Highly-efficient THz generation using nonlinear plasmonic metasurfaces

This content has been downloaded from IOPscience. Please scroll down to see the full text.
2017 J. Opt. 19 104001
(http://iopscience.iop.org/2040-8986/19/10/104001)

View the table of contents for this issue, or go to the journal homepage for more

Download details:

IP Address: 128.62.209.130
This content was downloaded on 04/09/2017 at 23:10

Please note that terms and conditions apply.

You may also be interested in:

Nonlinear silicon photonics
M Borghi, C Castellan, S Signorini et al.

New frontiers in quantum cascade lasers: high performance room temperature terahertz sources
Mikhail A Belkin and Federico Capasso

Design criteria for ultrafast optical parametric amplifiers
C Manzoni and G Cerullo

Giant spectral transformations in plasmon enhanced difference-frequency generation with polychromatic light
Franklin Che, Sergey A Ponomarenko and Michael Cada

Roadmap on optical metamaterials
Augustine M Urbas, Zubin Jacob, Luca Dal Negro et al.

Intense terahertz radiation and their applications
H A Hafez, X Chai, A Ibrahim et al.

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

Enhanced third harmonic generation with graphene metasurfaces
Boyuan Jin, Tianjing Guo and Christos Argyropoulos

A review of metasurfaces: physics and applications
Hou-Tong Chen, Antoinette J Taylor and Nanfang Yu
Highly-efficient THz generation using nonlinear plasmonic metasurfaces

M Tymchenko\textsuperscript{1}, J S Gomez-Diaz\textsuperscript{2}, J Lee\textsuperscript{3}, M A Belkin\textsuperscript{1} and A Alù\textsuperscript{1}

\textsuperscript{1} Dept. of Electrical & Computer Engineering, The University of Texas at Austin, Austin, TX 78701, United States of America
\textsuperscript{2} Dept. of Electrical & Computer Engineering, University of California, Davis, CA 95616, United States of America
\textsuperscript{3} School of Electrical & Computer Engineering, Ulsan National Institute for Science and Technology, Ulsan 687-798, Republic of Korea

E-mail: alu@mail.utexas.edu

Received 21 April 2017, revised 21 June 2017
Accepted for publication 28 July 2017
Published 29 August 2017

Abstract
Nonlinear metasurfaces loaded with multi-quantum-well (MQW) heterostructures constitute a rapidly progressing class of optical devices that combine high nonlinear generation efficiency with an ultrathin profile. Here, we introduce and discuss terahertz (THz) difference-frequency generation (DFG) using MQW-based plasmonic metasurfaces and present a comprehensive theory for their rigorous electromagnetic analysis. We explicitly take into account complex phenomena associated with the local intensity saturation of intersubband transitions and identify fundamental upper-bounds for DFG conversion efficiency. Using this framework, we design and analyze a nonlinear DFG metasurface providing giant DFG nonlinear response and conversion efficiency up to 0.01\% at 5.8 THz. Such metasurface can be used to generate 0.15 mW of THz power using pump intensities in the kW cm\textsuperscript{-2} range. We envision that such DFG metasurfaces can become a platform for uncooled, compact, and highly-efficient continuous-wave THz sources.

Keywords: nonlinear optics, metasurfaces, difference frequency generation, THz sources

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

1. Introduction

The terahertz (THz) frequency band, ranging from roughly 0.3 to 10 THz, offers unique opportunities for numerous applications, such as non-invasive imaging systems for inspection and screening, wireless communications with up to terabit-per-second data rates, spectroscopy, chemical and biomedical sensing, radio astronomy, and many others \cite{1–6}. These possibilities are not yet being fully exploited due to the lack of compact and efficient THz sources able to operate at room temperature. The power generated by vacuum electronic devices, such as traveling-wave tubes and backwards-wave oscillators \cite{7}, solid-state electronic devices, such as transistors, Schottky diode multipliers and Gunn oscillators \cite{2}, rolls off with frequency owing to transit-time and resistance-capacitance effects. As a result, even with the best devices, the available power generated above 1 THz is well below the milliwatt level. THz gas lasers and free-electron lasers are extremely expensive, bulky, and complex to build. Other common approaches to obtain THz beams are based on difference-frequency generation (DFG) and parametric oscillations in nonlinear crystals, such as LiNbO\textsubscript{3}, ZnTe, ZnSe, and quartz \cite{8–10}. In these setups, up to 1 W of power has been obtained in the range from 0.7 to 3 THz. Unfortunately, such setups are not only large and bulky, but also require high-intensity pump beams. Much more compact yet efficient THz radiation sources, such as p-Ge and quantum cascade lasers (QCLs), have a major drawback that they operate only at cryogenic temperatures \cite{2, 9, 11}. Typical output power of state-of-the-art THz QCLs ranges from tens of milliwatts in CW operation regime to a few watts in pulsed mode \cite{12, 13}. Recently, another promising approach based on DFG inside mid-infrared QCLs has been reported, with broad tunability and up to few milliwatts of peak THz power at room...
temperature [14–17]. Currently, the most widely employed methods to generate THz beams are based on ultrashort optical pulses, such as photoconductive antennas and optical rectification [9], delivering dozens or even hundreds of microwatts of radiated power. These techniques have been incorporated in many commercial CW and time-domain THz spectroscopy systems. However, the latter only achieve high performance for frequencies below approximately 3 THz. Despite all the innovative efforts and significant improvements in this research field over the past few years, the active quest for efficient, compact, and uncooled THz sources continues.

