

О скорости ударной ионизации в прямозонных полупроводниках

А. Н. Афанасьев¹⁾, А. А. Грешнов, Г. Г. Зегря

Физико-технический институт им. А.Ф. Иоффе РАН, 194021 Санкт-Петербург, Россия

Поступила в редакцию 20 февраля 2017 г.

В рамках 14-зонной $k\text{-}p$ модели исследована интенсивность процессов ударной ионизации в прямозонных полупроводниках и получены аналитические выражения для темпа ударной ионизации. Показано, что вблизи энергетического порога скорость процесса определяется суммой изотропного вклада, кубического по отстройке от порога, и сильно анизотропного квадратичного, возникающего лишь в меру взаимодействия с далекими зонами. Сопоставление этих вкладов в условиях усреднения по невырожденному изотропному распределению неравновесных электронов с некоторой эффективной температурой T^* показывает, что именно кубический, а не традиционно используемый квадратичный вклад доминирует для прямозонных полупроводников с $E_g < 1 - 1.5 \text{ эВ}$ вплоть до $T^* = 300 \text{ К}$, и это должно учитываться при расчетах приборных характеристик устройств, использующих эффект лавинного умножения носителей.

DOI: 10.7868/S0370274X17090077

Введение. Явление ударной ионизации в полупроводниках, состоящее в рождении электрон-дырочной пары в результате кулоновского взаимодействия высокоэнергетичного электрона проводимости с электронами валентной зоны (рис. 1), лежит в основе функционирования большого семей-

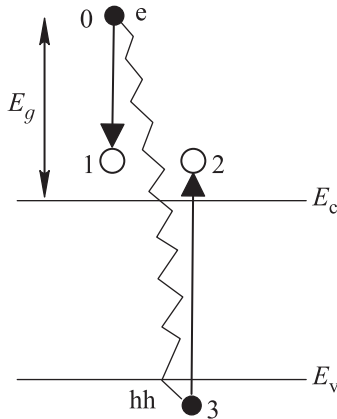


Рис. 1. Схема элементарного акта ударной ионизации

ства устройств современной электроники, таких как лавинно-пролетные диоды (ИМРАТТ), лавинные фотоприемники (АРД) [1], а также транзисторы с полевым контролем ударной ионизации (I-MOS) [2], в которых экспериментально достигнута крутизна подпороговой части ВАХ на уровне 5 мВ/дек при $T = 400 \text{ К}$, что позволяет в разы увеличивать скорость переключения. С точки зрения эффективности применения ударной ионизации в качестве фи-

зического принципа работы приборов наиболее удачным является использование в качестве областей лавинного умножения рассматриваемых в данной работе прямозонных полупроводников с относительно небольшой шириной запрещенной зоны $E_g \lesssim 1 \text{ эВ}$, в которых энергетический порог лишь немного отличается от величины щели: $E_{th} \approx E_g(1 + 2\mu)$ [3], где $\mu = m_e/m_{hh} \ll 1$. Одним из существенных факторов, затрудняющих развитие данных устройств, является недостаточная разработанность теоретического описания происходящих в них процессов ударной ионизации. Вследствие отсутствия последовательно полученных аналитических выражений для полного темпа ударной ионизации, большинство "инженерных" расчетов в настоящее время выполняются на основе упрощенных формул со свободно подгоняемыми параметрами, что не может не отражаться на адекватности получаемых результатов. В литературе существует неопределенность не только относительно значения префактора C , но и степени n в соотношении

$$W(E) = C(E - E_{th})^n, \tag{1}$$

описывающем зависимость темпа процесса от энергии налетающего электрона вблизи порога ударной ионизации E_{th} . Хотя в целом наиболее популярны представления о квадратичной зависимости $W(E)$ ($n = 2$), предложенной из общих соображений Л.В. Келдышем еще в 1959 г. [4], авторы работ [5, 6] пришли к выводу, что результаты численных расчетов $W(E)$ для тройного соединения $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ наилучшим образом описываются с помощью таких степеней, как $n = 2.5$ [5], $n = 4.3$ [5] и $n = 5.6$ [6].

