

ГЕТЕРОПЕРЕХОД, ВОЗНИКАЮЩИЙ НА ГРАНИЦЕ СКАЧКООБРАЗНОГО ИЗМЕНЕНИЯ КОНЦЕНТРАЦИИ СВОБОДНЫХ НОСИТЕЛЕЙ В ОДНОРОДНОМ ПО СОСТАВУ ПОЛУПРОВОДНИКЕ

Бычковский Д. Н., Константинов О. В., Царенков Б. В.

Показано, что сужение запрещенной зоны полупроводника при легировании происходит из-за двух главных причин:

— понижения энергии основного носителя заряда за счет кулоновского взаимодействия его с газом остальных основных носителей и примесных ионов;

— понижения энергии неосновного носителя за счет кулоновского взаимодействия его с газом основных носителей (до сих пор учитывалась только эта причина). Показано, что понижение энергии основного носителя точно такое же, как и понижение энергии неосновного. Найденное нами кулоновское понижение энергии основного (или неосновного) носителя в точности совпадает с известным выражением Дебая—Хюккеля, полученным для заряда в однородной плазме электролита.

В нашем рассмотрении в отличие от теории Дебая—Хюккеля берется за основу структура со скачкообразным изменением концентрации легирующей примеси. Заряд, находящийся в слабо легированной области, втягивается в сильно легированную область за счет притяжения к собственному изображению. Находится работа, которую при этом производит заряд; она оказывается равной вышеупомянутой энергии кулоновского взаимодействия в однородно легированном полупроводнике.

Таким образом, при скачкообразном легировании полупроводниковой структуры там возникает плазменный гетеропереход с одинаковыми по величине разрывами краев разрешенных зон.

Полученное нами значение сужения запрещенной зоны при легировании (или инжекции плазмы) отличается от известного в литературе значения.

Введение. Известны несколько моделей сужения запрещенной зоны полупроводника при легировании его донорной или акцепторной примесью, дающей мелкие примесные уровни.

Наиболее ранняя модель [1] заключается в том, что с ростом концентрации легирующей примеси происходят расщепление примесных уровней в примесную зону и слияние последней с дном зоны проводимости (для донорной примеси) или с потолком валентной зоны (для акцепторной). Эффект слияния примесной зоны с основной разрешенной зоной можно истолковывать как сужение запрещенной зоны. Эта модель в настоящее время перестала использоваться в связи с тем, что она дает слишком малые значения величины сужения запрещенной зоны [2], по порядку величины не сильно отличающиеся от энергии связи донора (или акцептора).

Другая модель сужения запрещенной зоны при легировании принадлежит Бонч-Бруевичу [3], Келдышу [4] и Кейну [5]. В ней сужение обуславливается флуктуациями электростатического потенциала вследствие нерегулярности распределения понижованных атомов примеси. Возникают «хвосты» плотности состояний зоны проводимости и валентной зоны, простирающиеся в запрещенную зону. Поскольку края зоны проводимости и валентной зоны идут при этом в пространстве параллельно друг другу, сужение запрещенной зоны можно получить, если учесть туннелирование электронов или дырок через потенциальные барьеры. Плотность состояний в хвосте оказывается при этом существенно отличающейся от той, которая имеется в идеальном кристалле. В рассматриваем

мой модели флуктуационное сужение запрещенной зоны ΔE_g при нулевой температуре ($T=0$) будет

$$\Delta E_g = (e^2/\epsilon L_D) f, \quad (1)$$

где e — заряд электрона, ϵ — диэлектрическая проницаемость полупроводника, L_D — дебаевская длина экранирования при $T=0$:

$$L_D = (\epsilon E_F / 6\pi N e^2)^{1/2}, \quad (2)$$

где E_F — энергия Ферми, N — концентрация основных носителей заряда. Безразмерный фактор f , входящий в (1), дается следующим соотношением:

$$f = (\pi N L_D^3)^{1/2}. \quad (3)$$

Он растет с ростом концентрации примесей, оставаясь при этом, однако, много меньше единицы: например, в случае кремния $f=0.04$ при $N=10^{17}$ см⁻³ и $f=0.13$ при $N=10^{19}$ см⁻³. Экспериментальные результаты свидетельствуют о том, что реальное сужение запрещенной зоны при легировании оказывается порядка $e^2/\epsilon L_D$, что существенно превосходит величину, даваемую (1). Таким образом, в настоящее время флуктуационный механизм уже более не рассматривается в литературе, связанной с вопросом о сужении запрещенной зоны при легировании. В современных монографиях [6, 7] причиной сужения считается кулоновское взаимодействие электрона с дыркой. Подобный механизм обсуждался Рогачевым [8], впервые наблюдавшим эффект сужения запрещенной зоны при инжекции электронно-дырочной плазмы. Обсуждавшийся им вклад кулоновского взаимодействия носит также название корреляционной кулоновской энергии.

Согласно модели, предложенной в [9], механизм сужения заключается в кулоновском взаимодействии неосновного носителя с газом основных носителей. Кулоновское притяжение будет уменьшать энергию активации образования электронно-дырочной пары или, иными словами, запрещенную зону полупроводника, согласно вычислениям [9], на величину

$$\Delta E_g = 3/4 e^2 / \epsilon L_D, \quad (4)$$

где L_D — дебаевская длина. Для вырожденного газа она дается формулой (2), а для невырожденного

$$L_D = (\epsilon k T / 4\pi N e^2)^{1/2}. \quad (5)$$

Цель нашей работы — рассмотрение изменения запрещенной зоны полупроводника при легировании, которое возникает за счет кулоновского взаимодействия основного носителя заряда с газом остальных основных носителей и примесных ионов, а также пересмотрение того вклада в сужение запрещенной зоны, который обусловлен взаимодействием неосновного носителя с газом основных носителей. Сужение запрещенной зоны при легировании важно учитывать в следующих прикладных задачах: 1) при оптимальном выборе степени легирования эмиттера и базы в биполярном транзисторе, силовом тиристоре, светодиоде и лазере для получения максимально возможного коэффициента инжекции эмиттера; 2) для точного задания энергии кванта, излучаемого в электролюминесцентном приборе.

1. Основные результаты

В настоящей работе показано, что взаимодействие основного носителя с газом остальных основных носителей и примесных ионов (например, электрона в полупроводнике n -типа) приводит к понижению дна зоны проводимости на величину

$$\Delta E_c = 1/2 e^2 / \epsilon L_D. \quad (6)$$

Кроме того, взаимодействие неосновного носителя с основными носителями (дырками в полупроводнике n -типа) приводит к повышению потолка валентной зоны на величину

$$\Delta E_v = \frac{1}{2} e^2 / \epsilon L_D. \quad (7)$$

Таким образом, полная величина сужения запрещенной зоны

$$\Delta E_g = \Delta E_c + \Delta E_v = e^2 / \epsilon L_D. \quad (8)$$

Понижение дна зоны проводимости (6) в литературе по полупроводникам вообще не упоминается. Что же касается повышения потолка валентной зоны (7), то, во-первых, оно отождествлялось с полным изменением ширины запрещенной зоны, а, во-вторых, при его расчете [9] была допущена ошибка, и в результате полная величина сужения принималась равной тому, что дается формулой (4),