In a related context, a novel class of nonlinear ultrathin metasurfaces consisting of periodically arranged plasmonic resonators placed on top of multi-quantum-well (MQW) substrates has recently been introduced [18–22]. MQWs are engineered artificial media consisting of semiconductor layers grown one over the other to form a stack. MQWs provide a direct way to engineering quantum levels (subbands) for electrons or holes moving transversely to the layers, control the transition rates, their population, and doping. Alongside numerous fascinating advances achieved with MQWs, they have been shown to possess an intrinsic second-order nonlinear susceptibility associated with electronic intersubband transitions that is several orders of magnitude larger than any other condensed matter system [23, 24]. This fact makes them one of the most attractive nonlinear materials. However, the electronic intersubband nonlinearity can be excited only by electric fields polarized perpendicular to the MQW layers [23, 25, 26], which makes it impossible to harness the giant nonlinear response from MQWs if the incident pump beam (or beams) is impinging orthogonally to semiconductor layers. Nevertheless, this issue has been successfully overcome in ultrathin metasurfaces by placing subwavelength plasmonic resonators specifically designed to simultaneously exhibit resonances at pump and nonlinearly generated frequencies of interest [18, 20–22]. In such MQW-loaded metasurfaces, the suitably designed resonators play multiple important roles. Firstly, they convert a transversely polarized pump wave(s) to local electric field polarized orthogonally to MQW layers. Secondly, plasmonic resonances provide high local field intensities, usually higher than the incident wave intensity. Finally, they efficiently outcouple the nonlinear signal to free-space. Ultrathin plasmonic metasurfaces loaded with only 400 nm thick MQWs have been shown to provide the record-breaking effective second-order nonlinear susceptibility of $1.2 \times 10^6 \text{pm V}^{-1}$ [11] for mid-infrared second harmonic generation (SHG), which to our knowledge is the largest nonlinear susceptibility ever reported in a condensed matter system for a three-wave frequency-mixing process in this wavelength range, and which has led to a conversion efficiency of 0.075% at pump intensities of only 15 kW cm$^{-2}$ [21]. For comparison, a conventional nonlinear medium like LiNbO$_3$ in a similar ultrathin geometry and similar pumping intensity would produce about 8 orders of magnitude weaker SHG signal. Furthermore, the flat geometry and robust field control provided by plasmonic resonators permits easy manipulation of the nonlinear field and enables advanced functionalities, such as steering and focusing of nonlinear beams, or generating vortex beams with polarization-dependent angular momentum, which were previously unavailable in nonlinear optics [27–32].

To date, only SHG nonlinear processes have been investigated for this class of plasmonic metasurfaces. Here, we propose, design and rigorously analyze a novel MQW-loaded ultrathin nonlinear metasurface aimed at highly-efficient DFG in continuous-wave regime at THz frequencies. DFG is a well-known second-order nonlinear process occurring when two pump beams with close frequencies $\omega_1$ and $\omega_2$ impinge on a $\chi^{(2)}$ nonlinear medium and generate nonlinear polarization currents radiating at the difference-frequency $\omega_3 = \omega_1 - \omega_2$ [8]. Building upon our previous theoretical and experimental efforts on SHG plasmonic metasurfaces [18, 22, 33], in this paper we develop a unique theoretical framework enabling us to analyze and design DFG ultrathin structures, taking into consideration complex phenomena such as saturation and losses. We also identify the upper bound for DFG conversion efficiency achievable in this class of nonlinear devices. We then design an efficient DFG nonlinear metasurface able to deliver hundreds of $\mu$Ws of power at around 5.8 THz using watt-level external pumps. We stress that this is the first analysis of flat ultrathin nonlinear devices for DFG based on hybrid multi-quantum well metasurfaces, enabling this level of DFG efficiencies. The proposed platform is robust, compact, operates at room temperature, is not constrained by phase-matching issues, and can deliver high-power continuous-wave THz beams using continuous-wave pumping with intensities in the kW cm$^{-2}$ range, safely below materials damage threshold.

