

## ОБОБЩЕНИЕ МЕТОДА ЭФФЕКТИВНОЙ МАССЫ ДЛЯ ПОЛУПРОВОДНИКОВЫХ СТРУКТУР С АТОМАРНО РЕЗКИМИ ГЕТЕРОПЕРЕХОДАМИ

Э. Е. Тахтамиров\*, В. А. Волков†

Институт радиотехники и электроники Российской академии наук  
103907, Москва, Россия

Поступила в редакцию, 28 июня 1999 г.

Метод огибающих функций Кона—Латтинжера обобщен на случай гетероструктур с атомарно резкими гетеропереходами на основе согласованных по постоянной решетки слоев родственных полупроводников с симметрией цинковой обманки. Для электронных состояний вблизи Г-точки в (001)-гетероструктурах выведено однозонное уравнение эффективной массы, учитывающее как пространственную зависимость последней, так и эффекты, связанные с атомарной резкостью гетеропереходов. Выявлен малый параметр, по степеням которого можно проклассифицировать различные вклады в это уравнение. Для дырочных состояний учтены лишь основные вклады в эффективный гамильтониан, обусловленные резкостью гетероперехода. Выведено выражение для параметра, определяющего смешивание состояний тяжелых и легких дырок в центре 2D-зоны Бриллюэна.

PACS: 73.20.Dx, 71.25.Cx, 73.40.Kp

### 1. ВВЕДЕНИЕ

Метод эффективной массы Латтинжера—Кона [1, 2] широко применяется для описания электронных состояний во внешних полях, плавно изменяющихся на масштабах постоянной решетки  $a$ . Хотя оригинально метод, базирующийся на формализме огибающих функций, применим лишь к однородным полупроводникам, его различные модификации интенсивно используются и для описания электронных состояний в полупроводниковых гетероструктурах. В последние годы возобновилась дискуссия о применимости метода эффективной массы для описания электронных и дырочных состояний в реальных наноструктурах [3–22]. Предложено много различных модификаций метода эффективной массы, актуальных в случае пространственно-зависящей эффективной массы  $m(\mathbf{r})$ . Существуют два способа построения приближения эффективной массы для гетероструктур. 1) Вывод эффективного гамильтониана для огибающих функций, определенного во всем координатном пространстве. Из интегрирования уравнения эффективной массы, в которое входит этот гамильтониан, вблизи гетерограницы можно (в случае необходимости) получить граничные условия для огибающих функций. 2) Вывод или, что бывает гораздо чаще, постулирование феноменологических граничных условий для огибающих функций на гетерогранице. При этом используются соображения симметрии, непрерывность плотности потока вероятности и т. п.

\*E-mail: gibbon@royal.net

†E-mail: VoVA@mail.cplire.ru

(впрочем, этих соображений, как правило, недостаточно для однозначного определения граничных условий). Второй способ применим в случае резких гетерограниц (во всех моделях, в рамках которых проводилось получение таких граничных условий, рассматривались математически резкие гетерограницы между левым и правым материалами). Молчаливо предполагается, что огибающая функция слева (справа) от гетерограницы удовлетворяют тому же уравнению, что и в объемном случае для левого (правого) материалов. При этом возникает весьма тонкая (и не обсужденная до сих пор) проблема превышения точности: граничные условия должны быть справедливы с той же точностью, с которой справедливы уравнения для огибающих функций.

Ниже мы будем следовать первому подходу, в рамках которого удастся корректно рассмотреть проблему точности (см. разд. 3).

Известно, что на пути построения единого уравнения для огибающих функций возникают две основные проблемы. Первая — это проблема упорядочения операторов импульса в операторе кинетической энергии (обусловленная некоммутативностью оператора импульса и функции  $m(\mathbf{r})$ ), от формы которого могут существенно зависеть и получаемые решения уравнения эффективной массы [18]. Вторая проблема заключается в том, что эффективный потенциал вблизи гетерограницы, как правило, не является плавной функцией на масштабах порядка  $a$ . Сомнению подвергается тогда сама корректность использования дифференциальных уравнений в методе огибающих функций. Обсудим эти проблемы в указанном порядке.

## 1.2. Учет пространственной зависимости эффективной массы

Необходимым условием применимости «однозонных» уравнений для огибающих функций (подразумевается одно уравнение, справедливое вблизи дна невырожденной зоны проводимости, или система уравнений для вырожденной валентной зоны), используемых в методе эффективной массы, является «мелкость» рассматриваемых состояний: их энергия, отсчитанная от края зоны, должна быть мала по сравнению с межзонной энергией. Поэтому, имея в виду однозонный вариант метода эффективной массы, мы ограничиваемся рассмотрением гетероструктур, состоящих из родственных материалов, когда разрывы зон малы по сравнению с характерными ширинами запрещенных зон; это, как правило, означает, что слабо отличаются и другие зонные параметры полупроводников. Рассмотрим первую проблему, возникающую уже и для гетероструктур с плавным на масштабах  $a$  изменением химического состава. Выберем за потенциал нулевого приближения потенциал кристаллической решетки, мысленно продолженной на все пространство, одного из материалов структуры (это не единственный выбор, см. [7]), а отличие потенциалов решеток остальных полупроводников от базисного будем считать малым возмущением. Следуя подходу Латтинжера и Кона и получая многозонную  $\mathbf{k}\mathbf{p}$ -систему уравнений (см., например, [3]), можно затем попытаться разрешить проблему о правильном порядке некоммутирующих операторов в операторе кинетической энергии для однозонных уравнений. Но здесь возникает еще одна проблема.

Сведение многозонной системы уравнений к однозонному уравнению эффективной массы достигается исключением из многозонной  $\mathbf{k}\mathbf{p}$ -системы малых огибающих функций в пользу больших посредством некоей процедуры. Сделаем небольшое отступление и воспользуемся формальной аналогией между релятивистским уравнением Дирака и многозонной  $\mathbf{k}\mathbf{p}$ -системой уравнений для огибающих функций [23], которая проще всего прослеживается в двухзонном приближении (зона проводимости и невырожденная валентная). В релятивистской теории есть два подхода к получению уравнения для

мелких состояний электрона. Один из них состоит в исключении малой позитронной компоненты волновой функции методом подстановки. При этом получим либо точное уравнение для электронной компоненты, которое не является уравнением на собственные значения ([24, гл. XX, § 28]), либо приближенное уравнение, за эрмитовостью которого нужно следить отдельно [25]. Другой подход — преобразование типа Фолди—Вотхойзена, приближенное унитарное преобразование уравнения Дирака ([24, гл. XX, § 33]).

В нашем случае первый подход сравнительно просто реализуется только в двухзонном приближении (см., например, [26]). При рассмотрении вклада далеких зон (а это необходимо, в частности, для правильного описания вклада тяжелых дырок) возникает ряд проблем. Так, авторам [14, 20] удалось учесть лишь некоторые из членов, являющихся поправками первого порядка к «стандартному» уравнению Кона—Латтинжера с пространственно-независящей эффективной массой (малым параметром является отношение характерного разрыва зон к характерной межзонной энергии). Однако рассматривая полученное, например, в [14] выражение для пространственно-зависящей эффективной массы, можно показать, что в эффективную массу края зоны проводимости полупроводника, не являющегося базисным, не вошли межзонные матричные элементы потенциала возмущения, полученные с использованием блоховских функций края зоны базисного кристалла (см. далее разд. 4.1). Легко видеть, что это равносильно мало оправданному пренебрежению различием межзонных матричных элементов оператора импульса, или, что эквивалентно, различием блоховских функций для составляющих структуру материалов.

Отсюда следует, что проблеме учета далеких зон следует уделить особое внимание. Попытки же ее решения с помощью прямого исключения малых огибающих функций методом подстановки помимо своей трудоемкости приводят, в конечном итоге, к неэрмитовому уравнению, решения которого требуется еще правильно интерпретировать.

