile of the TE mode showing a lateral far-field FWHM divergence angle of ∼4.5°.

intrinsically support whispering gallery modes. The energy of the modes is partially scattered by a wavelength-sized notch located on the boundary of the laser cavity and the scattered light is refracted by one half-side of the cavity boundary to form collimated emission.

In both approaches, we tried to achieve an optimal trade-off between low laser threshold and efficient and directional emission. To this end, we minimize modifications to the laser waveguides or cavities: the plasmonic structures are separated from the gain media; deformations and defects of the microcavities are carefully controlled. As a result, the temperature performance of our ridge-waveguide devices is preserved and the quality factors of the deformed microcavity devices are the highest among microcavity devices with comparable wavelengths.

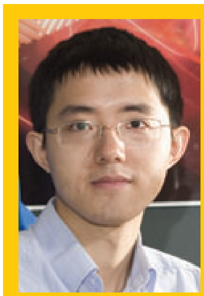
For the purpose of demonstration, the plasmonic structures are defined by focused-ion beam milling on the cleaved substrates of the lasers. Similar designs will work, according to our simulations, if the plasmonic patterns are in the plane of the laser resonators. Some recent experimental work has demonstrated in-plane generation and propagation of surface waves using QCLs [126, 127]. We envision inexpensive, high-throughput fabrication of in-plane plasmonic structures. For example, imprint lithography [128] seems to be suitable for the purpose; it creates patterns by mechanical deformation of imprint resist and subsequent processes. Nanoskiving [129] in combination with replica molding [130] has been developed to produce planar arrays of single- or multi-component (e. g., metals, semiconductors, insulators) structures with subwavelength features and high aspect ratio [131]. A soft-lithography based pattern transfer technique, "decal transfer" [132], has been demonstrated for convenient and reliable fabrication of plasmonic structures on non-conventional surfaces such as the facet of optical devices or curved surfaces.

In this paper we only discussed far-field engineering of infrared light sources, but the possibilities in the near-field are also numerous, such as optical antennas [133, 134] and active sensors [135]. The paper discussed controlling laser beam properties in the spatial domain; one can also engineer the emission spectra of QCLs by incorporating novel quasi-periodic distributed feedback structures [136] or deterministic aperiodic structures (DAPs) [137] into laser resonators. The spatial Fourier spectra of the DAPs can be engineered to control the multiple scattering and localization of light, which is promising for creating novel lasers with tailorable spectral and spatial emission characteristics.

Acknowledgements. We gratefully acknowledge the following people who contributed to the work presented in this paper: R. Blanchard, M. A. Kats, J. A. Fan, C. Pflügl, L. Diehl, M. Maiuri, M. Geiser, C. Yan, M. A. Belkin, S. P. Khanna, L. Li, A. G. Davies, E. H. Linfield, T. Edamura, S. Furuta, M. Yamanishi, H. Kan, M. Fischer, A. Wittmann, J. Faist, J. Unterhinninghofen, M. Hentschel, J. Wiersig, and N. Antoniou. We acknowledge funding from AFOSR under contract No. FA9550-09-0505-DOD, and FA9550-08-1-0047, the EPSRC (UK), and Hamamatsu Photonics (Japan). The authors acknowledge support from the National Science Foundation, Harvard Nanoscale Science and Engineering Center (NSEC), and the Center for Nanoscale Systems (CNS) at Harvard University. Harvard CNS is a member of the National Nanotechnology Infrastructure Network (NNIN). The computations in this Letter were run on the Odyssey cluster supported by the Harvard Faculty of Arts and Sciences (FAS) Sciences Division Research Computing Group.

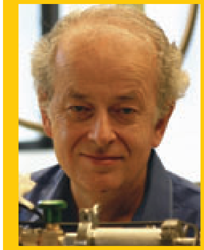
Received: 19 March 2011, **Revised:** 30 May 2011, **Accepted:** 23 June 2011

Key words: Mid-infrared and terahertz quantum cascade laser, microcavity laser, beam engineering, plasmonics, metasurface, collimation, whispering gallery modes, laser resonator.