2. Difference frequency generation with ultrathin plasmonic metasurfaces

The schematic of the nonlinear plasmonic metasurface we propose here is shown in figure 1(a): two normally impinging pump Gaussian beams with identical beam waist diameters $w$ and angular frequencies $\omega_1 = 2\pi \times 37 \text{ THz}$ and $\omega_2 = 2\pi \times 31.2 \text{ THz}$ impinge normally onto the metasurface and generate a difference-frequency signal in the THz band (corresponding to $\omega_3 = \omega_1 - \omega_2 = 2\pi \times 5.8 \text{ THz}$) that is subsequently radiated to free space in the backward direction. The metasurface shares the same fundamental principles as the ones aimed at SHG that we previously demonstrated both theoretically and experimentally [18, 21, 22, 28]; we couple electronic intersubband transitions in MQWs with electromagnetic resonances provided by subwavelength metallic inclusions. Despite the connection between SHG and DFG processes, the analysis and design of DFG nonlinear metasurfaces face some important challenges that we are carefully addressing in this work, such as the need to design resonant structures able to resonate at three largely separated frequencies, the optimization of MQWs aimed to DFG and associated saturation mechanisms, and a clear analysis of the entire electromagnetic process.

The unit-cell of the proposed metasurface is shown in figure 1(b). Here, 300 nm thick gold mace-shaped resonators are placed on top of a 800 nm thick MQW semiconductor layer backed by an opaque gold ground plate. The epi-transfer of the MQW layer onto a gold ground plate is achieved via an episide-down wafer bonding followed by a MQW substrate removal [18]. A 50 nm thick platinum layer between the MQWs and the gold back plate is added to prevent gold diffusion into the MQW structure during the wafer bonding process. The MQWs are then etched around the gold resonators. As has been experimentally demonstrated [21, 28], larger effective nonlinear responses and better polarization control can be achieved by judiciously removing the nonlinear material around the resonators. The shape of the resonators and the unit-cell are optimized to provide strong $z$-polarized fields in MQWs when the structure is excited at all three frequencies. At $\omega_1$ and $\omega_2$, the unit-cell provides similar dipole-resonance-like field enhancement for $x$-polarized impinging radiation, inducing vertically polarized electric fields underneath the resonator that can engage the nonlinearity in the MQW substrate. At the DFG frequency, $\omega_3$, the structure is designed to provide a strong dipolar resonance for $y$-polarized field that again induces vertical fields in the MQW substrate. Due to reciprocity, this property ensures that the nonlinearly generated vertical DFG polarization currents in the substrate can efficiently radiate in free-space a $y$-polarized beam. The corresponding spatial distributions of normalized $z$-polarized induced fields in MQWs when the structure is excited at all three frequencies. At $\omega_1$ and $\omega_2$, the unit-cell provides similar dipole-resonance-like field enhancement for $x$-polarized impinging radiation, inducing vertically polarized electric fields underneath the resonator that can engage the nonlinearity in the MQW substrate. At the DFG frequency, $\omega_3$, the structure is designed to provide a strong dipolar resonance for $y$-polarized field that again induces vertical fields in the MQW substrate. Due to reciprocity, this property ensures that the nonlinearly generated vertical DFG polarization currents in the substrate can efficiently radiate in free-space a $y$-polarized beam. The corresponding spatial distributions of normalized $z$-polarized induced fields, crucial to understanding of the operation of our metasurface, are shown in figure 1(c). It is also important to mention that while the MQW permittivity $\varepsilon_{MQW,ij}$ is nearly...
isotropic and constant at both pump frequencies, around the DFG frequency it is very dispersive and can exhibit an anisotropic response due to the confinement of electron motion across the MQW layers (in the z-direction). Since the electrons are free to move along semiconductor layers, the in-plane components of the permittivity $\varepsilon_{\text{MQW},xx} = \varepsilon_{\text{MQW},yy}$ can be computed using the Drude–Lorentz model, following the approach in [18]. The permittivity in the vertical direction $\varepsilon_{\text{MQW},zz}$ depends on quantum transitions between electron states in the MQW heterostructure and on the electron population in these states, which is controlled by pumping [18]. The MQW heterostructure may, in principle, be tailored to avoid strong absorption in the z-direction at the target THz frequency [34]. In this paper, for simplicity, we use the isotropic 3D Drude–Lorentz permittivity model [18] to assess a worst-case scenario of strong THz absorption in the z-direction. The permittivity and the corresponding simulated absorption spectra are shown in figure 1(d). In reality, we expect that the response will be actually more lenient in terms of absorption, due to the expected anisotropy.