Ниже мы будем следовать второму подходу, т. е. применять унитарное преобразование, исключаяющее малые огибающие функции [1, 27]. Поскольку мы рассматриваем гетероструктуры, состоящие из родственных материалов, стандартный метод эффективной массы будет играть роль первого приближения. Учет пространственной зависимости эффективной массы приведет к необходимости рассмотреть поправки к стандартной теории, причем следует учитывать все поправки одного порядка малости, не допуская превышения точности.

Чтобы понять, какие поправки следует учитывать, обратимся к релятивистской аналогии с гипотетическим уравнением Дирака, содержащим неоднородную щель  $2m(\mathbf{r})c^2$ , где  $c$  есть скорость света в вакууме (см. Приложение 1). Обычное однозонное уравнение эффективной массы является аналогом нерелятивистского уравнения Шредингера. Существенно однако, что эффективная масса в двухзонном приближении пропорциональна локальной ширине запрещенной зоны  $E_g(\mathbf{r})$  (это справедливо, если эффективная масса формируется, в основном,  $kp$ -взаимодействием), а ее относительное изменение  $\delta m/m \simeq \delta E_g/E_g$ . В силу того что поправка к кинетической энергии, описывающая пространственную зависимость эффективной массы, имела бы «релятивистский» характер, искомые уравнения для гетероструктур будут аналогичны уравнению Шредингера со всеми релятивистскими поправками, как обычными (вклад непараболичности закона дисперсии, пропорциональный  $p^4$ , где  $p$  — оператор импульса; вклад спин-орбитального взаимодействия и так называемый член Дарвина, пропорциональный второй производной потенциальной энергии), так и новой псевдорелятивистской

поправкой, описывающей  $\delta m(\mathbf{r})$ . Естественно, настоящие рассуждения справедливы при описании состояний, энергия которых, отсчитанная от края зоны любого из составляющих структуру материалов, порядка разрыва зоны. Случай очень малого разрыва зоны, когда разрыв мал по сравнению с энергией состояний, достаточно тривиален: в зависимости от энергии рассматриваемых состояний учет пространственной зависимости эффективной массы может потребовать учета членов со все более и более высокими степенями оператора импульса. Мы не будем рассматривать такую ситуацию. В этом смысле введение в уравнение эффективной массы члена, пропорционального четвертой степени оператора импульса, является необходимым условием рассмотрения ее пространственной зависимости. Заметим, что для однородных полупроводников уравнение эффективной массы, аналогичное уравнению Шредингера с первыми релятивистскими поправками, было обсуждено еще в [27, § 27].

Типичным недостатком прежних работ, посвященных обобщению метода эффективной массы на электронные состояния в гетероструктурах, является учет в рамках теории возмущений лишь части членов одного порядка малости. Так, в [3–22] учитывалась пространственная зависимость параметров эффективной массы, но пренебрегалось поправками  $\sim p^4$ . В работах [3, 7, 10, 14], посвященных выводу однозонных уравнений для огибающих функций непосредственно из многозонной  $\mathbf{k}\mathbf{p}$ -системы, основной ошибкой, приведшей к неправильному результату, являлась некорректная оценка, согласно которой вклад членов  $\mathbf{k}\mathbf{p}$ -взаимодействия (т. е. членов  $\hbar \mathbf{k} \mathbf{p}_{nn'}$  /  $m_0$ , где  $m_0$  — масса свободного электрона, а  $\mathbf{p}_{nn'}$  — межзонный матричный элемент импульса) порядка вклада членов потенциальной энергии (разницы в периодических потенциалах составляющих структуру полупроводников, рассмотренной по теории возмущений). В случае плавного гетероперехода корректная процедура вывода однозонного уравнения эффективной массы вблизи дна зоны проводимости с учетом всех указанных выше вкладов была проведена в [28].

## 1.2. Учет атомарной резкости гетеропереходов

Вторая проблема, требующая внимательного изучения — неплавность реальных гетеропереходов, когда переход от одного материала к другому происходит на масштабах порядка  $a$ . В этом случае, во-первых, многозонная  $\mathbf{k}\mathbf{p}$ -система Лейблера [3], где плавность потенциала была необходимым условием ее справедливости, нуждается в уточнении, а во-вторых, усложняется проблема перехода в  $\mathbf{g}$ -представление из ограниченного первой зоной Бриллюэна  $\mathbf{k}$ -пространства [1]. Необходимо также проанализировать следствия унитарного преобразования, исключающего далекие зоны. Важно оценить ошибку, допускаемую на каждом шаге. Оценка этой ошибки либо даст нам уверенность в отсутствии превышения точности, либо бросит вызов справедливости приближения эффективной массы. В известных нам работах, посвященных рассмотрению резких гетеропереходов, отсутствует такого рода оценка. В [14, 20], например, лишь указывается на ее малость, а в [19] были сделаны приближения, точность которых не оценивалась.

Таким образом, можно сформулировать следующие этапы построения приближения эффективной массы для гетероструктур: а) получение многозонной  $\mathbf{k}\mathbf{p}$ -системы уравнений для огибающих функций, должным образом учитывающей возможную резкость гетероперехода; б) сведение ее к однозонным уравнениям при помощи унитарного преобразования в  $\mathbf{k}$ -пространстве, переход в  $\mathbf{g}$ -представление и преобразование полученного уравнения к дифференциальной форме; в) оценка точности этих преобразований. Следуя этой схеме, в разд. 2 реализован этап а). Уравнения включают вклады,

связанные с неплавностью гетероперехода на масштабах порядка  $a$ , которые рассматриваются в рамках подхода, подобного использованному в [27] для описания короткодействующей части примесного потенциала. В разд. 3 показано, что единое во всем пространстве дифференциальное уравнение для огибающих функций для резких гетеропереходов существует, а его точность определяется именно процедурой перехода к однозонным уравнениям в  $\mathbf{g}$ -пространстве. Однозонные уравнения рассмотрены в разд. 4.1 (зона проводимости) и в разд. 4.2 (валентная зона). Раздел 4.3 посвящен гетероструктурам со сверхтонкими слоями. Показано, что дополнительные вклады в «стандартное» уравнение эффективной массы можно проклассифицировать по степеням малого параметра  $\bar{k}a_m$ , где  $\bar{k}$  есть характерное значение квазиимпульса состояния, а  $a_m$  порядка постоянной решетки. В разд. 5 построена иерархическая схема уравнений эффективной массы,  $n$ -ый уровень которой отвечает учету этих дополнительных вкладов вплоть до  $(\bar{k}a_m)^n$ . Нулевой уровень иерархии ( $n = 0$ ) соответствует «стандартному» уравнению эффективной массы с пространственно-независящими параметрами. На первом уровне иерархии каждая гетерограница дает дополнительный  $\delta$ -функциональный вклад в потенциальную энергию. И только на втором уровне иерархии проявляется пространственная зависимость эффективной массы, а также поправки, связанные со слабой непараболическостью спектра и гетероповерхностные члены спин-орбитального взаимодействия. На более высоких уровнях иерархии возникают нелокальные вклады, и однозонные дифференциальные уравнения эффективной массы не существуют. Результаты сравниваются с выводами других авторов. Краткие сообщения о полученных результатах были опубликованы в работах [29–31].

## 2. МНОГОЗОННАЯ $\mathbf{k}\mathbf{p}$ -СИСТЕМА УРАВНЕНИЙ ДЛЯ ОГИБАЮЩИХ ФУНКЦИЙ В СЛУЧАЕ РЕЗКОГО (001) ГЕТЕРОПЕРЕХОДА

Рассмотрим гетеропереход, образованный из родственных согласованных по постоянной решетке полупроводников со структурой цинковой обманки. Уравнение Шредингера без релятивистских поправок (которые будут учтены ниже) и в отсутствие внешних потенциалов имеет обычный вид:

$$\left( \frac{\mathbf{p}^2}{2m_0} + U(\mathbf{r}) \right) \Psi(\mathbf{r}) = \epsilon \Psi(\mathbf{r}).$$

Здесь  $U(\mathbf{r}) \equiv U$  — кристаллический потенциал гетероструктуры. Вначале мы будем использовать следующую модель этого потенциала:

$$U = U_1 + G(z)[U_2 - U_1] \equiv U_1 + G(z)\delta U, \quad (1)$$

а  $U_1 \equiv U_1(\mathbf{r})$  и  $U_2 \equiv U_2(\mathbf{r})$  являются периодическими (с одинаковым периодом) продолженными на всю структуру потенциалами левого и правого материалов соответственно, ось  $z$  направлена перпендикулярно плоскости гетероперехода,  $G(z)$  — формфактор гетероперехода:

$$G(z)|_{z < -d} = 0, \quad G(z)|_{z > d} = 1;$$

ширина переходной области гетероперехода есть  $2d$  (неодномерность функции  $G(z)$  будет учтена ниже).

