HOT-ELECTRON FLUCTUATIONS AND TRANSISTOR PERFORMANCE: HOT-PHONON EFFECTS

A. Matulionis [a] and H. Morkoç [b]

[a] *Center for Physical Sciences and Technology, A. Goštauto 11, 01108 Vilnius, Lithuania*
E-mail: matulionis@pfi.lt
[b] *Virginia Commonwealth University, 601 West Main street, Richmond, VA 23284, USA*

The frequency performance of a GaN heterostructure field-effect transistor is discussed in terms of plasmon-assisted dissipation of LO-mode heat accumulated by non-equilibrium longitudinal optical phonons (hot phonons). Hot phonons cause additional scattering of electrons and also facilitate defect formation. Plasmon–LO-phonon resonance is experimentally resolved in a wide range of experiments such as fluctuations, dissipation, hot-electron transport, transistor frequency performance, transistor reliability, and transistor phase noise.

Keywords: fluctuations, hot electrons, hot phonons, field-effect transistors

PACS: 73.50.Fq, 73.50.Td, 73.61.Ey, 73.63.Hs

1. Introduction

Hot-electron fluctuations originate from scattering and are a key in understanding fast and ultrafast kinetic processes in conductive channels subjected to high electric fields [1]. A channel with two-dimensional electron gas (2DEG) is the main component of a heterostructure field-effect transistor (HFET). A channel in a gallium nitride HFET is located at the hetero-interface between a nominally undoped GaN and an AlN and/or its alloys [2]. The initial 2DEG density in an as-grown channel depends on the heterostructure composition: that is, an InAlN alloy barrier supports channels with the highest 2DEG density [3]. Hot-electron effects are most important in the active part of the channel located underneath the gate electrode where the electric field and current concentrate. High density of supplied electric power causes heat accumulation aggravated by build-up of non-equilibrium distribution of longitudinal optical (LO) phonons briefly referred to as "hot phonons" [4]. The hot phonons stay "imprisoned" inside the 2DEG channel until the associated heat is converted into acoustic phonons and other vibrations [5]. Since the 2DEG is formed by the mobile electrons owing to polarization charges, no doping is needed for channel formation. Thus, the hot-electron and hot-phonon effects are not masked by impurity scattering in GaN 2DEG channels. Because of strong Fröhlich interaction in GaN, emission and reabsorption of LO phonons by hot electrons play the main role at high electric fields while electron interaction with acoustic phonons manifests itself predominately at low electric fields.

GaN HFETs demonstrate excellent frequency performance at microwave and millimetre-wave frequencies: its unity-gain cut-off frequency exceeds 300 GHz [6, 7]. The record achievements are supported by advanced short-gate technology and low intensity of impurity scattering. The main disadvantage of short-gate HFETs is low generated power at millimetre-wave frequencies.

The improvement of HFET performance can be expected if the channel contained more electrons. Simultaneously, charging of parasitic and other capacitances would take less time in benefit for frequency performance. A large body of reported maximum cut-off frequencies is observed at some optimal gate voltage: a stronger negative bias is needed when the initial 2DEG density is higher [8, 9] and vice versa [10–12]. While the optimal gate voltage is a linear function of the initial 2DEG density in as-grown channels, the optimal 2DEG density in the active part of the channel (underneath the gate) is found independent of the gate length and the initial 2DEG density [13]. The resonance-type dependences of the cut-off frequency, phase noise, and reliability of HFETs on the active 2DEG density correlate with similar resonance-type dependences of the hot-electron drift velocity, fluctuations and dissipation of LO-mode heat on the initial 2DEG density in as-grown channels [14].

The resonant dependence of the cut-off frequency on the 2DEG density necessitates reconsidering physical backgrounds of HFET operation. The novel fluctuation-based approach has been proposed [15] in order to explain the resonant performance of nitride HFETs. This paper discusses the resonant phenomenon in terms of hot-phonon lifetime.

2. Plasmon-assisted decay of hot phonons

The most direct technique for measuring hot-phonon lifetime is pump–probe Raman scattering [16]. A femtosecond pulse of heavily absorbed laser radiation generates high-energy electrons, they emit hot phonons, and the phonon decay is detected by a femtosecond probe pulse of another laser beam through measurement of time-dependent intensity of anti-Stokes line. The lifetime decreases as the electron density increases in bulk GaN [17]. The plasmon-assisted decay of hot phonons is a plausible explanation for the decrease of the lifetime when the plasmon energy approaches that of the LO-phonon [18].