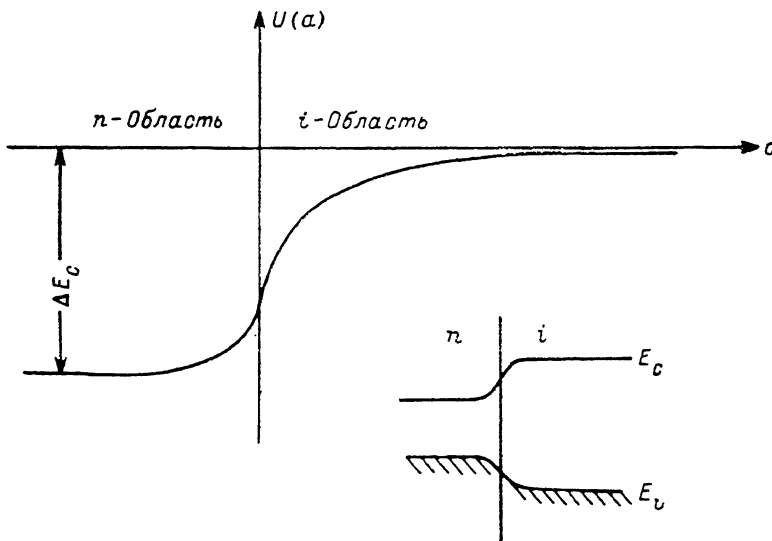


Рис. 1. Зависимость потенциальной энергии $U(a)$ от расстояния a между зарядом и поверхностью легированного полупроводника (n -типа).

На вставке показано изменение краев зоны проводимости и валентной зоны вблизи границы легированной и чистой областей полупроводника. Гетеропереход имеет одинаковые разрывы в этих зонах.

а не (8), как это следует из нашей работы. Важно подчеркнуть также, что принципиальное отличие наших результатов от известных состоит в том, что при легировании полупроводника (например, n -типа) происходят как повышение потолка валентной зоны, так и равное ему понижение дна зоны проводимости, тогда как ранее фактически предполагалось, что повышается только край валентной зоны. Равенство сдвигов краев зон — это простой физический факт, означающий, что энергия заряда, внесенного в плазму, понижается на одну и ту же величину независимо от знака заряда. Этот факт никак нельзя считать ранее не известным. Понижение энергии, описываемое формулами (6) и (7), было получено еще в 1923 г. Дебаем и Хюккелем в их классической работе, посвященной физике электролитов. Несложная теория Дебая и Хюккеля воспроизводится в книге Ландау и Лифшица [10], где показывается, что энергия заряда, внесенного в плазму, понижается на величину $e^2/2\epsilon L_D$, что и отражено в формулах (6), (7). В [10] упоминается о том, что этот результат соответствует первому члену разложения энергии плазмы в ряд по степеням взаимодействия между зарядами, и указано на то, что старшие члены разложения могут быть исследованы с помощью методики, развитой Веденовым и Ларкиным [11]. Таким образом, задача о понижении энергии заряда в плазме имеет давнюю историю и достаточно детально проработана. Возникает вопрос, почему же эти классические результаты до сих пор не были использованы в физике полупроводников? Нам представляется, что это связано с отсутствием наглядной физической интерпретации в методе расчета Дебая—Хюккеля. Поэтому одна из задач нашей работы

состояла в том, чтобы дать методику расчета, обладающую таким свойством. Согласно обсуждениям Веденова и Ларкина [11], понижение энергии Дебая—Хюккеля в термных теории многих тел есть кулоновская корреляционная энергия. Мы рассматриваем такую энергию как работу сил электрического изображения, которую совершает заряд при переносе его из нелегированного полупроводника в легированный. На рис. 1 показана потенциальная энергия $U(a)$ взаимодействия заряда с изображением как функция расстояния a от заряда до границы раздела. Когда заряд находится в слабо легированной области ($a > 0$), при больших a энергия $U(a)$ дается хорошо известным выражением

$$U(a) = -e^2/4\epsilon a. \quad (9)$$

По мере приближения заряда к границе раздела, когда расстояние в нелегированном материале до поверхности становится сравнимым с дебаевской дли-