Естественно считать возмущением потенциал  $G(z)\delta U$ . В качестве базиса разложения волновой функции мы используем полный ортонормированный набор функций Кона—Латгинжера  $\{u_{n0}e^{ikr}\}$ :

$$\Psi(\mathbf{r}) = \sum_{n'} \int \mathcal{F}_{n'}(\mathbf{k}') e^{ik'r} u_{n'0} d^3 k', \quad (2)$$

где  $u_{n0} \equiv u_{n0}(\mathbf{r})$  является периодической блоховской амплитудой для края  $\epsilon_{n0}$   $n$ -ой зоны левого кристалла в  $\Gamma$ -точке зоны Бриллюэна (в нерелятивистском пределе),

$$\left( \frac{\mathbf{p}^2}{2m_0} + U_1 \right) u_{n0} = \epsilon_{n0} u_{n0}.$$

Суммирование в (2) проводится по всем зонам, а интегрирование, здесь и далее, если не оговорено обратное, — по зоне Бриллюэна;  $\mathcal{F}_n(\mathbf{k})$  является огибающей функцией для  $n$ -ой зоны в  $\mathbf{k}$ -представлении. Следуя стандартной процедуре [1], можно получить систему  $\mathbf{k}\mathbf{p}$ -уравнений [3]:

$$\begin{aligned} & \left( \epsilon_{n0} + \frac{\hbar^2 \mathbf{k}^2}{2m_0} \right) \mathcal{F}_n(\mathbf{k}) + \sum_{n'} \frac{\hbar \mathbf{p}_{nn'}}{m_0} \mathcal{F}_{n'}(\mathbf{k}) + \\ & + \sum_{n'} \int \mathcal{M}_{nn'}(\mathbf{k}, \mathbf{k}') \mathcal{F}_{n'}(\mathbf{k}') d^3 k' = \epsilon \mathcal{F}_n(\mathbf{k}), \\ & \mathcal{M}_{nn'}(\mathbf{k}, \mathbf{k}') = \sum_j C_j^{nn'} \mathcal{G}(k_z - k'_z + K_{zj}) \delta(\mathbf{k}_{||} - \mathbf{k}'_{||} + \mathbf{K}_{||j}). \end{aligned} \quad (3)$$

Здесь  $\mathbf{p}_{nn'} = \langle n | \mathbf{p} | n' \rangle$  и  $C_j^{nn'} = \langle n | \delta U e^{ik_j r} | n' \rangle$ , а матричные элементы периодического оператора  $\mathbf{f}$  определяются следующим образом:

$$\langle n | \mathbf{f} | n' \rangle = \frac{(2\pi)^3}{\Omega} \int_{\text{cell}} u_{n0}^* \mathbf{f} u_{n'0} d^3 r,$$

где  $\Omega$  — объем элементарной ячейки;  $\mathbf{k}_{||} = (k_x, k_y, 0)$ , а  $K_{zj}$  и  $\mathbf{K}_{||j}$  являются, соответственно, перпендикулярной и параллельной плоскости гетерограницы компонентами вектора  $\mathbf{K}_j$  обратной решетки,  $\mathcal{G}(k_z)$  — фурье-преобразование функции  $G(z)$ . Проанализируем выражение для матричных элементов  $\mathcal{M}_{nn'}(\mathbf{k}, \mathbf{k}')$ :

$$\begin{aligned} \mathcal{M}_{nn'}(\mathbf{k}, \mathbf{k}') &= \sum_{j (\mathbf{K}_{||j}=0)} C_j^{nn'} \mathcal{G}(k_z - k'_z + K_{zj}) \delta(\mathbf{k}_{||} - \mathbf{k}'_{||}) + \\ &+ \sum_{j (\mathbf{K}_{||j} \neq 0)} C_j^{nn'} \mathcal{G}(k_z - k'_z + K_{zj}) \delta(\mathbf{k}_{||} - \mathbf{k}'_{||} + \mathbf{K}_{||j}). \end{aligned} \quad (4)$$

Вторая сумма в (4) описывает процессы переброса в двумерной зоне Бриллюэна, когда проекции любой пары векторов  $\mathbf{k}$  и  $\mathbf{k}'$  из объемной зоны Бриллюэна на плоскость гетероперехода удовлетворяют условию  $\mathbf{k}'_{||} - \mathbf{k}_{||} = \mathbf{K}_{||j} \neq 0$ . Для гетероперехода произвольной ориентации такие процессы переброса существуют. Однако для интересующей нас ориентации (001) их вклад в искомые уравнения для огибающих функций исчезает (см. Приложение 2):

$$\mathcal{M}_{nn'}(\mathbf{k}, \mathbf{k}') = \delta(k_{||} - k'_{||}) \left[ \mathcal{G}(k_z - k'_z) \delta U_{nn'} + \sum_{j \neq 0} C_j^{nn'} \mathcal{G}(k_z - k'_z + K_j) \right], \quad (5)$$

где введены обозначения:  $K_j = (4\pi/a)j$ ;  $j = \pm 1, \pm 2, \dots$ ;  $\delta U_{nn'} = C_0^{nn'}$ . Если  $G(z)$  является достаточно плавной функцией,  $a \ll d$ , и нас интересуют состояния с  $\bar{k}_z \ll 2\pi/a$ , где  $\bar{k}_z$  есть характерное значение квазиимпульса состояния, мы можем пренебречь вторым слагаемым в квадратных скобках выражения (5) и получить известный набор уравнений для огибающих функций [3, 28]. В случае же атомарно резкого гетероперехода можно действовать в духе метода, использовавшегося в [27] для описания короткодействующего примесного потенциала. Введем функцию  $G'(z) \equiv dG(z)/dz$ , локализованную на гетерогранице,  $|z| \leq d$ . Тогда для  $j \neq 0$  имеем

$$\begin{aligned} \mathcal{G}(k_z - k'_z + K_j) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} G(z) \exp(-i(k_z - k'_z + K_j)z) dz = \\ &= \frac{1}{2\pi i} \frac{1}{k_z - k'_z + K_j} \int_{-\infty}^{+\infty} G'(z) \exp(-i(k_z - k'_z + K_j)z) dz = \\ &= \frac{1}{2\pi i K_j} \left( 1 - \frac{k_z - k'_z}{K_j} + \dots \right) \int_{-d}^d G'(z) \exp(-iK_j z) [1 - i(k_z - k'_z)z + \dots] dz, \quad (6) \end{aligned}$$

и сумму в (5) можно записать в виде разложения по степеням  $(k_z - k'_z)$ :

$$\sum_{j \neq 0} C_j^{nn'} \mathcal{G}(k_z - k'_z + K_j) = \sum_{s=0,1,2,\dots} \frac{(k_z - k'_z)^s}{2\pi} D_{snn'}. \quad (7)$$

Константы в разложении (7) имеют вид

$$\begin{aligned} D_{0nn'} &= \sum_{j \neq 0} C_j^{nn'} \frac{1}{iK_j} \int_{-d}^d G'(z) \exp(-iK_j z) dz, \\ D_{1nn'} &= \sum_{j \neq 0} C_j^{nn'} \frac{1}{iK_j} \int_{-d}^d G'(z) \exp(-iK_j z) \left( -\frac{1}{K_j} - iz \right) dz, \dots \end{aligned}$$

Настоящий подход принципиально позволяет рассматривать даже математически резкие гетеропереходы, так как необходимая сходимость коэффициентов  $D_{l_{nn'}}$  обеспечивается свойством:  $C_j^{nn'} \rightarrow 0$  при  $K_j \rightarrow \infty$  (впрочем, физически реализуемый гетеропереход не может быть математически резким).

