scientific reports

OPEN

Terahertz radiation generation from metallic electronic structure manipulated by inhomogeneous DC-fields

H. Lin✉ & C. P. Liu✉

We propose a feasible, high-efficiency scheme of primary terahertz (THz) radiation source through manipulating electronic structure (ES) of a metallic film by targeted-designed DC-fields configuration. The DC magnetic field is designed to be of a spatially inhomogeneous strength profile, and its direction is designed to be normal to the film, and the direction of the DC electric field is parallel to the film. Strict quantum theory and numerical results indicate that the ES under such a field configuration will change from a 3D Fermi sphere into a highly-degenerate structure whose density-of-state curve has pseudogap near Fermi surface. Wavefunctions' shapes in this new ES are space-asymmetric, and the width of pseudogap near Fermi surface, as well as magnitudes of transition matrix element, can be handily controlled by adjusting parameter values of DC fields. Under available parameter values, the width of the pseudogap can be at milli-electron-volt level (corresponding to THz radiation frequency), and the magnitude of oscillating dipole can be at $10^{-9} C * m$-level. In room-temperature environment, phonon in metal can pump the ES to achieve population inversion.

The consensus that electronic structure (ES) of matter determines its optical property guides people to tailor/design suitable ES for achieving various optical applications. Compared with those secondary radiation sources[1–22], primary radiation sources based on suitable ES are more efficient[23–43]. Here, the phrase "secondary radiation sources" refers to that radiation generation arises from nonlinear optics effect or nonlinear interaction between driving radiation beam (at other wavelength) and driven emitter materials. The phrase "primary radiation sources" refers to that the driving is not radiation or electromagnetic wave. Impurity doping, alloying, growing/synthesizing various heterostructures such as semiconductor quantum well and superlattice are popular methods of preparing novel ESs[23–43]. These methods can be categorized into chemical manipulation of ES, which refers to change matter's chemical constituent. Clearly, the chemical manipulation can lead to "permanent" change in the ES of matter.

In contrast, fields manipulation is to modify the ES having a "volatile" change, which refers to the ES recovering to its primitive form when removing the manipulating fields. For example, De Hass–Van Alphen effect is an example of DC-(magnetic)-fields-manipulation of metal's ES, and field-effect-transistor (FET) is an example of DC-(electric)-fields-manipulation of semiconductor carriers' ES[44,45]. Such a "volatile" change is more appropriate for many applications because of its simplicity, efficiency and economy. The fields-manipulation is more flexible in achieving tunable solid-state radiation source because it does not need to change materials for achieving another center-frequency output.

Obvious application prospect of radiations at terahertz (THz) band[46] promotes intensive investigations on achieving THz primary radiation sources. Scarcity of materials with ES characterized by THz-level gap leads to so-called "THz gap"[47]. Main limit of above-mentioned band-structure-engineering[23–43] is that the designed ES comprises too many levels which often form narrow bands or so-called mini-bands. In room-temperature environment, radiation generation from such an ES in which two mini-bands spaced by a meV-level gap will have a too broad spectrum undesirable for many applications. Moreover, the dipole matrix element in such an ES is of finite enhancement than those in atomic and molecular ES.

The advantage of the fields-manipulating metallic ES over band-structure-engineering of semiconductor carriers[23–43] is that larger charge density, as well as larger dipole magnitude, can be achieved in the former. Usually, the density of free electrons in a metal is at $10^{22\sim23}$ cm^{-3}-level, higher several order of magnitude than that of

State Key Laboratory of High Field Laser Physics, Shanghai Institute of Optics and Fine Mechanics, P. O. Box 800-211, Shanghai 201800, China. ✉email: linhai@siom.ac.cn; chpliu@siom.ac.cn

carriers in semiconductor (about 10^{10} cm^{-3}). Therefore, manipulating free electrons in a metal is more hopeful to achieve higher output power than manipulating carriers in a semiconductor.

In this work, for more practical purpose, we investigate manipulating the ES of a metallic film by a special DC-fields configuration in which the magnetic field is space-inhomogeneous. It is well-known that a homogeneous DC magnetic field B_0 can lead to an ES in which the general form of 3D wavefunctions is a product of a 2D plane-wave (in terms of (z, x)) and a 1D harmonic-oscillator wavefunction (in terms of y). As shown latter, orthogonality among 1D harmonic-oscillator wavefunctions enables the dipole matrix elements for the transition between two Landau levels in such an ES to be confined at 0 and hence prohibits dipole radiation generation from the transition between two Landau levels. In contrast, targeted using inhomogeneous $B_0 e_z \sim x$ or y can lead to an ES whose 3D wavefunctions have 2D localized parts losing orthogonality. In such an ES, the above-mentioned prohibition on dipole radiation generation can be removed, and hence a practical radiation source is possible.

Theory and method

Electronic structure manipulated by homogeneous DC-fields, revisiting De Hass–Van Alphen effect.
According to many textbooks[44,45], free electrons under homogeneous DC magnetic field can be described by a class of 3D eigenstate wavefunctions whose general form is a product of a 2D plane-wave $exp(ik_z z + ik_x x)$ and a 1D harmonic-oscillator wavefunction $\psi_n^{1D}(y)$. Here, if 3D eigenstate wavefunctions of a quantum system are of a general form which is a product of a 1D extended-state wavefunction and a 2D localized-state one, i.e. $\psi_{k_z,n}^{3D}(x, y, z) = exp(ik_z z)\psi_n^{2D}(x, y)$, it is not certain for these 3D wavefunctions being orthogonal mutually. For example, if the 2D wavefunctions $\psi_n^{2D}(x, y)$ are governed by a 2D Hamiltonian operator as follows:

$$E_n \psi_n^{2D} = H^{2D} \psi_n^{2D} = \left[-\alpha \left(\partial_{xx} + \partial_{yy} \right) + \beta_x \partial_x + \beta_y \partial_y + V(x, y) \right] \psi_n^{2D}, \qquad (1)$$

where α, β_x and β_y are constants and the "potential-energy" function V does not contain differential operators such as ∂_x, it is easy to prove that orthogonality exists among them. It is well-known that orthogonality exists among harmonic-oscillator wavefunctions, and same conclusion is also hold for plane-waves.

In such a ES, dipole matrix elements $\int \left[\psi_{k_z,n}^{3D} \right]^* x, y, z \left[\psi_{k_z+q,m}^{3D} \right] dxdydz$ to be confined at 0 if $q \neq 0$ and $m \neq n$. For generating a photon with momentum $\hbar q$ and energy $\hbar qc$, momentum conservation and energy conservation demand two levels having a difference in their wavevectors. Thus, $\int \left[\psi_{k_z,n}^{3D} \right]^* x, y \left[\psi_{k_z+q,m}^{3D} \right] dxdydz$ are $= 0$ because of orthogonality among plane-waves, and $\int \left[\psi_{k_z,n}^{3D} \right]^* z \left[\psi_{k_z+q,m}^{3D} \right] dxdydz$ is $= 0$ because of orthogonality among harmonic-oscillator wavefunctions. In short, if 2D, as well as 1D, localized-state wavefunctions have orthogonality, these dipole matrix elements are 0, which prohibits dipole radiation generation.

In contrast, 2D localized-state wavefunctions can loose orthogonality for 2D Hamiltonian operator in suitable forms. For example, if the 2D Hamiltonian operator reads

$$H^{2D} = \left[-\alpha \left(\partial_{xx} + \partial_{yy} \right) + f_x \partial_x + f_y \partial_y + V(x, y) \right], \qquad (2)$$

where f_x and f_y are functions of variables (x, y) rather than two constants, its eigenstate 2D wavefunctions can be not orthogonal mutually. Once 2D localized-state wavefunctions loose orthogonality, these dipole matrix elements, such as $\int \left[\psi_{k_z,n}^{3D} \right]^* z \left[\psi_{k_z+q,m}^{3D} \right] dxdydz$, can be $\sim q^{-2} \neq 0$ for cases with $q \neq 0$ and $m \neq n$ in which $\int \left[\psi_{k_z,n}^{3D} \right]^* x, y \left[\psi_{k_z+q,m}^{3D} \right] dxdydz = 0$ still exist. This removes a prohibition on dipole radiation generation.

Electronic structure manipulated by inhomogeneous DC-fields configuration.
Above discussion enlightens us to design/tailor the Hamiltonian governing a quantum system. Hamiltonian in following form has a 2D-part meeting the general form Eq. (2)

$$H^{3D} = \frac{1}{2m} \left[p_x^2 + \left(p_y - eB_0 \beta * yx \right)^2 + p_z^2 \right] + E_0 y, \qquad (3)$$

where inhomogeneous DC magnetic field and its vector potential read

$$B = (0, 0, B_z) = \left(0, 0, B_0 * \beta * y \right); \qquad (4)$$

$$A = \left(0, A_y, 0 \right) = \left(0, B_0 * \beta * y * x, 0 \right), \qquad (5)$$

B_0, β and E_0 are constants. Although Hall current, which is common for metal under $E \times B$ configuration, can exist (because B is along z-axis and E is along y-axis in this configuration and hence a Hall current is along x-axis), this Hall current represents a secondary effect and is too weak compared with controlling current in solenoid responsible for producing B. Usually, perturbation Hamiltonian associated with such a Hall current (which is a consequence of another stronger perturbation Hamiltonian), is taken as too weak to being added into total Hamiltonian.

