Resonant plasmonic terahertz detection in vertical graphene-base hot-electron transistors

V. Ryzhii\textsuperscript{1,2}, T. Otsuji\textsuperscript{1}, M. Ryzhii\textsuperscript{3}, V. Mitin\textsuperscript{4}, and M. S. Shur\textsuperscript{5}

\textsuperscript{1}Research Institute of Electrical Communication, Tohoku University, Sendai 980-8577, Japan
\textsuperscript{2}Center for Photonics and Infrared Engineering, Bauman Moscow State Technical University and Institute of Ultra High Frequency Semiconductor Electronics of RAS, Moscow 111005, Russia
\textsuperscript{3}Department of Computer Science and Engineering, University of Aizu, Aizu-Wakamatsu 965-8580, Japan
\textsuperscript{4}Department of Electrical Engineering, University at Buffalo, SUNY, Buffalo, New York 14460-1920, USA
\textsuperscript{5}Department of Electrical, Computer, and System Engineering and Physics, Applied Physics, and Astronomy, Rensselaer Polytechnic Institute, Troy, New York 12180, USA

We analyze dynamic properties of vertical graphene-base hot-electron transistors (GB-HETs) and consider their operation as detectors of terahertz (THz) radiation using the developed device model. The GB-HET model accounts for the tunneling electron injection from the emitter, electron propagation across the barrier layers with the partial capture into the GB, and the self-consistent oscillations of the electric potential and the hole density in the GB (plasma oscillations), as well as the quantum capacitance and the electron transit-time effects. Using the proposed device model, we calculate the responsivity of GB-HETs operating as THz detectors as a function of the signal frequency, applied bias voltages, and the structural parameters. The inclusion of the plasmonic effect leads to the possibility of the HET-GBT operation at the frequencies significantly exceeding those limited by the characteristic RC-time. It is found that the responsivity of GB-HETs with a sufficiently perfect GB exhibits sharp resonant maxima in the THz range of frequencies associated with the excitation of plasma oscillations. The positions of these maxima are controlled by the applied bias voltages. The GB-HETs can compete with and even surpass other plasmonic THz detectors.

I. INTRODUCTION

Recently, vertical hot-electron transistors (HETs) with the graphene base (GB) and the bulk emitter and collector separated from the base by the barrier layers - the hot-electron graphene-base transistors (GB-HETs) - made of SiO\textsubscript{2} and Al\textsubscript{2}O\textsubscript{3} were fabricated and studied \cite{1-4}. These HETs are fairly promising devices despite their modest characteristics at the present. Similar devices can be based of GL heterostructures with the hBN, WS\textsubscript{2}, and other barrier layers \cite{5-8}. The history of different versions of HETs, including those with the thin metal base and the quantum-well (QW) base, in which the carriers are generated from impurities or induced by the applied voltages, as well as HETs with resonant-tunneling emitter, is rather long (see, for example, Refs. \cite{9-15}). Figure 1 shows schematically the GB-HET structure and its band diagrams. Depending on the GB doping, the base-emitter and collector-base voltages, $V_B$ and $V_C$, and the thicknesses of barrier layers separating the base from the emitter and collector, $W_E$ and $W_C$, respectively, the GB can be filled either with electrons or holes [compare Figs. 1(b) and 1(C)].

In GB-HETs (as in HETs), a significant fraction of the electrons injected from the top n-type emitter contact due to tunneling through the barrier top crosses the GB and reaches the collector, while the fraction of the electrons captured into the GB can be rather small. The tunneling electrons create the emitter-collector current $J_C$. The variations of the base-emitter voltage $V_B$ result in the variation of $J_C$.

In the case of the GB with a two-dimensional electron gas (2DEG), the GB-HET transistors can also be referred to as the N-n-N heterostructure HETs with the GB of n-type. In another case, when the GB comprises a two-dimensional hole (2DHG), the GB-HET transistors can also be called the N-p-N heterostructure bipolar transistors (HBTs) with the GB of p-type. The electrons captured into the GB support the emitter-base current $J_B$. The transistor gain in the common-emitter configuration is equal to $g_0 = J_C/J_B = (1 - p)/p$, where $p$ is the average capture probability of the capture of electrons into the GB during their transit. One of the main potential advantages of GB-HETs is the high-speed operation associated with the combination of a short transit time (due to the vertical structure), a high gain $g_0$ (because of a low probability, $p$, of the hot electron capture into the GB), and a low GB resistance (owing to a high mobility of the carriers in graphene). As shown recently \cite{17}, $p$ in the graphene heterostructures with proper inter-graphene barrier layers can be very small. Very small values of $p$ in GB-HETs can lead to a fairly high gain. The use of the undoped base with the induced carriers, leads to the exclusion of the scattering of the hot electrons crossing the GB with the impurities and to an increase of the mobility of the carriers localized in the GB.

By analogy with HETs with the QW-base made of the standard materials \cite{18,19}, GB-HETs can exhibit resonant response to the incoming signals associated with the excitation of the plasma oscillations. The resonant
plasma wave frequencies are determined by the characteristic plasma wave velocity $s$ (which increases with the carrier density $\Sigma_0$) and the lateral sizes of the GB $2L$, while the quality factor of the plasma resonances is mainly limited by the carrier momentum relaxation time $\tau$ associated with the scattering on impurities, various imperfections, and phonons. Due to the specific features of the carrier statistics and dynamics in the graphene layers [20], the plasma velocity $s > v_W$, where $v_W \approx 10^6$ cm/s is the characteristic velocity of the Dirac energy spectrum, and $s$ can markedly exceed that in the QW heterostructures. This promotes the realization of with the plasmonic resonances in the terahertz (THz) range even in the GB-HETs with fairly large lateral sizes. The possibility of achieving the elevated carrier mobilities in the GB can enable sharp plasmonic resonances in GB-HETs at room temperatures or above.

