

Конкуренция между изотропным и сильно анизотропным вкладами в темп ударной ионизации в прямозонных полупроводниках

© А.Н. Афанасьев, А.А. Грешнов, Г.Г. Зегря

Физико-технический институт им. А.Ф. Иоффе Российской академии наук,
194021 Санкт-Петербург, Россия

E-mail: afanasiev.an@mail.ru

Поступила в Редакцию 27 августа 2024 г.

В окончательной редакции 7 сентября 2024 г.

Принята к публикации 26 ноября 2024 г.

Показано, что темп процесса межзонной ударной ионизации в прямозонных кубических полупроводниках со слабым и сильным спин-орбитальным расщеплением валентной зоны является сильно анизотропным при низких эффективных температурах распределения электронов T и становится изотропным при увеличении T . Такое поведение связано с механизмом ударной ионизации, обеспеченным взаимодействием состояний электронов и тяжелых дырок через далекие зоны, которое исчезает в некоторых высокосимметричных направлениях распространения начального электрона, таких как [100] и [111]. При $T = 300$ К темп ударной ионизации в узкозонных полупроводниках InSb, InAs, GaSb и $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ изотропен, а для материалов со средней шириной запрещенной зоны типа InP, GaAs и CdTe изотропный и анизотропный вклады сравнимы. Предложено простое и обоснованное аналитическое обобщение формулы Келдыша, которое может быть использовано при моделировании работы устройств, использующих ударную ионизацию.

Ключевые слова: ударная ионизация, прямозонный полупроводник, **kp**-модель, горячие носители, численное моделирование.

DOI: 10.61011/FTP.2024.11.59485.5877

1. Введение

Явление межзонной ударной ионизации, заключающееся в рождении электрон-дырочной пары в результате кулоновского взаимодействия между горячим электроном проводимости и электронами валентной зоны (рис. 1, *a*), играет важную роль в функционировании многих устройств современной электроники. В некоторых из них, например полупроводниковых диодах и полевых транзисторах, лавинный пробой, вызванный ударной ионизацией, ограничивает диапазон рабочих напряжений. Поэтому традиционно ударная ионизация воспринимается как негативный эффект. В то же время умножение носителей в результате ударной ионизации лежит в основе работы лавинно-пролетных диодов (ИМПАТТ), лавинных фотодиодов (APD) [1] и транзистора с полевым контролем ударной ионизации (I-MOS) [2], который обладает крайне резкой сток-затворной характеристикой. В частности, была экспериментально продемонстрирована работа I-MOS с наклоном подпороговой части ВАХ ~ 5 мВ/дек. при $T = 400$ К, что позволяет существенно снизить скорость переключения устройства по сравнению с МДП транзистором.

Численное моделирование физических процессов, происходящих внутри полупроводниковых приборов, стало неотъемлемой частью разработки устройств. Однако зачастую физические модели, которые используются в программах моделирования, являются феноменологическими и содержат в себе большое число подгоночных параметров. Поскольку на практике выходные характеристики устройств, использующих ударную

ионизацию, зависят от многих подробностей, например конкретного вида зонной структуры материала и особенностей процессов рассеяния, определяющих вид неравновесной функции распределения, расчет ВАХ на основе реалистичной модели зонной структуры сложен с концептуальной и технической точек зрения. Поэтому наиболее популярным способом моделирования является метод Монте-Карло [3–13], однако его результаты зависят от конкретного вида зависимости микроскопического темпа ударной ионизации $W(E)$, от энергии горячего электрона, который ее инициирует. Феноменологически темп ионизации растет степенным образом с превышением энергии над пороговой:

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

Конкретный вид параметров n и C , а также границы применимости (1) устанавливаются на основе квантово-механических расчетов. Наиболее популярной [14] является квадратичная зависимость ($n = 2$), впервые полученная Келдышем [15] более полувека назад исходя из соображений об объеме фазового пространства, приходящегося на конечные состояния в элементарном акте ударной ионизации (рис. 1, *a*). Существующие оценки C на основе правила сумм [16] дают значения, существенно превышающие результаты численных расчетов на основе 30-зонной **kp**-модели [17]. В некоторых источниках (см. работу [18], с. 511) рекомендуется подгонять префактор в (1) для согласования с экспериментом при фиксированной степени $n = 2$. Аналитически коэффициент C перед квадратичной зависимостью вида (1) был

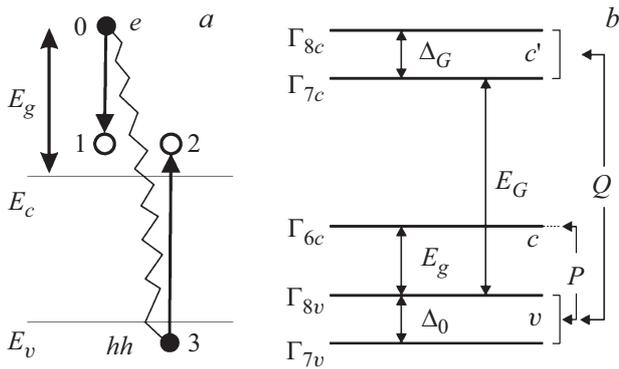


Рис. 1. *a* — схема элементарного процесса межзонной ударной ионизации; *b* — схема 14-зонной *кp*-модели.

рассчитан только для случая узкозонных кубических полупроводников [19].

Сложность описания квадратичного вклада в темп ударной ионизации связана с тем, что в модели изотропных зон, простейшим вариантом которой является 8-зонная *кp*-модель, учитывающая взаимодействие *s*-состояний зоны проводимости и *p*-состояний валентной зоны, префактор *C* равен нулю. Поэтому для вычисления кулоновского матричного элемента, определяющего вид префактора в (1), необходимо учитывать взаимодействие с далекими зонами [19–21], и величина квадратичного вклада оказывается малой для случая узкозонных полупроводников. Это подтверждается результатами численных расчетов [22], которые указывают на то, что главный вклад в темп ударной ионизации узкозонных полупроводников является кубическим ($n = 3$), а не квадратичным. Аналитическое выражение для кубического вклада было впервые получено Гельмонтом с соавт. [23]. Для прямозонных полупроводников со средней величиной запрещенной зоны типа GaSb, In_{0.53}Ga_{0.47}As, InP, In_{0.52}Al_{0.48}As, GaAs и CdTe обоснованный аналитический вид $W(E)$ отсутствует и при Монте-Карло моделировании процессов ударной ионизации выражение (1) используется с произвольными степенями *n*, например $n = 2.5$ и $n = 4.3$ в [24], $n = 5.2$ [3], $n = 3$ [4,6,25], $n = 3.9$ [26], $n = 1.85$ [27]. В некоторых теоретических работах [28,29] рассматривалась оптимизация численного расчета зонной структуры полупроводников для реалистичного моделирования процессов ударной ионизации в устройствах, однако интеграция вычислений зонной структуры с моделированием Монте-Карло слишком сложна для практического использования. На данный момент большинство программ моделирования работы устройств в лучшем случае использует (1) со свободно подгоняемыми параметрами *n* и *C*, что приводит к неконтролируемым результатам.