Рассмотрим поправки, связанные с резкостью гетероперехода. Простые оценки показывают, что члены, пропорциональные  $D_{0nn'}$ ,  $D_{1nn'}$ ,  $\dots$ , могут дать поправки, по порядку величины не большие, чем  $a\bar{k}_z$ ,  $(a\bar{k}_z)^2$ ,  $\dots$ , соответственно. Нашей целью является получение однозонных уравнений с пространственно-зависящими параметрами эффективной массы, что достигается при учете поправок порядка  $(\lambda\bar{k}_z)^2$  к стандартному

приближению. Здесь введена характерная «двухзонная» длина  $\lambda = \hbar(2mE_g)^{-1/2}$ . Для GaAs, например,  $\lambda \approx 6 \text{ \AA}$ . Мы будем пользоваться малостью параметра  $\tilde{k}_z d$ , что позволит нам записать конечное уравнение в достаточно простой форме (см. далее). Таким образом, три величины, имеющие размерность длины, ( $a$ ,  $d$  и  $\lambda$ ) в комбинации с  $\tilde{k}_z$  образуют три параметра, малость которых используется в настоящем методе. На наш взгляд, ситуация, когда  $a \lesssim d \lesssim \lambda$ , представляется наиболее реальной, реализующейся в полупроводниковых гетероструктурах с резкими гетеропереходами. Тогда параметр  $\lambda \tilde{k}_z$  можно считать главным малым параметром задачи и ограничиться поэтому в (7) членами с  $s = 0$  и  $s = 1$ .

В результате многозонная система  $k$ -р-уравнений (3) принимает следующий вид:

$$0 = \left( \epsilon_{n0} - \epsilon + \frac{\hbar^2 \mathbf{k}^2}{2m_0} \right) \mathcal{F}_n(\mathbf{k}) + \sum_{n'} \frac{\hbar \mathbf{p}_{nn'} \cdot \mathbf{k}}{m_0} \mathcal{F}_{n'}(\mathbf{k}) + \sum_{n'} \int \left[ \mathcal{G}(k_z - k'_z) \delta U_{nn'} + \frac{1}{2\pi} D_{0nn'} + \frac{k_z - k'_z}{2\pi} D_{1nn'} \right] \mathcal{F}_{n'}(k'_z, \mathbf{k}_{||}) dk'_z. \quad (8)$$

Здесь мы выделили разные вклады потенциала возмущения: вклад плавной части описывается первым слагаемым в квадратных скобках, а вклад резкой — вторым и третьим слагаемыми.

### 2.1. Учет 3D-характера формфактора

Рассмотрим теперь вместо (1) более реалистичный вид гетеропотенциала:

$$U = U_1 + g(\mathbf{r}_{||}, z) \delta U, \quad (9)$$

где  $\mathbf{r}_{||} = (x, y, 0)$ . По определению,  $g(\mathbf{r}_{||}, z) \big|_{z < -d} = 0$  и  $g(\mathbf{r}_{||}, z) \big|_{z > d} = 1$ , а функция  $g(\mathbf{r}_{||}, z)$  периодична по  $\mathbf{r}_{||}$ . Для нашего случая (001) гетероструктуры элементарные векторы трансляции в плоскости гетеропотенциала есть  $\mathbf{a}_1 = (1, 1, 0)a/2$  и  $\mathbf{a}_2 = (-1, 1, 0)a/2$ . Узлы обратной двумерной решетки с основными векторами  $\mathbf{b}_1 = (1, 1, 0)2\pi/a$  и  $\mathbf{b}_2 = (-1, 1, 0)2\pi/a$  являются проекциями узлов обратной трехмерной решетки на плоскость (001).

Разложим  $g(\mathbf{r}_{||}, z)$  в 2D-ряд Фурье:

$$g(\mathbf{r}_{||}, z) = \sum_l G_l(z) \exp(i\mathbf{K}_l \mathbf{r}_{||}),$$

где индекс суммирования  $l$  определен так, что векторы  $\mathbf{K}_l$  определяют узлы указанной двумерной обратной решетки, и

$$G_l(z) = \frac{1}{\Omega_{||}} \int_{\substack{2D \\ \text{cell}}} g(\mathbf{r}_{||}, z) \exp(-i\mathbf{K}_l \mathbf{r}_{||}) d^2 r_{||},$$

интегрирование проводится по элементарной ячейке 2D-решетки площадью  $\Omega_{||} = a^2/2$ . Обозначая  $G(z) \equiv G_0(z)$ , получим для потенциала возмущения

$$g(\mathbf{r}_{||}, z) \delta U = G(z) \delta U + \delta U \sum_{l \neq 0} G_l(z) \exp(i\mathbf{K}_l \mathbf{r}_{||}). \quad (10)$$

Видно, что простейшая модель (1) учитывает первое слагаемое в (10). Функции  $G_l(z)$  при  $l \neq 0$  отличны от нуля лишь в области  $z \in [-d, d]$ . Поэтому левая часть (3) будет включать дополнительную сумму интерфейсных вкладов:

$$\sum_{n'} \int \mathcal{M}_{nn'}^{\parallel}(\mathbf{k}, \mathbf{k}') \mathcal{F}_{n'}(\mathbf{k}') d^3 k',$$

где

$$\mathcal{M}_{nn'}^{\parallel}(\mathbf{k}, \mathbf{k}') = \sum_{l \neq 0; j} \frac{C_j^{nn'}}{(2\pi)^3} \int_{\text{all space}} G_l(z) \exp(i\mathbf{K}_l \mathbf{r}_{\parallel}) \exp(-i(\mathbf{k} - \mathbf{k}' + \mathbf{K}_j) \mathbf{r}) d^3 r.$$

Для состояний с  $|k_x| + |k_y| < \pi/a$  (см. Приложение 2) мы получим

$$\begin{aligned} \mathcal{M}_{nn'}^{\parallel}(\mathbf{k}, \mathbf{k}') &= \delta(\mathbf{k}_{\parallel} - \mathbf{k}'_{\parallel}) \sum_{l \neq 0; j} \frac{C_j^{nn'}}{2\pi} \delta_{\mathbf{K}_l, \mathbf{K}_{lj}} \int_{-d}^d G_l(z) \exp(-i(k_z - k'_z + K_{zj})z) dz = \\ &= \delta(\mathbf{k}_{\parallel} - \mathbf{k}'_{\parallel}) \sum_{j(\mathbf{K}_{lj} \neq 0)} \frac{C_j^{nn'}}{2\pi \Omega_{\parallel}} \int_{-d}^d dz \int_{\substack{2D \\ \text{cell}}} d^2 r_{\parallel} g(\mathbf{r}_{\parallel}, z) \exp(-i\mathbf{K}_{lj} \mathbf{r}_{\parallel}) \exp(-i(k_z - k'_z + K_{zj})z). \end{aligned}$$

Для плавного гетероперехода в сумме надо оставить только члены с  $K_{zj} = 0$  и обрабатывать выражение стандартным образом [3], так что в  $\mathbf{r}$ -представлении уравнения для огибающих функций будут включать дополнительный интерфейсный потенциал. Заметим, что в то время как плавная часть  $G(z)\delta U$  обеспечивает смешивание состояний одной и той же симметрии кристаллической решетки (т. е. локальная симметрия плавной части потенциала возмущения  $G(z)\delta U$  совпадает с симметрией объемного кристалла), рассматриваемый нами дополнительный потенциал обеспечивает смешивание и различных по симметрии состояний. Нас интересует случай, когда  $\bar{k}_z d \ll 1$ , и это позволяет нам не разделять эффекты этого дополнительного потенциала отдельно на вклады плавной и резкой частей, а использовать разложение

$$\mathcal{M}_{nn'}^{\parallel}(\mathbf{k}, \mathbf{k}') = \delta(\mathbf{k}_{\parallel} - \mathbf{k}'_{\parallel}) \sum_{s=0,1,2,\dots} \frac{(k_z - k'_z)^s}{2\pi} D_{snn'}^{\parallel}, \quad (11)$$