The application of the ac voltages $\pm \delta V_\omega/2$ at the signal frequency $\omega$, associated with the incoming radiation (received by an antenna) to the side contacts to the GB, results in the variation of the local ac potential difference $\delta \phi_\omega = \delta \phi_\omega(x)$ (the axis $x$ is directed in the GB plane) between the GB and the emitter contact. The emitter, base, and collector current densities $j_E$, $j_B$, and $j_C$ include the dc components $j_{E,0}$, $j_{B,0}$, and $j_{C,0}$ (determined by the applied bias voltages $V_B$ and $V_C$), the ac components $\delta j_{E,\omega}$, $\delta j_{B,\omega}$, and $\delta j_{C,\omega}$ (proportional to $\delta V_\omega$, and (due to the nonlinear dependence of the tunneling injection current on the local potential difference between the GB and the emitter) the rectified dc components $\overline{\delta j_{E,\omega}}$, $\overline{\delta j_{B,\omega}}$, and $\overline{\delta j_{C,\omega}}$ (proportional to $|\delta V_\omega|^2$, i.e., to the intensity of the incoming radiation received by an antenna. The net rectified current $\overline{\delta J_{C,\omega}}$ can serve as the output current in the GB-HETs radiation detectors. The detector responsivity $R_\omega \propto \overline{\delta J_{C,\omega}}/|\delta V_\omega|^2$. Due to the possibility of the plasmonic resonances, the rectified component (the detector output signal) can be resonantly large, similar to that in the HET detector [21] and other plasmonic THz detectors using different transistor structures, including those incorporating graphene [22][24].

In this paper, develop the GB-HET device model and evaluate the GB-HET characteristics as a radiation detector of radiation, in particular, in the THz range of signal frequencies.

II. DEVICE MODEL AND RELATED EQUATIONS

We consider a GB-HET with the highly conducting tunneling emitter and collector of the n-type and the undoped GB (as shown in Fig. 1) with a 2DHG induced by the applied bias voltages creating the N-p-N structure with the band diagram shown in Fig. 1(b). The GB-HET device model accounts for the tunneling injection of hot-electrons from the emitter to the barrier layer (above the barrier top), their propagation across the barrier layers, partial capture of hot-electrons into the GB, and the excitation of the self-consistent oscillations of the electric potential and the hole density in the GB (plasma oscillations). The quantum capacitance and the electron transit-time effects are also taken into account.

We assume that the bias voltages $V_B$ and $V_C$ are applied between the GL-base and the emitter and between the collector and the emitter contacts, respectively. In the framework of the gradual channel approximation [35], which is valid if $W_E, W_C \ll 2L$, the density of the two-dimensional hole gas (2DHG)$\Sigma = \Sigma(x, t)$ in the GL-base and its local potential $\varphi = \varphi(x, t)$ (counted from the potential of the emitter) are related to each other as

$$\Sigma = \frac{\kappa}{4\pi e} \left( \frac{\varphi - \mu/e - V_{bi}}{W_E} + \frac{\varphi - \mu/e - V_{bi} - V_C}{W_C} \right), \quad (1)$$

where $\kappa$ is the dielectric constant of these layers, $e$ is the hole charge, $V_{bi}$ is the built-in voltage between the contact material and an undoped GL, and $\mu$ is the 2DHG Fermi energy in the GB. In a degenerate 2DHG, $\mu = h v_W \sqrt{\pi \Sigma}$, where $h$ is the Planck constant. The dependence of the right side of Eq. (1) on the hole Fermi energy is interpreted as the effect of quantum capacitance [36][37].

The ac voltages $\pm \delta V_\omega/2$ and $\omega$ are applied between the side GB contacts connected with an antenna, so that the ac potential of the GB $\delta \phi_\omega = \delta \phi_\omega(x)$ obeys the following (asymmetric) conditions:

$$\delta \phi_\omega |_{x=\pm L} = \pm \delta V_\omega/2. \quad (2)$$
The side contacts can serve as the slot wave guide transforming the incoming THz radiation signals being received by an antenna into the ac voltage [see Fig.1(a)]. The ac component of the collector-emitter voltage $\delta V_{C,\omega}$ can also arise due to the ac potential drop across the load resistance. However, in the GB-HET with the wiring under consideration, $\delta V_{C,\omega}$ can be disregarded providing that the GB-HET structure is symmetrical (see below). For GB-HETs with the degenerate 2DHG, the relation between the variation of the hole density and the potential, $\delta \Sigma_o$ and $\delta \varphi_o$ (the ac components) are, as follows from Eq. (1), can be expressed via the net capacitance per unit area $C = C_q C_{quant}/(C_{quant} + C_g)$, which accounts for the geometrical capacitance $C_q = (C_e + C_C) = (\kappa/4\pi)(W_e^{-1} + W_C^{-1})$ (with $C_e \propto W_e^{-1}$ and $C_C \propto W_C^{-1}$ being the geometrical emitter and collector capacitances, respectively) and the quantum capacitance $[36, 37]$ $C_{quant} = (2\pi^2/\sqrt{\Sigma_0}/\sqrt{\hbar v_W}) = (2\pi^2 \mu_0/\hbar v_W^2)$, where $\Sigma_0$ and $\mu_0$ are the pertinent dc values of the hole density and the Fermi energy.

$$e \delta \Sigma_o = C \delta \varphi_o.$$  

(3)

Considering the electron tunneling from the emitter to the states above the top of the barrier (through the triangular barrier), the emitter electron tunneling current density $j_E$ can be presented as:

$$j_E = j_E^0 \exp \left(-\frac{F}{F_E}\right).$$  

(4)

Here $F = (a\sqrt{m} \Delta^{3/2}/e\hbar)$, is the characteristic tunneling field, $\Delta$ is the activation energy for electrons in the contact, $m$ is the effective electron mass in the barrier material, $a \sim 1$ is a numerical coefficient, $F_E = (\varphi - V_{bi})/W_E$ is the electric field in the emitter barrier, and $j_E^0$ is the maximum current density which can be provided the emitter contact. As follows from Eqs. (2) and (3), the ac component of the net ac currents $\delta j_{E,\omega}$ and $\delta j_{C,\omega}$ are, respectively, given by

$$\delta j_{E,\omega} = j_{E,0} \frac{F}{F_{E,0}} \delta F_{E,\omega} = \sigma_E \delta F_{E,\omega}.$$  

(5)

$$\frac{\delta j_{E,\omega}}{j_{E,0}} \approx \frac{1}{2} \left( \frac{F}{F_{E,0}} \right)^2 - \left( \frac{F}{F_{E,0}} \right) \left( \frac{\delta F_{E,\omega}}{F_{E,0}} \right)^2 \approx \frac{\sigma_E F}{4} \left( \frac{\delta F_{E,\omega}}{F_{E,0}} \right)^2.$$  