Corresponding experimental setup is illustrated as Fig. 1. A pair of Helmholtz square-cross-section coils is on purpose mis-aligned to be not co-axial and hence produces a magnetic "slope", which refers to B_z^2 is the function of y. The metallic film is separated from two electrode plates by two gaps, which might be filled with high-permittivity insulator or not. According to detailed description elsewhere[48], it is feasible to attain a dimensionless
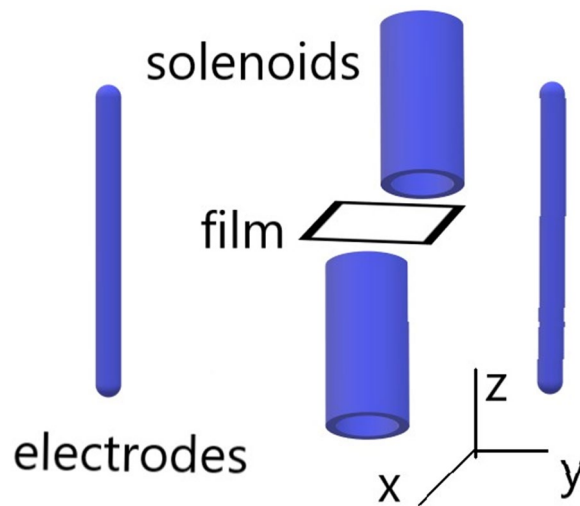
Figure 1. Sketch of field configuration and experimental setup. Note that two gaps might be filled with high-permittivity insulator, which can also act as an "holder" holding the film. If the "holder" is undertaken by thin dielectric substrate, suitably adjusting charge current in coils/solenoids can still warrant the metallic film to feel a magnetic slope. Note that cross-sections of solenoids are of square shape, which is not reflected in the scheme (due to the limit of plot software) and hence have to be stated clearly here.

slope β up to μm^{-1}-level or higher. Actually, numerous theoretical investigations on quantum Hall effect (QHE) are based on above Hamiltonian H at $B_0 * \beta * y \equiv constant$-case[49].

Electronic wave-function can be solved from

$$i\hbar \partial_t \psi = H^{3D} \psi; \tag{6}$$

in following general form

$$\psi = exp\left[ik_z z - i\frac{E_k}{\hbar}t\right]\phi(x, y), \tag{7}$$

where E_k is space-independent. Dividing both sides of Eq. (6) with ϕ, we obtain

$$r.h.s = \alpha * \left\{ \left[\partial_{xx} ln\phi + (\partial_x ln\phi)^2\right] + \left[\partial_y(\partial_y ln\phi + iwxy) + (\partial_y ln\phi + iwxy)^2\right] - sy \right\} + \alpha * k_z^2; \tag{8}$$

where constants are defined as

$$w = -\frac{eB_0\beta}{\hbar} = -[B_0]_{Tesla} * [\beta]_{\mu\text{m}^{-1}} * \frac{1.6}{1.054} \times 10^3 \mu\text{m}^{-3};$$

$$s = -2i\frac{m_e E_0}{B_0 \hbar \beta} = -2i * \frac{9.11}{1.054} \times 10^{-3} * [E_0]_{volt/meter}/[B_0]_{Tesla}/[\beta]_{\mu\text{m}^{-1}}; \tag{9}$$

$$\alpha = \frac{\hbar^2}{2m_e}\frac{1}{\mu\text{m}^2} = \frac{1.055^2}{2*0.91*1.602} \times 10^{-7} \text{ eV} = 1/2.88 \times 10^{-7} \text{ eV}.$$

Here, $[B_0]_{Tesla}$ is the value of B_0 in unit of *Tesla*, and other similar symbols are explained in the same manner.

By introducing new variables

$$y - x/s = u + v; \; x + y/s = u - v; \tag{10}$$

$$u = \frac{s+1}{2s}y + \frac{s-1}{2s}x; \; v = \frac{s-1}{2s}y - \frac{s+1}{2s}x; \tag{11}$$

we have

$$y = \frac{s+1}{s^2+1}su + \frac{s-1}{s^2+1}sv; \tag{12}$$

$$x = \frac{s-1}{s^2+1}su - \frac{s+1}{s^2+1}sv; \tag{13}$$

3

$$xy = \frac{s^4 - s^2}{\left[s^2 + 1\right]^2}\left[u^2 - v^2\right] - \frac{4s^3}{\left[s^2 + 1\right]^2}uv; \tag{14}$$

$$\partial_y = \frac{s+1}{2s}\partial_u + \frac{s-1}{2s}\partial_v; \tag{15}$$

$$\partial_x = \frac{s-1}{2s}\partial_u - \frac{s+1}{2s}\partial_v; \tag{16}$$

and can verify

$$
\begin{aligned}
d_v u &= \left[\frac{s*(s-1)}{s*s+1}d_y - \frac{s*(s+1)}{s*s+1}d_x\right]\left[\frac{s+1}{2s}y + \frac{s-1}{2s}x\right] \\
&= \frac{s*(s-1)}{s*s+1}*\frac{s+1}{2s} - \frac{s*(s+1)}{s*s+1}*\frac{s-1}{2s} = 0,
\end{aligned} \tag{17}
$$

which suggests independence between u and v. Thus, we re-write Eq. (8) as

$$R.H.S = \frac{r.h.s}{\alpha} - k_z^2 = \partial_x M + M^2 + \partial_y N + N^2 + \eta*\left[\partial_x N - \partial_y M\right]; \tag{18}$$

where $\eta = \frac{s}{iw}$ and

$$N = \partial_y ln\phi + iwxy; \quad M = \partial_x ln\phi; \tag{19}$$

$$\partial_x N - \partial_y M = iwy. \tag{20}$$

Thus, finally we have

$$E_k/\alpha - k_z^2 = R.H.S = \left[\partial_u(N+M) + \frac{1}{2}(N+M)^2\right] + \left[\partial_v(N-M) + \frac{1}{2}(N-M)^2\right]; \tag{21}$$

We can express $N + M$ and $N - M$ as binary power series

$$N - M = \sum_{i,j} f_{i,j}u^i v^j; \tag{22}$$

$$N + M = \sum_{i,j} g_{i,j}u^i v^j; \tag{23}$$

where coefficients $f_{i,j}$ and $g_{i,j}$ are complex-valued constants. These expansion coefficients can be solved from above equations through following strict procedure: for Eqs. (19, 20), order-by-order comparing their terms proportional to constant u^1 and v^1, we can obtain 6 algebra equations. For Eq. (21), 2 algebra equations can be obtained by order-by-order comparing its terms proportional to u^1 and v^1. Moreover, $\partial_x Eq.(21) = 0$ and $\partial_y Eq.(21) = 0$ also lead to 2 algebra equations. Note that 2 formulas $g_{0,0} = i(k_y + k_x)$ and $f_{0,0} = i(k_y - k_x)$ can be derived from definitions Eq. (19). Thus, 12 algebra equations and formulas (see "Appendix") can determine solutions of these these low-order expansion coefficients $g_{1,0}, g_{0,1}, f_{1,0}, f_{0,1}, f_{1,1}, g_{1,1}$. In this strict procedure, $g_{0,2}, g_{2,0}, f_{2,0}, f_{0,2}$ act as intermediate variables. Finally, $E_k - E_k^0$ is a function of $g_{1,0}$ and $f_{0,1}$ and hence can be known after these low-order expansion coefficients are solved according to this procedure (whose detailed formulas are presented at "Appendix" section).

According to formulas in "Appendix" Eq. (83), E_k is still $\sim k_z^2$ but has a more complicated dependence on $k_{\parallel}^2 \equiv k_x^2 + k_y^2, \theta \equiv arctan\left(\frac{k_y}{k_x}\right)$. This will lead to corresponding variation in density-of-state (DOS) curves. Moreover, all wavefunctions shape $|\psi_k|$ are space-asymmetric and hence correspond to dipole moment $d_x \equiv \int |\psi_k|^2 xdxdy \neq 0 \neq \int |\psi_k|^2 ydxdy \equiv d_y$.

Density-of-state of the ES. When external fields are absent, conduction-band of a metal corresponds to a Fermi sphere ES in 3D k-space and has a top at Fermi surface $|k| = k_F$ and a bottom at $|k| = 0$. Usually for light metals, their k_F are at $10^4 \mu m^{-1}$-level[44,45]. In the De Hass-Van Alphen effect, the ES becomes a set of 2D cylindric surfaces in 3D k-space and each cylindric surface corresponds to a sub-band[44,45]. Above result implies that, under inhomogeneous DC-fields configuration, the ES changes from a 3D Fermi sphere in which $E_k^0 \sim k_x^2 + k_y^2 + k_z^2$ and $0 \leq |k| < k_F$ exist, to a corrugated 3D Fermi sphere.

As shown in Fig. 2, each 2D cross-section (on the $k_x - k_y$ plane) of the 3D Fermi sphere can be divided into 6 segments, 3 corresponding to positive energy shift and 3 for negative. For $\theta = \frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{7\pi}{4}$, the energy shift ΔE can be $\pm\infty$ and correspond to nearly zero DOS, or $\frac{d\theta}{d\Delta E} \sim 0$ if $\theta = \frac{\pi}{4}, \frac{3\pi}{4}, \frac{5\pi}{4}, \frac{7\pi}{4}$.

Thus, for minimizing total energy, some regions on the $k_x - k_y$ plane, which are "unoccupied" or outside 2D Fermi circle, will be occupied because of negative energy shift over them, and some regions, which are "occupied" or inside 2D Fermi circle, will be unoccupied because of positive energy shift over them. For each $|k_{\parallel}|$-value,
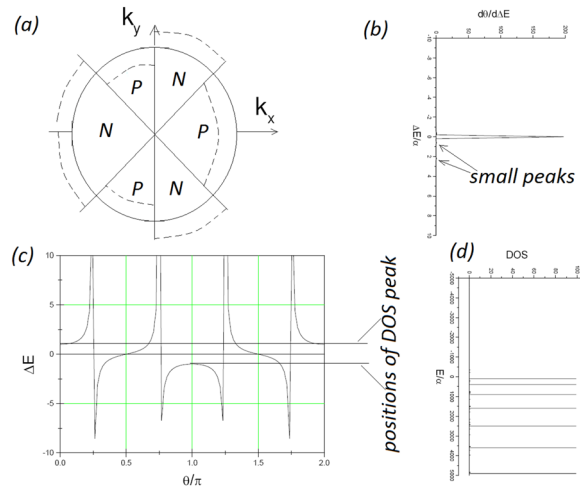
Figure 2. (**a**) Sketch illustrating energy shift varying over the $2D$ Fermi circle, where P (N) represents positive (negative) energy shift. The solid circle represents "unperturbed" $2D$ Fermi circle, and dashed lines represent realistic 2D Fermi surface. (**c**) is the dependence of the energy shift ΔE on $\theta \equiv arctan\left(\frac{k_y}{k_x}\right)$, where the definition of ΔE can be referred to Eq. (83). (**b**) is an example of DOS curve, $\sim \frac{d\theta}{d\Delta E}$, associated with the $\theta - \Delta E$ curve in (**c**). (**d**) illustrating that the superposition of DOS-curves like that in (**c**) can form DOS curve with pseudogap.

the dependence of the energy shift on θ, according to Eq. (83), will lead to a DOS curve with pseudogap (see Fig. 2c,b) because the DOS can arrive at its maximum at two θ-values: $\theta = 0$ and $\theta = \pi$. If each $|k_\parallel|$-value has such a DOS curve with pseudogap, their superpositions, as shown in Fig. 2d, will be possible to remain a pseudogap.