¹⁾e-mail: afanasiev.an@mail.ru

Кроме того, как отмечали Гельмонт и соавторы в работе [7], для наиболее узкозонных полупроводников с прямой запрещенной зоной, в частности, рассмотренного ими соединения $\text{Cd}_{0.2}\text{Hg}_{0.8}\text{Te}$ с $E_g \approx 0.1$ эВ, квадратичный вклад в $\mathcal{W}(E)$ представляет лишь академический интерес, а на деле зависимость $\mathcal{W}(E)$ является кубической, и для нее было приведено явное выражение (к сожалению, без подробностей его вывода).

В данной работе представлен микроскопический анализ процессов ударной ионизации в прямозонных полупроводниках, в результате которого была получена обоснованная функциональная зависимость темпа ударной ионизации \mathcal{W} от $E - E_{\text{th}}$. Показано, что скорость процесса определяется суммой изотропного вклада, кубического по отстройке от порога, и сильно анизотропного квадратичного, возникающего лишь в меру взаимодействия с далекими зонами. Сопоставление этих вкладов в реалистичных условиях показывает, что доминирование кубического вклада возникает, начиная с $E_g \sim 1$ эВ, и это обстоятельство должно учитываться при численном моделировании полупроводниковых приборов, эксплуатирующих эффект лавинного умножения носителей.

Теория. Для расчета скорости ударной ионизации будем использовать золотое правило Ферми, определяющее темп элементарного процесса ударной ионизации в результате кулоновского взаимодействия горячего электрона зоны проводимости с заполненной валентной зоной:

$$W_{\{0,3\} \rightarrow \{1,2\}} = \frac{2\pi}{\hbar} \left| \langle \alpha_1 \alpha_2 | \hat{V} | \alpha_0 \alpha_3 \rangle \right|^2 \delta(\Delta \mathcal{E}), \quad (2)$$

где нижние индексы нумеруют состояния согласно рис. 1, \hat{V} – оператор кулоновского взаимодействия, $\Delta \mathcal{E} = E_1 + E_2 - E_3 - E_0$ ²⁾ отражает баланс энергии, а $\alpha_i = \{\mathbf{k}_i, \xi_i\}$ – набор квантовых чисел: волнового вектора и спинового индекса. Темп ударной ионизации для заданного состояния налетающего электрона 0 (фигурирующий в левой части формулы (1)) определяется как

$$\mathcal{W} = \tau^{-1} = \sum_{\alpha_1, \alpha_2, \alpha_3} W_{\{0,3\} \rightarrow \{1,2\}}. \quad (3)$$

Поскольку электроны, способные к ударной ионизации, описываются экспоненциальным хвостом функции распределения [8], затухающим на характерном масштабе порядка нескольких десятков мэВ, в практическом отношении для описания темпа ударной

ионизации достаточно иметь выражение для $\mathcal{W}(E)$ лишь вблизи энергетического порога, где оно складывается из конкурирующих квадратичного и кубического вкладов,

$$\mathcal{W} = \mathcal{W}_2 + \mathcal{W}_3 = A(E - E_{\text{th}})^2 + B(E - E_{\text{th}})^3, \quad (4)$$

причем их сопоставимость, как будет показано ниже, обуславливается относительной малостью коэффициента A в меру отношения ширины запрещенной зоны E_g к расстоянию до далеких зон. Заметим также, что поскольку разница в пороговых энергиях для процессов ударной ионизации с участием тяжелых и легких дырок близка к ширине запрещенной зоны, вкладом последних можно пренебречь ввиду экспоненциальной малости по параметру E_g/T , где T – эффективная температура неравновесной функции распределения горячих электронов.