(6)

Here $j_{E,0} = j_E^0 \exp(-F/F_{E,0})$, $F_{E,0} = (V_E - V_{bi} - \mu_0/e)/W_E$, and $\sigma_E = j_0(F/F_{E,0})$ are the emitter dc current density (in the absence of the ac signals), the dc electric field in the emitter barrier, and the emitter differential conductance, respectively. The dc hole Fermi energy in the GB $\mu_0$ obeys the following equation:

$$\mu_0^2 = \frac{\kappa \hbar^2 v_W^2}{4e} \left( V_E - V_{bi} - \mu_0/e \right) \left( \frac{1}{W_E} + \frac{1}{W_C} \right) - V_C.$$  

(7)

The ac electric field components perpendicular to the GB plane in the emitter barrier, as well as in the collector barrier, are respectively given by

$$\delta F_{E,\omega} = \frac{C_{quant}}{C_{quant} + C_g} \delta \varphi_o,$$  

(8)

$$\delta F_{C,\omega} = -\frac{C_{quant}}{C_{quant} + C_g} \delta \varphi_o.$$  

(9)

Using the Shockley-Ramo theorem, one can find the ac component the emitter current densities coming to the base and the collector:

$$\delta j_{EB,\omega} = \sigma_{EB,\omega} \delta F_{E,\omega}, \quad \delta j_{EC,\omega} = \sigma_{EC,\omega} \delta F_{E,\omega}. \quad (10)$$

Here

$$\sigma_{EB,\omega} = \sigma_E \left[ \frac{1}{W_E} \int_0^{W_E} dz e^{i\omega z/v} ight. \left. - (1 - p) \int_{W_E + W_C}^{W_E} dz e^{i\omega z/v} \right] \left( \frac{1 - p}{W_E} \right) e^{i\omega \tau_E} (e^{i\omega \tau_C} - 1) \frac{1}{\tau_C}, \quad (11)$$

$$\sigma_{EC,\omega} = \sigma_E \left( (1 - p) \int_{W_E}^{W_E + W_C} dz e^{i\omega z/v} \right) \frac{1 - p}{\tau_C}, \quad (12)$$

where $v$ is the drift velocity of the hot electrons crossing the barriers above their tops (which is assumed to be constant) and $\tau_E = W_E/v$ and $\tau_C = W_C/v$ are the electron transit time across the emitter and collector barrier layers. The axis $z$ is directed perpendicular to the GB plane. Since under the boundary conditions (2), $\delta \varphi_o(x) = -\delta \varphi_o(-x)$ (see below) and, hence, $\delta F_{E,\omega}(x) = -\delta F_{E,\omega}(x)$, the net ac currents

$$\delta J_{EB,\omega} = \int_{-L}^{L} dx \delta j_{EB,\omega} = \sigma_{EB,\omega} \int_{-L}^{L} dx \delta F_{E,\omega} = 0, \quad (13)$$

The ac hole Fermi energy in the GB $\mu_0$ obeys the following equation:
\[ \delta J_{EC,\omega} = \int_{-L}^{L} dx \delta j_{EC,\omega} = \sigma_{EC,\omega} \int_{-L}^{L} dx \delta F_{E,\omega} = 0. \]  

Simultaneously for the rectified components of the dc current densities from the emitter to the GB and from the emitter to the collector one obtains

\[ \delta J_{EB,\omega} \propto \frac{p \sigma_F E}{4} \left( \frac{\delta F_{E,\omega}}{F_{E,0}} \right)^2 = p \Gamma |\delta \varphi_{\omega}|^2, \]  

where

\[ \delta J_{EC,\omega} \propto \frac{(1-p) \sigma_F E}{4} \left( \frac{\delta F_{E,\omega}}{F_{E,0}} \right)^2 = (1-p) \Gamma |\delta \varphi_{\omega}|^2, \]

\[ \Gamma = \frac{\sigma_F E}{4 F_{E,0} W} \frac{C_{\text{quant}}}{(C_{\text{quant}} + C_g)^2}. \]

Consequently, the net rectified components of the dc current densities from the emitter to the GB and from the emitter to the collector are given by

\[ \delta J_{EB,\omega} = \int_{-L}^{L} dx \delta j_{EB,\omega} \propto p \Gamma \int_{-L}^{L} dx |\delta \varphi_{\omega}|^2, \]

\[ \delta J_{EC,\omega} = \int_{-L}^{L} dx \delta j_{EC,\omega} \propto (1-p) \Gamma \int_{-L}^{L} dx |\delta \varphi_{\omega}|^2, \]

respectively.

The ac hole current along the GB is given by

\[ \delta J_{BB,\omega} = -\sigma_{BB,\omega} \frac{d \delta \varphi_{\omega}}{dx} \bigg|_{x=L}, \]

where the lateral ac conductivity of the GB \( \sigma_{BB,\omega} \) is given by (see, for example, Refs. [35-42]

\[ \sigma_{BB,\omega} = \frac{i e^2 \mu_0}{\pi \hbar (\omega + i \tau)}. \]

III. PLASMA OSCILLATIONS AND RECTIFIED CURRENT

To calculate the rectified current components using Eqs. (15) and (16), one needs to find the spatial distributions of the ac potentials in the GB \( \delta \varphi_{\omega} = \delta \varphi_{\omega}(x) \).