The pseudogap can be tuned, under available values of controllable parameters, to be at meV-level. According to Eq. (83), the fact $s^2 < 0$ (see Eq. 9) enables the width of the pseudogap at each $k_\parallel$-value, or $\frac{8ws^4}{[s^2+1]^2} * \frac{1}{|k_\parallel|}$, to be large enough. Moreover, the maximum of allowed $k_\parallel$-value corresponds to the narrowest among pseudogaps. Thus, as shown in Figs. 3 and 4 choosing suitable s^2-value, i.e., choosing suitable values of these controllable parameters, can lead to the superposition of these DOS curves over all allowed $k_\parallel$-values having a pseudogap with a desirable width. Because $1\,\text{meV} = 10^{-3}\,\text{eV} = 1.6 \times 10^{-19-3}\,\text{J}$ energy gap corresponds to a photon frequency $\nu = \frac{1.6 \times 10^{-19-3}J}{6.63 \times 10^{-34}J*S} = 0.25 \times 10^{12}S^{-1} = 0.25\,\text{THz}$, results presented in Figs. 3 and 4 indicate that the pseudogap can be tuned from 10^{-4} to 10^{-2} eV, and hence the central frequency of generated radiations can be tuned from 25 GHz to 2.5 THz. For example, up-right panel in Fig. 3 indicates a 5 meV pseudogap and hence implies a 1.25 THz central frequency, down-right panel in Fig. 4 indicates a 2 meV pseudogap and hence implies a 0.5 THz central frequency.

Space-asymmetric wavefunction shape $|\phi_{k_x,k_y}|$ implies each k-state corresponding to a dipole moment $d_{x,y} \equiv \int |\psi_k|^2 x, y dx dy$. According to "Appendix", the lowest order expansion terms of $Reln\phi_{k_x,k_y}$ is $\sim xy$. Thus, values of $d_{x,y}$ are dependent of a parameter $h_{1,1} \equiv \left[\partial_x \partial_y ln\phi\right]|_{x=0,y=0}$. Larger $|h_{1,1}|$-value will correspond to larger dipole moment. As reflected by Fig. 5, larger $k_\parallel$-values, which correspond to states near Fermi surface, correspond to larger $|h_{1,1}|$-values.

According to familiar formula for radiation generation from quantum transition[44,45], the radiation strength I, or transition matrix element, is proportional to both the product of DOS of initial state and that of final state and the product of dipole moment of initial state and that of final state:

$$I(\nu) \sim \sqrt{d_E}DOS(E) * \sqrt{d_{E+\hbar\nu}}DOS(E + \hbar\nu). \tag{24}$$

Note that high dipole moment and high DOS can be simultaneously achieved at some θ-regimes. For example, Fig. 5a indicates that $h_{1,1}$ can be large enough at $\theta = 0.5\pi$, and Fig. 2c indicates that $\theta = 0.5\pi$ corresponds to a very flat segment of the $\theta - \Delta E$ curve and hence to a large DOS-value. Therefore, the corrugated ES described above, which has a tunable width of pseudogap and sufficiently high dipole moments at states near peaks around pseudogap in DOS curves, is very favorable to powerful radiation output.

The partition function of the ES, which is taken as a canonical ensemble, reads

$$Z = \sum_i g_i exp(-E_i/K_B T), \tag{25}$$

where K_B is Boltzmann's constant. The ground state of such an ES is to fill the complicated 3D body and macroscopically corresponds to a static dipole (because all $(k_z, k_\parallel, \theta)$-states are of "polarized" shapes).

Parameters in laser rate equations. The new ES has a main feature: wavefunction $|\psi_{k_x,k_y,k_z} >= |\psi_{k_x,k_y}(x,y)|exp(ik_x x + ik_y y)exp(ik_z z)exp(i\omega_k t)$ have a space-varying shape $|\psi_{k_x,k_y}(x,y)|$. The
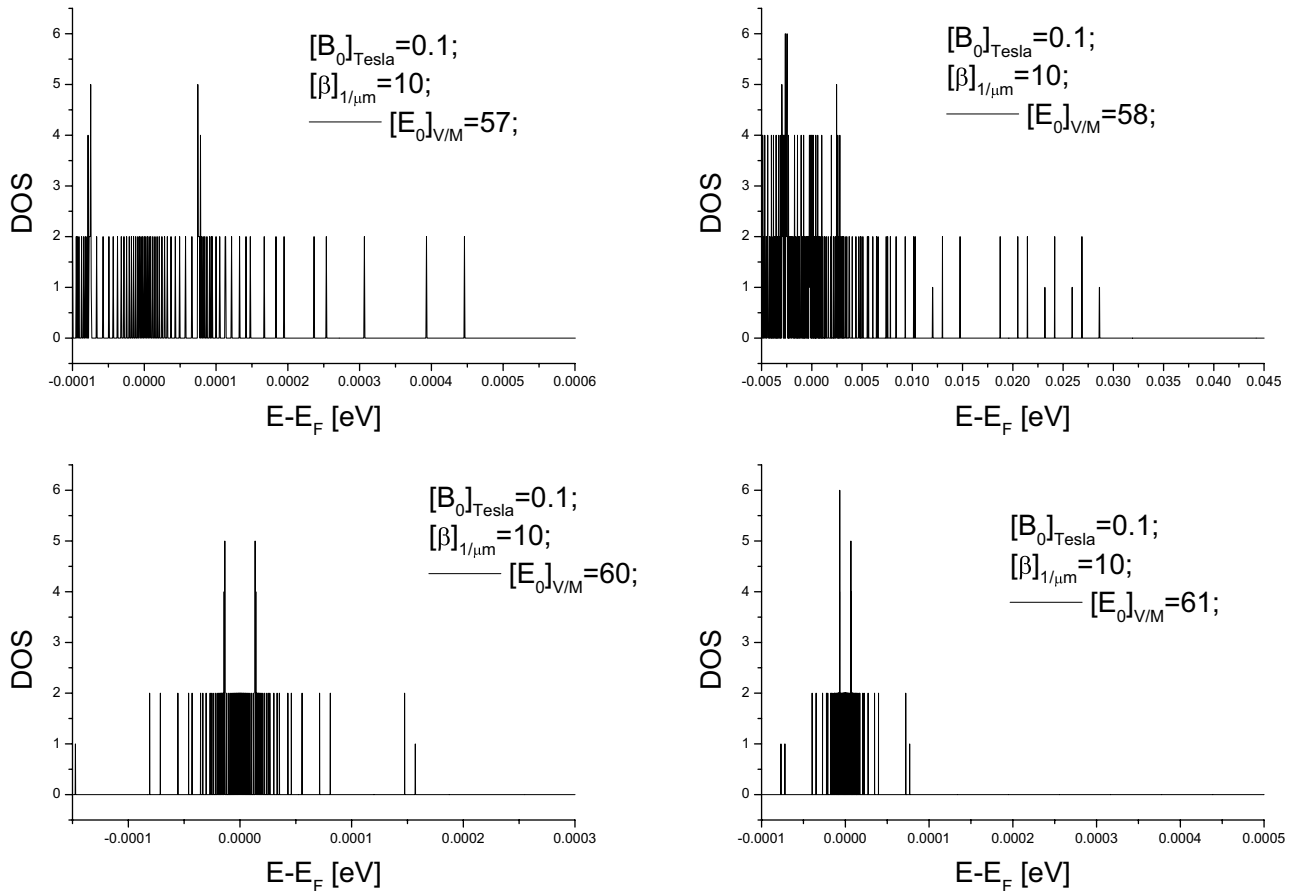
Figure 3. Example illustrating the dependence of the pseudogap width of DOS curve on controllable parameters.

ES of a quantum system determines its values of key parameters in rate equations of radiation generation. For example, the spontaneous emission lifetime $t_{spon,12}$ is defined as

$$1/t_{spon,12} = A_{12} \equiv \frac{e^2 \omega_0^3 n^3}{3\pi c^3 \hbar \varepsilon} |<\psi_1|r|\psi_2>|^2, \tag{26}$$

where $\hbar\omega_0 \equiv E_2 - E_1 > 0$, and other parameters are of respective customary definitions[48]. The value of A_{12} depends on not only energies of two levels, E_2 and E_1, but also space shapes of their wavefunctions, $|\psi_2|$ and $|\psi_1|$.

It is well-known that the dipole matrix element of an ES is proportional to the characteristic scale of wavefunctions of the ES, $|<\psi_1|r|\psi_2>| \equiv \xi^{-3} \int |\psi_1(r/\xi)|r|\psi_2(r/\xi)|d^3r \sim \xi$. For small quantum systems such as atoms and molecules, their wavefunctions have small space extension or highly localized. For atomic wavefunctions, these exist $|\psi^{atom}| \sim exp(-r/\xi)$ and $\xi \sim 0.053$ nm or Bohr radius (see many textbooks). As a consequence, the value of dipole matrix element of an atom $|<\psi_1|x,y,z|\psi_2>| \sim \xi$ will be about at nm-level and hence corresponding $t_{spon,12}$-values, for $|E_2 - E_1|$ at eV-level, is $\sim 10^{-9}$ s (see Eq. 8.3.12 at page 168 of Ref.[50]). Similar results hold for small molecules. Moreover, in semiconductor quantum well and superlattice, wavefunctions of carriers are of larger ξ-value, usually tens nm, than atomic and molecular wavefunctions. Impurity levels also have larger ξ-value because of larger effective electronic mass in Hydrogen-like model on impurity. In contrast, extended-state wavefunctions in the ES of a metal is of larger ξ-values, as well as larger values of dipole matrix element, than those in quantum well and superlattice.