To analyze our proposed DFG nonlinear metasurface, let us first consider illumination by a normally incident double-frequency field

$$\mathbf{E}_{\text{inc}}(\mathbf{r}, t) = E_{\text{inc}}^{(1)}(\mathbf{r}) e^{i \omega_1 t} + E_{\text{inc}}^{(2)}(\mathbf{r}) e^{i \omega_2 t} + \text{c.c.},$$

where ‘c.c.’ denotes complex conjugate, the subscript ‘inc’ indicates that the field is impinging upon the metasurface from free-space and $\mathbf{r}$ is the position vector. Then, an effective average nonlinear polarization current oscillating at $\omega_1 = \omega_1 - \omega_2$ induced in the structure can be expressed as [8]

$$P_{\text{eff}, \omega_1 - \omega_2} = \varepsilon_0 \sum_{m,n} \chi^{(2)}_{\text{eff,lin}}(\omega_1 - \omega_2; \omega_1, \omega_2) E_{\text{inc},m}^{(1)} E_{\text{inc},n}^{(2)},$$

where $\varepsilon_0$ is the free-space permittivity, $m$ and $n$ are indices for polarization axes $x$ or $y$ and the asterisk denotes the complex conjugate. Note that to be consistent with other publications on resonant intersubband nonlinearities in semiconductor heterostructures, equation (2) drops the ‘permutation factor’ of 2 used in [8] and, correspondingly, we will not use the factor of $1/2$ used in [8] in the expression for THz DFG nonlinearity $\chi^{(2)}_{\text{eff}}$ presented below. In equation (2), we adopt an effective intensity-dependent nonlinear susceptibility tensor $\chi^{(2)}_{\text{eff,lin}}(I^{(1)}, I^{(2)})$, similar to the one introduced in [21], that homogenizes the nonlinear response of the entire structure, taking phase-matching within the resonator and the influence of saturation and loss into account. Assuming that the conversion efficiency of the nonlinear process is low (below a few percent), one can apply the Lorentz reciprocity theorem at $\omega_1 = \omega_1 - \omega_2$ [18, 22, 35] to derive this effective nonlinear susceptibility as

$$\chi^{(2)}_{\text{eff,lin}}(I^{(1)}, I^{(2)}; \mathbf{r}) = \frac{1}{V} \int_V \chi^{(2)}_{\text{zz}}(I^{(1)}, I^{(2)}; \mathbf{r}) \times \frac{E_{\text{inc},m}^{(1)} E_{\text{inc},n}^{(2)} - E_{\text{inc},n}^{(1)} E_{\text{inc},m}^{(2)}}{E_{\text{inc},m}^{(1)} E_{\text{inc},n}^{(2)} - E_{\text{inc},n}^{(1)} E_{\text{inc},m}^{(2)}} d^2 \mathbf{r},$$

where $E_{\text{inc}}^{(m)}$ is the z-component of the electric field induced in the unit-cell by a $m$-polarized impinging plane wave of amplitude $E_{\text{inc},m}^{(z)}$ oscillating at frequency $\omega_i$ ($i = 1, 2, 3$), and $V$ is the unit-cell volume. It is important to keep in mind that $\chi^{(2)}_{\text{zz}}(I^{(1)}, I^{(2)}; \mathbf{r})$ is the only non-zero intrinsic nonlinear susceptibility tensor element of the unit-cell, and it is a function of the local intensities of the $z$-polarized fields $I^{(1)}$ and $I^{(2)}$ induced in the MQW. These intensities are given as [8, 36]

$$I^{(i)} = 2n_{\text{MQW}} c \varepsilon_0 |E_{\text{inc}}^{(i)}|^{2}, \quad i = 1, 2,$$

and they can be related to the external pump intensities $I^{(1)}$ and $I^{(2)}$ as

$$I^{(i)} = I^{(i)} n_{\text{MQW}} \left| \frac{E_{\text{inc}}^{(i)}}{E_{\text{inc},m}^{(z)}} \right|^2, \quad i = 1, 2,$$

where $n_{\text{MQW}}$ is the refraction index of MQWs at $\omega_i$. From equation (3), it is evident that the amplitude of the DFG signal is proportional to the spatial overlap integral among the fields at $\omega_1$, $\omega_2$, and $\omega_3 = \omega_1 - \omega_2$ within the MQW volume. In the specific case of the resonators shown in figure 1, the largest effective susceptibility tensor element is $\chi^{(2)}_{\text{eff,lin}}$, i.e., when both pump waves impinge on the structure with $x$-polarization, and the generated THz beam is $y$-polarized, as expected due to reciprocity. The spatial overlap over the MQWs volume—the real part of the integrand in equation (3)—is illustrated for this particular polarization combination in figure 1(b) (without considering saturation effects). It is also important to mention that $\chi^{(2)}_{\text{eff}}$ processes cannot be realized in geometries possessing centrosymmetry [8]. The asymmetric design of the resonator employed here clearly satisfies this criterion.