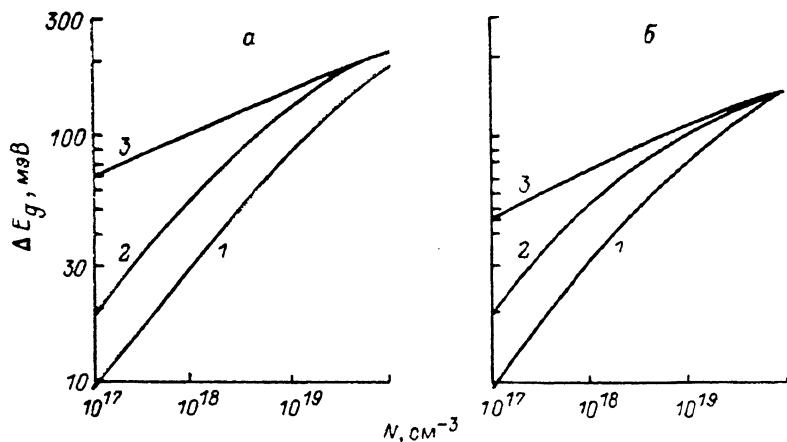


Рис. 2. Зависимость ширины запрещенной зоны от концентрации свободных носителей в кремнии n - (а) и p -типа (б).

$T, K: 1 - 300, 2 - 77, 3 - 0.$

ной в легированном материале, энергия взаимодействия перестает быть чисто кулоновской. На самой поверхности ($a=0$) энергия взаимодействия конечна:

$$U(0) = -e^2/3\epsilon L_D \quad (10)$$

и определяется дебаевской длиной L_D в легированной области. После того как заряд пересек границу легированной области и вошел в нее ($a < 0$), он продолжает втягиваться в эту область на расстояниях порядка L_D . Когда расстояние становится больше L_D , потенциальная кривая выходит на плоскую часть:

$$U(a) = U(0) - \frac{e^2}{6\epsilon L_D} + \frac{e^2}{4\epsilon |a|} e^{-|a|/L_D}. \quad (11)$$

При $|a| \gg L_D$ последнее экспоненциальное слагаемое в (11) исчезает, и перепад ΔU потенциальной энергии внутри легированной области будет

$$\Delta U = U(0) - \frac{e^2}{6\epsilon L_D}. \quad (12)$$

В итоге полная кулоновская энергия втягивания заряда в проводник ΔE_c дается соотношением

$$\Delta E_c = -[U(0) + \Delta U] = + \frac{e^2}{2\epsilon L_D}. \quad (13)$$

Это и есть кулоновская энергия (7) Дебая—Хюккеля. Симметрия кулоновской энергии для электрона и дырки вполне очевидна: любой из этих зарядов наводит в легированной области изображение противоположного знака при одинаковой конфигурации зарядов. Такова наглядная картина образования плазменного

гетероперехода на границе легированной и нелегированной областей полупроводника.

Физическая картина кулоновского понижения энергии заряда в однородной плазме в принципе не отличается от рассмотренной выше модели втягивания заряда изображением. В однородной плазме заряд отталкивает от себя одноимен-

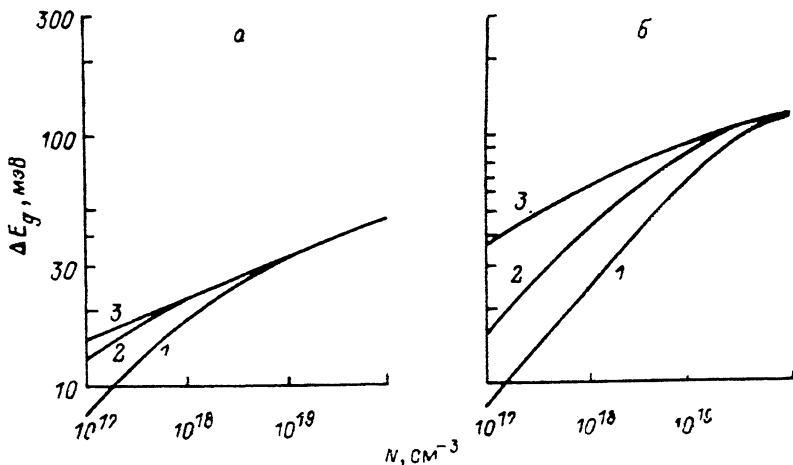


Рис. 3. Зависимость ширины запрещенной зоны от концентрации свободных носителей в арсениде галлия *n*- (а) и *p*-типа (б).