Nanfang Yu received the B. S. degree in electronic information science and technology from the Department of Electronics, Peking University, Beijing, China, in 2004, and the Ph. D. degree in engineering sciences from Harvard University, Cambridge, MA, in June 2009. His Ph. D. research advisor is Prof. Federico Capasso. Currently, he is a researcher associate in the School of Engineering and Applied Sciences at Harvard University, Cambridge, MA. His research interest includes quantum cascade lasers, plasmonics, metamaterials, and infrared imaging and spectroscopy.

Qi Jie Wang received the B. E. degree in electrical engineering from the University of Science and Technology of China (USTC), Hefei, China, in 2001, and the Ph. D. degree in electrical and electronic engineering from Nanyang Technological University, Singapore, in 2005. He was with School of Engineering and Applied Science, Harvard University, MA, USA, as a postdoctoral researcher in Jan. 2007. He is now an assistant professor with School of Electrical and Electronic Engineering and School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore. His current research interests include mid-infrared and Terahertz quantum cascade lasers, nanophotonics, nanomaterials, and fiber optics and photonics.



Federico Capasso is the Robert L. Wallace Professor of Applied Physics at Harvard University, which he joined in 2003 after a 26 years career at Bell Labs where he rose from postdoc to Vice President for Physical Research. He holds a Doctor of Physics degree from the University of Rome, Italy, 1973. His research includes the design of new artificial materials and novel devices, plasmonics, nanophotonics, and quantum electrodynamics. He pioneered band-structure engineering as a technique to design new heterostructure materials and devices. He is co-inventor of the quantum cascade laser, a fundamentally new light source, which has now been commercialized. Recent achievements include the first measurement of repulsive Casimir force and wavefront engineering of lasers using plasmonics.

References

[1] C. Gmachl, F. Capasso, D. L. Sivco, and A. Y. Cho, Rep. Prog. Phys. **64**, 1533–1601(2001).

[2] A. E. Siegman, Proc. SPIE **1868**, 2–12 (1993).

[3] J. F. Nye and M. V. Berry, Proc. R. Soc. Lond. A. **336**, 165–190 (1974).

[4] M. Padgett, J. Courtial, and L. Allen, Phys. Today **57**, 35–40 (2004).

[5] C. G. B. Garrett, W. Kaiser, and W. L. Bond, Phys. Rev. **124**, 1807–1809 (1961).

[6] S.-X. Qian, J. B. Snow, H.-M. Tzeng, and R. K. Chang, Science **231**, 486–488 (1986).

[7] S. L. McCall, A. F. J. Levi, R. E. Slusher, S. J. Pearton, and R. A. Logan, Appl. Phys. Lett. **60**, 289–291 (1992).

[8] V. S. Ilchenko, A. A. Savchenkov, A. B. Matsko, and L. Maleki, Phys. Rev. Lett. **86**, 3168–3171 (2001).

[9] A. Chiasera, Y. Dumeige, P. Féron, M. Ferrari, Y. Jestin, G. Nunzi Conti, S. Pelli, S. Soria, and G. C. Righini, Laser Photonics Rev. **4**, 457–482 (2009).

[10] A. Yariv and P. Yeh, Photonics: Optical Electronics in Modern Communications, 6th ed. (Oxford University Press, Oxford, 2007), Chap. 12, pp. 586–587; Chap. 16, pp. 719–720.

[11] D. Hofstetter, J. Faist, M. Beck, and U. Oesterle, Appl. Phys. Lett. **75**, 3769–3771 (1999).

[12] W. Schrenk, N. Finger, S. Gianordoli, L. Hvozdara, G. Strasser, and E. Gornik, Appl. Phys. Lett. **77**, 2086–2088 (2000).

[13] D. Hofstetter, J. Faist, M. Beck, A. Müller, and U. Oesterle, Physica E **7**, 25–28 (2000).

[14] C. Pflügl, M. Austerer, W. Schrenk, S. Golka, G. Strasser, R. P. Green, L. R. Wilson, J. W. Cockburn, A. B. Krysa, and J. S. Roberts, Appl. Phys. Lett. **86**, 211101 (2005).

[15] J. A. Fan, M. A. Belkin, and F. Capasso, Opt. Express **14**, 11672–11680 (2006).

[16] S. Kumar, B. S. Williams, Q. Qin, A. W. M. Lee, and Q. Hu, Opt. Express **15**, 114–128 (2007).