Using equation (3), we can evaluate the metasurface DFG conversion efficiency $\eta_{\text{DFG}}$ when the two plane waves with intensities $I^{(1)}$ and $I^{(2)}$ impinge on it. Following the same procedure as in [22], we find

$$\eta_{\text{DFG}}(I^{(1)}, I^{(2)}) = \frac{I^{(1)} - I^{(2)}}{I^{(1)} + I^{(2)}} = \frac{1}{8} \frac{I^{(1)} I^{(2)}}{I^{(1)} + I^{(2)}} \times \eta_0 (k_1 - k_2)^2 h^2 |\chi^{(2)}_{\text{eff,lin}}(I^{(1)}, I^{(2)})|^2,$$

where $k_i = \omega_i / c$, $\eta_0$ is a free-space impedance, and $h$ is the MQW thickness. This expression can be straightforwardly extended to Gaussian beams or other types of pump excitation [8, 22]. We note that one could omit the factor $h^2$ in equation (6) by introducing a surface nonlinear susceptibility tensor $\chi^{(2)}_{\text{eff}} = h \chi^{(2)}_{\text{eff,lin}}$, which describes the nonlinear response per surface area for an ideally thin metasurface. In what follows we avoid this step to allow a straightforward comparison between $\chi^{(2)}_{\text{eff}}$ and intrinsic $\chi^{(2)}_{\text{zz}}$ of the MQW substrate.

3. Design of the MQW structure for DFG

The MQW structure considered in this paper is shown in figure 2. The design is based on the concept presented in [34], modified for the case of In$_{0.53}$Ga$_{0.47}$As/In$_{0.48}$Al$_{0.52}$As grown on an InP substrate. The layers’ thicknesses and doping levels

M Tymchenko et al
are indicated in figure 2, and the intersubband transition energies, dipole moments and state lifetimes for the first three energy levels are given in table 1. The single-period heterostructure shown in figure 2 is repeated 34 times to create an 800 nm thick MQW nonlinear medium with giant intersubband $\chi^{(2)}_{zzz}$. The thickness of Al$_{0.48}$In$_{0.52}$As barriers separating adjacent MQW units is chosen to be 8 nm to reduce energy anticrossing between states in the adjacent MQW units to below 0.5 meV.

The intersubband nonlinearity for THz DFG of the MQW structure is computed assuming pump frequencies in resonance with 1–2 and 1–3 transitions, using the expression [34, 37]

$$\chi^{(2)}_{zzz}(\omega_3 = \omega_1 - \omega_2) \approx \frac{1}{\varepsilon_0} \left( \frac{N_i - N_j}{E_1 - E_j} \right) \times \left( \frac{q^2}{E_1} \right) \times N_i - N_j \left( E_1 - E_j \right),$$

where $q_i$ is the electron charge, $N_i$ are the population densities in the electron states $i = 1, 2$ and 3, see figure 2, $q_i z_i$, $E_{ij}$, and $\Gamma_{ij}$ are the dipole matrix element, energy and broadening of the transition between states $i$ and $j$, respectively, see table 1. Given the average MQW doping of $N_e = 3.9 \times 10^{17}$ cm$^{-3}$, and assuming that in the absence of excitation the electron population remains in the ground state, we obtain $\chi^{(2)}_{zzz} = 2.8 \times 10^6$ pm V$^{-1}$, an impressive number.

Table 1 also lists the non-radiative relaxation times $\tau_{ij}$ between states 1, 2 and 3. The relaxation times between states 2 and 1 and 3 were computed using longitudinal-optical phonon scattering rates and the scattering between states 1, 2 and 3. The relaxation times between states 1, 2 and 3, see equation [3]. The intersubband transition energies, dipole moments and state lifetimes for the first three energy levels are given in table 1. The layer sequence in nm is 4/3.4/2.7/9.1/4, where Al$_{0.48}$In$_{0.52}$As barriers are shown in bold, and the 9.1 nm In$_{0.53}$Ga$_{0.47}$As well is doped to $1 \times 10^{18}$ cm$^{-3}$.

SHG nonlinear metasurfaces [22], we have rigorously taken into account the influence of saturation effects on the nonlinear response, identified the upper bound for SHG efficiency, and predicted their electromagnetic response with great accuracy. This framework has also unequivocally confirmed that saturation effects must be considered at each design step, including MQW band-engineering and resonator design. In this section, we extend the saturation analysis introduced in [22] to estimate realistic conversion efficiencies in DFG nonlinear metasurfaces, and establish fundamental upper bounds for the performance of these devices.

Let us start by rewriting the $zzz$-component of the intersubband DFG nonlinear susceptibility of MQWs given in equation (7) as

$$\chi^{(2)}_{zzz}(I_{z^1}, I_{z^2}) \approx \psi [S_1^{DFG}(I_{z^1}, I_{z^2}) \zeta_{13} + S_2^{DFG}(I_{z^1}, I_{z^2}) \zeta_{12}],$$

where $S_i^{DFG}(I_{z^1}, I_{z^2}) \approx (N_i - N_j)/N_e$ and $S_i^{DFG}(I_{z^1}, I_{z^2}) \approx (N_i - N_j)/N_e$ are the two saturation factors that depend on carrier populations in each subband, $N_i$ with $i = 1, 2, 3$, normalized to a bulk MQW doping level $N_e$. In equation (8), the parameters $\psi$, $\zeta_{13}$, and $\zeta_{21}$ are given as