Для последующего анализа удобно воспользоваться представлением кулоновского взаимодействия в виде интеграла Фурье [9], а волновые векторы участвующих в процессе частиц представить в виде суммы пороговых значений \mathbf{k}_i^{th} и малых отстроек $\mathbf{q}_i = \mathbf{k}_i - \mathbf{k}_i^{\text{th}} \ll \mathbf{k}_i^{\text{th}}$. С учетом малости величин q_i выражение для $\mathcal{W}(E)$ преобразуется к виду

$$\mathcal{W} = \frac{\pi \hbar F \left(\frac{\Delta_0}{E_g} \right)}{12 m_e E_g^2} \left(\frac{4\pi e^2}{\varepsilon} \right)^2 \int \frac{d^3 q_1 d^3 q_2}{(2\pi)^6} [\mathcal{I}_{cv}(\mathbf{q}_1, \mathbf{q}_3) + \mathcal{I}_{cv}(\mathbf{q}_2, \mathbf{q}_3)] \delta \left(q_1^2 + q_2^2 - \frac{2m_e(E_0 - E_{\text{th}})}{\hbar^2} \right), \quad (5)$$

где $\mathcal{I}_{cv}(\mathbf{q}_i, \mathbf{q}_j) = I_{cv}(\mathbf{k}_i(\mathbf{q}_i), \mathbf{k}_j(\mathbf{q}_j))$, I_{cv} обозначают просуммированные по спиновым переменным квадраты интегралов перекрытия блоховских функций тяжелой дырки и конечного электрона, $\mathbf{q}_3 = \mathbf{q}_1 + \mathbf{q}_2 - \mathbf{q}_0$ выражается через переменные интегрирования с помощью закона сохранения импульса, причем $q_0 = \left(\frac{\partial E_0}{\partial k_0} \right)_{\text{th}}^{-1} (E_0 - E_{\text{th}})$ и сонаправлен с \mathbf{k}_0^{th} , Δ_0 – величина спин-орбитального расщепления валентной зоны, а

$$F(x) = \frac{(1+x)^2(1+x/3)^3}{(1+7x/9+x^2/6)(1+2x/3)^2(1+x/2)}. \quad (6)$$

Из формулы (5) следует, что последовательные показатели степени в разложении $\mathcal{W}(E)$ по $E - E_{\text{th}}$ соответствуют разложению \mathcal{I}_{cv} по \mathbf{q}_i . В частности, квадратичный вклад определяется интегралами перекрытия для пороговой конфигурации участвующих в процессе частиц $\mathcal{I}_{cv}(\mathbf{0}, \mathbf{0}) = I_{cv}(\mathbf{k}_i^{\text{th}}, \mathbf{k}_j^{\text{th}})$, в соответствии с которой для изотропного закона дисперсии

²⁾В формулах, задействующих индексы состояний других частиц, энергию начального электрона мы будем дополнять индексом “0” (E_0 вместо E).

волновые векторы в начальных и конечных состояниях должны быть коллинеарны [10], а импульс налетающего электрона в пределе $\mu = m_e/m_{hh} \ll 1$ отдается преимущественно дырке [3]. Соответственно, вклады с более высокой степенью (в том числе и кубический) определяются неколлинеарным взаиморасположением волновых векторов. Поэтому для определения коэффициентов A и B в формуле (4) достаточно выполнить расчет вкладов в \mathcal{I}_{cv} , происходящих из разложения по отстройкам волновых векторов от пороговых (\mathbf{q}_i). При этом мы будем использовать 14-зонную $\mathbf{k} \cdot \mathbf{p}$ модель [11], явно учитывающую шестикратно вырожденную вторую зону проводимости, которая расположена на несколько электронвольт выше E_c (рис. 2). Если же ограничиться

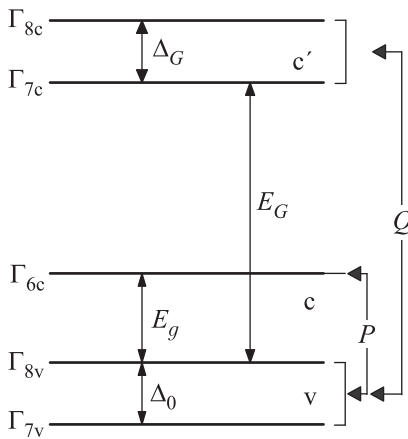


Рис. 2. Графическое представление 14-зонной $\mathbf{k} \cdot \mathbf{p}$ модели.