Данная статья посвящена изучению обоснованного вида $W(E)$ в прямозонных кубических полупроводниках с малой и средней величиной запрещенной зоны в практически значимых условиях, когда эффективная

температура неравновесного распределения электронов составляет порядка нескольких десятков мэВ. Получены явные аналитические выражения для коэффициентов при квадратичном и кубическом вкладах. Приведена оценка температуры кроссовера T^* , при которой темпы генерации носителей, связанные с обоими вкладами, становятся равны в условиях модельного изотропного классического распределения неравновесных электронов. Полученные результаты дают качественное объяснение и количественный критерий доминирования кубического вклада при комнатной температуре в узкозонных полупроводниках, тогда как при средней величине запрещенной зоны оба вклада сравнимы.

2. Квантово-механическая теория темпа межзонной ударной ионизации

2.1. Общее выражение для темпа ударной ионизации

Обычно электроны проводимости рассматриваются как квазичастицы, не взаимодействующие с электронами валентной зоны. В рамках такого подхода обратный к ударной ионизации процесс оже-рекомбинации (*снс*-типа) можно представить как результат взаимодействия двух электронов проводимости, при котором один из них переходит в высоколежащее состояние в зоне проводимости, а другой — в незаполненное состояние в валентной зоне. Однако, как можно показать в приближении Хартри-Фока, процесс межзонной ударной ионизации определяется кулоновским взаимодействием между горячим электроном проводимости и всеми электронами валентной зоны. Поэтому, чтобы получить полный темп ударной ионизации для конкретного состояния горячего электрона „0^c“ (см. схему на рис. 1, *a*), парциальные темпы элементарных процессов, описываемые

$$W = \frac{2\pi}{\hbar} \left\langle \alpha_1 \alpha_2 \left| \frac{e^2}{|\mathbf{r}_1 - \mathbf{r}_2|} \right| \alpha_0 \alpha_3 \right\rangle^2 \delta(\Delta E), \quad (2)$$

необходимо просуммировать по возможным начальным состояниям „3^c“ электрона в валентной зоне и конечным состояниям „1^c“ и „2^c“ электронов в зоне проводимости. Здесь $\Delta E = E_1 + E_2 - E_0 - E_3$ отражает баланс энергии, а $\alpha_i = \{\mathbf{k}_i, \xi_i\}$ обозначает совокупность квантовых чисел — волнового вектора \mathbf{k}_i и проекции полного момента ξ_i для состояний в зоне проводимости ($i = 0, 1$ или 2) и в валентной зоне ($i = 3$). Мы рассматриваем только процессы с участием состояний тяжелых дырок, поскольку состояния легких и спин-отщепленных дырок лежат намного ниже по энергии для волновых векторов, превосходящих пороговый (для полупроводников с малой и средней величиной запрещенной зоны), так что соответствующие процессы ударной ионизации связаны с очень горячими и немногочисленными электронами.

Параметры зонной структуры (взяты из работ [31–35]) полупроводников с малой и средней величиной запрещенной зоны в рамках используемой 14-зонной \mathbf{kp} -модели и соответствующие величины безразмерных параметров β (17) и $x = \Delta_0/E_g$, а также эффективной температуры кроссовера (28)

	E_g , эВ**	Δ_0 , эВ*	P , эВ·Å	E_G , эВ	Q , эВ·Å	β^{**}	x^{**}	T^* , К***
InSb	0.235	0.81	9.63	3.18	7.83	7.53	4.31	2.8
InAs	0.418	0.38	9.01	4.48	7.72	1.13	0.89	10.2
GaSb	0.81	0.76	9.69	3.11	8.25	0.41	0.93	140
In _{0.53} Ga _{0.47} As	0.817	0.324	9.81	4.51	8.25	0.22	0.33	75
InP	1.423	0.107	8.45	4.78	7.88	0.01	0.03	464
In _{0.52} Al _{0.48} As	1.545	0.295	9.09	4.51	8.25	0.04	0.15	516
GaAs	1.519	0.341	9.88	4.54	8.68	0.06	0.17	440
CdTe	1.61	0.95	9.5	5.4	7.87	0.22	0.56	313

Примечание. * — при $T = 0$ К; ** — при $T = 300$ К; *** — вычислены с учетом температурной зависимости ширины запрещенной зоны в соответствии с данными из работ [32,33,36].

Используя фурье-представление кулоновского потенциала и производя интегрирование блоховских амплитуд при нулевом \mathbf{k} , выражение (2) может быть переписано [16,18] в виде

$$W = \frac{2\pi}{\hbar} \left(\frac{4\pi e^2}{\kappa} \right)^2 \frac{I_{cc}(\alpha_0\alpha_1)I_{cv}(\alpha_2, \alpha_3)}{|\mathbf{k}_0 - \mathbf{k}_1|^4} \delta_{\Delta\mathbf{k},0} \delta(\Delta E), \quad (3)$$

где $\Delta\mathbf{k} = \mathbf{k}_0 + \mathbf{k}_3 - \mathbf{k}_1 - \mathbf{k}_2$ отражает баланс импульса, а квадраты интегралов перекрытия блоховских функций I_{cc} и I_{cv} могут быть записаны с помощью векторов состояний $|F\rangle$ в базе блоховских функций точки Γ первой зоны Бриллюэна $u_n^{(0)}(\mathbf{r})$ как $I_{c,c/v}(\alpha_i, \alpha_j) = |\langle F_{\alpha_i} | F_{\alpha_j} \rangle|^2$.