где

$$\begin{aligned} D_{0nn'}^{\parallel} &= \sum_{j(\mathbf{K}_{lj} \neq 0)} \frac{C_j^{nn'}}{\Omega_{\parallel}} \int_{-d}^d dz \int_{\substack{2D \\ \text{cell}}} d^2 r_{\parallel} g(\mathbf{r}_{\parallel}, z) \exp(-i\mathbf{K}_j \mathbf{r}), \\ D_{1nn'}^{\parallel} &= \sum_{j(\mathbf{K}_{lj} \neq 0)} \frac{C_j^{nn'}}{\Omega_{\parallel}} \int_{-d}^d dz \int_{\substack{2D \\ \text{cell}}} d^2 r_{\parallel} g(\mathbf{r}_{\parallel}, z) \exp(-i\mathbf{K}_j \mathbf{r}) (-iz), \dots \end{aligned}$$

В разложении (11) нужно сохранить только первых два члена; слагаемые, пропорциональные  $D_{0nn'}^{\parallel}$  и  $D_{1nn'}^{\parallel}$ , могут дать поправки порядка  $\bar{k}_z d$  и  $(\bar{k}_z d)^2$  соответственно.

Мы показали, что учет трехмерности формфактора, см. (9), не представляет особого труда для анализа, а теперь сделаем одно важное замечание, которое позволит нам пользоваться именно простой моделью (1). Функция  $g(\Gamma_{||}, z)$  имеет более низкую симметрию, чем  $G(z)$ , — она инвариантна относительно преобразований симметрии из точечной группы  $C_{2v}$ . Полный же потенциал возмущения в обеих моделях, (1) и (9), обладает одной и той же симметрией, как точечной ( $C_{2v}$ ), так и трансляционной в плоскости гетероперехода. Информация о симметрии  $C_{2v}$  будет сохранена, однако, только при условии учета вклада резкой части потенциала  $G(z)\delta U$  в модели (1). Поэтому использование (9) вместо (1) не даст ничего качественно нового, а лишь приведет к перенормировке некоторых параметров, а именно, тех, которые являются пренебрежимо малыми для плавного гетероперехода в модели (1). Выражения для этих параметров станут очень громоздкими, поэтому в дальнейшем мы ограничимся моделью (1).

## 2.2. Учет релятивистских вкладов

Рассмотрим теперь релятивистские эффекты. Ограничимся лишь спин-орбитальным взаимодействием. Остальные релятивистские вклады лишь повлияют на значения констант, которые мы получим. Будем считать, что в рамках теории возмущений характерный параметр спин-орбитального взаимодействия, а также разница таковых параметров левого и правого кристаллов меньше или порядка характерного разрыва зон. Разложение полной волновой функции, как и раньше, дается выражением (2). Опуская промежуточные выкладки, приведем полученную довольно громоздкую кр-систему уравнений с учетом спин-орбитального взаимодействия:

$$\begin{aligned}
 & \left( \epsilon_{n0} - \epsilon + \frac{\hbar^2 \mathbf{k}^2}{2m_0} \right) \mathcal{F}_n(\mathbf{k}) + \sum_{n'} \frac{\hbar \mathbf{p}_{nn'} \mathbf{k}}{m_0} \mathcal{F}_{n'}(\mathbf{k}) + \\
 & + \sum_{n'} \delta U_{nn'} \int \mathcal{G}(k_z - k'_z) \mathcal{F}_{n'}(k'_z, \mathbf{k}_{||}) dk'_z + \sum_{n'} \frac{\hbar \langle n | [\nabla U_1 \mathbf{p}] | n' \rangle \sigma}{4m_0^2 c^2} \mathcal{F}_{n'}(\mathbf{k}) + \\
 & + \sum_{n'} \frac{\hbar \langle n | [\nabla \delta U \mathbf{p}] | n' \rangle \sigma}{4m_0^2 c^2} \int \mathcal{G}(k_z - k'_z) \mathcal{F}_{n'}(k'_z, \mathbf{k}_{||}) dk'_z + \\
 & + \sum_{n'} \frac{\hbar \langle n | [\mathbf{n} \delta U \mathbf{p}] | n' \rangle \sigma}{4m_0^2 c^2} \int i(k_z - k'_z) \mathcal{G}(k_z - k'_z) \mathcal{F}_{n'}(k'_z, \mathbf{k}_{||}) dk'_z + \\
 & + \sum_{n'} \int \frac{\hbar^2 \langle n | [\nabla \delta U \mathbf{k}'] | n' \rangle \sigma}{4m_0^2 c^2} \delta(\mathbf{k}_{||} - \mathbf{k}'_{||}) \mathcal{G}(k_z - k'_z) \mathcal{F}_{n'}(\mathbf{k}') d^3 k' + \\
 & + \sum_{n'} \int \left( \frac{1}{2\pi} D_{0nn'} + \frac{k_z - k'_z}{2\pi} D_{1nn'} \right) \mathcal{F}_{n'}(k'_z, \mathbf{k}_{||}) dk'_z + \\
 & + \sum_{n'} \int \left( \frac{1}{2\pi} S_{0nn'} + \frac{k_z - k'_z}{2\pi} S_{1nn'} \right) \sigma \mathcal{F}_{n'}(k'_z, \mathbf{k}_{||}) dk'_z + \\
 & + \sum_{n'} \int \frac{1}{2\pi} \delta(\mathbf{k}_{||} - \mathbf{k}'_{||}) [\hbar \mathbf{k}' \sigma] \mathbf{B}_{0nn'} \mathcal{F}_{n'}(\mathbf{k}') d^3 k' = 0. \tag{12}
 \end{aligned}$$

Векторы  $S_{0nn'}$ ,  $S_{1nn'}$  и  $V_{0nn'}$  имеют следующий вид:

$$S_{0nn'} = \sum_{j \neq 0} \frac{\hbar \langle n | [\nabla (e^{iK_j z} \delta U) \mathbf{p}] | n' \rangle}{4iK_j m_0^2 c^2} \int_{-d}^d G'(z) e^{-iK_j z} dz,$$

$$S_{1nn'} = - \sum_{j \neq 0} \frac{\hbar \langle n | [\nabla (e^{iK_j z} \delta U) \mathbf{p}] | n' \rangle}{4K_j m_0^2 c^2} \int_{-d}^d G'(z) e^{-iK_j z} dz -$$

$$- \sum_{j \neq 0} \frac{\hbar \langle n | e^{iK_j z} [\nabla \delta U \mathbf{p}] | n' \rangle}{4iK_j^2 m_0^2 c^2} \int_{-d}^d G'(z) e^{-iK_j z} dz,$$

$$V_{0nn'} = \sum_{j \neq 0} \frac{\hbar \langle n | \nabla (e^{iK_j z} \delta U) | n' \rangle}{4iK_j m_0^2 c^2} \int_{-d}^d G'(z) e^{-iK_j z} dz.$$

Здесь  $\mathbf{p}$  — единичный вектор вдоль оси  $z$ ,  $\mathbf{p}G'(z) \equiv \nabla G(z)$ , а  $\sigma$  — матрицы Паули. В левой части (12) четвертое слагаемое описывает спин-орбитальное взаимодействие в потенциале базисного полупроводника; пятое, шестое и седьмое слагаемые обусловлены плавной частью потенциала возмущения. Слагаемые, пропорциональные  $S_{0nn'}$ ,  $S_{1nn'}$  и  $V_{0nn'}$ , вызваны резкостью потенциала. В [28], рассматривая состояния зоны проводимости в гетероструктурах с плавными гетеропереходами, мы пренебрегли шестым и седьмым слагаемыми в левой части (12) как малыми. Учитывалось, что во втором порядке теории возмущений они совместно с  $\hbar \mathbf{k} p_{nn'}/m_0$  дадут лишь поправку порядка  $(\lambda \bar{k}_z)^2 m/m_0$ , которая в этом случае мала в меру малости эффективной массы по сравнению с  $m_0$ . Для дырочных же состояний  $m/m_0$  может не являться малым параметром.