[17] E. Mujagić, L. K. Hoffmann, S. Schartner, M. Nobile, W. Schrenk, M. P. Semtsiv, M. Wienold, W. T. Masselink, and G. Strasser, Appl. Phys. Lett. **93**, 161101 (2008).

[18] E. Mujagić, C. Schwarzer, W. Schrenk, Y. Yao, J. Chen, C. Gmachl, and G. Strasser, Opt. Eng. **49**, 111113 (2010).

[19] L. Mahler, A. Tredicucci, F. Beltram, C. Walther, J. Faist, B. Witzigmann, H. E. Beere, and D. A. Ritchie, Nature Photon. **3**, 46–49 (2009).

[20] I. Vurgaftman and J. R. Meyer, IEEE J. Quantum Electron. **38**, 592–602 (2002).

[21] R. Colombelli, K. Srinivasan, M. Troccoli, O. Painter, C. F. Gmachl, D. M. Tennant, A. M. Sergent, D. L. Sivco, A. Y. Cho, and F. Capasso, Science **302**, 1374–1377 (2003).

[22] Y. Chassagneux, R. Colombelli, W. Maineult, S. Barbieri, H. E. Beere, D. A. Ritchie, S. P. Khanna, E. H. Linfield, and A. G. Davies, Nature **457**, 174–178 (2009).

[23] Y. Chassagneux, R. Colombelli, W. Maineult, S. Barbieri, S. P. Khanna, E. H. Linfield, and A. G. Davies, Appl. Phys. Lett. **96**, 031104 (2010).

[24] Q.-Y. Lu, W.-H. Guo, W. Zhang, L.-J. Wang, J.-Q. Liu, L. Li, F.-Q. Liu, and Z.-G. Wang, Appl. Phys. Lett. **96**, 051112 (2010).

[25] M. I. Amanti, M. Fischer, G. Scalari, M. Beck, and J. Faist, Nature Photon. **3**, 586–590 (2009).

[26] J. U. Nöckel and A. D. Stone, Nature **385**, 45–47 (1997).

[27] C. Gmachl, F. Capasso, E. E. Narimanov, J. U. Nöckel, A. D. Stone, J. Faist, D. L. Sivco, and A. Y. Cho, Science **280**, 1556–1564 (1998).

[28] G. D. Chern, H. E. Tureci, A. D. Stone, R. K. Chang, M. Kneissl, and N. M. Johnson, Appl. Phys. Lett. **83**, 1710–1712 (2003).

[29] R. Audet, M. A. Belkin, J. A. Fan, B. G. Lee, K. Lin, F. Capasso, E. E. Narimanov, D. Bour, S. Corzine, J. Zhu, and G. Höfler, Appl. Phys. Lett. **91**, 131106 (2007).

[30] M. Hentschel, T.-Y. Kwon, M. A. Belkin, R. Audet, and F. Capasso, Opt. Express **17**, 10335–10343(2009).

[31] M. Hentschel, Q. J. Wang, C. Yan, F. Capasso, T. Edamura, and H. Kan, Opt. Express **18**, 16437–16442 (2010).

[32] J. U. Nöckel, A. D. Stone, and R. K. Chang, Opt. Lett. **19**, 1693–1695 (1994).

[33] R. G. Kouyoumjian, Proc. IEEE **53**, 864–876 (1965).

[34] D. A. Hill, Electromagnetic Fields in Cavities: Deterministic and Statistical Theories (John Wiley & Sons, Inc., New Jersey, 2009), pp. 243.