Shorter lifetimes cause higher electron drift velocity [19]. Correspondingly, a higher cut-off frequency can be expected for HFETs fabricated from structures with higher electron density unless the plasmon energy exceeds that of the LO-phonon. The bulk electron density for plasmon–LO-phonon energy crossover is ~$1 \cdot 10^{19}$ cm^{-3} in GaN.

The Raman pump–probe technique is difficult to apply to a 2DEG channel with high equilibrium density of electrons, in particular, when the electron density exceeds the value for plasmon–LO-phonon energy crossover. The highest electron densities per unit volume tend to 10^{20} cm^{-3} for GaN 2DEG channels of interest for HFETs [15].

An alternative technique is based on electron fluctuations [19] (for details see Refs. [20–22]). Figure 1(a) illustrates the observed density dependent hot-phonon lifetime for 2DEG channels (symbols) located in GaN and InGaAs layers [23]. Note that the lifetime is the shortest at the 2DEG density of ~$6.5 \cdot 10^{12}$ cm^{-2} and ~$2.7 \cdot 10^{12}$ cm^{-2} for the nitride and the arsenide 2DEG channels, respectively (fitted resonance curves). The resonance takes place in the vicinity of the plasmon–LO-phonon energy crossover, which is in qualitative agreement with the expectation: the crossover density is higher in nitrides because the LO-phonon energy is higher and the plasmon energy is lower as compared with arsenides.

The resonance density shifts when the electron temperature increases and the electron gas occupies more volume because of thermal expansion (Fig. 1(b), stars) [15]. In particular, the resonance density is $1 \cdot 10^{13}$ cm^{-2}
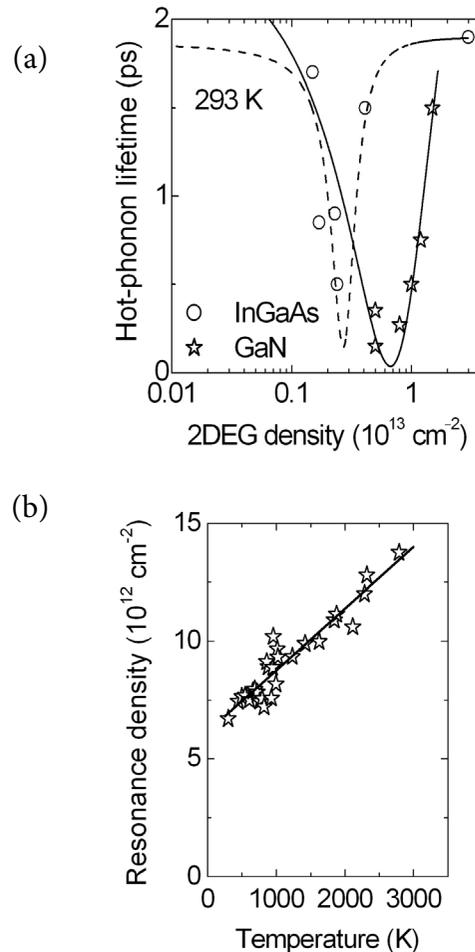

Fig. 1. Plasmon-assisted decay of hot phonons: (a) resonance dependence on electron density for GaN and InGaAs 2DEG channels [23], (b) dependence on hot-electron temperature of resonance density for GaN 2DEG channel (stars) [15].

when the electron temperature is in the vicinity of 1400 K in the GaN 2DEG channel.

The electron temperature increases with the supplied electric power. Correspondingly, the hot-phonon lifetime depends on the supplied electric power as illustrated in Fig. 2 [24]. While the lifetime is ~300 fs at low levels of the supplied power, the minimum value of ~30 fs is reached at the power of ~20 nW/electron, and the values above 100 fs are found at high power levels. The shortest hot-phonon lifetime is observed at ~700 K electron temperature. At this temperature, the resonance 2DEG density is estimated to be $8 \cdot 10^{12}$ cm^{-2}.