Законы сохранения энергии и импульса, фигурирующие в (3) в виде дельта-функций, накладывают ограничения на возможные значения волновых векторов \mathbf{k}_i . Это приводит к пороговым условиям для ударной ионизации, которые с учетом непараболичности дисперсии начального электрона при $\mu = m_e/m_{hh} \ll 1$ имеют вид

$$E_{th} = E_e(\mathbf{k}_0^{th}) = E_g(1 + 2\mu), \quad (4)$$

$$\mathbf{k}_0^{th} = \mathbf{k}_g(1 + F_0(\Delta_0/E_g)\mu), \quad (5)$$

$$\mathbf{k}_3^{th} = -\mathbf{k}_0^{th}(1 - 2\mu), \quad (6)$$

$$\mathbf{k}_1^{th} = \mathbf{k}_2^{th} = \mu\mathbf{k}_g, \quad (7)$$

где

$$F_0(x) = \frac{3/2}{F_1(x)F_2(x)},$$

$$k_g = \frac{2}{\hbar} \sqrt{F_1(\Delta_0/E_g)m_e E_g}$$

— волновой вектор электронов с энергией $E_0(\mathbf{k}_g) = E_g$ и

$$F_1(x) = \frac{(1 + 2x/3)(1 + x/2)}{(1 + x)(1 + x/3)}, \quad (8)$$

$$F_2(x) = \frac{(1 + x)^2(1 + \frac{x}{3})^3}{(1 + \frac{7}{9}x + \frac{x^2}{6})(1 + \frac{2}{3}x)^2(1 + \frac{x}{2})}. \quad (9)$$

Значения функций $F_1(\Delta_0/E_g)$ и $F_2(\Delta_0/E_g)$ равны единице для обоих пределов $\Delta_0 \ll E_g$ и $\Delta_0 \gg E_g$. Поскольку \mathbf{kp} -взаимодействие между зоной проводимости Γ_{6c} (обозначена „с“ на рис. 1, *b*) и валентными зонами Γ_{8v} и Γ_{7v} (или „v“ на рис. 1, *b*) не дает вклада в дисперсию тяжелых дырок, малость μ эквивалентна $E_g/E_G \ll 1$, где E_G обозначает минимальное расстояние между зоной v и зонами, дающими вклад в обратную массу тяжелой дырки (вторая зона проводимости c' в рамках 14-зонной \mathbf{kp} -модели, используемой в данной работе (см. рис. 1, *b* и таблицу). Величины спин-орбитальных расщеплений зон c' и v также малы по сравнению с расстоянием $c' - v$, $\Delta_{0,G}/E_G \ll 1$ [30] (см. таблицу).

На практике функция распределения горячих электронов, способных инициировать ударную ионизацию, простирается на гораздо меньшем масштабе (например, 25 мэВ), чем пороговая энергия E_{th} . Поэтому для описания ударной ионизации удобно ввести „надпороговые“ компоненты волновых векторов $\mathbf{k}_i = \mathbf{k}_i - \mathbf{k}_i^{th}$ и рассматривать лишь область $(E - E_{th})/E_{th} \ll 1$. При таких предположениях темп ударной ионизации, вызванной электроном в состоянии α_0 , представим в виде

$$W = \frac{\pi\hbar F_2(\frac{\Delta_0}{E_g})}{12m_e E_g^2} \left(\frac{4\pi e^2}{\kappa} \right)^2 \int \frac{d^3 q_1 d^3 q_2}{(2\pi)^6} [\tilde{I}_{cv}(\mathbf{q}_1, \mathbf{q}_3) + \tilde{I}_{cv}(\mathbf{q}_2, \mathbf{q}_3)] \delta\left(q_1^2 + q_2^2 - \frac{2m_e(E_0 - E_{th})}{\hbar^2}\right), \quad (10)$$

где

$$\tilde{I}_{cv}(\mathbf{q}_i, \mathbf{q}_3) = \sum_{\xi_i, \xi_3} I_{cv}(\mathbf{k}_i^{th} + \mathbf{q}_i, \xi_i; \mathbf{k}_3^{th} + \mathbf{q}_1 + \mathbf{q}_2 - \mathbf{q}_0, \xi_3) \quad (11)$$

обозначает межзонный интеграл перекрытия, просуммированный по проекциям полного момента (на направление \mathbf{k}_3) состояний тяжелой дырки $\xi_3 = \pm \frac{3}{2}$ и конечных электронов $\xi_i = \pm \frac{1}{2}$. Предполагается, что надпороговая компонента волнового вектора начального электрона $|\mathbf{q}_0| = \left(\frac{\partial E_0}{\partial \mathbf{k}_0}\right)_{th}^{-1} (E_0 - E_{th})$ коллинеарна \mathbf{k}_0^{th} . Выражение (10) показывает, что энергетическая (и угловая)

зависимость темпа ударной ионизации W определяется поведением квадрата интеграла перекрытия I_{cv} вблизи порога $\mathbf{q}_{1,2} = 0$. Поскольку I_{cv} выражает степень подмешивания состояний зоны проводимости и валентной зоны друг к другу, конкретный вид интеграла перекрытия существенно зависит от используемой модели зонной структуры полупроводника.

2.2. Кубический вклад в темп ударной ионизации

Минимальный базис такой модели состоит из восьми блоховских функций $u_n^{(0)}(\mathbf{r})$ в Γ -точке зоны Бриллюэна: двух s -типа и шести p -типа. Взаимодействие, которое необходимо учесть, — это прямое \mathbf{kp} -взаимодействие между s - и p -состояниями, описываемое единственным „кейновским“ матричным элементом P [37]. Данная 8-зонная модель хорошо описывает дисперсию электронов и легких дырок в узкозонных полупроводниках, однако тяжелые дырки остаются бездисперсными. Вблизи порога $k_i \ll k_3$ для конечных состояний ($i = 1, 2$) и явное выражение для квадрата матричного элемента в рамках 8-зонной модели принимает вид

$$I_{cv}(\alpha_i, \alpha_3) = \frac{P^2 |[\mathbf{k}_i \times \mathbf{k}_3]|^2}{2E_g^2 k_3^2} \delta_{|\xi_i - \xi_3|, 1}. \quad (12)$$

Матричный элемент (12) обращается в нуль для коллинеарных волновых векторов, поэтому с учетом пороговых условий (6) и (7) \tilde{I}_{cv} может быть представлен в виде

$$\tilde{I}_{cv}(\mathbf{q}_i, \mathbf{q}_3) = \frac{P^2 q_{i\perp}^2}{E_g^2} = \frac{\hbar^2 q_{i\perp}^2}{2m_e} \frac{1 + \frac{\Delta_0}{E_g}}{E_g + \frac{2}{3} \Delta_0}, \quad (13)$$

где $\mathbf{q}_{i\perp}$ обозначает компоненту $\mathbf{q}_{i,2}$ в плоскости, перпендикулярной волновому вектору начального электрона. С учетом явного вида \tilde{I}_{cv} (13) темп ударной ионизации W (10) преобразуется к кубическому по $E - E_{th}$ вкладу [19]:

$$W_3(E) = B(E - E_{th})^3, \quad (14)$$

$$B = \frac{\omega_B^*}{18E_g^3} \frac{E_g + \Delta_0}{E_g + \frac{2}{3} \Delta_0} F_2\left(\frac{\Delta_0}{E_g}\right), \quad (15)$$

где $\omega_B^* = \frac{m_e e^4}{2\hbar^3 \kappa^2}$ обозначает боровскую частоту электронов проводимости. В предельном случае бесконечной величины спин-орбитального расщепления Δ_0 (соответствующем 6-зонной \mathbf{kp} -модели) данный ответ сводится к результату (2) из работы [23], тогда как учет конечности приводит к дополнительному множителю, равному $2/3$ при $\Delta_0 \rightarrow 0$.