The wavefunctions' shape also affects electron–phonon coupling constant. Electron–phonon interaction can be represented by a perturbation Hamiltonian density[45]

$$\begin{aligned} H_{e-ph}^1 &= u \cdot \nabla V(r - R_n) = \sum_n d_n \cdot \nabla V(r - R_n) \\ &= A_q \sum_n cos(q \cdot R_n - \omega t)e \cdot \nabla V(r - R_n), \end{aligned} \tag{27}$$

where $d_n = A_q cos(q \cdot R_n - \omega t)$ is the displacement of the n-th atom and V is atomic potential. Corresponding scattering probability, or electron–phonon coupling constant $g_{1,2}$, reads[45]
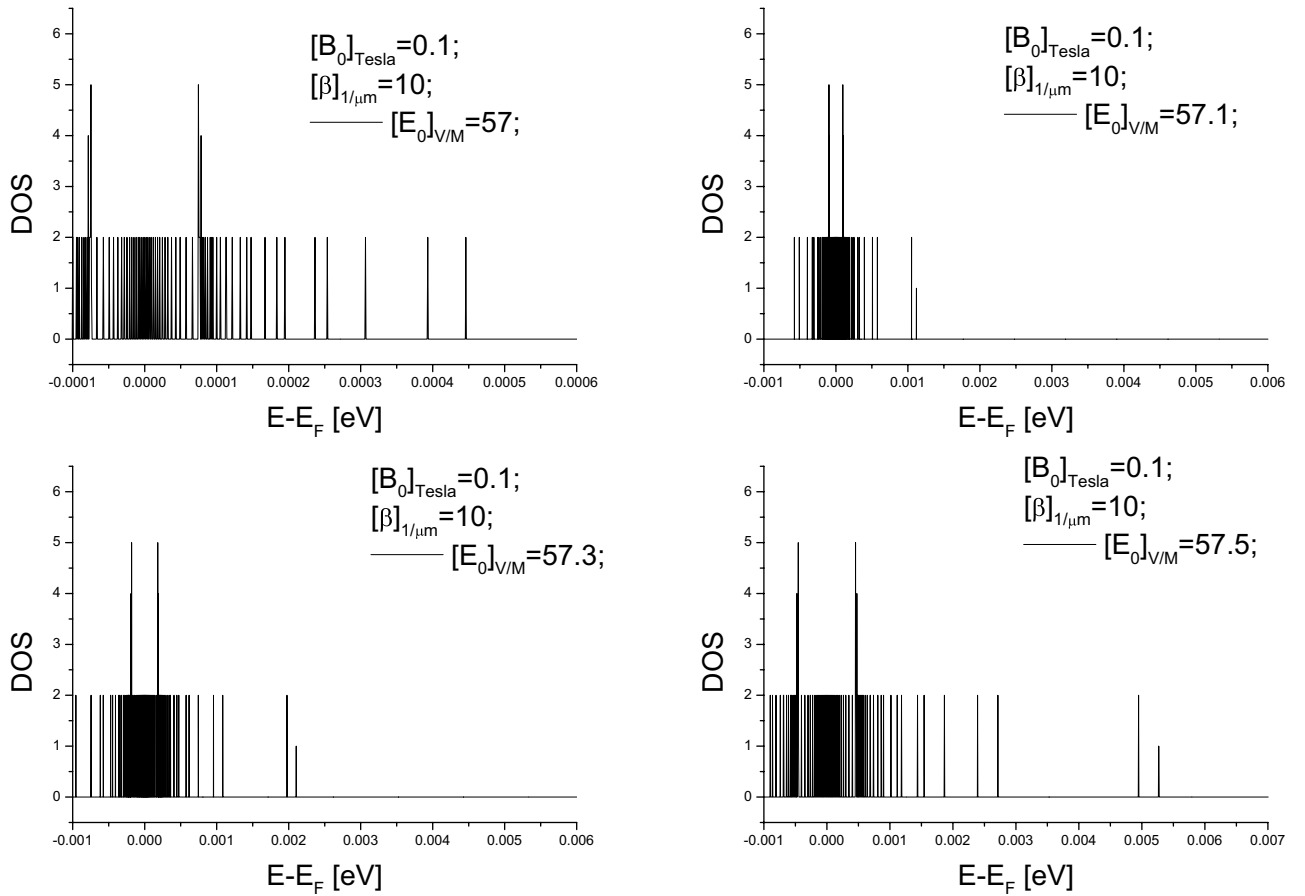
6

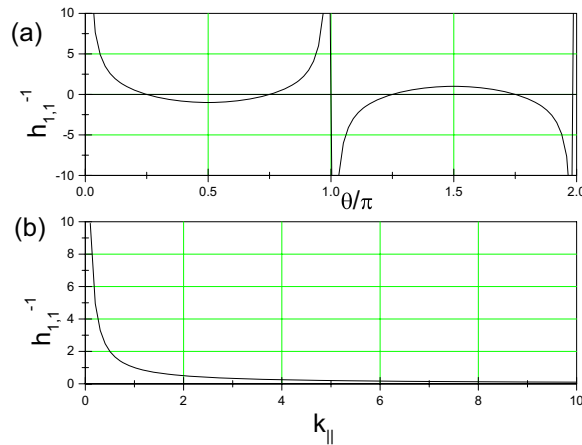**Figure 4.** Example illustrating meV-level pseudogap width.

Figure 5. (**a**) $h_{1,1}^{-1}$ VS $\theta \equiv arctan(k_y/k_x)$ under a given $k_\parallel \equiv \sqrt{k_x^2 + k_y^2}$ and (**b**) $h_{1,1}^{-1}$ VS $k_\parallel$ under a given θ, where $h_{1,1} \equiv \left[\partial_x \partial_y ln\phi\right]|_{x=0,y=0}$ and hence $h_{1,1}$ is proportional to the dipole moment $\int |\phi|^2 x dx dy$ or $\int |\phi|^2 y dx dy$.

$$
\begin{aligned}
M_{2,1} &= g_{2,1}^2 \equiv |<\psi_1|u \cdot \nabla V_i|\psi_2>|^2 \\
&= \frac{2\pi^2}{\hbar}\left\{|<2|\frac{A_q}{2}\sum_n exp(iq \cdot R_n)e \cdot \nabla V(r - R_n)|1>|^2\delta\left[E_{k'-E_k-\hbar\omega}\right]\right. \\
&\quad \left. +|<2|\frac{A_q}{2}\sum_n exp(-iq \cdot R_n)e \cdot \nabla V(r - R_n)|1>|^2\delta\left[E_{k'-E_k+\hbar\omega}\right]\right\} \\
&= M_{1,2};
\end{aligned}
\tag{28}
$$

where phonon number A_q^2 can be obtained from Bose–Einstein distribution function: $g_{BE}(\omega_q; T) \equiv \frac{1}{exp(\hbar\omega_q/K_BT)-1}$. Clearly, $M_{1,2}$ depends on phonon number A_q^2 and space shapes of $|\psi_1>$ and $|\psi_2>$.

Although both the electron–phonon coupling constant and the spontaneous emission lifetime are affected by the space shapes of wavefunctions, their dependencies on the shapes are different. This implies that with the ES changing (i.e., shapes of wavefunctions changing), the ratio among these coefficients in rate equations of radiation generation (see below) also change and hence so do its solutions[51,52]. Moreover, there is difference between the scattering by longitudinal-wave acoustic phonon $u_{lo} = exp(iq_xx + i\omega_q^{lo}t)e_x$ and that by transverse-wave one $u_{tr} = exp(iq_xx + i\omega_q^{tr}t)e_y$. In the former case, $u \cdot \nabla V \sim x * x$ while in the latter case, $u \cdot \nabla V \sim x * y$ (because ionic potential V is a function of $x^2 + y^2 + z^2$). Thus, the difference between two electron–phonon coupling constants is dependent on shapes of wavefunctions

$$g_{2,1}^{tr} \sim \int |\psi_1|iqxy|\psi_2|dxdy; \tag{29}$$

$$g_{2,1}^{lo} \sim \int |\psi_1|iqx^2|\psi_2|dxdy; \tag{30}$$

Likewise, the spontaneous emission coefficient A_{12} also depends on shapes of wavefunctions in a different manner

$$A_{12} \sim \int |\psi_1|\sqrt{x^2 + y^2}|\psi_2|dxdy. \tag{31}$$

The rate equations of radiation generation based on 2-level read

$$d_tP_2 = P_1P_2 * \sigma - A * P_2; \tag{32}$$

$$d_tP_1 = - P_1P_2 * \sigma; \tag{33}$$

where P_2 and P_1 are populations of two levels. Here, because conventional components for achieving amplified radiation, such as reflecting mirrors or resonant cavity[52], are not used, there is no feedback usage of generated radiation and hence the population inversion is merely due to the phonon pump, i.e., not amplified. The output radiation is indeed an enhanced spontaneous emission from an inverted equilibrium population. Because no radiation is input, such an enhanced spontaneous emission can also be viewed as an infinite-fold amplification, i.e., the enhancement in spontaneous emission is viewed as infinitely amplifying 0 input radiations. Therefore, we borrow the phrase "stimulated emission", despite not very accurate, to refers to this enhancement in spontaneous emission.

Equations (32, 33) can be derived in following way. If phonon scattering is not taken into account, populations of two levels are

$$P_2^0 = N(E_2) * g_{FD}(E_2; T); \tag{34}$$

$$P_1^0 = N(E_1) * g_{FD}(E_1; T), \tag{35}$$

where $N(E_2)$ and $N(E_1)$ are DOS of two levels, respectively. During Δt the number of electrons transiting from E_2 to E_1, due to phonon-scattering, is[45]

$$b = g_{FD}(E_2; T) * N(E_2) * M_{2,1} * N(E_1) * (1 - g_{FD}(E_1; T)) * \Delta t. \tag{36}$$

where the factor $(1 - g_{FD}(E_1; T))$ represents un-occupation probability of the E_1-state and hence means that the transition is prohibited if the E_1-state has been occupied. Likewise, during Δt the number of electrons transiting from E_1 to E_2 is[45]

$$a = g_{FD}(E_1; T) * N(E_1) * M_{1,2} * N(E_2) * (1 - g_{FD}(E_2; T)) * \Delta t. \tag{37}$$

Therefore, after taking into account the phonon scattering, populations of two levels read

$$P_2 = P_2^0 - b + a; \tag{38}$$

$$P_1 = P_1^0 + b - a, \tag{39}$$

where

$$\frac{a - b}{\Delta t} = N(E_1) * M_{1,2} * N(E_2) * g_{FD}(E_2) * g_{FD}(E_1)* $$
$$* \left[exp(E_2/K_BT) - exp(E_1/K_BT) \right] > 0 \tag{40}$$

Equations (38, 39) can be re-written as Eqs. (32, 33) if introducing

$$\sigma = M_{1,2} * \left[exp(E_2/K_BT) - exp(E_1/K_BT) \right] \tag{41}$$

.