For this purpose we use the hydrodynamic equations for the hole transport along the GB [43, 44] (see also Refs. [21, 22]) coupled with the Poisson equation solved using the gradual channel approximation [i.e., using Eq. (1)]. Linearizing the hydrodynamic equations and Eq. (1) and taking into account that the ac component of the hole Fermi energy is expressed via the variation of their density and the potential, we arrive at the following equation for the ac component of the GB potential \( \delta \varphi_{\omega} \) (compare with Refs. [16, 21, 22] which should be solved with boundary conditions given by Eq. (2):

\[ \frac{d^2 \delta \varphi_{\omega}}{dx^2} + \frac{\omega + i \nu}{s^2} \delta \varphi_{\omega} = 0. \]

Here \( s \) is the characteristic velocity of the plasma waves in the gated graphene layers, which is given by \( s = \sqrt{e^2 \Sigma_0/m C} \propto \Sigma_0^{1/4} \) [18], \( m_0 = \mu_0 v_F^2 \propto \sqrt{\Sigma_0} \) being the so-called “fictitious” effective hole (electron) mass in graphene layers [27], \( \nu = (\sigma_{BB,\omega}/W E C_g) \) and \( \nu = 1/\tau + \nu \), where \( \tau \) is the hole momentum relaxation time in the 2DHG and \( \nu = \nu_{\text{visc}} + \nu_{\text{rad}} \) is associated with the contribution of the 2DHG viscosity to the damping (see, for example, Ref. [21]) and with the radiation damping of the plasma oscillations. The latter mechanism is associated with the recoil that the holes in the GB feel emitting radiation (the pertinent term in the force acting on the holes is referred to as the Abraham-Lorentz force or the radiation reaction force [46, 47]). Taking into account that the viscosity damping rate is proportional to the second spatial derivative, for the gated plasmons with the acoustic-like spectrum \( \nu_{\text{visc}} \propto \zeta \omega^2/s^2 \) (\( \zeta \) is the 2DHG viscosity), and considering that \( \nu_{\text{rad}} \propto (2e^2/3mc^3)\omega^2 \) [45, 47], we put \( \nu = 1/\tau + \eta \omega^2 \), where \( \eta \) is the pertinent damping parameter, and \( c \) is the speed of light. If the viscosity damping surpasses the radiation damping, one can set \( \eta \approx \zeta/s^2 \).

The characteristic plasma-wave velocity \( s \) is determined by \( \Sigma_0 \) (as well as the thicknesses of the emitter and barrier layers \( W_E \) and \( W_C \) [20]) and, hence, can be changed by the variations of the bias voltages \( V_E \) and \( V_B \). Due to this the characteristic plasmonic frequency \( \Omega = \pi s/L \) can be effectively controlled by these voltages.

Equation (22) with Eq. (2) yield

\[ \delta \varphi_{\omega} = \frac{\sin(\sqrt{\omega + i \nu}/(\omega + i \nu) L/s)}{\sin(\sqrt{\omega + i \nu}/(\omega + i \nu) L/s) \sqrt{2}} \frac{d \varphi_{\omega}}{dx}. \]

One can see from Eq. (23) that the spatial dependence of \( \delta \varphi_{\omega} \) is rather complex. In particular, when \( \omega \) approaches to \( n \Omega \), where \( n = 1, 2, 3, ..., \) this distribution can be oscillatory with fairly high amplitude of the spatial oscillations when \( \Omega \gg \nu \).

Substituting \( \delta \varphi_{\omega} \) given by Eq. (23) to Eqs. (15) and (16), we arrive at the following:

\[ \delta J_{EB,\omega} \propto \frac{p \Gamma}{\sin(\sqrt{\omega + i \nu}/(\omega + i \nu) L/s)^2} \left| \frac{d \varphi_{\omega}}{dx} \right|^2, \]
\[
\delta J_{\text{EC,}\omega} \simeq (1 - p) \Gamma \\
\times \left| \frac{\sin(\sqrt{\omega + i\nu}(\omega + i\nu) x/s)}{\sin(\sqrt{\omega + i\nu}(\omega + i\nu) L/s)} \right|^2 \frac{\delta V_\omega}{4}, \quad (25)
\]

\[
\delta J_{\text{EB,}\omega} \simeq p \Gamma \\
\times \int_{-L}^{L} dx \left| \frac{\sin(\sqrt{\omega + i\nu}(\omega + i\nu) x/s)}{\sin(\sqrt{\omega + i\nu}(\omega + i\nu) L/s)} \right|^2 \delta V_\omega, 
\]

\[
\delta J_{\text{EC,}\omega} \simeq (1 - p) \Gamma \\
\times \int_{-L}^{L} dx \left| \frac{\sin(\sqrt{\omega + i\nu}(\omega + i\nu) x/s)}{\sin(\sqrt{\omega + i\nu}(\omega + i\nu) L/s)} \right|^2 \frac{\delta V_\omega}{4}, \quad (27)
\]

As follows from Eqs. (24) - (27), \( \delta J_{\text{EC,}\omega} = [(1 - p)/p] \delta J_{\text{EB,}\omega} \) and \( \delta J_{\text{EC,}\omega} = [(1 - p)/p] \delta J_{\text{EB,}\omega} \) (as for the pertinent dc current in the absence of the THz signals), so that \( \delta J_{\text{EC,}\omega} \gg \delta J_{\text{EB,}\omega} \).

Using Eqs. (20) and (24), for the ac current between the based contacts we obtain

\[
\delta J_{\text{BB,}\omega} = -ie^2 \mu_0 \cot(\sqrt{\omega + i\nu}(\omega + i\nu) L/s) \\
\times \sqrt{\frac{\sqrt{\omega + i\nu} \delta V_\omega}{2}}, \quad (28)
\]

Due to the symmetry of the GB-HET structure and the asymmetric spatial distribution of \( \delta \phi_\omega \), there is no rectified component of the lateral current in the GB-base, i.e., \( \delta J_{\text{BB,}\omega} = 0 \).

At very low signal frequencies \( \omega \ll \nu \), \( |\nu| \simeq 4\pi \rho \sigma_E / \kappa \), from Eqs. (23) and (22) we obtain

\[
\delta \phi_\omega \simeq \frac{\sinh(\sqrt{\nu} \omega L/s)}{\sinh(\sqrt{\nu} \omega L/s)} \frac{\delta V_\omega}{2}, \quad (29)
\]

\[
\delta J_{\text{EC,}\omega} \simeq (1 - p) \Gamma L \left( \frac{2\sqrt{\nu} \omega L/s}{\sinh(\sqrt{\nu} \omega L/s)} - 1 \right) \frac{\delta V_\omega^2}{4} \\
\simeq (1 - p) \Gamma L \frac{\delta V_\omega^2}{6}. \quad (30)
\]