2.3. Квадратичный вклад в темп ударной ионизации

Таким образом, квадратичный вклад в темп ударной ионизации, связанный с величиной межзонного

интеграла перекрытия на пороге, остается за рамками минимальной модели зонной структуры. Для описания квадратичного вклада необходимо использовать более сложные модели, учитывающие взаимодействие с далекими зонами и понижение сферической симметрии до кубической (группы O_h/T_d), в частности 14-зонную \mathbf{kp} -модель (расширенную модель Кейна [30]). В этой модели помимо \mathbf{kp} -взаимодействия состояний валентной зоны и зоны проводимости явным образом (в первом порядке по k) учитывается взаимодействие состояний валентной зоны с шестью дополнительными блоховскими состояниями симметрии Γ_{7c} и Γ_{8c} , лежащими на несколько эВ выше E_c (вторая зона проводимости, c' на рис. 1, b). Сила $c'-v$ взаимодействия описывается матричным элементом Q , имеющим значение порядка P (см. таблицу). Отсутствие центра инверсии в полупроводнике допускает взаимодействие между зонами c и c' , описываемое матричными элементами P' и Δ' , величина которых на порядок меньше P , Q и Δ_0 соответственно [38].

Для того чтобы учесть дополнительное взаимодействие между c - и v -зонами по теории возмущений, удобно разделить полный \mathbf{kp} -гамильтониан на главную часть $H_0(\mathbf{k})$, представляющую собой гамильтониан минимальной 8-зонной модели и энергии c' состояний при $k = 0$, и возмущение $V(\mathbf{k})$, описывающее $c'-v$ взаимодействие. Шесть собственных состояний H_0 при $\mathbf{k} = \mathbf{k}_3^{th} \simeq -\mathbf{k}_g$, соответствующих c' -зоне, лежат далеко по энергии от остальных восьми состояний электронов с энергией $E_e = E_v + 2E_g$, тяжелых дырок при $E_{hh}^{(0)} = E_v$, а также состояний легких и спин-отщепленных дырок с энергиями

$$E_{lh/so} = E_v - \frac{E_g}{2} \left(1 + x \pm \sqrt{\frac{x^3 + x^2 - x + 3}{x + 3}} \right), \quad (16)$$

где $x = \Delta_0/E_g$. Выражение (16) указывает, что минимальное расстояние по энергии между тяжелыми дырками и другими ветвями валентной зоны превышает $\min(E_g, \Delta_0)/2$ для $x \geq 1$, что соответствует случаю узкозонных полупроводников (см. таблицу). Следовательно, невозмущенное состояние тяжелой дырки является невырожденным в этом случае и может быть применен соответствующий тип теории возмущений. Конкретный метод вычисления многозонных функций Блоха, которому мы следуем, описан в *Приложении I*. Однако истинный параметр, который не должен быть малым, чтобы метод теории возмущений, изложенный в *Приложении I*, был справедлив, это

$$\beta = \frac{\Delta_0 E_g}{6Q^2 k_g^2} = \frac{\Delta_0 P^2 E_g}{12Q^2 E_g^2} \frac{E_g + \Delta_0/3}{E_g + \Delta_0/2}. \quad (17)$$

Соотношения (П.1.1)–(П.1.3) показывают, что \mathbf{kp} -члены возмущения $V(-\mathbf{k}_g)$ воздействуют на невозмущенное состояние тяжелой дырки $|F_{hh}^{(0)}\rangle$ дважды, что приводит к возникновению множителя $\propto (Qk_g)^2 \propto E_g^2$ и двух энергетических знаменателей, ассоциированных с функциями

ми Грина. Первый знаменатель определяется расстоянием между зонами c' и v , равным E_G , а второй — расстоянием между состояниями тяжелых дырок и электронов либо легких дырок или спин-отщепленных дырок. Когда Δ_0 становится много меньше E_g , спин-отщепленные дырки ведут себя подобно тяжелым и энергетический зазор между ними при конечном волновом векторе $\mathbf{k} = \mathbf{k}_3^{\text{th}} \simeq -\mathbf{k}_g$, равный $E_v - E_{so}$, согласно (16), стремится к значению при $k = 0$, равному Δ_0 . Из этого можно сделать вывод, что описанный метод теории возмущений не работает при $\beta \ll 1$ и его применение может приводить к расходимости при $\Delta_0 \rightarrow 0$. Однако последнего не происходит из-за того, что в этом пределе невозмущенное состояние спин-отщепленных дырок не перекрывается с s -состояниями, поскольку они трансформируются во вторую ветвь тяжелых дырок. Поэтому при рассмотрении полупроводников со средней шириной запрещенной зоны, когда $x = \Delta_0/E_g \ll 1$ (или же $\beta \ll 1$) становится малым параметром (см. таблицу) в дополнение к $m_e/m_{hh} \ll 1$, теория возмущений должна учитывать вырождение состояний тяжелых дырок. В этом случае $|F_{hh}^{(0)}\rangle$ в Приложении I будет иметь смысл правильной волновой функции нулевого приближения, которая соответствует верхней ветви тяжелых дырок, расщепленных \mathbf{kr} -взаимодействием между c - и v -зонами.