$$\psi = \frac{q^3}{\varepsilon_0 (E_1 - E_2 - j\Gamma_{23})},$$

4. MQWs saturation and upper bound on DFG efficiency

Experimental measurements have clearly indicated that the efficiency of nonlinear metasurfaces is severely affected by the saturation of MQWs even at moderate pump intensities [18]. The reason is that for efficient polarization conversion plasmonic resonators need to be made highly resonant. With the increase of pump intensity, the enhanced fields around the resonators deplete the ground MQW subband of charge carriers, quenching the nonlinear generation from MQWs in the saturated area. In our recent theoretical analysis developed for...
The populations of different subbands can be computed from the steady-state solution of coupled-rate equations:

$$\frac{\partial N_i}{\partial t} = \frac{\alpha_{i3}(\omega_i)I_{i3}^{\omega_i}}{h\omega_1} - \frac{\alpha_{i3}(\omega_i)I_{i3}^{\omega_i}}{h\omega_2} + \frac{\alpha_{i1}(\omega_i)I_{i1}^{\omega_i}}{h\omega_1} + \frac{N_i - N_{j0}}{\tau_2} + \frac{N_i - N_{j0}}{\tau_3} = 0,$$

$$\frac{\partial N_j}{\partial t} = \frac{\alpha_{j1}(\omega_j)I_{i1}^{\omega_i}}{h\omega_1} - \frac{\alpha_{j2}(\omega_j)I_{i2}^{\omega_i}}{h\omega_2} + \frac{N_j - N_{j0}}{\tau_2} = 0,$$

for which the carrier conservation condition $N_i + N_2 + N_j = N_i$ holds, $N_{j0}$ is the carrier concentration in the $j$th subband in thermal equilibrium, and $\tau_2$ and $\alpha_j$ are the relaxation time and the absorption coefficient between $i$ and $j$ subbands, respectively. The latter is given by [38]

$$\alpha_j(\omega_i) = \frac{N_j - N_j}{h\omega_1} = \frac{(N_j - N_j)J_n^{\omega_i}z_j^2}{\epsilon_0 n_{MQW}^2c[(E - E_j)^2 + (\Gamma_j)^2]}.$$  

Neglecting the effect of optical heating, and assuming that the Fermi level is just above the bottom subband, we can set $N_{j0} = N_{j0} \approx 0$. Solving equations (11)–(13), we can then derive the saturation factors in closed form:

$$S_{1D}^{DFG}(I_{i1}^{\omega_i}, I_{i2}^{\omega_i}) \approx \frac{2E_{i1}^{\omega_i}E_{i2}^{\omega_i}}{(I_{i1}^{\omega_i} + I_{i2}^{\omega_i})^2(2 + C_{12}) + 2E_{i1}^{\omega_i}E_{i2}^{\omega_i}(1 + C_{12})} \left(1 + \frac{\tau_1}{\tau_{12}}\right).$$

$$S_{2D}^{DFG}(I_{i1}^{\omega_i}, I_{i2}^{\omega_i}) \approx \frac{I_{i1}^{\omega_i}I_{i2}^{\omega_i} + I_{i2}^{\omega_i}I_{i1}^{\omega_i}(2 + C_{12}) + 2E_{i1}^{\omega_i}E_{i2}^{\omega_i}(1 + C_{12})} \left(1 + \frac{\tau_1}{\tau_{12}}\right) \left(1 + \frac{\tau_1}{\tau_{12}}\right).$$

where $I_{i1}^{\omega_i}$ and $I_{i2}^{\omega_i}$ are saturation currents defined as

$$I_{i1}^{\omega_i} = \frac{N_i h\omega_1}{2\alpha_{i1}(\omega_i)\tau_1},$$

and $C_{12}$ relates the coupling between the first and the second subbands, $N_i = N_{i2}N_i$.

$$C_{12} = \frac{I_{i1}^{\omega_i}I_{i2}^{\omega_i} + I_{i2}^{\omega_i}I_{i1}^{\omega_i} + 2E_{i1}^{\omega_i}E_{i2}^{\omega_i}(1 + C_{12})}{I_{i1}^{\omega_i}I_{i2}^{\omega_i} + I_{i2}^{\omega_i}I_{i1}^{\omega_i} + 2E_{i1}^{\omega_i}E_{i2}^{\omega_i}(1 + C_{12})}.$$

Note that $C_{12} \approx 1$ when these subbands are strongly coupled, a scenario that occurs when pump beam has high intensity and $(I_{i1}^{\omega_i}/R_{i1}^{\omega_i} + I_{i2}^{\omega_i}/R_{i2}^{\omega_i})^{-1} \ll 1$.