Since in reality \( |\nu| \ll \nu \), there is an intermediate range of frequencies \( |\nu| \ll \omega < \nu \). Assuming that \( |\nu| \ll \nu \), \( \Omega^2 / \omega < \nu \), from Eq. (27) we arrive at

\[
\delta J_{\text{EC,}\omega} \simeq (1 - p) \Gamma L \left( \frac{s}{\nu_\omega} \right) \frac{\delta V_\omega^2}{4} \\
\propto (1 - p) \sqrt{\frac{\Omega^2}{\nu_\omega}} \delta V_\omega^2. \quad (31)
\]

**IV. GB-HET DETECTOR RESPONSIVITY**

**A. Current responsivity**

The rectified current \( \delta J_{\text{EC,}\omega} \) can be considered as the output signal used for the detection of electromagnetic radiation (in particular, THz radiation). The current detector responsivity using this output signal is defined as

![Image](image-url)
where $H$ is the GB-HET lateral size in the direction along the contacts to the GB, $I_\omega$ is intensity of the incident radiation, and $S$ is the antenna aperture. The latter is given by $S = \lambda^2 G / 4 \pi$, where $G$ is the antenna gain, $\lambda$ is the speed of light in vacuum. Taking into account that $I_\omega = c E^2 / 8 \pi$, where $E$ is the radiation electric field in vacuum, and estimating $\delta V_\omega$ as $\delta V_\omega = \lambda E / \pi$, one can arrive at

$$|\delta V_\omega|^2 = \frac{8 \lambda^2 I_\omega}{\pi c} = \frac{32 \pi c}{\omega^2} I_\omega.$$  

(33)

Considering Eqs. (27), (32), and (33), for the current responsivity (in A/W units) we find

$$R_\omega = \frac{R}{L} \int_{-L}^{L} dx \left| \frac{\sin[(\omega + i \nu)(\omega + i \nu)|x|/s]}{\sin[(\omega + i \nu)(\omega + i \nu)L/s]} \right|^2,$$  

(34)

where

$$R \approx \frac{8(1 - p) \Gamma L H}{cG} = \rho \frac{LH}{W_E^2},$$

(35)

$$\rho = \frac{8(1 - p) j_0}{cG} \left( \frac{F}{F_{E,0}} \right)^2.$$  

(36)

As follows from Eq. (34), at relatively long hole momentum relaxation times $\tau$, the GB-HET responsivity exhibits a series of very sharp and high peaks. It is instructive that in the GB-HETs under consideration, the height of the resonant peaks does not decrease with an increasing peak index (as in some other devices using the plasmonic resonances). Although some decrease in the peaks height with the increasing resonance index attributed to the effect of viscosity and radiative damping takes place, the pertinent effect is relatively weak (see below).

In the limiting cases $\omega \ll \nu$, $|\nu|$ and $|\nu| \ll \nu, \Omega^2 / \omega < \nu$, corresponding to Eqs. (25) and (26), one obtains

$$R_\omega = R_0 \approx \frac{2}{3} R$$

(37)

and

$$R_\omega \approx R \left( \frac{s}{L} \sqrt{\frac{2}{\nu \omega}} \right) = R \sqrt{\frac{2 \Omega^2}{\pi^2 \nu \omega}} < R,$$  

(38)

respectively. The quantities $R$ and $R_\omega$ depend on the geometrical and quantum capacitances and, therefore, on the barrier layers thicknesses:

$$R_\omega \propto R \propto \frac{1}{W_E^2} \left( \frac{C_{\text{quant}}}{C_{\text{quant}} + C_g} \right)^2 = \frac{1}{W_E^2} \left[ 1 + \frac{\kappa}{4 \pi C_{\text{quant}} \left( W_E^{-1} + W_C^{-1} \right)} \right]^{-2}.$$  

(39)

For the doping level of the emitter contact $N_D = 10^{18}$ cm$^{-3}$, assuming that the thermal electron velocity $v_T = 10^7$ cm/s, we obtain $j_E = 1.6 \times 10^6$ A/cm$^2$. Setting $\Delta = 0.2$ eV and the effective mass in the barrier layer $m \approx 2.5 \times 10^{-28}$ g, we arrive at $F \sim 2 \times 10^6$ V/cm. Setting also $F_{E,0} = 5 \times 10^5$ V/cm (to provide the hole density in the GB about of $\Sigma = 10^{12}$ cm$^{-2}$), we find $j_{E,0} \approx 2.9 \times 10^4$ A/cm$^2$ and $\sigma_E \approx 0.23$ Ohm$^{-1}$ cm$^{-1}$. At these parameters, one obtains also $C_{\text{quant}} \approx 2.6 \times 10^6$ cm$^{-1}$. Using these data and setting in addition $p \approx 1$, $W_E W_C / (W_E + W_C) = 5$ nm, $\kappa = 4$, and $G = 1.5$, from Eq. (36) we obtain $\rho \approx 2 \times 10^{-4}$ A/W. If $L = 1.5$ $\mu$m and $H = 10$ $\mu$m, Eqs. (35) - (36) yield $R \approx 30$ A/W and $R_0 \approx 20$ A/W. These parameters are also used in the estimates below.

In the case of high quality factor of the plasma resonances $\Omega/\nu$, the quantities $\delta J_{EB,\omega}$ and $\delta J_{EC,\omega}$ and, hence, the responsivity as functions of the signal frequency $\omega$ described by Eqs. (27) and (34) exhibit sharp resonance peaks at $\omega \approx n \Omega$, attributed to the resonant excitation of the plasma oscillations (standing plasma waves). The peak width is primarily determined by the frequency $\nu$. Indeed, for $\Omega/\nu \gg 1$ Eq. (34) yields

$$\max R_\omega \approx R_0 \sim \frac{3}{2} R_0 \left( \frac{2 \Omega^2}{\pi^2 \nu} \right)^2 \gg R_0.$$  

(40)

Using the above estimate for $R_0$, the peak values of the responsivity at $\Omega = 5/3$ THz and $\tau = 1$ ps is approximately equal to $\max R_\omega \approx R_0 \approx 1.33 \times 10^4$ A/W.