В результате возведения интеграла перекрытия в квадрат и суммирования по ξ явное выражение для главного вклада в \tilde{I}_{cv} , определяющего квадратичное слагаемое в темпе ударной ионизации, принимает вид

$$\tilde{I}_{cv}(0, 0) = \frac{8E_g^2}{E_G^2} \frac{Q^4}{P^4} K(\mathbf{u}, \beta) \frac{1+x/2}{1+x/3}, \quad (18)$$

где $K(\mathbf{u}, \beta)$ обозначает кубический инвариант, который может быть записан в терминах параметра β и инвариантных многочленов 4-го и 6-го порядков

$$I(\mathbf{u}) = u_x^2 u_y^2 + u_x^2 u_z^2 + u_y^2 u_z^2, \quad (19)$$

$$J(\mathbf{u}) = u_x^2 u_y^2 u_z^2. \quad (20)$$

Здесь $\mathbf{u} = \mathbf{k}_0/k_0$ характеризует направление распространения начального электрона по отношению к кристаллографическим осям. Явный вид кубического инварианта $K(\mathbf{u}, \beta)$ приведен в Приложении II. Для больших ($\beta \rightarrow \infty$) и малых ($\beta \rightarrow 0$) значений β анизотропия квадратичного вклада описывается

$$K_\infty(\mathbf{u}) = I(1-3J) \quad (21)$$

и

$$K_0(\mathbf{u}) = K_\infty(\mathbf{u}) - I^2 + 3J + \frac{I^2(1-4I) - J(2-9I)}{\sqrt{I^2 - 3J}}, \quad (22)$$

соответственно. Ранее подобный (18) (для предела $\beta \gg 1$) результат был получен в работе [20] в терминах параметров Латтинжера γ_2 и γ_3 и в работе [19] в

рамках 14-зонной \mathbf{kr} -модели при $\Delta_0 \rightarrow \infty$ и использован для анализа оже-рекомбинации и ударной ионизации соответственно. Подставляя $\tilde{I}_{cv}(\mathbf{q}_i, \mathbf{q}_3)$ вида (18) в (10), можно получить аналитическое выражение для квадратичного вклада в темп ударной ионизации:

$$W_2(E, \mathbf{u}) = A(E - E_{\text{th}})^2, \quad (23)$$

$$A = \frac{3}{4} \frac{\omega_B^*}{E_G^2} \frac{Q^4}{P^4} K(\mathbf{u}, \beta) \frac{E_g + \frac{1}{2}\Delta_0}{E_g + \frac{1}{3}\Delta_0} F_2\left(\frac{\Delta_0}{E_g}\right). \quad (24)$$

С учетом наличия E_G^2 в знаменателе квадратичный вклад (24) имеет 2-й порядок малости по параметру $\mu = m_e/m_{hh}$, что приводит, как показано далее, к его конкуренции с кубическим вкладом (15) в полупроводниках с малой и средней величиной запрещенной зоны при характерном превышении энергии горячих электронов над порогом $E - E_{\text{th}}$ порядка нескольких десятков мэВ. Темп ионизации, описываемый W_2 , сильно зависит от ориентации направления движения горячего электрона относительно кристаллографических осей. В обоих случаях сильного $\beta \gg 1$ и слабого $\beta \ll 1$ спин-орбитального расщепления валентной зоны квадратичный вклад исчезает в высокосимметричных направлениях [100] и [111]. Однако анизотропия W_2 , описываемая $K_\infty(\mathbf{u})$ и $K_0(\mathbf{u})$, различается: в последнем случае квадратичный вклад дополнительно исчезает в направлении [110]. В плоскостях (111) квадратичный вклад становится изотропным при сильном спин-орбитальном взаимодействии, поскольку $I(\mathbf{u}_{(111)}) = 1/4$, тогда как $K_0(\mathbf{u}_{(111)})$ воспроизводит нетривиальную угловую зависимость $J(\mathbf{u})$. В полупроводниках группы T_d отсутствие центра инверсии и спин-орбитальное взаимодействие приводят к дополнительному вкладу в W_2 , не исчезающему в главных кристаллографических направлениях. Однако такой вклад мал по параметру $\Delta_{c'v}/\Delta_0$, где $\Delta_{c'v}$ — величина недиагонального спин-орбитального $c'-v$ -взаимодействия [30], поэтому с практической точки зрения этот эффект несуществен.

3. Обсуждение результатов

Для того чтобы сравнить значимость двух вкладов в полный темп ударной ионизации

$$W_{\text{tot}}(E, \mathbf{u}) = A(\mathbf{u})(E - E_{\text{th}})^2 + B(E - E_{\text{th}})^3, \quad (25)$$

мы рассматриваем ансамбль невырожденных электронов, выведенных из равновесия приложением электрического поля, и вычисляем темпы генерации носителей R_2 и R_3 [соответствующие темпам ударной ионизации (24) и (15)], усредненные по направлениям электрического поля. Поскольку такое усреднение эквивалентно усреднению по направлениям начального электрона \mathbf{u} при предполагаемой изотропности распределения, темп генерации носителей может быть записан как

$$R_i = \bar{W}_i N_0, \quad (26)$$

$$\bar{W}_i = \int_{E_{th}}^{+\infty} \frac{dE}{T} \frac{d\mathbf{u}}{4\pi} W(E, \mathbf{u}) \exp\left(-\frac{E - E_{th}}{T}\right), \quad (27)$$

где $N_0 = D(E_{th})\bar{\delta}f(E_{th})T$ — неравновесная концентрация горячих электронов над порогом ударной ионизации, \bar{W}_i — темп ударной ионизации, усредненный по направлениям распространения начальных электронов и их распределению, $D(E)$ — плотность состояний в зоне проводимости, $\delta f(E)$ — неравновесная часть функции распределения, T — эффективная температура распределения, которая определяется энергией электрона, набираемой на длине свободного пробега eEl или ее комбинацией с энергией оптического фонона $\hbar\omega_{opt}$ [16]. После интегрирования $(E - E_{th})^n$ с больцмановским распределением мы приходим к следующему выражению для температуры кроссовера, при которой $\bar{W}_2 = \bar{W}_3$ (или $R_2 = R_3$):

$$T^* = \frac{\bar{A}}{3B} = 8 \frac{Q^4}{P^4} \frac{E_g^3}{E_G^2} \bar{K}(\beta) F_1\left(\frac{\Delta_0}{E_g}\right). \quad (28)$$

Значение усредненного (по направлениям \mathbf{u}) кубического инварианта при произвольном β лежит между двумя пределами при бесконечной и нулевой спин-орбите $\bar{K}_\infty < \bar{K}(\beta) < \bar{K}_0$, при этом $\bar{K}_\infty = 2/35$, а $\bar{K}_0 = 0.069$. Поведение $\bar{K}(\beta)$, рассчитанное на основе (П.П.1)–(П.П.6), представлено на рис. 2.