The resonance has also been resolved in drift velocity experiments on gateless GaN channels [25]. Symbols in Fig. 3 illustrate the electron drift velocity measured for different channels at the same applied electric field of 60 kV/cm. The results are fitted with the resonance curve: the highest drift velocity is found at a 2DEG
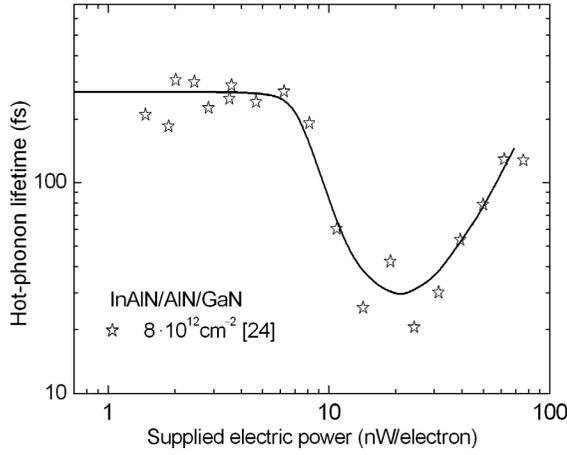
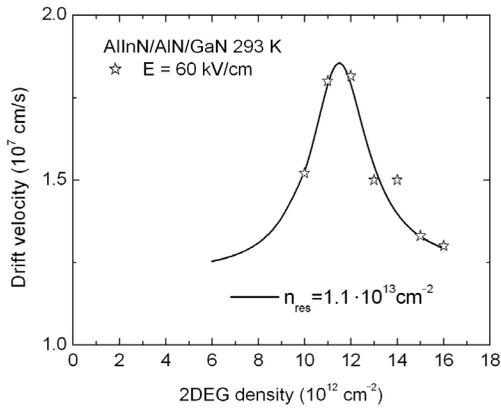Fig. 2. Dependence of hot-phonon lifetime on supplied power [24].

Fig. 3. Resonance dependence of electron drift velocity on 2DEG density for GaN 2DEG channel [25].

density of $1.1 \cdot 10^{13}$ cm^{-2}. The maximum of the electron drift velocity has also been observed at $9.5 \cdot 10^{12}$ cm^{-2} in an independent experiment [26]. The resonance density depends on electric field [25], in qualitative agreement with the effect of thermal expansion of electron gas illustrated in Fig. 1(b).

3. Plasmon-assisted resonance in GaN HFETs

The highest cut-off frequency of a given HFET is reached when certain optimal voltage is applied to the gate electrode. Consider the recent short-gate GaN HFETs [6–12]: the 300 GHz milestone has been reached and exceeded (Fig. 4). When electron density is higher, a stronger negative bias is needed on the gate. For example, the optimal gate voltage is $V_g^{\mathrm{opt}} = -0.5$ V when the initial 2DEG density is $n_{\mathrm{init}} = 1.2 \cdot 10^{13}$ cm^{-2} in the as-grown channel, but $V_g^{\mathrm{opt}} = -5.6$ V when
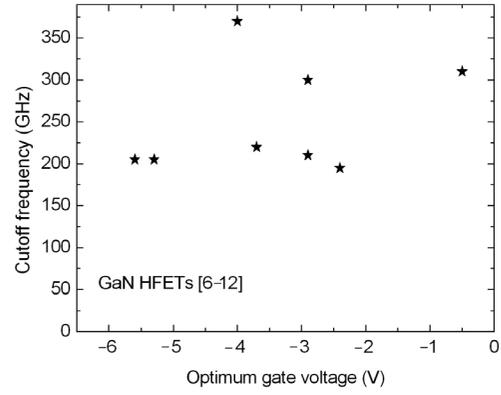

Fig. 4. Unity-gain cut-off frequency against the optimum gate voltage for the best short-gate GaN HFETs fabricated from heterostructures with different as-grown 2DEG density symbols [6–12]).

$n_{\mathrm{init}} = 2.4 \cdot 10^{13}$ cm^{-2} (Fig. 4). The negative gate voltage reduces the 2DEG density in the active part of the channel underneath the gate.

Let us estimate the electron density in the channel under the gate as follows:

$$n_g = n_{\mathrm{init}} - (1/e)\, C_g \left| V_g \right|, \tag{1}$$

where V_g is the gate voltage, C_g is the gate capacitance, n_{init} is the initial 2DEG density in the as-grown channel, and e is the elementary charge. An HFET is most vulnerable when a negative voltage is applied to the gate electrode and, after Eq. (1), the electron density underneath the gate is reduced. Correspondingly, the current and the electric field concentrate in the active part of the channel under the gate where the electron density is the lowest and the resistivity is the highest. It is then natural to expect that the channel degradation will most likely originate underneath the gate when an HFET is electrically stressed with high drain and gate voltages.