$T, K: 1 - 300, 2 - 77, 3 - 0.$

ные подвижные заряды и взаимодействует с «оголившимися» при этом неподвижными зарядами. Если же подвижные заряды разноименны с рассматриваемым, то заряд создает вокруг себя сгусток этих зарядов, происходит такое же понижение энергии, как и в первом случае. Однако от этих соображений труднее перебростить мостик к расчету, чем для модели втягивания.

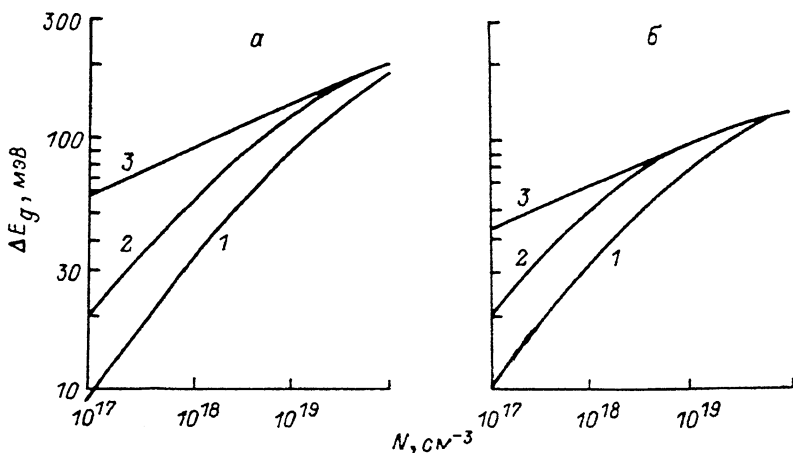


Рис. 4. Зависимость ширины запрещенной зоны от концентрации свободных носителей в фосфиде галлия *n*- (а) и *p*-типа (б).

$T, K: 1 - 300, 2 - 77, 3 - 0.$

Сужение запрещенной зоны определяется лишь подвижными зарядами в легированной области, концентрация которых входит в дебаевскую длину L_D . При произвольной степени вырождения носителей в сильно легированной области длина L_D определяется следующим выражением:

$$L_D = \left[\frac{\varepsilon k T}{4\pi N e^2} \frac{\mathcal{F}_{1/2}(\eta)}{\mathcal{F}_{-1}(\eta)} \right]^{1/2}. \quad (14)$$

Здесь

$$\eta = \frac{E_F}{kT}, \quad N = N_{c,v} \mathcal{F}_{1/2}(\eta), \quad N_{c,v} = 2 \left(\frac{m_{c,v} kT}{2\pi\hbar^2} \right)^{3/2}, \quad (15)$$

$N_{c,v}$ — плотности состояний, m_c и m_v — эффективные массы плотности состояний. Функции $\mathcal{F}_{1/2}$ и $\mathcal{F}_{-1/2}$ — интегралы Ферми [12], определенные тем, что в отсутствие вырождения ($\eta < 0$, $|\eta| \gg 1$)

$$\mathcal{F}_{1/2}(\eta) = \mathcal{F}_{-1/2}(\eta) \approx \text{кр } \eta.$$

С помощью формул (8) и (14) можно получить величину сужения запрещенной зоны при различных концентрациях подвижных носителей и температурах для различных полупроводников. На рис. 2 построено семейство зависимостей