Температурная зависимость ширины запрещенной зоны $E_g(T)$ приводит к нелинейной зависимости отношения \bar{W}_3/\bar{W}_2 от эффективной температуры T , а (28) становится трансцендентным уравнением. Используя параметры зонной структуры из таблицы и эмпирические температурные зависимости ширин запрещенной зоны (в частности, для CdTe из работы [32] (аппроксимация Manoogian–Wooley) и из работ [33,36] (аппроксимация Varshni) для других соединений), были рассчитаны зависимости $\bar{W}_3(T)/\bar{W}_2(T)$ для полупроводников с малой и средней запрещенной зоной (рис. 3), а также температуры кроссовера (см. таблицу). При низких эффек-

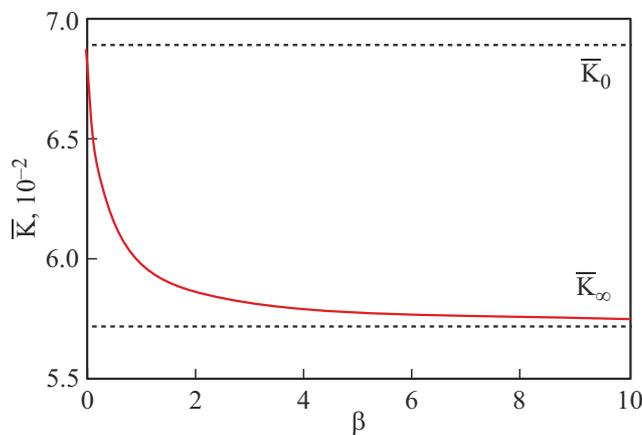


Рис. 2. Среднее значение кубического инварианта (П.П.1) при произвольных параметрах β (17).

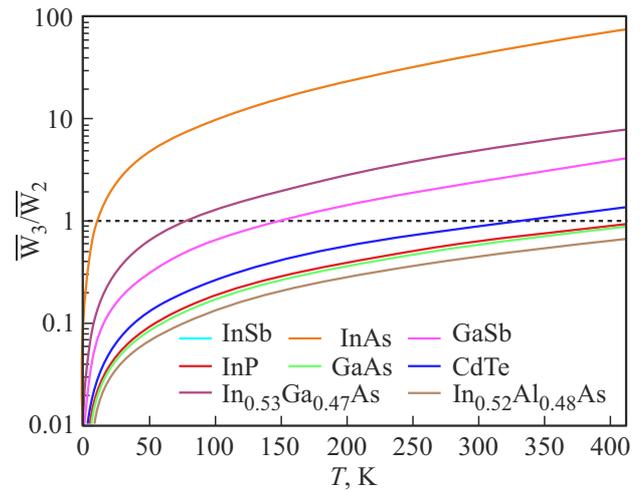


Рис. 3. Конкуренция между усредненными кубическим и квадратичным вкладами в темп ударной ионизации для полупроводников, представленных в таблице, при различных эффективных температурах распределения горячих электронов над порогом. Точки пересечения сплошных кривых с пунктирной линией соответствуют эффективным температурам кроссовера. (Цветной вариант рисунка представлен в электронной версии статьи).

тивных температурах темп ударной ионизации любого полупроводника описывается квадратичным вкладом и является сильно анизотропным. С увеличением T величина изотропного кубического вклада быстро растет и при комнатной температуре он доминирует над $W_2(E, \mathbf{u})$ в узкозонных полупроводниках типа InSb, InAs, GaSb и $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$, а при средней величине запрещенной зоны, например в InP, GaAs и CdTe, оба вклада сравнимы. В $\text{In}_{0.52}\text{Al}_{0.48}\text{As}$ кроссовер наступает при значительно большей температуре порядка 500 К.

4. Заключение

В заключение кратко обсудим точность полученных аналитических выражений. Поскольку при описании ударной ионизации в прямозонных полупроводниках главная трудность, а следовательно, и неточность заключается в получении корректного выражения для интегралов перекрытия начальных и конечных состояний, которые определяются конкретным видом многозонных волновых функций, были проведены расчеты усредненного темпа ударной ионизации на основе численной диагонализации 14-зонной $\mathbf{k}\mathbf{p}$ -модели. Для снижения сложности расчетов в (10) мы опустили несущественную зависимость вычисляемых интегралов перекрытия от \mathbf{q}_3 , использовали приближение главного порядка по $\mu \ll 1$ для пороговых волновых векторов (5)–(7) и разложение $\tilde{I}_{cv}(\mathbf{q}_{1,2}, \mathbf{q}_3)$ по $q_{1,2}/k_g \ll 1$, а также осуществили аналитическое интегрирование полного темпа по величинам $\mathbf{q}_{1,2}$ и q_0 . В результате 9-кратное инте-

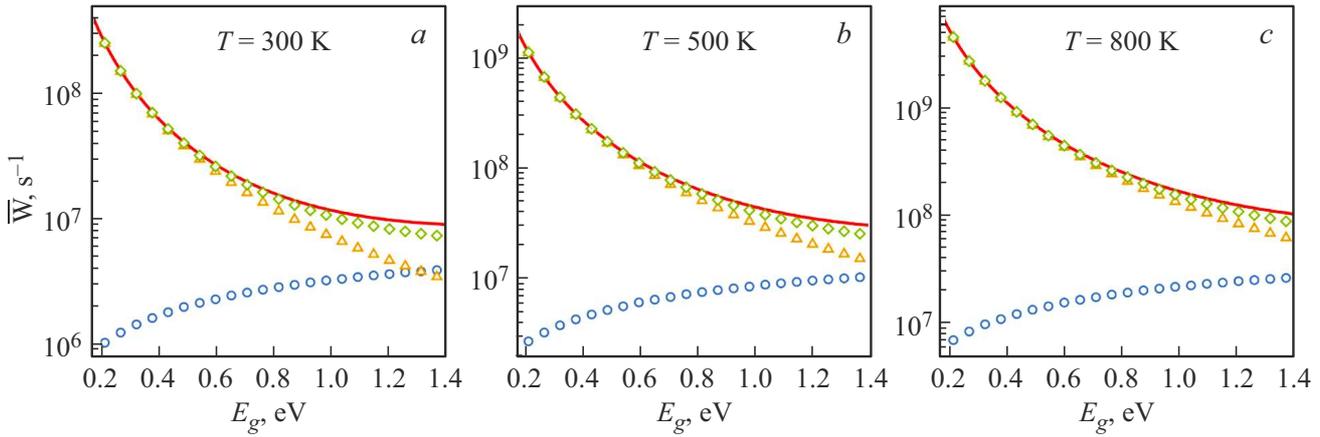


Рис. 4. Зависимость усредненного темпа ударной ионизации от ширины запрещенной зоны при различной эффективной температуре носителей: 300 (a), 500 (b) и 800 К (c). Сплошные красные кривые соответствуют полным темпам $\bar{W}_{tot} = \bar{W}_2 + \bar{W}_3$, вычисленным на основе аналитических выражений (15) и (24); маркеры отвечают численно рассчитанному полному темпу (зеленые ромбы), парциальным квадратичному (голубые кружки) и кубическому (оранжевые треугольники) вкладам.