учетом лишь $\mathbf{k} \cdot \mathbf{p}$ взаимодействия между зоной проводимости (c) и валентной зоной (v), полученный интеграл перекрытия будет соответствовать нулевому значению квадратичного вклада, так как блоховские функции тяжелой дырки и электрона ортогональны при коллинеарных волновых векторах в рамках восьмизонной модели Кейна, не включающей слагаемые латтинжеровского типа [12, 13]. При этом выражение для коэффициента B имеет вид

$$B = \frac{\omega_B^*}{18E_g^3} \frac{E_g + \Delta_0}{E_g + \frac{2}{3}\Delta_0} F\left(\frac{\Delta_0}{E_g}\right), \quad (7)$$

которое в пределе $\Delta_0 \gg E_g$ переходит в приведенное в работе [7] (здесь $\omega_B^* = m^*e^4/2\kappa^2\hbar^3$ – боровская частота для электронов зоны проводимости). Заметим, что кубический вклад не зависит от направления волнового вектора налетающего электрона, поскольку он происходит от сферически симметричной части 14-ти зонного $\mathbf{k} \cdot \mathbf{p}$ гамильтониана.

Для расчета квадратичного вклада удобно воспользоваться малостью отношения величин E_g, Δ_0, Δ_G к расстоянию между валентной зоной (v) и второй зоной проводимости (c'), и в результате расчета \mathcal{I}_{cv} по теории возмущений мы приходим к следующему выражению для коэффициента при квадратичном вкладе:

$$A = \frac{4}{3} \frac{\omega_B^*}{E_G^2} \frac{Q^4}{P^4} J_4(\mathbf{u}) \frac{E_g + \frac{1}{2}\Delta_0}{E_g + \frac{1}{3}\Delta_0} F\left(\frac{\Delta_0}{E_g}\right), \quad (8)$$

где P и Q – матричные элементы оператора импульса между блоховскими функциями зон c - v и c' - v соответственно (рис. 2), $J_4(\mathbf{u}) = I_1(\mathbf{u}) - 3I_1^2(\mathbf{u})$ – комбинация кубических инвариантов четвертого порядка $I_1(\mathbf{u}) = u_x^2u_y^2 + u_x^2u_z^2 + u_y^2u_z^2$, $\mathbf{u} = \mathbf{k}_0/k_0$ – единичный вектор в направлении движения начального электрона. Данный вклад возникает в меру подмешивания к блоховским функциям тяжелой дырки состояний далеких зон, которое в рамках 14-ти зонной модели описывается блоком $\mathbf{k} \cdot \mathbf{p}$ -взаимодействия валентной зоны со второй зоной проводимости (c'), имеющим кубическую симметрию. Вследствие этого квадратичный вклад приобретает нетривиальную угловую зависимость, продемонстрированную на рис. 3. В частности, квадратичный вклад исчезает, если инициирующий акт ударной ионизации электрон налетает в кристаллографических направлениях [100] или [111]. Также заметим, что коэффициент A содержит в знаменателе вторую степень E_G и является в этом смысле малым по сравнению с коэффициентом B , поэтому конкурентность полученных вкладов, доминирование квадратичного или кубического, зависит от конкретного распределения горячих электронов по энергии.