[35] M. Born and E. Wolf, Principles of Optics, 7th ed. (Cambridge University Press, Cambridge, UK, 1999), Chap. 3, pp. 119.
[36] V. A. Borovikov and B. Y. Kinber, Geometrical Theory of Diffraction (The Institution of Electrical Engineers, London, UK, 1994), Chap. 3, pp. 30.
[37] P. W. Anderson, Phys. Rev. **109**, 1492–1505 (1958).
[38] D. S. Wiersma, P. Bartolini, A. Lagendijk, and R. Righini, Nature **390**, 671–673 (1997).
[39] H. Cao, Lasing in Disordered Media, in: Progress in Optics, Vol. 45, edited by E. Wolf (Elsevier, Amsterdam, 2003), pp. 317–370.
[40] N. Yu, J. Fan, Q. J. Wang, C. Pflügl, L. Diehl, T. Edamura, M. Yamanishi, H. Kan, and F. Capasso, Nature Photon. **2**, 564–570 (2008).
[41] N. Yu, R. Blanchard, J. Fan, T. Edamura, M. Yamanishi, H. Kan, and F. Capasso, Appl. Phys. Lett. **93**, 181101 (2008).
[42] N. Yu, R. Blanchard, J. Fan, Q. J. Wang, C. Pflügl, L. Diehl, T. Edamura, M. Yamanishi, H. Kan, and F. Capasso, Opt. Express **16**, 19447–19461 (2008).
[43] N. Yu, Q. J. Wang, M. A. Kats, J. A. Fan, S. P. Khanna, L. Li, A. G. Davies, E. H. Linfield, and F. Capasso, Nature Mater. **9**, 730–735 (2010).
[44] N. Yu, Q. J. Wang, C. Pflügl, L. Diehl, F. Capasso, T. Edamura, S. Furuta, M. Yamanishi, and H. Kan, Appl. Phys. Lett. **94**, 151101 (2009).
[45] N. Yu, M. A. Kats, C. Pflügl, M. Geiser, Q. J. Wang, M. A. Belkin, F. Capasso, M. Fischer, A. Wittmann, J. Faist, T. Edamura, S. Furuta, M. Yamanishi, and H. Kan, Appl. Phys. Lett. **95**, 161108 (2009).
[46] C. Yan, Q. J. Wang, L. Diehl, M. Hentschel, J. Wiersig, N. Yu, C. Pflügl, F. Capasso, M. A. Belkin, T. Edamura, M. Yamanishi, and H. Kan, Appl. Phys. Lett. **94**, 251101 (2009).
[47] Q. J. Wang, C. Yan, L. Diehl, M. Hentschel, J. Wiersig, N. Yu, C. Pflügl, M. A. Belkin, T. Edamura, M. Yamanishi, H. Kan, and F. Capasso, New J. Phys. **11**, 125018 (2009).
[48] Q. J. Wang, C. Yan, N. Yu, J. Unterhinninghofen, J. Wiersig, C. Pflügl, L. Diehla, T. Edamura, M. Yamanishi, H. Kan, and F. Capasso, Proc. Natl. Acad. Sci. USA **107**, 22407–22412 (2010).
[49] J. Faist, F. Capasso, D. L. Sivco, A. L. Hutchinson, and A. Y. Cho, Science **264**, 553–556 (1994).
[50] C. Sirtori, F. Capasso, J. Faist, A. L. Hutchinson, D. L. Sivco, and A. Y. Cho, IEEE J. Quantum Electron. **34**, 1722–1729 (1998).
[51] M. Razeghi, IEEE J. Sel. Top. Quantum Electron. **15**, 941–951 (2009).
[52] M. Razeghi, Proc. SPIE **7230**, 723011 (2009).
[53] A. Lyakh, R. Maulini, A. Tsekoun, R. Go, C. Pflügl, L. Diehl, Q. J. Wang, F. Capasso, and C. K. N. Patel, Appl. Phys. Lett. **95**, 141113 (2009).
[54] R. Maulini, A. Lyakh, A. Tsekoun, R. Go, C. Pflügl, L. Diehl, F. Capasso, and C. K. N. Patel, Appl. Phys. Lett. **95**, 151112 (2009).
[55] D. Hofstetter, M. Beck, T. Aellen, and J. Faist, Appl. Phys. Lett. **78**, 396–398 (2001).
[56] R. Köhler, A. Tredicucci, F. Beltram, H. E. Beere, E. H. Linfield, A. G. Davies, D. A. Ritchie, R. C. Iotti, and F. Rossi, Nature **417**, 156–159 (2002).
[57] S. Kumar, Q. Hu, and J. L. Reno, Appl. Phys. Lett. **94**, 131105 (2009).