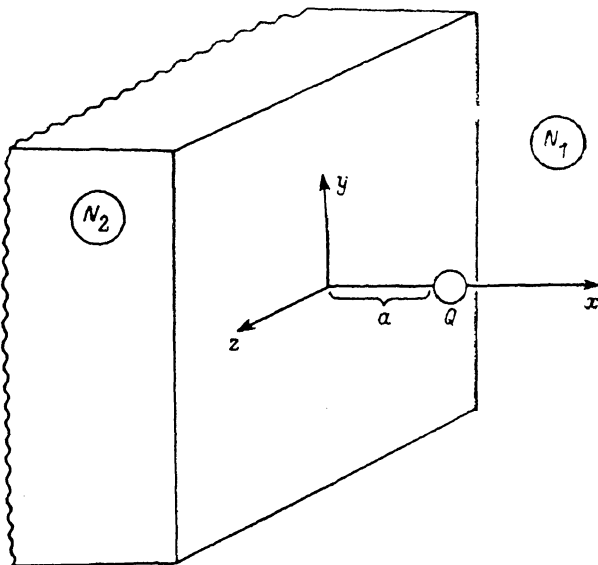


Рис. 5. Расположение осей координат на плоскости раздела двух областей полупроводника с различной концентрацией свободных носителей (слева — N_2 , справа — N_1).

$\Delta E_g(N)$ при температурах 300, 77 и 0 К (кривые 1, 2 и 3) для кремния n - и p -типа (рис. 2, а, б). С понижением температуры дебаевская длина L_D уменьшается, а энергия ΔE_g возрастает. Однако это изменение не беспредельно, поскольку происходит вырождение газа носителей. При этом кривая ΔE_g оказывается тем выше, чем больше эффект ая масса плотности состояний. У электронов кремния эффективная масса больше, чем у дырок ($m_c = 1.09m_0$, $m_v = 0.55m_0$), из-за шести долин в зоне проводимости. В случае арсенида галлия (рис. 3, а, б) ситуация обратная ($m_c = 0.07m_0$, $m_v = 0.47m_0$). Случай фосфида галлия (рис. 4, а, б; $m_c = 0.79m_0$, $m_v = 0.4m_0$) более похож на случай кремния; число долин здесь три. Рассматривая случай $T=0$, следует иметь в виду, что минимальная концентрация носителей должна быть больше критической концентрации для перехода Мотта, так как в противном случае произойдет вымораживание носителей на примесные центры.

2. Теоретическая модель

Рассмотрим плоскую границу двух сред с различными концентрациями подвижных носителей N_1 и N_2 ($N_1 \ll N_2$) и соответственно двумя различными постоянными экранирования $\kappa_1 = L_{D1}^{-1}$ и $\kappa_2 = L_{D2}^{-1}$, где L_{D1} и L_{D2} — дебаевские длины экранирования. Пусть заряд Q находится на расстоянии a от границы раздела в среде 1. Расположение осей координат показано на рис. 5. Уравнение Пуассона, описывающее электростатический потенциал $\phi(r)$, создаваемый зарядом Q и его изображением, имеет следующий вид:

$$\nabla^2 \varphi(r) = \begin{cases} \kappa_1^2 \varphi(r) - 4\pi Q \delta(r-a)/\varepsilon, & x > 0, \\ \kappa_2^2 \varphi(r), & x < 0. \end{cases} \quad (16)$$

Совершим над этим уравнением преобразование Фурье по координатам y и z , лежащим в плоскости границы раздела ($x=0$), согласно обычному определению:

$$\varphi_q(x) = \int_{-\infty}^{\infty} \int dy dz \varphi(r) \exp[i(-i(q_y y + q_z z))]. \quad (17)$$

Тогда уравнение (16) примет следующую форму для фурье-образа:

$$\frac{d^2 \varphi_q}{dx^2} = \begin{cases} s_1^2 \varphi_q(x) - 4\pi Q \delta(x-a)/\varepsilon, & x > 0, \\ s_2^2 \varphi_q(x), & x < 0, \end{cases} \quad (18)$$

$$s_1^2 = \kappa_1^2 + q^2, \quad s_2^2 = \kappa_2^2 + q^2, \quad q^2 = q_y^2 + q_z^2. \quad (19)$$