Phase noise is a sensitive tool for monitoring defect formation. A set of nearly identical GaN HFETs was subjected to the same drain voltage for a long period of time at different fixed gate voltages applied to each HFET [27]. Figure 5 illustrates the change in phase noise caused by the stress and measured before and after the stress at low drain and gate voltages.

As expected, the phase noise is intense at the strongest negative bias on the gate during the stress (see Fig. 5, star at $n_g = 0.34 \cdot 10^{13}$ cm^{-2}). However, even a 2 dB stronger noise is not expected but observed when the gate bias is weak ($n_g > 1.3 \cdot 10^{13}$ cm^{-2}). The lowest damage is observed near $n_g = 1 \cdot 10^{13}$ cm^{-2}.

The experimental dependence on the average 2DEG density n_g (Fig. 5, stars) can be fitted with a simple
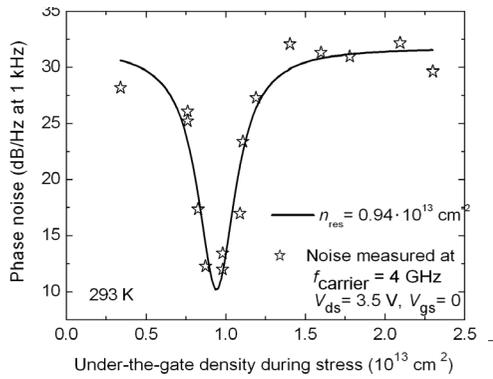
Fig. 5. Resonance dependence of phase noise on average 2DEG density underneath the gate in electrically stressed GaN HEMT [27].

resonance curve. The optimal 2DEG density for the slowest degradation is $0.94 \cdot 10^{13}$ cm^{-2} (curve). The minimum is explained by the lowest number of generated defects – hundred times lower when the HFET is stressed under near-the-resonance conditions. The plasmon-assisted decay of hot phonons suggests a plausible explanation: if the defects are generated during the stress by hot phonons, the hot-phonon effect is weaker when their number is lower, that is, when the lifetime is shorter. The estimated resonance density is in agreement with the data of Fig. 1(b).

Figure 6 illustrates the optimal 2DEG density underneath the gate deduced from the cut-off frequency experiments on GaN HFETs (Fig. 4). The results are plotted against the optimal gate voltage responsible for the best frequency performance. A comparison of
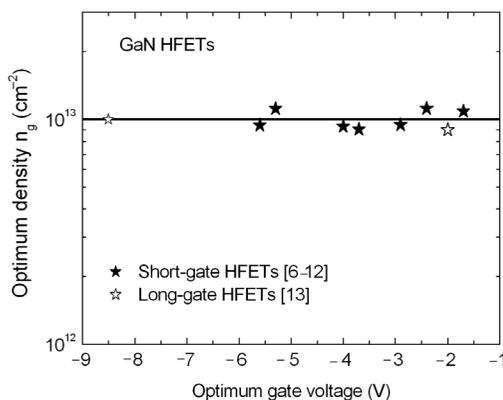

Fig. 6. Optimum under-the-gate density against optimum gate voltage for GaN HFETs fabricated from heterostructures with different as-grown 2DEG density: short-gate devices with cut-off frequency exceeding 200 GHz (closed symbols [6–12]), long-gate devices with cut-off frequency below 15 GHz (open symbols [13]).

open symbols for the long-gate GaN HFETs and closed symbols for the best short-gate GaN HFETs demonstrates that the results are nearly independent of the gate length and the initial 2DEG density. Thus, the conditions for the best frequency performance are more or less the same despite different gate lengths, gate voltages and cut-off frequencies.

The HFET cut-off frequency is measured as a function of the average 2DEG density n_g in the active part of the channel of a GaN HFET [26]. The hot-electron drift velocity is extracted, and the dependence on the density is compared with the resonance curve. The estimated resonance density is $0.95 \cdot 10^{13}$ cm^{-2}. The resonance is explained in terms of the hot-phonon lifetime if dependence of the resonance density on the electron temperature is taken into account (Fig. 1(b)).