гирование по $q_{0,1,2}$ было сведено к усреднению темпа по направлениям волнового вектора горячих электронов. Эволюция получившихся значений усредненного темпа ударной ионизации от E_g (остальные параметры зонной структуры соответствуют InAs) при эффективных температурах $T = 300, 500, 800$ К представлена на рис. 4. Простое аналитическое выражение $\bar{W}_{tot} = 2\bar{A}T^2 + 6BT^3$ для темпа ударной ионизации, полученное из (25) и (27), хорошо согласуется с численными результатами в широком диапазоне E_g вплоть до 1.4 эВ. В частности, соответствующая величина среднего отклонения составляет 7%, а максимальное расхождение ограничено 18% для всех температур. Величины аналитической и численной „ширин запрещенной зоны кроссовера“ [когда $\bar{W}_2(E_g) = \bar{W}_3(E_g)$] также близки.

Расхождение между аналитическими и численными результатами связано с выражением (24), которое недооценивает квадратичный вклад в особенности для широкозонных полупроводников, когда главный параметр малости в нашей теории $\mu = m_e/m_{hh}$ приближается к единице. Тем не менее 5%-ое согласование с численными результатами для всего диапазона рассмотренных ширин запрещенной зоны может быть достигнуто за счет включения в аналитические выражения поправок высших порядков (вплоть до 2-го) по E_g/E_G к (15) и (24). Можно также ожидать, что в случае сильной анизотропии распределения горячих электронов в сильном электрическом поле [39], угловая зависимость скорости генерации носителей будет следовать анизотропии суммарного темпа ударной ионизации (25). Таким образом, полученное простое аналитическое обобщение (25) привычной формулы Келдыша для темпа ударной ионизации в прямозонных полупроводниках может быть использовано при моделировании работы полупроводниковых устройств.

Конфликт интересов

Авторы заявляют, что у них нет конфликта интересов.

Приложение I

Метод теории возмущений

для многозонной функции тяжелой дырки

Для того чтобы рассчитать многозонную волновую функцию состояния тяжелой дырки по теории возмущений, удобно ввести „невозмущающую“ функцию Грина $G_0(E) = (E - H_0)^{-1}$. Тогда поправка 1-го порядка к невозмущенному состоянию тяжелой дырки $|F_{hh}^{(0)}\rangle$ может быть представлена в виде

$$|F_{hh}^{(1)}\rangle = \lim_{E \rightarrow \tilde{E}_v} G_0(E, -\mathbf{k}_g) V(-\mathbf{k}_g) |F_{hh}^{(0)}\rangle. \quad (\text{П.1.1})$$

В соответствии с определением V и G_0 в разд. 2 $|F_{hh}^{(1)}\rangle$ принадлежит к c' подпространству базисных функций и ортогональна восьми состояниям, относящимся к зонам c и v . Поэтому приближение 1-го порядка не дает вклад в интеграл перекрытия $\tilde{I}_{cv}(0, 0)$, как и поправка 1-го порядка к энергии $E_{hh}^{(1)} = 0$. Поправка 2-го порядка к энергии тяжелой дырки E_{hh} имеет вид

$$E_{hh}^{(2)} = \lim_{E \rightarrow \tilde{E}_v} \langle F_{hh}^{(0)} | V(-\mathbf{k}_g) | F_{hh}^{(1)}(E) \rangle \quad (\text{П.1.2})$$

и соответствующая поправка к многозонной волновой функции $|F_{hh}\rangle$ может быть переписана как

$$|F_{hh}^{(2)}\rangle = \lim_{E \rightarrow \tilde{E}_v} G_0(E, -\mathbf{k}_g) \times \left[V(-\mathbf{k}_g) |F_{hh}^{(1)}(E)\rangle - E_2(E) |F_{hh}^{(0)}\rangle \right]. \quad (\text{П.1.3})$$

Выражение (П.1.2) определяет энергию тяжелой дырки и соотношение между ее массой и параметрами 14-зонной

модели Q и E_G , а (П.1.3) приводит к главному приближению для межзонного ($c-v$) интеграла перекрытия,

$$\langle F_e(\mathbf{k}_i^{\text{th}}, \xi_i) | F_{hh}(\mathbf{k}_3^{\text{th}}, \xi_3) \rangle \simeq \langle F_e^{(0)}(\mu\mathbf{k}_g, \xi_i) | F_{hh}^{(2)}(-\mathbf{k}_g, \xi_3) \rangle, \quad (\text{П.1.4})$$

где $|F_e^{(0)}\rangle$ обозначает чистую s функцию, описывающую низкоэнергетичные состояния конечных электронов в однозонном приближении, волновой вектор которых мал по сравнению с волновыми векторами начальных состояний (0 и 3). Поэтому поправки к $|F_e^{(0)}\rangle$ не фигурируют в (П.1.4) в главном порядке по $\mu = m_2/m_{hh}$.

Приложение II Кубический инвариант, описывающий анизотропию квадратичного вклада

Выражение для кубического инварианта, описывающего анизотропию квадратичного вклада в темп ударной ионизации при произвольном β , имеет вид

$$K(\mathbf{u}, \beta) = K_1(\mathbf{u}, \beta) + K_2(\mathbf{u}, \beta), \quad (\text{П.1.1})$$

$$K_1(\mathbf{u}, \beta) = \frac{K_1(K_\infty - K_1) + \beta K_2 + \beta^2(K_\infty + K_1)}{\sqrt{K_1 + \beta^2}(\beta + \sqrt{K_1 + \beta^2})}, \quad (\text{П.1.2})$$

$$K_2(\mathbf{u}, \beta) = \frac{K_2 + \beta(K_\infty - K_1)}{\beta + \sqrt{K_1 + \beta^2}}, \quad (\text{П.1.3})$$

$$K_\infty(\mathbf{u}) = I - 3I^2, \quad (\text{П.1.4})$$

$$K_1(\mathbf{u}) = I^2 - 3I, \quad (\text{П.1.5})$$

$$K_2(\mathbf{u}) = -4I^3 + I^2 + 9IJ - 2J, \quad (\text{П.1.6})$$

где $I(\mathbf{u})$ и $J(\mathbf{u})$ обозначают инвариантные многочлены 4-го и 6-го порядков $I(\mathbf{u})$ соответственно (разд. 2.3).