Radiation generation from 3-level model and 2-level model.

Many authors have taken into account phonon scattering in radiation generation from semiconductor[36,43,51]. For metal film held by dielectric "holder", environment temperature can be felt by the film due to the connection between the "holder" and vacuum chamber wall. Because of the difference between g^{tr} and g^{lo}, radiation generation is easy to involve in three levels: one from a state below the pseudogap and two above. An electron initially at a state below-pseudogap is scattered by a high-energy longitudinal-wave phonon to a state above-pseudogap and then is scattered by a low-energy transverse-wave phonon to another state above-pseudogap, and then return to the state below-pseudogap by emitting a photon. Because both phonon and photon have linear dispersion relation and their phase velocities satisfy $c_{tr} \equiv \frac{\omega^{tr}}{q} < c_{lo} \equiv \frac{\omega^{lo}}{q} < c$ (because of different values of elastic constants for longitudinal-wave and transverse-wave[53]). As illustrated in Fig. 6, this inequality relation $\frac{\omega^{tr}}{q} < \frac{\omega^{lo}}{q} < c$ can warrant a closed energy–momentum conservation process among phonon and photon. Moreover, two electronic states within a same parabolic band cannot ensure the difference between their momentum and that between their energy to satisfy a linear dispersion relation hold by photon, but two states from different parabolic bands can. These features well illustrate why the beating between longitudinal-wave phonon and transverse-wave one can allow an inter-band transition. Why this beating cannot cause a 3D Fermi sphere ES to generate photon can also be explained (because any two electronic states are in a same parabolic band $E_k \sim k^2$). Moreover, as previously mentioned, because different parabolic bands correspond to different 2D shaped-parts of 3D wavefunctions, losing orthogonality among 2D shaped-parts also favor the dipole matrix elements for the transition $(k_z + q, m) \rightarrow (k_z, n)$ to be $\neq 0$. Namely, there are two prohibitions on the dipole radiation generation, one is the orthogonality of 3D wavefunctions and the other is parabolic dispersion relation of the 1D plane-wave part of 3D wavefunctions.

Because sound speed in many metals can be $10^{3\sim4}$ m/s-level, if $c_{lo} = 5 \times 10^3$ m/s and $c_{tr} = 3 \times 10^3$ m/s, there will be $\frac{k_3}{q_2} = \frac{c_{lo} - c_{tr}}{c - c_{tr}} \sim 2 \times 10^{-5}$ and hence yielding a photon $\hbar k_3 c \sim$ meV will need a phonon $\hbar q_2 c_{lo} \sim \frac{c - c_{tr}}{c_{lo} - c_{tr}} \frac{c_{lo}}{c}$ meV ~ 2.5 meV , which can be supplied sufficiently for most metals in room-temperature environment 300 K.

Rate equations of radiation generation based on 3-level model read

$$d_t P_3 = P_1 P_3 * \sigma_{13} - P_3 P_2 * \sigma_{23}; \tag{42}$$

$$d_t P_2 = P_3 P_2 * \sigma_{23} - A * P_2; \tag{43}$$

$$d_t P_1 = - P_1 P_3 * \sigma_{13}. \tag{44}$$

As shown in Eq. (28), the dependence of $M_{1,2}$ on phonon strength $A_q^2 \sim g_{BE}(\omega_q; T)$ determines $A_{k_3-k_1}^2 < A_{k_3-k_2}^2$ (because $g_{BE}(\omega_{q_{13}}; T) < g_{BE}(\omega_{q_{23}}; T)$ if $q_{13} > q_{23}$), and hence $M_{1,3} < M_{2,3}$ or $\sigma_{13} < \sigma_{23}$.

As previously mentioned, eV-level gap between atomic levels can lead to $A \sim 10^9$ s^{-1}[50]. Thus, even if assuming meV-level gap exists in atomic ES, corresponding A will be $\sim 10^0$ because of $A \sim \omega_0^3$ as shown in Eq. (26). Luckily, in the ES above-described, the ξ-value is $\sim 2\pi/\sqrt{k_x^2 + k_y^2} \sim \mu$m, about 10^5-folds of that of atomic one. Thus, the A-value in the new ES can reach $\sim 10^{10}$ s^{-1}. Moreover, for many metals with a Fermi sphere ES, phonon scattering can lead to a relaxation-time about $10^{-13} \sim 10^{-14}$ s or $\sigma \sim 10^{13} \sim 10^{14}$ s^{-1}[45]. The new ES enables these σ-values to be kept at this level.

Actually, above-mentioned two characteristic time scales: $1/A$ and $1/\sigma$, represent respectively the strength of electron-photon interaction/coupleing and that of electron-phonon one, which are expectation values of two perturbation energies on electrons $\int \psi_{k+q}^* H_{e-photon} \psi_k d^3 r$ and $\int \psi_{k+q}^* H_{e-phonon} \psi_k d^3 r$, where $H_{e-photon} = eE \cdot r$ and $H_{e-phonon} = u \cdot \nabla V$. Here, under room temperature environment, ionic displacement u can be estimated, according to usual values of elastic constant K of metals $\sim 10^2$ GPa $= 10^{11}$ J/m^3[44] for yielding elastic energy $\frac{1}{2} K a^2 * u \sim 1$ meV $= 1.6 \times 10^{-22}$ J, where $a \sim 100 pm$ represents the size of a cell, to be $\sim 10^{-2} a \sim 1 pm$. Such a displacement, 10^{-2} folds of inter-atom distance a, can correspond to a 10 meV-level energy perturbation on an electron because the inner electric field V/a in most metals can be at $V/[100 pm]$-level. Namely, for most metals in 300K environment, $\int \psi_{k+q}^* H_{e-photon} \psi_k d^3 r < \int \psi_{k+q}^* H_{e-phonon} \psi_k d^3 r$, especially $\int \psi_{k+q}^* H_{e-THz} \psi_k d^3 r << \int \psi_{k+q}^* H_{e-phonon} \psi_k d^3 r$, can be warranted.

Note that the energy increment of an excited electron is not the kinetic energy of a phonon, instead, it is the variation of Coulomb energy associated with the ionic displacement, $u \cdot \nabla V_c$, that contribute this energy increment. Although the phonons number is not too large, electronic excitation is still very obvious. Therefore, it is not necessary to seek for applying large-amplitude THz accoustic wave on the metal or inputting large mount of phonons. This greatly lessen experiment condition requirement. Otherwise, we have to seek for a powerful THz mechanics vibration source whose difficulty is illustrated as follows: to input THz accoustic wave into the metal, we utilize the well-known formula reflecting cantilever transverse oscillation frequency v inverse proportional to its extruded length L (see page 120–121 of ref.[53]), i.e. $v = \frac{\pi^2}{L^2} \sqrt{\frac{EI}{\rho S}}$, for a typical metal material with following representative values: elastic module $E = 2 \times 10^{11}$ Pa $= 2 \times 10^{11}$ kg $\times \frac{1}{m} \times \frac{1}{s^2}$, cross-section area $S = 1 \times 10^{-6}$ m^2, density $\rho = 2700$ kg/m^3, moment of inertia $I = ab^3/12$, $I/S \sim$ m^2, extruding length $L = 10^{-3}$ m,there will be $v \sim$ MHz. Namely, it is difficult to make an object whose material is common and size is in a macroscopic scale ($> \mu$m) to do a vibration with a frequency at THz. A man-made macroscopically-sized THz mechanics vibration source is nearly not available.

But the metal's "native" THz phonons are of sufficient mount at 300 K. Because the number of phonons at a specified frequency v is the product of Bose–Einstein distribution function g_{BE} at v and phonon DOS at v:
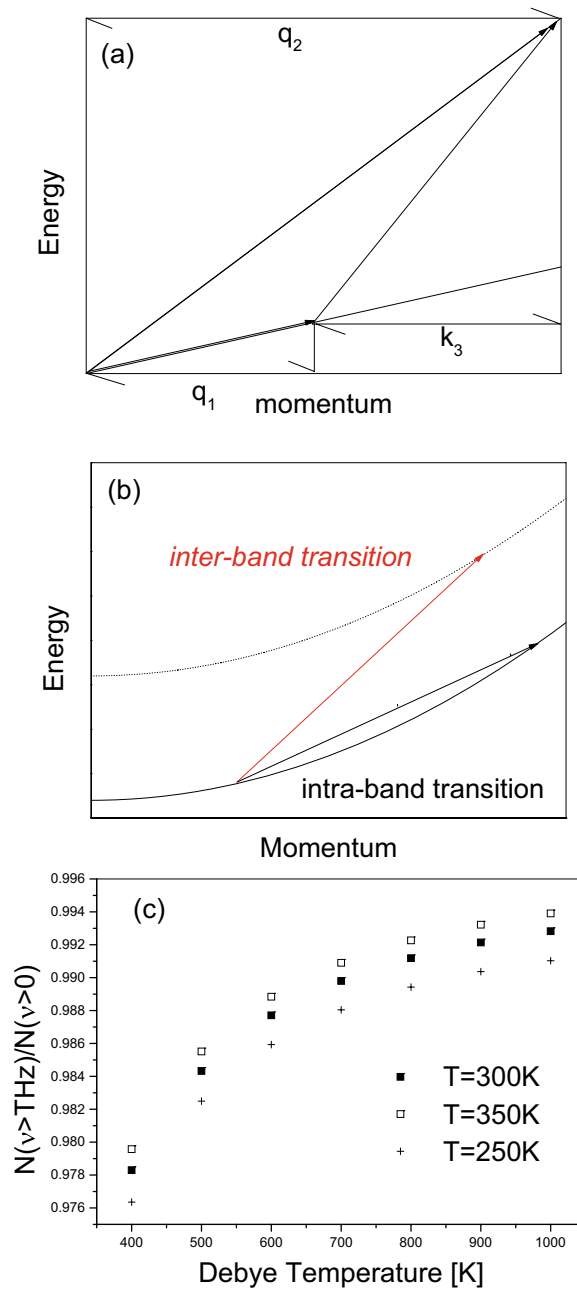
Figure 6. (**a**) A sketch illustrates that three waves with linear dispersion relation can form a closed energy–momentum conservation process. The slope of each straight-line segment represents its phase velocity. Because of $c_{tr} < c_{lo} < c$, $q_2 = q_1 + k_3$ and $c_{lo}q_2 = \omega_2 = \omega_1 + E_k = c_{tr}q_1 + ck_3$ can be satisfied, where c_{lo} and c_{tr} are phonon's phase velocities and c is light velocity. (**b**) A sketch illustrates that a transition between two electronic states k_4, k_5 within a same parabolic band cannot ensure the momentum difference $k_5 - k_4 = k_3$ and the energy difference $E_5 - E_4 = k_5^2 - k_4^2$ to satisfy a linear dispersion relation $\frac{E_5 - E_4}{k_5 - k_4} = const$, which can be ensured by a transition between two states from different parabolic bands or $E_5 = E_4 + 2k_3 * k_4 + k_3^2 + E_g$. (**c**) The dependence of the ratio between the number of phonons whose frequencies > 1 THz and total number of phonons on Debye temperature.