4. Discussion

A large and wide body of experimental data show resonance-type dependence in diverse phenomena such as HFET cut-off frequency, phase noise, and device damage (Figs. 5, 6). A resonance LC circuit is often used to explain resonances in the equivalent circuit approach. However, the same hypothetic LC circuit cannot explain the observed resonances in diverse experiments on frequency performance, phase noise, and device damage simultaneously. Therefore, the equivalent circuit approach is not a useful tool for understanding the main cause.

Supposing that there is some fundamental process behind the resonance observed in HFETs, it should manifest itself in any device. Therefore, considering the simplest device can be helpful. The results on gateless channels demonstrate similar resonances (Figs. 1–3). The hot-electron drift velocity (Fig. 3) is an example of this sort: the resonance is resolved at a fixed electric field when the drift velocity is measured for channels with different electron densities. The drift velocity data can be used to estimate the average time between collisions. The result is in the range of hundreds of femtoseconds. Consequently, ultrafast scattering is involved, and the Boltzmann equation for hot electrons seems to be the most suitable instrument for considering the resonance.

The Boltzmann equation accepts any hypothetic scattering mechanism. If a resonant scattering mechanism were inserted into the Boltzmann kinetic equation for electrons, the electron drift velocity would decrease in the vicinity of the resonance. This contradicts the experimental data (Fig. 3): the measured drift velocity is higher in channels having the resonant density. Thus, the approach of the Boltzmann kinetic equation for hot electrons fails.

Let us consider coupled kinetic equations for hot electrons and hot phonons. When a hypothetical resonance hot-phonon decay mechanism is inserted into the Boltzmann kinetic equation for hot phonons, the phonon mode occupancy decreases at the resonance and causes the associated reduced electron–phonon scattering in agreement with the experimental data on electron drift velocity (Fig. 3). Thus, a plausible explanation follows from coupling of hot electrons and hot phonons. Intense electron–LO-phonon interaction and weak coupling with the thermal bath causes formation of an almost isolated hot subsystem composed of hot electrons and hot phonons [19]. The main path for energy dissipation from this hot subsystem is conversion of the LO-mode heat into the LA-mode heat (Fig. 7).

Fluctuations possess important information about fast and ultrafast processes. In particular, the time constant of the conversion can be estimated from experiments on hot-electron fluctuations: the pioneering result of 350 fs for AlGaN/GaN 2DEG channels [19] is in excellent agreement with the result of 380 fs obtained two years later for a similar channel through sophisticated femtosecond pump–probe experiment on LO-phonon-assisted intersubband absorption [28]. These results are in accord with the electron-density-dependent hot phonon lifetime obtained from Raman experiment for *bulk* GaN [17] when the electron density in the Raman experiment approaches the average bulk density in the 2DEG channels [29].

The fluctuation technique is perfectly suited for measuring the hot-phonon lifetime in typical voltage-biased GaN-based channels for HFETs. The spatial resolution is high: the fluctuations of current originate where the current flows. Consequently, the technique is applicable to channels of nanometric thickness. The dependence on the electron density and the supplied electric power is essential. The observed non-monot-

onous dependence on the 2DEG density is explained in terms of plasmon-assisted conversion of the LO-mode heat into the LA-mode heat. The optimal conditions for the fastest conversion are associated with the LO-phonon–plasmon resonance. The signatures of the plasmon-assisted conversion of the LO-mode heat manifest themselves in device degradation experiments as well [26, 30]. The HFET degrades relatively slowly and operates faster when dissipation of the LO-mode heat is faster.

The optimal 2DEG density depends on the electron temperature controlled by the supplied power (Fig. 2 [24]). The observed ultrafast conversion of the LO-mode heat into LA-mode heat with subsequent drain into the remote heat sink helps in obtaining high drift velocity of the electrons at high electric fields (Fig. 3). Correspondingly, the frequency performance of GaN HFETs is improved when the active part of the transistor is operated under the near-resonance electron densities (Fig. 6). The ability to tune the LO-phonon lifetime is an important finding crucial for advanced GaN HFET technology.

Acknowledgements

References

[1] H.L. Hartnagel, R. Katilius, and A. Matulionis, *Microwave Noise in Semiconductor Devices* (Wiley, New York, 2001).