Обсуждение и выводы. В практическом отношении важную роль играет интегральная скорость ударной ионизации

$$\mathcal{R} = \sum_{\alpha_0} \mathcal{W}(E(\mathbf{k}_0)) \tilde{f}_{\alpha_0}, \quad (9)$$

имеющая смысл числа дополнительных электрон-дырочных пар, генерируемых в единицу времени в результате элементарных процессов ударной ионизации, и зависящая от высокоэнергетического хвоста неравновесной добавки к функции распределения $\tilde{f}_\alpha = f_\alpha - f_0(E(\mathbf{k}))$. Поскольку нечетные гармоники функции распределения (в том числе и главная, генерируемая электрическим полем), согласно формуле (9), дают нулевой вклад в величину \mathcal{R} , для сопоставления двух вкладов в интегральную скорость ударной ионизации мы сравним парциальные вклады $\mathcal{R}_{2,3}$, происходящие от двух слагаемых в формуле

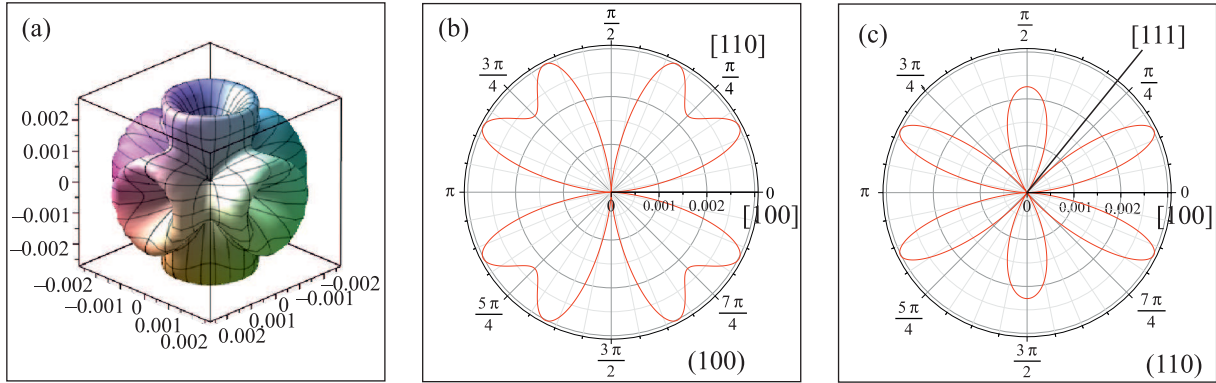


Рис. 3. (Цветной онлайн) Угловая зависимость квадратичного вклада в темп ударной ионизации $\mathcal{W}_2(\Theta_0, \varphi_0)$ в единицах ω_B^* для заданной энергии налетающего электрона $(E - E_{th})/E_g = \sqrt{3/2}$. (a) – Трехмерный график. (b) – Сечение плоскостью (100). (c) – Сечение плоскостью (110)

(4) при неравновесной добавке к функции распределения вида $\tilde{f}(E) = \mathcal{N} \exp(-E/T)$. Здесь T обозначает эффективную температуру распределения горячих электронов (совпадающую в случае слабого внешнего воздействия с температурой решетки). В этом случае выражения для вкладов \mathcal{R}_2 и \mathcal{R}_3 можно записать в виде

$$\mathcal{R}_i = D(E_{th}) \mathcal{N} \int_{E_{th}}^{\infty} \overline{W_i(E)} \exp(-E/T) dE, \quad (10)$$

где горизонтальная черта обозначает усреднение по углам, а $D(E)$ – плотность состояний. Используя величину среднего $\overline{J_4(\mathbf{u})} = 2/35$ и приравнявая получающиеся выражения для $\mathcal{R}_{2,3}$, для эффективной температуры кроссовера между двумя механизмами имеем

$$T^* = \frac{\overline{A}}{3B} = \frac{16 Q^4 E_g^3}{35 P^4 E_G^2} G\left(\frac{\Delta_0}{E_g}\right), \quad (11)$$

где $G(x) = \frac{(1+2x/3)(1+x/2)}{(1+x)(1+x/3)}$. Для параметров полупроводника InAs [11], обладающего запрещенной зоной около 0.4 эВ, данное выражение дает $T^* \approx 10$ К, соответственно при комнатной температуре квадратичный вклад в 50 раз меньше кубического. Как несложно видеть из формулы (11), доминирование кубического вклада над квадратичным при $T^* = 300$ К ожидается для большинства прямозонных полупроводников с $E_g \lesssim 1-1.5$ эВ, при этом с количественной точки зрения полное пренебрежение кубическим вкладом вряд ли оправдано и для более широкозонных материалов.