$n(\nu) = \frac{1}{\exp\left(\frac{\hbar\nu}{K_B T}\right) - 1} * \nu^2$, although low-energy phonon have very large g_{BE}-value, their low DOS-value ν^2 determine their number are near 0 at the $\nu \sim 0$ end. For most metals, their Debye temperature values Θ, (or their highest phonon frequencies), are $200 \sim 400$ K[44,45], thus, the total phonon number $\int_0^{\Theta K_B/T} n(\nu)d\nu$ can have a large fraction $\int_{1THz}^{\Theta K_B/T} n(\nu)d\nu$ whose frequencies are $> 1THz$, i.e. $\frac{\int_{1THz}^{\Theta K_B/T} n(\nu)d\nu}{\int_0^{\Theta K_B/T} n(\nu)d\nu} \gg 0$ if $T = 300$ K and $\Theta = 400$ K or higher. Fig. 6c indicates that this ratio is close enough to 1 at some typical situations. Moreover, the frequency
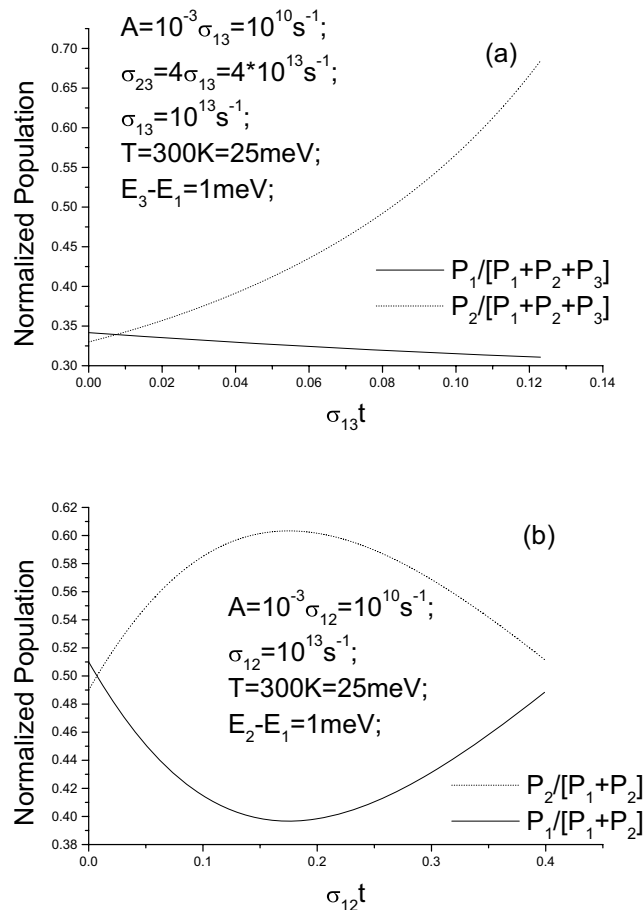
Figure 7. Examples of solutions of laser rate equations. (**a**) for 3-levels model and (**b**) for 2-levels model.

corresponded by maximum phonon DOS in many metals can be up to tens meV[44,45], therefore a metal's "native" THz phonons are in sufficient mount for pumping electrons.

By choosing suitable values of controllable parameters, rate equations for 3-level model can yield solutions satisfying population inversion. This can be referred to Fig. 7. Actually, even for 2-level model, suitable values of controllable parameters can also warrant population inversion even for 2-level model (see Fig. 7).

The output power of a mono-color dipole oscillation $e * d * sin(\omega t)$ is $P_w = \frac{e^2 d^2 \omega^4}{12\pi\varepsilon_0 c^3} = 2.84 \times 10^{-54} * [d]_m[\omega]_{Hz} Watt$, where $\varepsilon_0 = 8.85 \times 10^{-12} F/m$, $c = 3 \times 10^8 m/s$, $e = 1.6 \times 10^{-19} C$, dimensionless parameter $[d]_m$ is the value of d (dipole magnitude) in unit of meter, and $[\omega]_{Hz}$ is defined likewise. Usually, free electron density in a simple metal is at $10^{22\sim23}$ cm^{-3}-level, which corresponds to an inter-atom interval at nm-level. Because THz frequency is non-transparent relative to metal medium with such a $10^{22\sim23}$ cm^{-3}-level density , it is only a thin layer near the surface that can have contribution to THz output. For a film 1 cm $\times$ 1 cm $\times$ 0.01 cm, the thickness 0.01 cm is at the level of $\lambda_{THz} = 0.03$ cm, the wavelength of THz photon, and hence it is suitable to view the layer of a 0.01 cm-thickness as the "skin" layer of a THz EM wave.

Thus, a metallic film 1 cm $\times$ 1 cm $\times$ 10^{-2} cm can contain $10^{20\sim21}$ conduction-band electrons. For the new ES, states whose energies are within the range $[E_F - 0.05 * K_B T, E_F + 0.05 * K_B T]$ have obvious contribution to various quantum transition processes, and the number of these states is estimated to be proportional to $0.1 * K_B T/E_F$, which is ~ 2.5 meV/10 eV $\sim 10^{-4}$ for metals in room-temperature environment 300 K, and hence is $\sim 10^{16\sim17}$. Of course, if the width of the range is $0.01 * K_B T$, the number of these states will be $\sim 10^{15\sim16}$, still higher several orders of magnitude than that of carriers in semiconductor. According to above formula, it is feasible to an output power at *Watt*-level from a film 1 cm $\times$ 1 cm $\times$ 10^{-2} cm which can contribute 10^{15} electrons to participate oscillation with a *THz* frequency and hence yield $10^{15} *$ meV $= 1.6 \times 10^{-7}$ J energy over a time-scale at *ps*-level.

Moreover, it is easy to understand that intraband carrier acceleration by radiated THz electric field can lead to significant THz absorption and hence decrease the THz output power. Luckily, the THz-active layer is only of a thickness below the THz wavelength, ~ 0.01 cm $< \lambda_{THz}$. Before $E_{x,y}$ is significantly depleted by intraband acceleration, it has transmitted the thin layer. Moreover, because space-asymmetric wavefunctions' shape, intraband carrier acceleration to higher k_x-state might imply larger THz dipole moment and hence does not definitely imply "significant THz absorption". A definite answer will be the task of future works.

Beside, the coupling loss to free-space THz radiation is also a factor affecting THz output power. For outstanding main purpose, this work avoids to consider various complicated secondary effects/mechanisms affecting THz output power. For example, intraband carrier acceleration above-mentioned can "heat" carriers and

then "heat" ions, through carrier-ion collisions, and cause phonons energy rising and then in turn cause more THz EM generation. The main purpose is on the feasibility of fields-manipulating ES for achieving a specified wavelength EM output. These secondary effects/mechanisms will be content of future works.

Summary

Fields-manipulating the ES of a metal enables high-density free electrons to form a new "volatile" ES, which can couple with acoustic phonon to achieve population inversion between states spaced by meV-level pseudogap. The fact that mobile electrons in metal is higher several orders of magnitude than that in semiconductor determines the technical scheme presented here to be of marked application prospect.

Data availability

All data generated or analysed during this study are included in this published article (and its Supplementary information files).

Appendix

Detailed formulas are presented here. From Eqs. (10–12, 19, 20), we have

$$\partial_y ln\phi = \frac{s+1}{2s}\partial_u ln\phi + \frac{s-1}{2s}\partial_v ln\phi; \tag{45}$$

$$\partial_x ln\phi = \frac{s-1}{2s}\partial_u ln\phi - \frac{s+1}{2s}\partial_v ln\phi; \tag{46}$$

$$\frac{1}{2}\partial_u[N-M] - \frac{1}{2s}\partial_u[N+M] - \frac{1}{2}\partial_v[N+M] - \frac{1}{2s}\partial_v[N-M]$$
$$= iws\frac{s+1}{s^2+1}u + iws\frac{s-1}{s^2+1}v, \tag{47}$$

Substituting Eqs. (45, 46) back into Eqs. (19, 20), we have

$$\partial_u ln\phi = s\frac{s+1}{s^2+1}\big[N - iwxy\big] + s\frac{s-1}{s^2+1}M; \tag{48}$$

$$\partial_v ln\phi = -s\frac{s-1}{s^2+1}\big[N - iwxy\big] + s\frac{s+1}{s^2+1}M; \tag{49}$$

which "$\Longrightarrow$"

$$\partial_v\left\{\frac{s+1}{s^2+1}\big[N - iwxy\big] + \frac{s-1}{s^2+1}M\right\} = \partial_u\left\{-\frac{s-1}{s^2+1}\big[N - iwxy\big] + \frac{s+1}{s^2+1}M\right\} \tag{50}$$

$\Longrightarrow$

$$\partial_v\left\{\frac{s+1}{s^2+1}N + \frac{s-1}{s^2+1}M\right\} + \partial_u\left\{\frac{s-1}{s^2+1}N - \frac{s+1}{s^2+1}M\right\} = iw\left[\frac{s+1}{s^2+1}\partial_v(xy) + \frac{s-1}{s^2+1}\partial_u(xy)\right]; \tag{51}$$