Мы не рассматриваем линейные по  $\mathbf{k}$  вклады спин-орбитального взаимодействия, обусловленные потенциалом  $U_1$ . Они дадут поправки порядка  $(\lambda \bar{k}_z)^3$  (в третьем порядке совместно с двумя членами типа  $\hbar \mathbf{k} p_{nn'}/m_0$ ) подобно вкладу, ответственному за снятие спинового вырождения в зоне проводимости объемного полупроводника (такой малости вкладами мы пренебрегаем), а для валентной зоны известно, что в первом порядке вклад этих членов мал, а во втором порядке вместе с  $\hbar \mathbf{k} p_{nn'}/m_0$  они лишь перенормируют значения параметров эффективной массы.

Что касается вкладов от резкости гетероповерхностного потенциала в спин-орбитальное взаимодействие, то члены, пропорциональные  $S_{0nn'}$ , могут дать поправки порядка  $a \bar{k}_z$ , а члены  $S_{1nn'}$  и  $V_{0nn'}$  — порядка  $(a \bar{k}_z)^2$ .

Обобщение на случай многих гетеропереходов проводится тривиально. При этом удобно выбрать координаты гетерограниц так, чтобы между ними укладывалось целое число  $a/2$ , тогда фазовый множитель при каждом из разложений типа (7) будет равен единице.

### 3. ПРОБЛЕМА ПРЕОБРАЗОВАНИЯ УРАВНЕНИЙ ЭФФЕКТИВНОЙ МАССЫ ИЗ ПРОСТРАНСТВА КВАЗИИМПУЛЬСОВ В КООРДИНАТНОЕ ПРОСТРАНСТВО

В следующем разделе мы получим однозонные уравнения эффективной массы для зоны проводимости и валентной зоны. Но прежде всего мы обсудим пробле-

му, возникающую в методе огибающих функций и связанную с ограниченностью  $k$ -пространства. Рассмотрим следующее однозонное уравнение для огибающей функции  $f(k_z)$  в  $k$ -представлении:

$$\int H(k_z, k'_z) f(k'_z) dk'_z = \epsilon f(k_z), \quad (13)$$

причем  $k_z$  и  $k'_z$  ограничены зоной Бриллюэна. При преобразовании (13) в координатное представление мы получим, вообще говоря, интегральное уравнение. Проблема состоит в точности, с которой можно получить дифференциальное уравнение в  $g$ -пространстве. Рассмотрим уравнение, подобное (13), но в котором  $k_z$  и  $k'_z$  принадлежат всему обратному пространству:

$$\int_{-\infty}^{+\infty} H(k_z, k'_z) g(k'_z) dk'_z = \epsilon g(k_z). \quad (14)$$

Фурье-преобразование уравнения (14) с учетом (12) даст дифференциальное уравнение в  $g$ -представлении. Если бы функции  $g(k_z)$  обращались в нуль при  $k_z \notin$  зоне Бриллюэна, они являлись бы также решениями (13), и мы бы решили нашу проблему точно. В общем случае это не так. Но для того чтобы (13) и (14) были приблизительно эквивалентны, необходимо, чтобы  $g(k_z)$  были малы при  $k_z \notin$  зоне Бриллюэна. В теории плавных возмущений такая малость обеспечивается экспоненциально убывающими огибающими функциями в  $k$ -представлении, однако в случае резких возмущений огибающие функции будут лишь степенным образом убывающими функциями  $k_z$ . Так, если огибающая функция обладает одним разрывом, ее фурье-преобразование  $g(k_z) \propto (\delta\bar{g}/\bar{g})(k_z)^{-1}$  при больших  $k_z$  (когда затухнут экспоненциальные вклады, связанные с эффектами плавных полей), здесь  $(\delta\bar{g}/\bar{g})$  есть характерный относительный разрыв функции в  $g$ -представлении. Если рассмотреть, например, стандартные уравнения эффективной массы [1] с разрывными потенциалами, то вторые производные соответствующих огибающих функций будут разрывны с характерным относительным разрывом порядка единицы (опять же для состояний, энергия которых, отсчитанная от края зоны левого или правого материала, порядка разрыва зоны), а ошибка от использования дифференциальных уравнений будет порядка  $(\bar{k}_z/K)^3$ , где  $K$  — радиус зоны Бриллюэна по оси  $k_z$ .

В случае квантовой ямы шириной  $L$  можно рассмотреть два случая:  $\bar{k}_z L \gtrsim 1$  и  $\bar{k}_z L \ll 1$ . В первом случае ошибка такого же порядка, как и для отдельного гетероперехода, а во втором случае она может быть порядка  $(\bar{k}_z L)^{-1}(\bar{k}_z/K)^3$ . Это оценка сверху. Для симметричной квантовой ямы в зоне проводимости, например, ошибка зависит от знака произведения значений огибающей функции на гетерограницах, и для состояний второй  $2D$ -подзоны она завышена. В предельном случае узкой квантовой ямы,  $L \lesssim 1/K$ , ее потенциал можно заменить на  $\delta$ -функцию. Мы получим огибающую функцию с разрывной производной и ошибку порядка  $(\bar{k}_z/K)^2$ .

Выше, в (13), мы молчаливо считали, что гамильтониан  $H(k_z, k'_z)$  определен для всех  $k_z$  и  $k'_z$ , принадлежащих зоне Бриллюэна. Поскольку наша цель состоит в получении именно однозонных уравнений, мы должны учесть еще одно обстоятельство. В  $k$ -пространстве вблизи  $\Gamma$ -точки существует область  $\Lambda_1$ , в которой спектр состояний зоны проводимости объемного материала можно записать в виде ряда по степеням квазиимпульса (для состояний вырожденной зоны спектр определяется матрицей, элементы

которой являются такими рядами). Ряд сходится при  $|\mathbf{k}| < 1/2\lambda$ , что следует из двухзонного приближения (такая оценка справедлива, если эффективная масса формируется, в основном,  $\mathbf{k}\mathbf{r}$ -взаимодействием). Именно эту область описывает гамильтониан однозонного уравнения. Существует также область  $\Lambda_2$ , где взаимодействие состояний выделенных зон с удаленными не описывается этим рядом. В нашем случае резких гетеропереходов огибающие функции в  $\mathbf{k}$ -представлении убывают степенным образом, поэтому мы должны правильно описать и область  $\Lambda_2$ , что будет сделано в другом месте в связи с проблемой междолинного смешивания состояний в гетероструктурах. Здесь можно упомянуть, что если отношение характерного разрыва зон к энергетическому зазору между интересующими нас состояниями в области  $\Lambda_1$  и состояниями в области  $\Lambda_2$  является малым параметром  $\omega \ll 1$ , то ошибка от пренебрежения областью  $\Lambda_2$  будет меньше или порядка  $\omega(a\bar{k}_z)^2$ . Таким образом, эффективный радиус в  $\mathbf{k}$ -пространстве, определяющий точность сведения интегрального уравнения к дифференциальному, на самом деле не задается размерами зоны Бриллюэна по оси  $k_z$ , а зависит от того, какие зоны считаются далекими и «устраиваются» унитарным преобразованием. В нашем случае этот радиус есть величина порядка  $1/\lambda$ .