Figure 3 shows the dependence of the normalized GB-HET current responsivity calculated using Eq. (34) for several sets of parameters. As seen, the positions of the responsivity peaks shift toward higher frequencies when the plasma frequency increases, i.e., when $s$ increases and/or $L$ decreases because $\Omega \propto s / L$ (compare the curves "1" and "2"). A shortening of the momentum relaxation time $\tau$ leads to a smearing of the peaks (compare the curve "3" corresponding to $\tau = 1$ ps and the curve "4" corresponding to $\tau = 0.5$ ps). Figure 4 shows the lowering and broadening of the resonance peaks of the GB-HET current responsivity with the decreasing momentum relaxation time $\tau$ described by Eq. (34). As seen, at $\tau < 0.3 - 0.4$ ps, the responsivity peaks vanish, while at $\tau = 1$ ps they are fairly sharp and high.

**B. Voltage responsivity**

The variation of the dc current component $\delta J_{EC,\omega}$ cause by the ac signals results in a change of the volt-
The voltage responsivity of the GB-HET under consideration is independent of the signal frequency. For $W_E = W_C = W$, Eqs. (38) and (40) yield

$$R^V_\omega = R_\omega \frac{r_C}{1 + 2\pi C_{\text{quant}}W/\kappa}.$$  \(\text{(43)}\)

At $r_C \ll (1 + 2\pi C_{\text{quant}}W/\kappa)(W/LH\sigma_E) = \tau_C$, Eq. (43) yields the obvious formulas:

$$R^V_0 = R_0 r_C.$$  \(\text{(44)}\)

at low frequencies, and

$$R^V_\Omega = \frac{3R_0}{2\Omega} \left(\frac{2\Omega}{\pi \nu}\right)^2 r_C$$  \(\text{(45)}\)

at the plasma resonance $\omega = \Omega$.

For $W > (1 + 2\pi C_{\text{quant}}W/\kappa)(W/LH\sigma_E) = \tau_C$, from Eq. (43) we obtain, respectively,

$$R^V_\Omega \simeq R_0 \frac{(W^2/LH)}{\sigma_E W} \left(1 + \frac{2\pi C_{\text{quant}}W}{\kappa}\right).$$  \(\text{(46)}\)

and

$$R^V_\Omega \simeq \frac{3R_0}{2} \left(\frac{2\Omega}{\pi \nu}\right)^2 \frac{(W^2/LH)}{\sigma_E W} \left(1 + \frac{2\pi C_{\text{quant}}W}{\kappa}\right).$$  \(\text{(47)}\)

For $\tau = 1$ ps, $s = 5 \times 10^6$ cm/s, and $L = 1.5 \mu$m (as for the curve "1" in Fig. 3), so that $\Omega = 5/3$ THz, as well as $\sigma_E = 0.23$ A/V cm, $C_{\text{quant}} = 2.6 \times 10^6$ cm$^{-1}$ (see the estimate in the previous subsection), and $\kappa = 4$ from Eq. (47) we obtain the following estimate: $R^V_\Omega \simeq 2 \times 10^9$ V/W. For the above parameters one obtains $\tau_C \simeq 150$ Ohm. Even at smaller $r_C$, the voltage responsivity can be fairly large. Setting $r_C = 5-10$ Ohm, we obtain $R^V_\Omega \simeq 100-200$ V/W and $R^V_\Omega \simeq (6.6 - 13.2) \times 10^9$ V/W, respectively.
Considering Eqs. (42) and (39), we find the following dependences of the current and voltage responsivities \( R_\omega \) and \( R_\nu^V \) on the emitter and collector barriers thickness \( W \) (at not too large \( r_C \)):

\[
R_\nu^V = R_\omega r_C \propto \frac{4\pi^2 C_{\text{quant}}}{\kappa^2(1 + 2\pi C_{\text{quant}} W/\kappa)^2}.
\]  

(48)

According to Eq. (48), \( R_\omega \) and \( R_\nu^V \) markedly decrease in the range \( W > \kappa/2\pi C_{\text{quant}} \approx 10 \text{ nm} \).

V. DISCUSSION

As follows from Eqs. (34) and (40), an increase in \( \nu \) with increasing frequency \( \omega \) due to the reinforcement of the plasma oscillation damping associated with the viscosity and the radiative damping, might lead to the gradual lowering of the resonant peaks with their index \( n \). However, our estimates show that the contribution of these two mechanisms to the net damping is small compared to the damping associated with the hole momentum relaxation (collisional damping). Indeed, disregarding the radiative damping and assuming the 2DHG viscosity to be \( \zeta = 10 \text{ cm}^2/\text{s} \) (i.e., smaller than in the standard 2DEG and 2DHG in the GaAs based heterostructures [21, 43, 50]) and \( s = (2.5 - 5) \times 10^8 \text{ cm/s} \), we obtain \( \eta \simeq (4 - 16) \times 10^{-17} \text{ s} \). Hence, in the frequency range \( \omega/2\pi \leq 2 \text{ THz} \) (as in Fig. 2), we find \( \nu_{\text{visc}} = \eta \omega^2 \leq (6.3 - 25.3) \times 10^8 \text{ s}^{-1} \ll 1/\tau \). Therefore, the heights of the responsivity peaks in Fig. 2 in the curve "3" at \( \omega \simeq 0.8 \text{ and } 1.6 \text{ THz are virtually equal. However, the peaks corresponding to higher resonances with the frequencie in the range 5 - 10 \text{ THz can be markedly lowered and smeared, because in this range } \nu_{\text{visc}} \text{ can become comparable with } 1/\tau. \) For example, for the same values of \( \eta, \tau = 1 \text{ ps}, \omega/2\pi = 5 - 10 \text{ THz}, \) we obtain \( \nu_{\text{visc}} \tau \simeq 0.16 - 0.64 \).

Equation (42) describes the saturation of the voltage responsivity \( R_{\Omega}^V \) peak value with increasing load resistance \( r_C \). This is associated with the effect of the voltage drop across the load on the potential drop between the emitter and the base and, hence, the hole Fermi energy in the GB, which determine the injection current. Such an effect is due to the finite value of the GB quantum capacitance (see, Refs. [12, 13]) - if \( C_{\text{quant}} \) tends to infinity, the emitter-base voltage becomes independent of \( r_C \), and the saturation of the \( R_{\Omega}^V \sim r_C \) dependence vanishes.