[58] G. Scamarcio, M. Troccoli, F. Capasso, A. L. Hutchinson, D. L. Sivco, and A. Y. Cho, Electron. Lett. **37**, 295–296 (2001).
[59] M. Yamanishi, K. Fujita, T. Edamura, and H. Kan, Opt. Express **16**, 20748–20758 (2008).
[60] S. Kumar, C. W. I. Chan, Q. Hu, and J. L. Reno, Nature Phys. **7**, 166–171 (2011).
[61] H. A. Atwater, Sci. Am. **296**, iss. 4, 56–63 (2007).
[62] E. Ozbay, Science **311**, 189–193 (2006).
[63] S. A. Maier, Plasmonics: Fundamentals and Applications (Springer Science+Business Media LLC, New York, 2007), Chap. 5, pp. 65–88.
[64] D. J. Barber and I. C. Freestone, Archeometry **32**, 33–45 (1990).
[65] G. Mie, Ann. Phys. (Leipzig) **330**, 377–445 (1908).
[66] P. Bharadwaj, B. Deutsch, and L. Novotny, Adv. Opt. Photon. **1**, 438–483 (2009).
[67] L. Novotny and N. van Hulst, Nature Photon. **5**, 83–90 (2011).
[68] E. A. Ash and G. Nicholls, Nature **237**, 510–512 (1972).
[69] E. Betzig, J. K. Trautman, T. D. Harris, J. S. Weiner, and R. L. Kostelak, Science **251**, 1468–1470 (1991).
[70] F. Zenhausern, M. P. O'Boyle, and H. K. Wickramasinghe, Appl. Phys. Lett. **65**, 1623–1625 (1994).
[71] S. M. Nie and S. R. Emery, Science **275**, 1102–1106 (1997).
[72] K. Kneipp, Y. Wang, H. Kneipp, L. T. Perelman, I. Itzkan, R. R. Dasari, and M. S. Feld, Phys. Rev. Lett. **78**, 1667–1670 (1997).
[73] B. Knoll and F. Keilmann, Nature **399**, 134–137 (1999).
[74] B. Rothenhäusler and W. Knoll, Nature **332**, 615–617 (1988).
[75] J. R. Krenn, A. Dereux, J. C. Weeber, E. Bourillot, Y. Lacroute, J. P. Goudonnet, G. Schider, W. Gotschy, A. Leitner, F. R. Aussenegg, and C. Girard, Phys. Rev. Lett. **82**, 2590–2593 (1999).
[76] M. Quinten, A. Leitner, J. R. Krenn, and F. R. Aussenegg, Opt. Lett. **23**, 1331–1333 (1998).
[77] T. W. Ebbesen, H. J. Lezec, H. Ghaemi, T. Thio, and P. A. Wolf, Nature **391**, 667–669 (1998).
[78] H. J. Lezec, A. Degiron, E. Devaux, R. A. Linke, L. Martín-Moreno, F. J. García-Vidal, and T. W. Ebbesen, Science **297**, 820–822 (2002).
[79] J. B. Pendry, L. Martín-Moreno, and F. J. García-Vidal, Science **305**, 847–848 (2004).
[80] F. J. García-Vidal, L. Martín-Moreno, and J. B. Pendry, J. Opt. A: Pure Appl. Opt. **7**, S97-S101 (2005).
[81] J. B. Pendry, A. J. Holden, W. J. Stewart, and I. Youngs, Phys. Rev. Lett. **76**, 4773–4776 (1996).
[82] D. R. Smith, W. J. Padilla, D. C. Vier, S. C. Nemat-Nasser, and S. Schultz, Phys. Rev. Lett. **84**, 4184–4187 (2000).
[83] J. B. Pendry, Phys. Rev. Lett. **85**, 3966–3969 (2000).
[84] R. A. Shelby, D. R. Smith, and S. Schultz, Science **292**, 77–79 (2001).
[85] H. R. Stuart and D. G. Hall, Appl. Phys. Lett. **73**, 3815–3817 (1998).
[86] K. R. Catchpole and A. Polman, Opt. Express **16**, 21793–12800 (2008).
[87] A. Partovi, D. Peale, M. Wuttig, C. A. Murray, G. Zydzik, L. Hopkins, K. Baldwin, W. S. Hobson, J. Wynn, J. Lopata, L. Dhar, R. Chichester, and J. H.-J. Yeh, Appl. Phys. Lett. **75**, 1515–1517 (1999).
[88] X. L. Shi, L. Hesselink, and R. L. Thornton, Opt. Lett. **28**, 1320–1322 (2003).