Ниже мы получим уравнение для зоны проводимости, опуская детали унитарного преобразования в  $\mathbf{k}$ -пространстве и сразу переходя в  $\mathbf{r}$ -представление. Формально конечное уравнение будет представлять собой дифференциальное уравнение четвертого порядка, причем удовлетворяющая ему огибающая функция в самом неблагоприятном для точности случае разрывного потенциала будет иметь разрывную вторую производную с характерным разрывом порядка самой второй производной. Можно действовать иначе. Сводя дифференциальное уравнение четвертого порядка к физически эквивалентному второго порядка [28], мы получим разрывную огибающую функцию с характерным относительным разрывом порядка  $(\bar{k}_z \lambda)^2$ . Все это означает, что для единичного гетероперехода или широкой квантовой ямы точность полученного уравнения эффективной массы ограничивается учетом всех вкладов вплоть до членов порядка  $(\bar{k}_z \lambda)^3$ . В случае же узкой квантовой ямы уже при  $L \sim \lambda$  уравнение эффективной массы должно включать лишь первые поправки, связанные с эффектами резкости гетероперехода, а рассмотрение всех остальных поправок, в том числе приводящих к пространственной зависимости параметров эффективной массы, будет являться превышением точности. В таком случае удобен формализм короткодействующего потенциала, уже использовавшийся для получения разложения (11). Этому посвящен разд. 4.3.

## 4. ОДНОЗОННЫЕ УРАВНЕНИЯ

### 4.1. Зона проводимости

#### 4.1.1. Плавный гетеропереход

Переход от многозонной системы уравнений для огибающих функций к однозонному осуществляется стандартным образом [1, 27]. Для плавных гетеропереходов однозонное уравнение для огибающих функций для  $c$ -зоны (проводимости) было получено в [28]. В  $\mathbf{r}$ -представлении оно имеет вид

$$\epsilon_{c0} F_c(\mathbf{r}) + \frac{1}{2} m^{\alpha_1}(z) p m^{\beta_1}(z) p m^{\alpha_1}(z) F_c(\mathbf{r}) + \Gamma(z) \Delta U_c F_c(\mathbf{r}) +$$

$$+\alpha_0 \mathbf{p}^4 F_c(\mathbf{r}) + \beta_0 \left( \mathbf{p}_{||}^2 p_z^2 + p_x^2 p_y^2 \right) F_c(\mathbf{r}) + \eta [\mathbf{p} \mathbf{n}] \cdot \boldsymbol{\sigma} \Gamma'(z) F_c(\mathbf{r}) = \epsilon F_c(\mathbf{r}). \quad (15)$$

Разрыв зоны проводимости  $\Delta U_c$  и модифицированный формфактор гетероперехода  $\Gamma(z)$  определяются выражением

$$\Gamma(z) \Delta U_c = G(z) \delta U_{cc} + \sum_n' \frac{|\delta U_{cn}|^2}{\epsilon_{c0} - \epsilon_{n0}} G^2(z),$$

так что во всех малых поправках  $G(z)$  можно заменить на  $\Gamma(z)$ . Пространственно-зависящая эффективная масса дается выражением

$$m(z) = m_1 [1 + m_1 (\mu_1 - \mu_2) \Gamma(z)],$$

и  $m_1$  есть эффективная масса края зоны проводимости левого материала, а для  $m_2$ , являющейся эффективной массой края зоны проводимости правого материала, мы имеем

$$1/m_2 = 1/m_1 + \mu_2 - \mu_1.$$

Параметры  $\alpha_1$  и  $\beta_1$  определены следующим образом:

$$\alpha_1 = \frac{\mu_1}{2(\mu_2 - \mu_1)}, \quad 2\alpha_1 + \beta_1 = -1.$$

Параметры  $\mu_1$  и  $\mu_2$  даются выражениями

$$\mu_1 = \sum_n' 2 \frac{|\langle c | p_x | n \rangle|^2 \delta U_{cc}}{m_0^2 (\epsilon_{c0} - \epsilon_{n0})^2} - \sum_{n,l} 4 \frac{\langle c | p_x | n \rangle \langle n | p_x | l \rangle \delta U_{lc}}{m_0^2 (\epsilon_{c0} - \epsilon_{n0}) (\epsilon_{c0} - \epsilon_{l0})}, \quad (16)$$

$$\mu_2 = \sum_{n,l} 2 \frac{\langle c | p_x | n \rangle \delta U_{nl} \langle l | p_x | c \rangle}{m_0^2 (\epsilon_{c0} - \epsilon_{n0}) (\epsilon_{c0} - \epsilon_{l0})}. \quad (17)$$

В (15)  $\alpha_0$  и  $\beta_0$  — параметры непараболичности объемного спектра. Наконец, последний параметр, входящий в уравнение, есть

$$\eta = \sum_{n,l} \frac{\hbar^2 \langle c | p_z | n \rangle \langle n | [\nabla \delta U \mathbf{p}]_x | l \rangle \langle l | p_y | c \rangle}{4im_0^4 c^2 (\epsilon_{c0} - \epsilon_{n0}) (\epsilon_{c0} - \epsilon_{l0})}.$$

Во Введении было указано, что в [14] методом прямого исключения малых огибающих функций было получено неправильное выражение для эффективной массы края зоны проводимости полупроводника, не являющегося базисным. Это соответствует отсутствию в выражении (16) второй суммы, а в выражении (17) — членов с  $n \neq l$ . Таким образом, следует быть осторожным при использовании метода прямого исключения малых огибающих функций.

Связь огибающей функции зоны проводимости  $F_c(\mathbf{r})$  с полной волновой функцией дается выражением

$$\Psi(\mathbf{r}) = u_{c0} \{1 + 2Rm_1 (\Gamma(z) \Delta U_c + \epsilon_{c0} - \epsilon)\} F_c(\mathbf{r}) + \sum_n' \frac{u_{n0}}{(\epsilon_{c0} - \epsilon_{n0})} \left[ \frac{\hbar \langle n | \mathbf{p} | c \rangle \nabla}{im_0} + \delta U_{nc} \Gamma(z) + \sum_{l,\alpha,\beta} \frac{\langle n | p_\alpha | l \rangle \langle l | p_\beta | c \rangle}{m_0^2 (\epsilon_{c0} - \epsilon_{l0})} p_\alpha p_\beta \right] F_c(\mathbf{r}). \quad (18)$$

Здесь

$$R = \frac{1}{2} \sum_n' \frac{|(c|p_x|n)|^2}{m_0^2 (\epsilon_{c0} - \epsilon_{n0})^2},$$

а член  $2Rm_1(\Gamma(z)\Delta U_c + \epsilon_{c0} - \epsilon)$  в фигурных скобках выражения (18) происходит из члена  $-Rp^2$ , рассмотренного как возмущение с использованием стандартного уравнения эффективной массы. В квадратных скобках (18) пренебрежено слагаемым

$$\frac{\hbar \langle n | [\nabla U_1 \mathbf{p}] | c \rangle \sigma}{4m_0^2 c^2} + \Gamma(z) \frac{\hbar \langle n | [\nabla \delta U \mathbf{p}] | c \rangle \sigma}{4m_0^2 c^2}, \quad (19)$$

поскольку наибольший вклад в матричные элементы (19) вносит область потенциала вблизи атомных ядер, в которой оператор спин-орбитального взаимодействия можно записать в виде произведения операторов спина электрона и орбитального момента, а функция  $u_{c0}$  сферически-симметрична (орбитальный момент равен нулю).