Необходимо заметить, что хотя предлагаемая аналитическая теория темпа ударной ионизации при анализе квадратичного вклада использует в качестве малого параметра отношение ширины запрещенной зоны E_g к расстоянию от валентной зоны до второй зоны проводимости E_G и ее результаты кажутся

применимыми для произвольного отношения Δ_0/E_g , это не совсем так. Дело в том, что при уменьшении спин-орбитального расщепления валентной зоны дисперсионная ветвь спин-орбитально отщепленных дырок $E_{so}(\mathbf{k})$ приближается к ветви тяжелых дырок и в пределе $\Delta_0 \rightarrow 0$ совпадает с ней, образуя вторую ветвь тяжелых дырок. В таких условиях возникает вырождение, препятствующее применимости использованной в работе формы теории возмущений. Поэтому на самом деле формула (8) корректна при условии $\Delta_0/E_g \gg E_g/E_G$, что на практике может не выполняться для полупроводников из относительно легких элементов с запрещенной зоной более 1 эВ, например для наиболее популярного прямозонного полупроводника GaAs параметр $\beta = \Delta_0 E_G/E_g^2$ равен 0.8, а для InP – 0.25. Однако, как показывают проведенные нами численные расчеты, усредненное по углам значение коэффициента \overline{A} отличается от соответствующего формуле (8) не более чем на 10%, хотя угловая зависимость при этом существенно видоизменяется. На основании проведенного расчета можно утверждать, что и для указанных соединений кубический вклад в интегральный темп ударной ионизации сопоставим с квадратичным при эффективной температуре неравновесной функции распределения горячих электронов порядка 300 К. Также в ближайшее время мы планируем построить аналитическую теорию, описывающую квадратичный вклад в темп ударной ионизации для случая $\beta \ll 1$, трактуя при этом спин-орбитальное взаимодействие в качестве возмущения.

Работа выполнена при поддержке Российского научного фонда, грант # 17-72-20245.

1. S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*, 3rd ed. New Jersey: John Wiley & Sons. Inc. (2007).

2. K. Gopalakrishnan, P. B. Griffin, and J. D. Plummer, IEEE Trans. Electron. Dev. **52**, 69 (2005).
3. В. Н. Абакумов, В. И. Перель, И. Н. Ясиевич, *Безызлучательная рекомбинация в полупроводниках*, Издательство ПИЯФ, СПб. (1997).
4. Л. В. Келдыш, ЖЭТФ **37**, 713 (1959).
5. K. Y. Choo and D. S. Ong, J. Appl. Phys. **96**, 5649 (2004).
6. D. Harrison, R. A. Abram, and S. Brand, J. Appl. Phys. **85**, 8186 (1999).
7. B. Gelmont, K. Kim, and M. Shur, Phys. Rev. Lett. **69**, 1280 (1992).
8. A. P. Dmitriev, M. P. Mikhailova, and I. N. Yassievich, Phys. Stat. Sol. (b) **140**, 9 (1987).
9. Г. Г. Зегря, А. С. Полковников, ЖЭТФ **113**, 1491 (1998).
10. C. L. Anderson and C. R. Crowell, Phys. Rev. B **5**, 2267 (1972).
11. R. Winkler, *Spin-orbit Coupling Effects in two-dimensional Electron and Hole Systems*, Springer-Verlag (2003).
12. E. O. Kane, J. Phys. Chem. Solids **1**, 249 (1957).
13. А. А. Грешнов, Ю. Б. Васильев, Н. Н. Михайлов, Г. Ю. Васильева, Д. Смирнов, Письма ЖЭТФ **97**, 108 (2013).