[89] N. Yu, R. Blanchard, J. Fan, Q. Wang, C. Pflügl, L. Diehl, T. Edamura, S. Furuta, M. Yamanishi, H. Kan, and F. Capasso, IEEE Trans. Nanotechnol. **9**, 11–29 (2010).
[90] J. Zenneck, Ann. Phys. **23**, 846–866 (1907).
[91] J. P. Tetienne, R. Blanchard, N. Yu, P. Genevet, M. A. Kats, J. A. Fan, T. Edamura, S. Furuta, M. Yamanishi, H. Kan, and F. Capasso, New J. Phys. **13**, 053057 (2001).
[92] P. Lalanne and J. P. Hugonin, Nature Phys. **2**, 551–556 (2006).
[93] S. C. Kitson, W. L. Barnes, and J. R. Sambles, Phys. Rev. Lett. **77**, 2670–2673 (1996).
[94] R. Zia and M. L. Brongersma, Nature Nanotech. **2**, 426–429 (2007).
[95] J. A. Kong, Electromagnetic Wave Theory (EMW Publishing, Cambridge, USA, 2000), Chap. 4, pp. 510.
[96] http://www.lumerical.com/.
[97] H. C. Liu and F. Capasso, Intersubband Transitions in Quantum Wells: Physics and Device Applications I (Academic Press, San Diego, 2000), Chap. 1, pp. 8.
[98] M. A. Ordal, L. L. Long, R. J. Bell, S. E. Bell, R. R. Bell, R. W. Alexander, Jr., and C. A. Ward, Appl. Opt. **22**, 1099–1120 (1983).
[99] P. R. West, S. Ishii, G. V. Naik, N. K. Emani, V. M. Shalaev, and A. Boltasseva, Laser Photonics Rev. **4**, 795–808 (2010).
[100] I. Hamberg and C. G. Granqvist, J. Appl. Phys. **60**, R123-R159 (1986).
[101] J. Gómez Rivas, M. Kuttge, H. Kurz, P. Haring Bolivar, and J. A. Sánchez-Gil, Appl. Phys. Lett. **88**, 082106 (2006).
[102] X. Wang, A. A. Belyanin, S. A. Crooker, D. M. Mittleman, and J. Kono, Nature Phys. **6**, 126–130 (2010).
[103] N. Marschall, B. Fischer, and H. J. Queisser, Phys. Rev. Lett. **27**, 95–97 (1971).
[104] G. Borstel and H. J. Falge, Surface Phonon-Polaritons, in: Electromagnetic Surface Modes, edited by A. D. Boardman (John Wiley & Sons, Chichester, 1982), Chap. 6, pp. 219–248.
[105] V. V. Bryxin, D. N. Mirlin, and I. I. Reshina, Solid State Commun. **11**, 695–699 (1972).
[106] R. F. Wallis and J. J. Brion, Solid State Commun. **9**, 2099–2103 (1971).
[107] A. Mooradian and G. B. Wright, Phys. Rev. Lett. **16**, 999–1001 (1966)
[108] P. Perlin, J. Camassel, W. Knap, T. Taliercio, J. C. Chervin, T. Suski, I. Grzegory, and S. Porowski, Appl. Phys. Lett. **67**, 2524–2526 (1995).
[109] D. J. Lockwood, G. Yu, and N. L. Rowell, Solid State Commun. **136**, 404–409 (2005).
[110] G. Goubau, J. Appl. Phys. **21**, 1119–1128 (1950).
[111] W. Rotman, Proc. IRE **39**, 952–959 (1951).
[112] F. J. Zucker, Surface-Wave Antennas and Surface-Wave-Excited Arrays, in: Antenna Engineering Handbook, 2nd ed., edited by R. C. Johnson and H. Jasik (McGraw-Hill Book Company, New York, 1961), Chap. 12, pp. 12-7–8.
[113] F. J. Zucker, Surface-Wave Antennas, in: Antenna Theory, part 2, edited by R. E. Collin and F. J. Zucker (McGraw-Hill Book Company, New York, 1969), Chap. 21, pp. 331–338.