The THz detectors using a similar operation principle and InP double heterojunction bipolar transistors (DHBTs) were recently fabricated and studied experimentally [51, 53]. The estimated responsivities of the DHBTs in question for the non-resonant detection regime are somewhat smaller but of the same order of magnitude than those given by Eq. (43) and the pertinent estimates. However, the experimental values of the responsivity are much smaller than the values predicted above for the resonant detection (see Eqs. (44) and (46) and the estimates based on these equations]. Apart from the parasitic effects and the absence of any spatial coupling antenna, this can be attributed to the doping of the base in the InP-DHBTs, which inevitably leads to relatively a shorter hole momentum relaxation time \( \tau \) compared to that in the GB (where the 2DHG is induced by the applied voltages). Possibly, a higher probability of the hot electron capture into the InGaAs base (that had a relatively large thickness of 28 nm) in the DHBTs in comparison with the GB-HETs can be an additional factor.

The GB-HET current and voltage responsivities are determined by several characteristics: the characteristics of the tunneling emitter, geometrical characteristics of the GB-HET structure, materials of the emitter as well as the emitter and collector barrier layers, and applied bias voltages. The diversity of these factors enables the optimization of the GB-HETs operating as resonant plasmonic detectors, in particular, an increase in the responsivity in comparison with the values obtained in the above estimates. The resonant plasmonic THz detectors can based on not only the GB-HET structure shown in Fig. 1(a) (with an extra antenna connected to the GB side contacts), but also based on lateral structures with the GB contacts forming a periodic array.

The comparison of the GB-HETs [11, 14] and InP-DHBTs [51, 53] with the GB-HETs under consideration highlights the following advantages of the latter: (i) a longer momentum relaxation time of holes \( \tau \) in the GB; (ii) a higher plasma-wave velocity \( s \) that enables higher resonant plasma frequencies; (iii) a smaller capture probability of hot electrons into the GB and, consequently, larger (or even much larger) fraction of the hot electrons reaching the collector; (iv) coupling the incoming THz signal to the GB resulting in the absence of the ac current in the emitter-collector circuit and preventing the RC effects usually hindering the high-frequency operation.

VI. CONCLUSIONS

We developed an analytical model for vertical heterostructure HETs with the GB of the p-type sandwiched between the wide-gap emitter and collector layers and the N-type contacts. Using this model, we described the GB-HET dynamic properties and studied the GB-HET operation as detectors of THz radiation. The main features of the GB-HETs are high hole mobility in the GB, low probability the capture of the hot electrons injected from the emitter and crossing the GB, and the absence of the collector ac current. These features enable pronounced voltage-controlled plasmonic response of the GB-HETs to the incoming THz radiation, high hot-electron injection efficiency, and the elimination of the RC-limitations leading to elevated the GB-HET current and voltage responsivities in the THz range of frequencies, particularly at the plasmonic resonances at room temperature. This might provide the superiority of the GB-HET-based THz
detectors over other plasmonic THz detectors based on the standard heterostructures. Thus, the THz detectors based on the GB-HETs can be interesting for different applications.

Acknowledgments

The authors are grateful to D. Coquillat and F. Teppe for the information related to their experimental data and by the US Army Research Laboratory Cooperative Research Agreement, respectively.