[2] H. Morkoç, *Handbook of Nitride Semiconductors and Devices*, vol. 3 (Wiley-VCH, Weinheim, 2009).

[3] J. Kuzmik, IEEE Electron Device Lett. **22**, 510–12 (2001).

[4] P. Kocevar, Physica B & C **134**, 155–158 (1985).

[5] A. Matulionis, J. Liberis, I. Matulionienė, M. Ramonas, and E. Šermukšnis, Proc. IEEE **98**, 1118–26 (2010).

[6] D.S. Lee, X. Gao, S. Guo, D. Kopp, P. Fay, and T. Palacios. IEEE Electron Device Lett. **32**, 1525–27 (2011).

[7] Y. Yuanzheng et al. IEEE Electron Device Lett. **33**, 988–90 (2012).

[8] H. Sun, A.R. Alt, H. Benedickter, E. Feltin, J-F. Carlin, M. Gonschorek, N. Grandjean, and C.R. Bolognesi, IEEE Electron Device Lett. **31**, 957–959 (2010).

[9] S. Tirelli, D. Marti, H. Sun, A.R. Alt, J-F. Carlin, N. Grandjean, and C.R. Bolognesi, IEEE Electron Device Lett. **32**, 1364–1366 (2011).

[10] D.S. Lee, J.W. Chung, H. Wang, X. Gao, Sh. Guo, P. Fay, and T. Palacios, IEEE Electron Device Lett. **32**, 755–757 (2011).
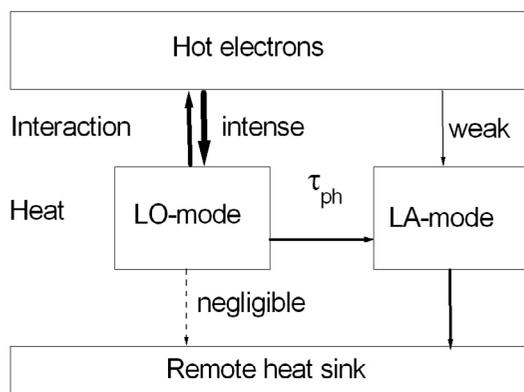

Fig. 7. Main path of heat dissipation supported by plasmon-assisted conversion of LO-mode heat into LA-mode heat.

[11] R. Wang, G. Li, O. Laboutin, Y. Cao, W. Johnson, G. Snider, P. Fay, D. Jena, and H. Xing, IEEE Electron Device Lett. **32**, 892–894 (2011).

[12] R. Wang, G. Li, J. Verma, B. Sensale-Rodriguez, T. Fang, J. Guo, Z. Hu, O. Laboutin, Y. Cao, W. Johnson, G. Snider, P. Fay, D. Jena, and H. Xing, IEEE Electron Device Lett. **32**, 1215–1217 (2011).

[13] A. Šimukovič, A. Matulionis, J. Liberis, E. Šermukšnis, P. Sakalas, F. Zhang, J.H. Leach, V. Avrutin, and H. Morkoç, Semicond. Sci. Technol. **28**, 055008 (2013).

[14] A. Matulionis, Semicond. Sci. Technol. **28**, 074007 (2013).

[15] A. Matulionis, J. Liberis, I. Matulionienė, E. Šermukšnis, J.H. Leach, M. Wu, and H. Morkoç, Physica Status Solidi A **208**, 30–36 (2011).

[16] J.A. Kash and J.C. Tsang, Nonequilibrium phonons in semiconductors, in: *Spectroscopy of Nonequilibrium Electrons and Phonons*, eds. C.V. Shank and B.P. Zakharchenya, Modern Problems in Condensed Matter, vol. 35 (Elsevier, Amsterdam, 1992), pp. 113–167.

[17] K.T. Tsen, J.G. Kiang, D.K. Ferry, and H. Morkoç, Appl. Phys. Lett. **89**, 112111 (2006).

[18] A. Dyson and B.K. Ridley, J. Appl. Phys. **103**, 114507 (2008).

[19] A. Matulionis, J. Liberis, I. Matulionienė, M. Ramonas, L.F. Eastman, J.R. Shealy, V. Tilak, and A. Vertiatchikh, Phys. Rev. B **68**, 035338 (2003).