Following similar considerations as in [22], we can prove that the conversion efficiency of the metasurfaces is fundamentally limited by these saturation effects. For this purpose, we assume the best possible scenario of a unit cell that is able to produce uniform $z$-polarized fields across the MQW volume at all frequencies. For such an idealized situation, the intensity-dependent effective susceptibility can be found as

$$\chi_{\text{eff},\text{yxx}}^{(2)}(I_{i1}^{\omega_i}, I_{i2}^{\omega_i}) = \psi_1^{(S_{1D}^{DFG})}(I_{i1}^{\omega_i}, I_{i2}^{\omega_i}) \frac{R_{i1}^{\omega_i}R_{i2}^{\omega_i}}{R_{i1}^{\omega_i}R_{i2}^{\omega_i} + 2E_{i1}^{\omega_i}E_{i2}^{\omega_i}(1 + C_{12})} + \psi_2^{(S_{2D}^{DFG})}(I_{i1}^{\omega_i}, I_{i2}^{\omega_i}) \frac{R_{i1}^{\omega_i}R_{i2}^{\omega_i}}{R_{i1}^{\omega_i}R_{i2}^{\omega_i} + 2E_{i1}^{\omega_i}E_{i2}^{\omega_i}(1 + C_{12})},$$

where $f_{i1}^{\omega_i} = E_{i1}^{\omega_i}/E_{inc}^{\omega_i}$ denotes the uniform field enhancement factor within the unit-cell at $\omega_i$, $p = 1, 2, 3$. Let us now assume a negligible dispersion of MQWs at pump frequencies, $n_{MQW}^{\omega_i} \approx n_{MQW}^{\omega_i} = n_{MQW}^{\omega_i}$ and identical pump intensities $I_{i1}^{\omega_i} = I_{i2}^{\omega_i} = I_{i1}^{\omega_i}$ that are much larger than the saturation intensity for transition between the levels 1 and 3, i.e. $I_{i1}^{\omega_i} \gg I_{i1}^{\omega_i}$. In this scenario, the effective intensity-dependent nonlinear susceptibility yields

$$\chi_{\text{eff},\text{yxx}}^{(2)}(I_{i1}^{\omega_i}) \approx \psi_1^{(S_{1D}^{DFG})}(I_{i1}^{\omega_i}) \frac{R_{i1}^{\omega_i}R_{i2}^{\omega_i}}{R_{i1}^{\omega_i}R_{i2}^{\omega_i} + 2E_{i1}^{\omega_i}E_{i2}^{\omega_i}(1 + C_{12})}$$

$$\times \frac{2E_{i1}^{\omega_i}E_{i2}^{\omega_i}}{(I_{i1}^{\omega_i} + I_{i2}^{\omega_i})^2(2 + C_{12}) + 2E_{i1}^{\omega_i}E_{i2}^{\omega_i}(1 + C_{12})} \left(1 + \frac{\tau_1}{\tau_{12}}\right).$$

Combining (19) with (6), we find that the DFG conversion efficiency approaches zero for input beams of extremely high intensities:

$$\lim_{I_{i1}^{\omega_i} \to \infty} \eta_{DFG}(I_{i1}^{\omega_i}) = 0,$$
Here, we assume that there are two pump beams having identical powers. The shaded area around the red curve denotes the bounds of \( \eta_{\text{DFG}} \), whereas the maximum generated intensity is upper-bounded by

\[
\max I^{\omega_1 \rightarrow \omega_2}(I^{\omega_1 \rightarrow \omega_2}) = \frac{1}{18} \frac{\eta_{\text{DFG}}(k_1 - k_2)^2 k^2}{n_{\text{MQW}}} \times \left[ \psi \zeta_{13} f_{\omega_1 \omega_2} \frac{I_{\omega_1}^{\omega_1}}{I_{\omega_1}^{\omega_1} + I_{\omega_2}^{\omega_2}} \left( 1 + \frac{\pi_1}{\pi_2} \right) \right]^2,
\]

(21)

Our analysis has profound implications for the design of nonlinear DFG metasurfaces: given \( f_{\omega_1 \omega_2} \), using equation (21) we can estimate the approximate maximum conversion efficiency the metasurface can provide before testing an actual design.

5. Results and performance

In this section, we apply our proposed theoretical framework to study the performance of the MQW-based nonlinear metasurface illustrated in figure 1. We consider a realistic MQW stack, defined by the parameters given in table 1, with a maximum intrinsic DFG susceptibility \( \chi_{\text{eff}}^{(2)} = 2.8 \times 10^6 \text{ pm V}^{-1} \). The effective permittivity of this MQW is shown in figure 1(d).

In figure 3, we show the influence of saturation on \( \chi_{\text{eff}}^{(2)} \) as a function of the local intensity, evaluated using equations (8)–(10) and (13)–(17). We also show the effective nonlinear susceptibility \( \chi_{\text{eff,crystal}}^{(2)} \), computed using equation (3) as a function of the external peak beam intensity. It is evident that, under external pump intensities of about 10 kW cm\(^{-2}\), saturation effects start to be noticeable, effectively quenching the DFG response at intensities above 1 MW cm\(^{-2}\).