Граничные условия состоят в непрерывности функции $\varphi_q(x)$ и ее производной при $x=0$ и равенстве нулю при $x \rightarrow \pm \infty$. Решение $\varphi_q(x)$ можно искать в виде суммы

$$\varphi_q(x) = \Phi_q(x) + \psi_q(x), \quad (20)$$

где Φ_q — решение однородного уравнения:

$$\Phi_q(x) = \begin{cases} A e^{-s_1 x}, & x > 0, \\ B e^{s_2 x}, & x < 0, \end{cases} \quad (21)$$

а ψ_q — решение неоднородного уравнения, описывающее поле точечного заряда в неограниченной среде 1:

$$\psi_q(x) = 2\pi Q e^{-s_1 |x-a|/\varepsilon s_1}. \quad (22)$$

Из условия «сшивания» решения (20) при $x=0$ находим константы интегрирования A и B :

$$A = -\frac{2\pi Q}{\varepsilon s_1} \frac{s_2 - s_1}{s_2 + s_1} e^{-s_1 a}, \quad B = A + \frac{2\pi Q}{\varepsilon s_1} e^{-s_1 a}. \quad (23)$$

Функция

$$\Phi_q^{(+)}(x) = A e^{-s_1 x} \quad (24)$$

описывает фурье-образ потенциала изображения заряда, наводимого в среде 2 и действующего в среде 1 на заряд Q . С помощью (24) нетрудно найти фурье-образ силы, действующей на заряд Q :

$$F_q^{(+)}(a) = -Q \partial \Phi_q^{(+)} / \partial x |_{x=a}. \quad (25)$$

Подставляя (23) и (24) в (25), получим, что

$$F_q^{(+)}(a) = -\frac{2\pi Q^2}{\varepsilon} \frac{s_2 - s_1}{s_2 + s_1} e^{-2s_1 a}. \quad (26)$$

Совершим над (26) обратное преобразование Фурье:

$$F^{(+)}(a) = \frac{1}{(2\pi)^2} \int \int dq_y dq_z F_q^{(+)}(a) \exp[i(q_y y + q_z z)]. \quad (27)$$

Таким образом, в цилиндрической системе координат

$$F^{(+)}(a) = -\frac{Q^2}{\varepsilon} \int_0^{\infty} q dq \frac{s_2 - s_1}{s_2 + s_1} e^{-2s_1 a}. \quad (28)$$

С помощью выражения для силы можно найти потенциальную энергию заряда Q в поле изображения как функцию расстояния a :

$$U^{(+)}(a) = \int_a^{\infty} F^{(+)}(a) da = -\frac{Q^2}{2\epsilon} \int_0^{\infty} q dq \frac{s_2 - s_1}{s_2 + s_1} e^{-2\epsilon s_1 a}. \quad (29)$$

Совершенно аналогично проводится расчет силы, когда заряд расположен в среде 2, т. е. при $a < 0$:

$$F^{(-)}(a) = -\frac{Q^2}{\epsilon} \int_0^{\infty} q dq \frac{s_2 - s_1}{s_2 + s_1} e^{2\epsilon s_1 a}. \quad (30)$$

Это дает энергию взаимодействия заряда с изображением при $a < 0$

$$U^{(-)}(a) = \int_a^0 F^{(-)}(a) da + U^{(+)}(0). \quad (31)$$

Интегрируя по a , получим

$$U^{(-)}(a) = -\frac{Q^2}{2\epsilon} \int_0^{\infty} \frac{q dq}{s_2} \frac{s_2 - s_1}{s_2 + s_1} (1 - e^{2\epsilon s_1 a}) + U^{(+)}(0). \quad (32)$$

Полный перепад потенциальной энергии будет равен

$$\begin{aligned} \Delta E_e^{(1,2)} &= -U^{(+)}(0) - U^{(-)}(-\infty), \\ \Delta E_e^{(1,2)} &= \frac{Q^2}{2\epsilon} \int_0^{\infty} q dq \left(\frac{1}{s_1} - \frac{1}{s_2} \right) = \frac{Q^2}{2\epsilon} (x_2 - x_1). \end{aligned} \quad (33)$$