[20] E. Šermukšnis, J. Liberis, and A. Matulionis, Lith. J. Phys. **47**, 491–498 (2007).

[21] A. Matulionis, Lith. J. Phys. **47**, 297–302 (2007).

[22] J. Liberis, I. Matulionienė, M. Ramonas, and L.F. Eastman, Hot phonons in high-power microwave HEMT and FET channels, in: *Advanced Semiconductor Materials and Devices Research: III-Nitrides and SiC*, ed. Ho-Young Cha (Transworld Research Network, Kerala, India, 2009), pp. 203–242.

[23] A. Matulionis, Phys. Status Solidi C **6**, 2834–2839 (2009).

[24] A. Matulionis, J. Liberis, I. Matulionienė, M. Ramonas, E. Šermukšnis, J.H. Leach, M. Wu, X. Ni, X. Li, and H. Morkoç, Appl. Phys. Lett. **95**, 192102 (2009).

[25] L. Ardaravičius, J. Liberis, O. Kiprijanovič, A. Matulionis, M. Wu, and H. Morkoç, Phys. Status Solidi Rapid Res. Lett. **5**, 65–67 (2011).

[26] J.H. Leach, C.Y. Zhu, M. Wu, X. Ni, X. Li, J. Xie, Ü. Özgür, H. Morkoç, J. Liberis, E. Šermukšnis, A. Matulionis, T. Paskova, E. Preble, and K.R. Evans, Appl. Phys. Lett. **96**, 133505 (2010).

[27] C. Kayis, R.A. Ferreyra, M. Wu, X. Li, Ü. Özgür, A. Matulionis, and H. Morkoç, Appl. Phys. Lett. **99**, 063505 (2011).

[28] Z. Wang, K. Reimann, M. Woerner, T. Elsaesser, D. Hofstetter, J. Hwang, W.J. Schaff, and L.F. Eastman, Phys. Rev. Lett. **94**, 037403 (2005).

[29] A. Matulionis, J. Phys. Condens. Matter **21**, 174203 (2009).

[30] J.H. Leach, M. Wu, X. Ni, J. Lee, Ü. Özgür, H. Morkoç, J. Liberis, E. Šermukšnis, A. Matulionis, H. Cheng, and C. Kurdak, Appl. Phys. Lett. **95**, 223504 (2009).

KARŠTŲJŲ ELEKTRONŲ FLIUKTUACIJOS IR TRANZISTORIŲ VEIKIMAS: KARŠTŲJŲ FONONŲ VAIDMUO

A. Matulionis [a], H. Morkoç [b]

[a] *Fizinių ir tecnologijos mokslų centras, Vilnius, Lietuva*
[b] *Virdžinijos Commonwealth universitetas, Ričmondas, JAV*

Santrauka

Galio nitrido įvairialyčių lauko tranzistorių sparta aptariama pasitelkus žinias apie karštųjų fononų skilimą į akustinius fononus dalyvaujant kitokių svyravimų kvantams. Karštaisiais fononais įprasta vadinti išilginius optinius fononus, kuriuos išspinduliuoja karštieji elektronai, taip atsikratydami savo perteklinės energijos. Kaupdamiesi dvimatėje protakoje, karštieji fononai trukdo elektronams judėti, taigi jie riboja tranzistoriaus veikos dažnį. Dėl šios priežasties labai spartus fononų skilimas gerina tranzistorių dažnio ir kitokias savybes. Vilniuje sukurtas fliuktuacinis metodas yra pagrindinis būdas karštųjų fononų skilimo pusamžiui matuoti dvimatėse protakose femtosekundinių trukmių ruože. Metodas buvo aprobuotas tiriant silicį, kur patikimų rezultatų yra gautų ir kitais metodais. Pastebėtas labai spartus rezonansinis plazmonų skatinamas karštųjų fononų virtimas akustiniais fononais. Tai tranzistorių veiklą gerinantis reiškinys – tranzistoriaus protaka atsikrato karštųjų fononų, padidėja elektronų dreifo greitis, sulėtėja tranzistorių senėjimas, sumažėja fazinis ir amplitudinis triukšmas, padidėja ribinis veikos dažnis. Eksperimentiniai karštųjų fononų pusamžio ir susijusių reiškinių tyrimo rezultatai aptariami kartu su proveržiu gerinant įvairialyčių labai sparčių GaN lauko tranzistorių savybes.