The power conversion efficiency \( \eta_{\text{DFG}} \) computed using equation (6) is shown in figure 4(a). Here and in what follows, we assume that the two pump beams carry the same net power \( P^{\omega_1 \rightarrow \omega_2} \), and that the total power delivered to the metasurface is \( 2P^{\omega_1 \rightarrow \omega_2} \). It is seen that the maximum DFG conversion efficiency corresponds to pump power of about 220 mW (peak intensity of only 100 kW cm\(^{-2}\)), with an impressive value of 0.01\%.

Even though this result is theoretical, this is a truly remarkable number for a sub-\( \mu \text{m} \)-thick system illuminated at such a low pump intensity, and it may potentially challenge even the most efficient state-of-the-art nonlinear optical setups [9]. For the reference, note that a LiNbO\(_3\) crystal of the same thickness would provide \( \eta_{\text{DFG}} \) of only \( 10^{-12} \) (i.e. \( 10^{-10}\%) \) [18, 21]. It is also evident that the achievable conversion efficiency noticeably deviates from the case of no saturation at intensities above 2 kW cm\(^{-2}\). With an increase of pump power beyond 100 kW cm\(^{-2}\), saturation effects start to dominate, and the DFG power conversion efficiency drops.

It is also instructive to study the dependence of DFG conversion efficiency versus the transition linewidths \( \Gamma_i \), which can be manipulated when designing the semiconductor heterostructure. In figure 4(a), the lightly colored region around the \( \eta_{\text{DFG}} \) curve corresponds to the case when all \( \Gamma_i \) are synchronously modified between 80% and 120% of the initial values from table 1 (smaller linewidths correspond to higher efficiency). It is seen that such arbitrary yet relatively large variations lead to less than one order of magnitude change in \( \eta_{\text{DFG}} \), indicating a good tolerance to MQW quality. Finally, the area shaded with orange above intensities of about 160 kW cm\(^{-2}\) indicates the region in which the optical damage to the metasurface may occur [39–41].

The total DFG power delivered by our proposed metasurfaces at 5.8 THz is shown in figure 4(b). Different from conversion efficiency, the actual radiated DFG power increases almost monotonically until much higher pump intensities of about 1.3 MW cm\(^{-2}\), reaching 0.25 mW. This is, again, a remarkable figure that highlights the practical potential of the proposed DFG nonlinear metasurface platform. For comparison, state-of-the-art QCLs are able to deliver about an order of magnitude larger power, but they...
must be operated at cryogenic temperatures. It is important to mention, however, that the peak DFG power lies deep in the region in which high dissipative losses in plasmonic resonators will likely lead to photodamage. Still, 0.15 mW of THz output power is achievable at pump intensities of about 140 kW cm\(^{-2}\), below the damage threshold.

6. Conclusions

In this work, we have introduced a novel platform for efficient generation of THz radiation using nonlinear DFG metasurfaces that couple intersubband transitions in MQWs to electromagnetic resonances engineered in ultrathin, subwavelength, plasmonic structures. We have developed a theoretical framework able to accurately characterize such devices, taking complex phenomena such as saturation and loss into account, and providing theoretical upper bounds for their electromagnetic response. Numerical simulations confirm the tremendous potential of nonlinear DFG metasurfaces, exhibiting extremely large conversion efficiencies up to \(10^{-3}\), and promising to deliver sub-nanowatt of THz power using pump beams with peak intensities in the \(10^{7}\) kW cm\(^{-2}\) range. Given the fact that such metasurfaces can operate at room temperature under continuous-wave pumping and that they do not require active cooling, they become an appealing alternative to other compact THz sources such as p-Ge lasers, QCLs, and THz sources based on on-cavity DFG in QCLs. The proposed platform is compact and robust to fabrication tolerances in the MQWs, and we expect that it can be largely improved by engineering optimized plasmonic resonators. Combined with their robust field control that can potentially lead to advanced waveform engineering functionalities [27, 28, 35], the proposed nonlinear DFG metasurfaces constitute an important milestone in the quest for compact and efficient THz radiation sources.

Acknowledgments

This work was partially supported by the Air Force Office of Scientific Research and the National Science Foundation.

References

[26] Capasso F, Sirtori C and Cho A Y 1994 Coupled quantum well semiconductors with giant electric field tunable nonlinear optical properties in the infrared IEEE J. Quantum Electron. 30 1313–26

30] Li G et al 2015 Continuous control of the nonlinearity phase for harmonic generations Nat. Mater. 14 607–12

31] Li G, Zhang S and Zentgraf T 2017 Nonlinear photonic metasurfaces Nat. Rev. Mater. 2 17010


39] Caldarola M et al 2015 Non-plasmonic nanoantennas for surface enhanced spectroscopies with ultra-low heat conversion Nat. Commun. 6 7915