Таким образом, высота потенциальной ступени на резком переходе ($x=0$) оказывается равной разности энергий, определяемых формулой Дебая—Хюккеля (7) для плазмы. Поэтому удобно переопределить потенциальную энергию $U(a)$, выбрав для нее в качестве нуля положение дна зоны проводимости в чистом полупроводнике:

$$\tilde{U}(a) = -\frac{Q^2}{2\epsilon} G(a), \quad (34)$$

$$G(a) = \begin{cases} G^{(+)}(a), & a > 0, \\ G^{(-)}(a), & a < 0, \end{cases} \quad (35)$$

где

$$G^{(+)}(a) = x_1 + \int_0^{\infty} \frac{q dq}{s_1} \frac{s_2 - s_1}{s_2 + s_1} e^{-2s_1 a}, \quad (36)$$

$$G^{(-)}(a) = G^{(+)}(0) + \int_0^{\infty} \frac{q dq}{s_2} \frac{s_2 - s_1}{s_2 + s_1} (1 - e^{2s_1 a}). \quad (37)$$

Интегралы $G^{(+)}(0)$ и $G^{(-)}(-\infty)$ вычисляются аналитически:

$$G^{(+)}(0) = x_1 + \frac{1}{3} (x_2 - x_1) \left(2 - \frac{x_1}{x_1 + x_2} \right), \quad G^{(-)}(-\infty) = x_2. \quad (38)$$

Итак, окончательно получаем

$$\tilde{U}(a) |_{a \rightarrow \infty} = -\frac{Q^2}{2\epsilon} x_1, \quad (39)$$

$$\tilde{U}(a) |_{a=0} = -\frac{Q^2}{2\epsilon} \left[x_1 + \frac{1}{3} (x_2 - x_1) \left(2 - \frac{x_1}{x_1 + x_2} \right) \right], \quad (40)$$

$$\tilde{U}(a) |_{a \rightarrow -\infty} = -\frac{Q^2}{2\epsilon} x_2. \quad (41)$$

Если среда 1 — чистый полупроводник, то $x_1=0$, и из формулы (40) получается приведенное ранее выражение (11).

В заключение авторы выражают благодарность В. В. Осипову за полезное обсуждение.

Список литературы

- [1] Stern F., Talley R. M. // Phys. Rev. 1955. V. 100. N 6. P. 1638—1642.
- [2] Kleppinger D. D., Lindholm F. A. // Sol. St. Electron. 1971. V. 14. N 5. P. 407—420.
- [3] Бонч-Бруевич В. Л. // ФТТ. 1962. Т. 4. В. 10. С. 2660—2674.
- [4] Келдыш Л. В., Прошко Г. П. // ФТТ. 1963. Т. 5. В. 12. С. 3378—3389.
- [5] Kane E. O. // Phys. Rev. 1963. V. 131. N 1. P. 79—88.
- [6] Зи С. Физика полупроводниковых приборов. Т. 1. М., 1984. 455 с.
- [7] Блихер А. Физика силовых биполярных и полевых транзисторов. Л., 1986. 356 с.
- [8] Аснин В. М., Рогачев А. А. // ФТТ. 1963. Т. 5. В. 6. С. 1730—1732.
- [9] Lanyon H. P. D., Tuft R. A. // Trans. Electron. Dev. 1976. V. ED-26. N 5. P. 1014—1022.
- [10] Ландау Л. Д., Лифшиц Е. М. Статистическая физика. М., 1976. 583 с.
- [11] Веденов А. А., Ларкин А. И. // ЖЭТФ. 1959. Т. 36. В. 4. С. 1133—1142.
- [12] Блэкмор Дж. Статистика электронов в полупроводниках. М., 1964. 392 с.

Физико-технический институт
им. А. Ф. Иоффе АН СССР
Ленинград

Получена 16.05.1990
Принята к печати 22.06.1990