$\Longrightarrow$

$$s\partial_v(N+M) + \partial_v(N-M) + s\partial_u(N-M) - \partial_u(N+M)$$
$$= iw\{[s+1]\left[-\frac{(s^2-1)s^2}{(s^2+1)^2}2v - \frac{2s^3}{(s^2+1)^2}2u\right] + [s-1]\left[\frac{(s^2-1)s^2}{(s^2+1)^2}2u - \frac{2s^3}{(s^2+1)^2}2v\right]\}$$
$$= \frac{2iw}{(s^2+1)^2}\{-\big[(s^2-1)s^2(s+1) + 2s^3(s-1)\big]v + \big[(s^2-1)s^2(s-1) - 2s^3(s+1)\big]u\} \tag{52}$$
$$= \frac{2iw}{(s^2+1)^2}\{\big[(1-s)(s^4 + 4s^3 + s^2)\big]v + \big[(1+s)(s^4 - 4s^3 + s^2)\big]u\}$$

From Eq. (47), its terms $\sim$ constant meet

$$\frac{1}{2}f_{1,0} - \frac{1}{2s}g_{1,0} - \frac{1}{2}g_{0,1} - \frac{1}{2s}f_{0,1} = 0; \Longrightarrow \big[sf_{1,0} - g_{1,0}\big] - \big[sg_{0,1} + f_{0,1}\big] = 0; \tag{53}$$

its $\sim u^1$ terms meet

$$\frac{1}{2}2f_{2,0} - \frac{1}{2s}2g_{2,0} - \frac{1}{2}g_{1,1} - \frac{1}{2s}f_{1,1} = iws\frac{s+1}{s^2+1};$$
$$\implies \left[f_{2,0} - \frac{1}{s}g_{2,0}\right] - \left[\frac{1}{2}g_{1,1} + \frac{1}{2s}f_{1,1}\right] = +iws\frac{s+1}{s^2+1}; \tag{54}$$

its $\sim v^1$ terms meet

$$\frac{1}{2}f_{1,1} - \frac{1}{2s}g_{1,1} - \frac{1}{2}2g_{0,2} - \frac{1}{2s}2f_{0,2} = iws\frac{s-1}{s^2+1};$$
$$\implies \left[g_{0,2} + \frac{1}{s}f_{0,2}\right] - \left[\frac{1}{2}f_{1,1} - \frac{1}{2s}g_{1,1}\right] = -iws\frac{s-1}{s^2+1}; \tag{55}$$

Likewise, from Eq. (52), its terms $\sim$ constant meet

$$\left[sg_{0,1} + f_{0,1}\right] + \left[sf_{1,0} - g_{1,0}\right] = 0; \tag{56}$$

its $\sim u^1$ terms meet

$$\left[sg_{1,1} + f_{1,1}\right] + \left[s*2f_{2,0} - 2g_{2,0}\right] = \frac{2iw}{\left(s^2+1\right)^2}\left[(s+1)\left(s^4 - 4s^3 + s^2\right)\right]; \tag{57}$$

its $\sim v^1$ terms meet

$$\left[s*2g_{0,2} + 2f_{0,2}\right] + \left[sf_{1,1} - g_{1,1}\right] = -\frac{2iw}{\left(s^2+1\right)^2}\left[(s-1)\left(s^4 + 4s^3 + s^2\right)\right]; \tag{58}$$

For the dispersion relation Eq. (21), its $\sim u^1$ terms

$$f_{1,1} + f_{0,0}f_{1,0} + 2g_{2,0} + g_{0,0}g_{1,0} = 0; \tag{59}$$

its $\sim v^1$ terms

$$g_{1,1} + g_{0,0}g_{0,1} + 2f_{0,2} + f_{0,0}f_{0,1} = 0; \tag{60}$$

$\partial_x Eq.(21)$ and $\partial_y Eq.(21)$ read

$$\frac{s-1}{2s}\left[\partial_{uu}(N+M) + (N+M)\partial_u(N+M)\right]$$
$$- \frac{s+1}{2s}\left[\partial_{vv}(N-M) + (N-M)\partial_v(N-M)\right] = 0; \tag{61}$$

$$\frac{s+1}{2s}\left[\partial_{uu}(N+M) + (N+M)\partial_u(N+M)\right]$$
$$- \frac{s-1}{2s}\left[\partial_{vv}(N-M) + (N-M)\partial_v(N-M)\right] = 0; \tag{62}$$

which means

$$\left[\partial_{uu}(N+M) + (N+M)\partial_u(N+M)\right] = 0; \tag{63}$$

$$\left[\partial_{vv}(N-M) + (N-M)\partial_v(N-M)\right] = 0; \tag{64}$$

whose constant terms satisfies

$$2g_{2,0} + g_{0,0}g_{1,0} = 0; \tag{65}$$

$$2f_{0,2} + f_{0,0}f_{0,1} = 0. \tag{66}$$

Moreover, because of

$$2ik_x = g_{0,0} - f_{0,0}; \quad 2ik_y = g_{0,0} + f_{0,0}, \tag{67}$$

there are

$$g_{0,0} = i\left(k_y + k_x\right); \quad f_{0,0} = i\left(k_y - k_x\right). \tag{68}$$

12 equations and formulas, (10 equations in Eqs. (53–60, 65–66) and 2 formulas in Eq. (68)), enable us to determine electronic structure through following procedure. Equations (53, 56) lead to

$$\left[sf_{1,0} - g_{1,0}\right] = 0 = \left[sg_{0,1} + f_{0,1}\right]; \tag{69}$$

Equations (54, 57, 65) lead to

$$\left[f_{2,0} - \frac{1}{s}g_{2,0}\right] = \frac{iw}{2}\frac{s+1}{\left(s^2+1\right)^2}\left[s^3 - 4s^2 + s + s^3 + s\right] = \frac{iw}{2}\frac{s+1}{\left(s^2+1\right)^2}\left[2s^3 - 4s^2 + 2s\right]; \tag{70}$$

$$\left[\frac{1}{2}g_{1,1} + \frac{1}{2s}f_{1,1}\right] = \frac{iw}{2}\frac{s+1}{\left(s^2+1\right)^2}\left[s^3 - 4s^2 + s - s^3 - s\right] = \frac{iw}{2}\frac{s+1}{\left(s^2+1\right)^2}\left[-4s^2\right]; \tag{71}$$

Equations (55, 58, 66) lead to

$$\left[g_{0,2} + \frac{1}{s}f_{0,2}\right] = -\frac{iw}{2}\frac{s-1}{\left(s^2+1\right)^2}\left[s^3 + 4s^2 + s + s^3 + s\right] = -\frac{iw}{2}\frac{s-1}{\left(s^2+1\right)^2}\left[2s^3 + 4s^2 + 2s\right]; \tag{72}$$

$$\left[\frac{1}{2}f_{1,1} - \frac{1}{2s}g_{1,1}\right] = -\frac{iw}{2}\frac{s-1}{\left(s^2+1\right)^2}\left[s^3 + 4s^2 + s - s^3 - s\right] = -\frac{iw}{2}\frac{s-1}{\left(s^2+1\right)^2}\left[4s^2\right]; \tag{73}$$

Equations (71, 73) yield

$$\begin{aligned}
f_{1,1} &= \frac{iws^2}{\left[s^2+1\right]^3}\left\{[s+1][-4s] - [s-1]\left[4s^2\right]\right\} \\
&= \frac{4iws^2}{\left[s^2+1\right]^3}\left\{-s^3 - s\right\} = -\frac{4iws^3}{\left[s^2+1\right]^2};
\end{aligned} \tag{74}$$

$$\begin{aligned}
g_{1,1} &= \frac{iws^2}{\left[s^2+1\right]^3}\left\{[s+1]\left[-4s^2\right] + [s-1][4s]\right\} \\
&= \frac{4iws^2}{\left[s^2+1\right]^3}\{-s^3 - s\} = -\frac{4iws^3}{\left[s^2+1\right]^2}.
\end{aligned} \tag{75}$$

Moreover, Equations (59, 60, 65, 66) lead to (the last "=" is due to Eq. (68))

$$f_{1,1} = -f_{0,0}f_{1,0} = -i\left(k_y - k_x\right)f_{1,0}; \tag{76}$$

$$g_{1,1} = -g_{0,0}g_{0,1} = -i\left(k_y + k_x\right)g_{0,1}, \tag{77}$$

Equations (69, 76, 77) yield

$$g_{1,0} = sf_{1,0} = \frac{is}{k_y - k_x}f_{1,1}; \tag{78}$$

$$f_{0,1} = -sg_{0,1} = -\frac{is}{k_y + k_x}g_{1,1}. \tag{79}$$

Substituting Eqs. (68, 78, 79) into Eq. (21), we have

$$\begin{aligned}
E_k/\alpha &= k_z^2 + g_{1,0} + \frac{1}{2}g_{0,0}^2 + f_{0,1} + \frac{1}{2}f_{0,0}^2 \\
&= k_z^2 + k_y^2 + k_x^2 + \frac{is}{k_y - k_x}f_{1,1} - \frac{is}{k_y + k_x}g_{1,1}.
\end{aligned} \tag{80}$$

Because s is an imaginary number, $f_{1,1}$ and $g_{1,1}$ are therefore complex-valued (see whether Eqs. (74, 75) contain terms $\sim s^{odd}$). Thus, $ImE_k = 0$ will demand

$$0 = \frac{Imf_{1,1}}{k_y - k_x} - \frac{Img_{1,1}}{k_y + k_x}; \tag{81}$$

$$\Longrightarrow$$

$$0 = \frac{0}{k_y - k_x} - \frac{0}{k_y + k_x}; \tag{82}$$

The energy shift, or so-called "self-energy", reads

$$ReE_k/\alpha - \left[k_z^2 + k_\parallel^2\right] = \frac{is}{k_y - k_x}Ref_{1,1} - \frac{is}{k_y + k_x}Reg_{1,1}$$

$$= \frac{4ws^4}{\left[s^2 + 1\right]^2} * \left[\frac{1}{k_y - k_x} - \frac{1}{k_y + k_x}\right] = \frac{8ws^4}{\left[s^2 + 1\right]^2} * \left[\frac{k_x}{k_y^2 - k_x^2}\right] \quad (83)$$

$$= \frac{8ws^4}{\left[s^2 + 1\right]^2} * \frac{1}{|k_\parallel|} * \frac{cos(\theta)}{cos(2\theta)}.$$

Note that because of $s^2 < 0$, the energy shift can be $\pm\infty$ if $s^2 = -1$. According to the definition of s (see Eq. 9), it is feasible to achieve $s^2 = -1$ under available values of those controllable parameters.