[1] W. Mehr, J. Ch. Scheytt, J. Dabrowski, G. Lipper, Y.-H. Xie, M. C. Lemme, M. Ostling, and G. Lupina, IEEE Electron Device Lett. 33, 691 (2012).
[2] B. D. Kong, C. Zeng, D. K. Gaskill, K.L. Wang, K. W. Kim, Appl. Phys. Lett. 101, 263112 (2012).
[3] S. Vaziri, G. Lupina, C. Henkel, A. D. Smith, M. Ostling, J. Dabrowski, G. Lipper, W. Mehr, and M. C. Lemme Nano Lett. 13, 1435 (2013).
[4] C. Zeng, E. B. Song, M. Wang, S. Lee, C. M. Torres, Jr., J. Tang, B. H. Weiler, and K. L. Wang, Nano Lett. 13, 2370 (2013).
[5] L. Britnell, R. V. Gorbachev, R. Jalil, B. D. Belle, F. Shedin, A. Mishchenko, T. Georgiou, M. I. Katsnelson, L. Eaves, S. V. Morozov, N. M. R. Peres, J. Leist, A. K. Geim, K. S. Novoselov, and L. A. Ponomarenko, Science, 335, 947 (2012).
[6] T. Georgiou, R. Jalil, B. D. Belle, L. Britnell, R. V. Gorbachev, S. V. Morozov, Y.-J. Kim, A. Cholmsia, S. J. Haigh, O. Makarovszky, L. Eaves, L. A. Ponomarenko, A. K. Geim, K. S. Novoselov, and A. Mishchenko, Nature Nanotechnology 8, 100 (2013).
[7] L. Britnell, R. V. Gorbachev, A. K. Geim, L. A. Ponomarenko, A. Mishchenko, M. T. Greenaway, T. M. Fromhøld, K. S. Novoselov, and L. Eaves, Nature Comm. 4, 1794 (2013).
[8] M. Liu, X. Yin, and X. Zhang, Nano Lett. 12, 1482 (2012).
[9] C. A. Mead, Proc. IRE, 48, 359 (1960).
[10] J. M. Shannon, IEEE J. Solid-State Electron Devices, 3, 142 (1979).
[11] M. Heiblum, D. C. Thomas, C. M. Knoedler, M. I. Nathan, Surf. Sci. 174, 478 (1986).
[12] S. Luryi, IEEE Electron Device Lett. 6, 178 (1985).
[13] S. Luryi, in High Speed Semiconductor Devices, edited by S. M. Sze (Wiley, New York, 1990), p.399.
[14] J. Xu and M. S. Shur, Double Base Hot Electron Transistor, United States Patent #4,901,122, February 13 (1990).
[15] M. S. Shur, R. Gaska, A. Bykhovski, M. A. Khan, and J. W. Yang, Appl. Phys. Lett. 76, 3298 (2000).
[16] M. Asada, et.al. Jpn. J. Appl. Phys. 47, 4375 (2008).
[17] V. Ryzhii, T. Otsui, M. Ryzhii, V. Yu. Aleshkin, A. A. Dubinov, V. Mitin, and M.S.Shur, J. Appl. Phys. 117, 154504 (2015).
[18] V. Ryzhii, Appl. Phys. Lett. 70, 2532 (1997).
[19] V. Ryzhii, Jpn. J. Appl. Phys. 37, 5937 (1998).
[20] V. Ryzhii, A. Satou, and T. Otsui, J. Appl. Phys. 101, 024509 (2007).
[21] M. I. Dyakonov and M. S. Shur, IEEE Trans. Electron Devices, 43, 1640 (1996).
[22] W. Knap, Y. Deng, S. Rumyantsev, J.-Q. Lu, M. S. Shur, C. A. Saylor, and L. C. Brunel, Appl. Phys. Lett. 80, 3433 (2002).
[23] X. G. Peralta, S. J. Allen, M. C. Vanke, N. E. Harff, J. A. Simmons, M. P. Lilly, J. L. Reno, P. J. Burke, and J. P. Eisenstein, Appl. Phys. Lett. 81, 1627 (2002).
[24] T. Otsui, M. Hanane and O. Ogawara, Appl. Phys. Lett. 85, 2119 (2004).
[25] J. Lusakowski, W. Knap, N. Dyakonova, L. Varani, J. Mateos, T. Gonzales, Y. Roelens, S. Bullaert, A. Cappy and K. Karpierz, J. Appl. Phys. 97, 064307 (2005).
[26] F. Teppe, W. Knap, D. Veksler, M. S. Shur, A. P. Dmitriev, Y. Yu. Kucharovski, and S. Rumyantsev, Appl. Phys. Lett. 87, 052105 (2005).
[27] V. Ryzhii, A.Satou, W.Knap, and M.S.Shur. J. Appl. Phys. 99, 084507 (2008).
[28] A. El Fatimy, F. Teppe, N. Dyakonova, W. Knap, D. Seliuta, G. Valusis, A. Shcherepetov, Y. Roelens, S. Bollaert, A. Cappy, and S. Rumyantsev, Appl. Phys. Lett. 89, 131926 (2006).
[29] J. Torres, P. Nouvel, A. Akwaoue-Ondo, L. Chusseau, F. Teppe, A. Shcherepetov, and S. Bollaert, Appl. Phys. Lett. 89, 201101 (2006).
[30] S.A. Boubanga Tombet, Y. Tanimoto, A. Satou, T. Suemitsu, Y. Wang, H. Minamide, H. Ito, D. V. Fateev, V.V. Popov, and T. Otsui, Appl. Phys. Lett., 104, 262104 (2014).
[31] Y. Kurita, G. Ducournau, D. Coquillat, A. Satou, K. Kobayashi, S.A. Boubanga-Tombet, Y.M. Meziani, V.V. Popov, W. Knap, T. Suemitsu, and T. Otsui Appl. Phys. Lett., 89, 201101 (2006).
[32] L. Vicarelli, M. S. Vitiello, D. Coquillat, A. Lombardo, A. C. Ferrari, W. Knap, M. Polini, V. Pellegrini, and A. Tredicucci, Nature Mat. 11, 865 (2012).
[33] V. Ryzhii, T. Otsui, M. Ryzhii, and M. S. Shur, J. Phys: D: Appl. Phys. 45, 302001 (2012).
[34] V. Ryzhii, A. Satou, T. Otsui, M. Ryzhii, V. Mitin, and M. S. Shur, J. Phys: D: Appl. Phys. 46, 315107 (2013).
[35] M. S. Shur Physics of Semiconductor Devices, (Prentice Hall, Mew Jersey, 1990).
[36] S. Luryi, Appl. Phys. Lett. 52, 501 (1988).
[37] T. Fang, A. Konar, H. Xing, and D. Jena, Appl. Phys. Lett. 91, 092109 (2007).
[38] L.A. Falkovsky and A.A. Varlamov, Eur. Phys. J. B 56, 281 (2007).
[39] V.P. Gusynin, S.G. Sharapov, Phys. Rev. B 73, 245411 (2006).
[40] J. Cserti, Phys. Rev. B 75, 033405 (2007).
[41] L. A. Falkovsky, J. Phys.: Conf. Series 129, 012004 (2008).
[42] A. H. Castro Neto, F Guinea, N.M. R Peres, K. S. Novoselov, A. K. Geim, Rev. Mod. Phys. 81, 109, 2009.
[43] D. Svintsov, V. Vyurkov, S. Yurchenko, T. Otsuji, and V. Ryzhii, J. Appl. Phys. 111, 083715 (2013).
[44] B. N. Narozhny, I. V. Gornyi, M. Titov, M. Schutt, and A. D. Mirlin Phys. Rev. B 91, 035414 (2015).
[45] M. A. Kats, N. Y. Geneved, Z. Gaburro, and F. Capasso, Opt. Exp. 19, 21748 (2011).
[46] L. D. Landau and E. M. Lifshitz, The Theory of Fields, (Pergamon, Oxford, 1971).
[47] D. J. Griffiths, Introduction to Electrodynamics, (Benjamin Cummings, 1999)
[48] R. E. Colin, it Antenna and Radiowave Propagation (McGraw-Hill New York, 1985).
[49] M. Miller, J. Schmalian and L. Fritz, Phys. Rev. Lett. 103, 025301 (2009).
[50] M. Mendoza, H. J. Herrmann, and S.ucci, Scientific Reports 3, 1052 (2013).
[51] D. Coquillat, V. Nodjiadjim, A. Konczykowska, M. Riet, N. Dyakonova, C. Consejo, F. Teppe, J. Godin, W. Knap, IRMMW-THz: Int. Conf. on Infrared, Millimeter, and Terahertz Waves, Tucson, AZ, USA, 2014.
[52] D. Coquillat, V Nodjiadjim, A Konczykowska, N Dyakonova, C Consejo, S Ruffenach, F Teppe, M Riet, A Muraviev, A Gutin, M Shur, J Godin, and W Knap,
[53] D. Coquillat, V. Nodjiadjim, A. Konczykowska, N. Dyakonova, C. Consejo, S. Ruffenach, F. Teppe, M. Riet, A. Muraviev, A. Gutin, M. Shur, J. Godin, W. Knap, Proc. 40th IRMMW-THz: Int. Conf. on Infrared, Millimeter, and Terahertz Waves (Hong Kong, China), 2015.