[114] I. V. Novikov and A. A. Maradudin, Phys. Rev. B **66**, 035403 (2002).
[115] D. K. Gramotnev and D. F. P. Pile, Appl. Phys. Lett. **85**, 6323–6325 (2004).
[116] S. I. Bozhevolnyi, V. S. Volkov, E. Devaux, J.-Y. Laluet, and T. W. Ebbesen, Nature **440**, 508–511 (2006).
[117] S. I. Bozhevolnyi, V. S. Volkov, E. Devaux, and T. W. Ebbesen, Phys. Rev. Lett. **95**, 046802 (2005).
[118] S. I. Bozhevolnyi, Opt. Express **14**, 9467–9476 (2006).
[119] S. I. Bozhevolnyi and J. Jung, Opt. Express **16**, 2676–2684 (2008).
[120] J. Wiersig and M. Hentschel, Phys. Rev. Lett. **100**, 033901 (2008).
[121] J. Wiersig, J. Opt. A, Pure Appl. Opt. **5**, 53–60 (2003).
[122] C. Gmachl, J. Faist, F. Capasso, C. Sirtori, D. L. Sivco, and A. Y. Cho, IEEE J. Quantum Electron. **33** 1567–73 (1997).
[123] S. Gianordoli, L. Hvozdara, G. Strasser, W. Schrenk, J. Faist, and E. Gornik, IEEE J. Quantum Electron. **36** 458–64 (2000).
[124] S.-B. Lee, J. Yang, S. Moon, J.-H. Lee, K. An, J.-B. Shim, H.-W. Lee, and S. W. Kim, Phys. Rev. A **75**, 011802(R) (2007).
[125] R. K. Luneburg, Mathematical Theory of Optics (University of California Press, Berkeley, 1964), Chap. 3, p. 132.
[126] A. Babuty, A. Bousseksou, J.-P. Tetienne, I. Moldovan-Doyen, C. Sirtori, G. Beaudoin, I. Sagnes, Y. De Wilde, and R. Colombelli, Phys. Rev. Lett. **104**, 226806 (2010).
[127] J.-P. Tetienne, A. Bousseksou, D. Costantini, R. Colombelli, A. Babuty, I. Moldovan-Doyen, Y. De Wilde, C. Sirtori, G. Beaudoin, L. Largeau, O. Mauguin, and I. Sagnes, Appl. Phys. Lett. **97**, 211110 (2010).
[128] S. Y. Chou, P. R. Krauss, and P. J. Renstrom, Science **272**, 85–87 (1996).
[129] Q. Xu, R. M. Rioux, M. D. Dickey, and G. M. Whitesides, Nano Lett. **41**, 1566 (2008).
[130] Y. Xia and G. M. Whitesides, Angew. Chem. Int. Ed. **37**, 550–575 (1998).
[131] D. J. Lipomi, M. A. Kats, P. Kim, S. H. Kang, J. Aizenberg, F. Capasso, and G. M. Whitesides, ACS Nano **4**, 4017–4026 (2010).
[132] E. J. Smythe, M. D. Dickey, G. M. Whitesides, and F. Capasso, ACS Nano **3**, 59–65 (2009).
[133] E. Cubukcu, N. Yu, E. J. Smythe, L. Diehl, K. B. Crozier, and F. Capasso, IEEE J. Sel. Top. Quantum Electron. **14**, 1448–1461 (2008).
[134] N. Yu, E. Cubukcu, L. Diehl, D. Bour, S. Corzine, J. Zhu, G. Höfler, K. B. Crozier, and F. Capasso, Opt. Express **15**, 13272–13281 (2007).
[135] M. A. Belkin, M. Lončar, B. G. Lee, C. Pflügl, R. Audet, L. Diehl, F. Capasso, D. Bour, S. Corzine, and G. Höfler, Opt. Express **15**, 11262–11271 (2007).
[136] L. Mahler, A. Tredicucci, F. Beltram, C. Walther, J. Faist, H. E. Beere, D. A. Ritchie, and D. S. Wiersma, Nature Photon. **4**, 165–169 (2010).
[137] J.-K. Yang, S. V. Boriskina, H. Noh, M. J. Rooks, G. S. Solomon, L. Dal Negro, and H. Cao, Appl. Phys. Lett. **97**, 223101 (2010).