References

1. Daniel, H. U., Maurer, B., Steiner, M. & Walther, H. Schottky diode mixer for visiblelaser light and microwave harmonics up to 0.43 THz. *Appl. Phys. Lett.* **41**, 313–316 (1982).
2. Evenson, K. M., Jennings, D. A. & Petersen, F. R. Tunable far-infrared spectroscopy. *Appl. Phys. Lett.* **44**, 576–579 (1984).
3. Chuang, S. L., Schmitt-Rink, S., Greene, B. I., Saeta, P. N. & Levi, A. F. J. Optical rectification at semiconductor surface. *Phys. Rev. Lett.* **68**, 102–105 (1992).
4. Brener, I. *et al.* Terahertz emission from electric field singularities in biased semiconductors. *Opt. Lett.* **21**, 1924–1926 (1996).
5. Shikata, J., Sato, M., Taniuchi, T., Ito, H. & Kawase, K. Enhancement of terahertz-wave output from LiNbO3 optical parametric oscillators by cryogenic cooling. *Opt. Lett.* **24**, 202–204 (1999).
6. Cote, D., Fraser, J. M., Decamp, M., Bucksbaum, P. H. & van Driel, H. M. THz emission from coherently controlled photocurrents in GaAs. *Appl. Phys. Lett.* **75**, 3959–3961 (1999).
7. Shi, W., Ding, Y. J., Fernelius, N. & Vodopyanov, K. Single-frequency terahertz source pumped by Q-switched fiber lasers based on difference-frequency generation in GaSe crystal. *Opt. Lett.* **32**, 949–942 (2007).
8. Johnston, M. B. *et al.* Generation of high-power terahertz pulses in a prism. *Opt. Lett.* **27**, 1935–1937 (2002).
9. McLaughlin, R. *et al.* Enhanced coherent terahertz emission from indium arsenide in the presence of a magnetic field. *Appl. Phys. Lett.* **76**, 2038–2040 (2000).
10. Suizu, K., Miyamoto, K., Yamashita, T. & Ito, H. High-power terahertz-wave generation using DAST crystal and detection using mid-infrared powermeter. *Opt. Lett.* **32**, 2885–2887 (2007).
11. Shvartsman, L. D. & Laikhtman, B. InAs/GaSb laser: Prospects for efficient terahertz emission. *Appl. Phys. Lett.* **93**, 131104 (2008).
12. Mu, X. D., Ding, Y. J. & Zotova, Y. B. Transition from photocurrent surge to resonant optical rectification for terahertz generation in p-InAs. *Opt. Lett.* **32**, 3321–3323 (2007).
13. Hwang, J. S., Lin, H. C., Lin, K. I. & Zhang, X. C. Terahertz radiation from InAlAs and GaAs surface intrinsic-structures and the critical electric fields of semiconductors. *Appl. Phys. Lett.* **87**, 121107 (2005).
14. Miyamoto, K., Minamide, H., Fujiwara, M., Hashimoto, H. & Ito, H. Widely tunable terahertz-wave generation using an N-benzyl-2-methyl-4-nitroaniline crystal. *Opt. Lett.* **33**, 252–254 (2008).
15. Zhang, H., Wahlstrand, J. K., Choi, S. B. & Cundiff, S. T. Contactless photoconductive terahertz generation. *Opt. Lett.* **36**, 223–225 (2011).
16. Darmo, J., Strasser, G., Muller, T., Bratschitsch, R. & Unterrainer, K. Surface-modified GaAs terahertz plasmon emitter. *Appl. Phys. Lett.* **81**, 871–873 (2002).
17. Kashiwagi, T. *et al.* A high- intrinsic Josephson junction emitter tunable from 0.5 to 2.4 terahertz. *Appl. Phys. Lett.* **107**, 082601 (2015).
18. Kadlec, F., Kuzel, P. & Coutaz, J. Optical rectification at metal surfaces. *Opt. Lett.* **29**, 2674–2676 (2004).
19. Awari, N. *et al.* Narrow-band tunable terahertz emission from ferrimagnetic Mn3-xGa thin films. *Appl. Phys. Lett.* **109**, 032403 (2016).
20. Jin, Y., Ma, X. F., Wagoner, G. A., Alexander, M. & Zhang, X. C. Anomalous optically generated THz beams from metal/GaAs interfaces. *Appl. Phys. Lett.* **65**, 682–685 (1994).
21. Korbly, S. E., Kesar, A. S., Sirigiri, J. R. & Temkin, R. J. Observation of frequency-locked cohernt terahertz Smith–Purcell radiation. *Phys. Rev. Lett.* **94**, 054803 (2005).
22. Berry, C. W. *et al.* High power terahertz generation using 1550nm plasmonic photomixers. *Appl. Phys. Lett.* **105**, 011121 (2014).
23. Potz, W. Infrared light emission from semiconductor double wells. *Appl. Phys. Lett.* **68**, 2553–2555 (1996).
24. Liu, R. B. & Zhu, B. F. Tunable terahertz emission from difference frequency in biased superlattices. *Appl. Phys. Lett.* **84**, 2730–2732 (2004).
25. Shimada, Y., Sekine, N. & Hirakawa, K. Terahertz emission due to interminiband resonant Zener tunneling in wide-miniband superlattices. *Appl. Phys. Lett.* **84**, 4926–4928 (2004).
26. Kohler, R. *et al.* High-intensity interminiband terahertz emission from chirped superlattices. *Appl. Phys. Lett.* **80**, 1867–1869 (2002).
27. Lee, A. W. M. *et al.* High-power and high-temperature THz quantum-cascade lasers based on lens-coupled metal-metal waveguide. *Opt. Lett.* **32**, 2840–2842 (2007).
28. Walther, C. *et al.* Quantum cascade lasers operating from 1.2 to 1.6 THz. *Appl. Phys. Lett.* **91**, 131122 (2007).
29. Turinkovic, D. *et al.* Ultra-broadband heterogeneous quantum cascade laser emitting from 2.2 to 3.2 THz. *Appl. Phys. Lett.* **99**, 191104 (2011).
30. Jasnot, F. R. *et al.* Direct surface cyclotron resonance terahertz emission from a quantum cascade structure. *Appl. Phys. Lett.* **100**, 102103 (2012).
31. Sirtori, C., Barbieri, S. & Colombelli, R. Wave engineering with THz quantum cascade lasers. *Nat. Photon.* **7**, 691–701 (2013).
32. Altukhov, I. V. *et al.* Towards quantum-well resonant-state terahertz laser. *Appl. Phys. Lett.* **79**, 3909–3911 (2001).
33. Lu, Q. Y., Slivken, S., Bandyopadhyay, N., Bai, Y. & Razeghi, M. Widely tunable room temperature semiconductor terahertz source. *Appl. Phys. Lett.* **105**, 201102 (2014).
34. Liang, G. Z. *et al.* Single-mode surface-emitting concentric-circular-grating terahertz quantum cascade lasers. *Appl. Phys. Lett.* **102**, 031119 (2013).
35. Deutsch, C. *et al.* High performance InGaAs/GaAsSb terahertz quantum cascade lasers operating up to 142K. *Appl. Phys. Lett.* **101**, 211117 (2012).
36. Smet, J. H., Fonstad, C. G. & Hu, Q. Intrawell and interwell intersubband transitions in multiple quantum wells for far-infrared source. *J. Appl. Phys.* **79**, 9305–9320 (1996).

37. Belkin, M. A. *et al.* Room temperature terahertz quantum cascade laser surce based on intracavity difference-frequency generation. *Appl. Phys. Lett.* **92**, 201101 (2008).
38. Lu, Q. Y., Bandyopadhyay, N., Slivken, S., Bai, Y. & Razeghi, M. Room temperature single-mode terahertz sources based on intra-cavity difference-frequency generation in quantum cascade lasers. *Appl. Phys. Lett.* **99**, 131106 (2011).
39. Belkin, M. A. *et al.* Terahertz quantum-cascade-laser source based on intracavity difference-frequency generation. *Nat. Photon.* **1**, 288–292 (2007).
40. Vijayraghavan, K. *et al.* Terahertz sources based on Cerenkov difference-frequency generation in quantum cascade laser. *Appl. Phys. Lett.* **100**, 251104 (2012).
41. Lu, Q. Y., Bandyopadhyay, N., Slivken, S., Bai, Y. & Razeghi, M. Widely tuned room temperature terahertz quantum cascade laser sources based on difference-frequency generation. *Appl. Phys. Lett.* **101**, 251121 (2012).
42. Vijayraghavan, K. *et al.* Broadly tunable terahertz generation in mid-infrared quantum cascade lasers. *Nat. Commun.* **4**, 1–7 (2013).
43. Odnoblyudov, M. A., Yassievich, I. N., Kagan, M. S., Galperin, Yu. M. & Chao, K. A. Population inversion induced by resonant states in semiconductors. *Phys. Rev. Lett.* **83**, 644–647 (1999).
44. see many textbooks on solid-state physics, such as Ashcroft, N. W. & Mermin, N. D., Solid State Physics. 270, (Harcourt Inc, 1976).
45. Huang, K. & Han, R. *Solid State Physics* (Higher-Education Press, China**, 1988).
46. Chan, W. L., Deibel, J. & Mittleman, D. M. Imaging with terahertz radiation. *Rep. Prog. Phys.* **70**, 1325–1380 (2007).
47. Sirtori, C. Bridge for the terahertz gap. *Nature* **417**, 132–133 (2002).
48. Lin, H. & Liu, C. P. Direct classic route of generating mono-color EM-pulse with attosecond-level duration. *Appl. Sci.* **9**, 362–370 (2019).
49. Prange, R. E. & Girvin, S. M. *The Quantum Hall Effect* (Springer, 1987).
50. Yariv, A. *Quantum Electronics* 3rd edn. (Wiley, 1988).
51. Preu, S. *et al.* Tunable, continuous-wave Terahertz photomixer sources and applications. *J. Appl. Phys.* **109**, 061301 (2011).
52. Svelto, O. *Principles of Lasers* 5th edn. (Springer, 2010).
53. Landau, L. D. & Lifschitz, E. M. *Elasticity". Chinese Translation Edition* (High Education Press, Beijing, 2009).

Acknowledgements

Author contributions

H.L. and C.P.L. wrote the main manuscript text and prepared figures. All authors reviewed the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to H.L. or C.P.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.