Список литературы

- [1] С. Зи. *Физика полупроводниковых приборов* (М., Мир, 1984).
- [2] K. Gopalakrishnan, P.B. Griffin, J.D. Plummer. *IEEE Trans. Electron Dev.*, **52**, 69 (2005).
- [3] S. Trumm, M. Betz, F. Sotier, A. Leitenstorfer, A. Schwanhäuser, M. Eckardt, O. Schmidt, S. Malzer, G.H. Döhler, M. Hanson, D. Driscoll, A.C. Gossard. *Appl. Phys. Lett.*, **88**, 132113 (2006).
- [4] S. Chen, G. Wang. *J. Appl. Phys.*, **103**, 023703 (2008).
- [5] F. Bertazzi, M. Moresco, E. Bellotti. *J. Appl. Phys.*, **106**, 063718 (2009).
- [6] C.K. Chia. *Appl. Phys. Lett.*, **97**, 073501 (2010).
- [7] E. Bellotti, F. Bertazzi. *J. Appl. Phys.*, **111**, 103711 (2012).
- [8] S. Shishehchi, F. Bertazzi, E. Bellotti. *J. Appl. Phys.*, **113**, 203709 (2013).
- [9] S. Ašmontas, R. Raguotis, S. Bumelienė. *Semicond. Sci. Technol.*, **28**, 025019 (2013).
- [10] K. Kodama, H. Tokuda, M. Kuzuhara. *J. Appl. Phys.*, **114**, 044509 (2013).
- [11] K. Ghosh, U. Singiseti. *J. Appl. Phys.*, **124**, 085707 (2018).
- [12] S. Ašmontas, S. Bumelienė, J. Gradauskas, R. Raguotis, A. Sužiedėlis. *Semicond. Sci. Technol.*, **34**, 075016 (2019).
- [13] S. Ašmontas, S. Bumelienė, J. Gradauskas, R. Raguotis, A. Sužiedėlis. *Sci. Rep.*, **10**, 10580 (2020).
- [14] M.V. Fischetti, S.E. Laux. *Phys. Rev. B*, **38**, 9721 (1988).
- [15] Л.В. Келдыш. *ЖЭТФ*, **37**, 713 (1959).
- [16] B.K. Ridley. *Quantum Processes in Semiconductors* (Oxford University Press, N.Y., 2013).
- [17] M.G. Burt, S. Brand, C. Smith, R.A. Abram. *J. Phys. C: Solid State Phys.*, **17**, 6385 (1984).
- [18] K.F. Brennan. *The Physics of Semiconductors: With Applications to Optoelectronic Devices* (Cambridge University Press, Cambridge, 1999).
- [19] А.Н. Афанасьев, А.А. Грешнов, Г.Г. Зегря. *Письма ЖЭТФ*, **105**, 586 (2017).
- [20] Б.Л. Гельмонт. *ЖЭТФ*, **75**, 536 (1978).
- [21] R. Redmer, J.R. Madureira, N. Fitzer, S.M. Goodnick, W. Schattke, E. Schöll. *J. Appl. Phys.*, **87**, 781 (2000).
- [22] A.R. Beattie, R.A. Abram, P. Scharoch. *Semicond. Sci. Technol.*, **5**, 738 (1990).
- [23] B. Gelmont, K.-S. Kim, M. Shur. *Phys. Rev. Lett.*, **69**, 1280 (1992).
- [24] K.Y. Choo, D.S. Ong. *J. Appl. Phys.*, **96**, 5649 (2004).
- [25] C.K. Chia, G.K. Dalapati. *IEEE Trans. Electron Dev.*, **60**, 3435 (2013).
- [26] D. Dolgos, A. Schenk, B. Witzigmann. *J. Appl. Phys.*, **111**, 073714 (2012).
- [27] I.C. Sandall, J.S. Ng, S. Xie, P.J. Ker, C.H. Tan. *Opt. Express*, **21**, 8630 (2013).
- [28] P. Scharoch, R.A. Abram. *Semicond. Sci. Technol.*, **3**, 973 (1988).
- [29] S. Brand, R.A. Abram. *J. Phys. C: Solid State Phys.*, **17**, L201 (1984).
- [30] R. Winkler. *Spin-Orbit Coupling Effects in Two-Dimensional Electron and Hole Systems* (Springer Verlag, Berlin–Heidelberg, 2003).
- [31] M. Cardona, N.E. Christensen, G. Fasol. *Phys. Rev. B*, **38**, 1806 (1988).
- [32] G. Fonthal, L. Tirado-Mejía, J. Marín-Hurtado, H. Ariza-Calderón, J. Mendoza-Alvarez. *J. Phys. Chem. Solids*, **61**, 579 (2000).
- [33] I. Vurgaftman, J.R. Meyer, L.R. Ram-Mohan. *J. Appl. Phys.*, **89**, 5815 (2001).
- [34] W.H. Lau, J.T. Olesberg, M.E. Flatté. *Electronic structures and electron spin decoherence in (001)-grown layered zincblende semiconductors* (2004). arXiv:condmat/0406201 [cond-mat.mes-hall]
- [35] J.-M. Jancu, R. Scholz, E.A. de Andrada e Silva, G.C. La Rocca. *Phys. Rev. B*, **72**, 193201 (2005).
- [36] *New Semiconductor Materials Database. Characteristics and Properties*. Ioffe Institute (<http://www.matprop.ru/>)
- [37] E.O. Kane. *J. Phys. Chem. Solids*, **1**, 249 (1957).
- [38] S. Richard, F. Aniel, G. Fishman. *Phys. Rev. B*, **70**, 235204 (2004).
- [39] A.P. Dmitriev, M.P. Mikhailova, I.N. Yassievich. *Phys. Status Solidi B*, **113**, 125 (1982).

Редактор Г.А. Оганесян

Competition between isotropic and strongly anisotropic terms in the impact ionization rate of narrow- and middle-gap cubic semiconductors

A.N. Afanasiev, A.A. Greshnov, G.G. Zegrya

loffe Institute,
194021 St. Petersburg, Russia

Abstract We report on the strong anisotropy of the inter-band process of impact ionization in direct-gap cubic semiconductors with either weak or strong spin-orbit coupling at low effective temperatures of electron distribution T , and the crossover to isotropic behavior with increasing T . Such anisotropy is related to specific mechanism of the impact ionization involving coupling of the electron and heavy hole states *via* remote bands, which is vanishing for some high-symmetry propagation directions of an initial electron, namely [100] and [111]. At room temperature impact ionization rate in narrow-gap semiconductors InSb, InAs, GaSb and $\text{In}_{0.53}\text{Ga}_{0.47}\text{As}$ is isotropic while in middle-gap InP, GaAs and CdTe both terms are comparable. We propose simple and justified analytic generalization of Keldysh formula for the impact ionization rate, which is suitable for incorporation into modelling software.