В [28] было показано, что при  $\bar{k}_z d \ll 1$  уравнение (15) может быть заменено эквивалентным уравнением, где функция Хэвисайда  $\theta(z - z_0)$  заменит функцию  $\Gamma(z)$ , причем координата  $z_0$ , задающая положение математически резкого гетероперехода, может быть выбрана произвольно в пределах  $-d \leq z_0 \leq d$ . Метод, использовавшийся в [28] для такого перехода, не является самым удобным. Существует более простой путь перехода к математически резкому гетеропереходу, основанный на следующей цепочке равенств, справедливых для операторов, действующих на плавные функции:

$$\begin{aligned} \Gamma(z) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk_x e^{ik_x z} \int_{-\infty}^{+\infty} \Gamma(z') e^{-ik_x z'} dz' = \theta(z - z_0) + \\ &+ \frac{1}{2\pi} \int_{-\infty}^{+\infty} dk_x e^{ik_x(z-z_0)} \int_{-d}^d (\Gamma(z') - \theta(z' - z_0)) e^{-ik_x(z'-z_0)} dz' \approx \theta(z - z_0) + \\ &+ \delta(z - z_0) \left[ \int_{-d}^d \Gamma(z) dz - (d - z_0) \right] + \delta'(z - z_0) \left[ \int_{-d}^d \Gamma(z) (z_0 - z) dz + \frac{(d - z_0)^2}{2} \right]. \end{aligned}$$

Принимая  $z_0 = 0$ , вместо (15) получим более удобный вид уравнения эффективной массы для плавного гетероперехода:

$$\begin{aligned} [\epsilon_{c0} - \epsilon + \theta(z)\Delta U_c + \Delta U_c \rho_0 \delta(z)] F_c(\mathbf{r}) + \frac{1}{2} m^{\alpha_2}(z) \mathbf{p} m^{\beta_2}(z) \mathbf{p} m^{\alpha_2}(z) F_c(\mathbf{r}) + \\ + \alpha_0 \mathbf{p}^4 F_c(\mathbf{r}) + \beta_0 (\mathbf{p}_{||}^2 p_x^2 + p_x^2 p_y^2) F_c(\mathbf{r}) + \eta [\mathbf{p} \mathbf{n}] \sigma \delta(z) F_c(\mathbf{r}) = 0. \end{aligned} \quad (20)$$

Здесь

$$m(z) = m_1 [1 + m_1 (\mu_1 - \mu_2) \theta(z)],$$

а параметр  $\alpha_2$  не только зависит от материалов гетероперехода, но и учитывает конечность его ширины через его зависимость от  $\Gamma(z)$ :

$$\alpha_2 = \frac{\mu_1 + 2\Delta U_c \hbar^{-2} \left[ d^2 - \int_{-d}^d 2\Gamma(z) z dz \right]}{2(\mu_2 - \mu_1)}$$

Равенство  $2\alpha_2 + \beta_2 = -1$  при этом сохраняется, а

$$\rho_0 = \int_{-d}^d \Gamma(z) dz - d.$$

Проделанный выше переход к математически резкому гетеропереходу диктуется лишь соображениями удобства, при этом явная форма функции  $\Gamma(z)$  проявляется лишь интегральным образом в величине двух параметров,  $\alpha_2$  и  $\rho_0$ .

#### 4.1.2. Резкий гетеропереход

Для включения в (20) поправок, связанных с неплавностью гетероперехода, следует в рамках стандартной процедуры [1, 27] учесть, что вклады, определяемые членами  $D_{0nn'}$ ,  $D_{1nn'}$  и  $B_{0nn'}$ , должны рассматриваться в первом порядке теории возмущений ( $S_{0cc} = S_{1cc} = 0$ , что следует из симметрии цинковой обманки), а вклады от  $D_{0nn'}$  и  $S_{0nn'}$  — еще и во втором порядке совместно с членами  $\hbar k p_{nn'}/m_0$ . Используя свойства симметрии, нетрудно получить дополнительный вклад  $H_{abr}$  в гамильтониан уравнения (15):

$$H_{abr} = D_{0cc}\delta(z) + \rho\delta'(z) + \tilde{\eta} [pn] \sigma\delta(z). \tag{21}$$

Здесь

$$\begin{aligned} D_{0cc} &= - \sum_{j \neq 0} \frac{\langle c | \delta U \cos(K_j z) | c \rangle}{K_j} \int_{-d}^d G'(z) \sin(K_j z) dz, \\ \rho &= \sum_{n, j \neq 0} \frac{\hbar \langle c | p_z | n \rangle \langle n | \delta U \sin(K_j z) | c \rangle}{i K_j m_0 (\epsilon_{c0} - \epsilon_{n0})} \int_{-d}^d G'(z) \cos(K_j z) dz + \\ &+ \sum_{j \neq 0} \frac{\langle c | \delta U \cos(K_j z) | c \rangle}{K_j} \int_{-d}^d G'(z) \left( \frac{\cos(K_j z)}{K_j} + z \sin(K_j z) \right) dz, \\ \tilde{\eta} &= \sum_{j \neq 0} \left( \sum_n \frac{\hbar \langle c | [\nabla(\sin(K_j z) \delta U) \times \mathbf{p}]_x | n \rangle \langle n | p_y | c \rangle}{2 K_j m_0^3 c^2 (\epsilon_{c0} - \epsilon_{n0})} - \right. \\ &\left. - \frac{\hbar \langle c | \nabla_z (\sin(K_j z) \delta U) | c \rangle}{4 K_j m_0^2 c^2} \right) \int_{-d}^d G'(z) \cos(K_j z) dz. \end{aligned}$$

Из сравнения выражения (21) с гамильтонианом уравнения (20) следует, что для зоны проводимости эффекты резкости гетероперехода не приводят к изменению формы

однозонного уравнения, а лишь перенормируют использующиеся в нем константы. С учетом (20) и (21) искомое уравнение для зоны проводимости принимает вид

$$(\epsilon_{c0} - \epsilon + \theta(z)\Delta U_c + d_1\delta(z))F_c(\mathbf{r}) + \frac{1}{2}m^\alpha(z)\mathbf{p}m^\beta(z)\mathbf{p}m^\alpha(z)F_c(\mathbf{r}) + \alpha_0\mathbf{p}^4F_c(\mathbf{r}) + \beta_0\left(\mathbf{p}_{\parallel}^2p_z^2 + p_z^2p_y^2\right)F_c(\mathbf{r}) + d_2[\mathbf{p}\mathbf{n}]\sigma\delta(z)F_c(\mathbf{r}) = 0. \quad (22)$$

Слагаемое  $\rho\delta'(z)$  в выражении (21) поглотилось оператором кинетической энергии, при этом для  $\alpha$  получаем

$$\alpha = \frac{\mu_1 + 2\Delta U_c\hbar^{-2} \left[ d^2 - \int_{-d}^d 2\Gamma(z)z dz \right] + 4\rho\hbar^{-2}}{2(\mu_2 - \mu_1)}$$

(здесь исправлена ошибка, допущенная в [29]), и  $2\alpha + \beta = -1$ , а  $d_1 = D_{0cc} + \Delta U_c\rho_0$  и  $d_2 = \eta + \tilde{\eta}$ .

Обсудим гамильтониан уравнения (22). Первое слагаемое представляет потенциальную энергию электрона в  $c$ -зоне. Возможность существования гетероинтерфейсного члена, пропорционального  $d_1$ , обсуждалась в [5]; видно, что этот вклад исчезает для математически резкого гетероперехода (в этом довольно далеком от реальности случае модели (1) и (9) тождественны). Второе слагаемое — квадратичный по импульсу пространственно-зависящий оператор кинетической энергии; в общем виде такая форма была предложена в [6]. Отметим, что параметр  $\alpha$  не является универсальной константой, а зависит как от материалов, так и от формы переходной области гетероперехода, что даже интуитивно кажется естественным (см. также [8]). Если случайным образом окажется, что  $m(z) = \text{const}$ , т.е.  $\mu_1 = \mu_2$ , в использующейся нами форме оператора кинетической энергии возникнут неопределенности вида  $1^\infty$ , которые легко раскрываются. Таких неопределенностей не будет, если воспользоваться другой эквивалентной формой записи этого оператора — выделить член, аналогичный релятивистскому члену Дарвина (см. [28]):

$$T_2 = \mathbf{p} \frac{1}{2m(z)} \mathbf{p} + \left( \frac{\mu_1\hbar^2}{4} + \Delta U_c \left[ \frac{d^2}{2} - \int_{-d}^d \Gamma(z)z dz \right] + \rho \right) \delta'(z).$$

Далее, третье и четвертое слагаемые в гамильтониане уравнения (22) описывают поправки на слабую непараболичность и определяются только объемными параметрами. Пятое слагаемое описывает интерфейсное спин-орбитальное взаимодействие (см., например, [32]), сила которого ( $d_2$ ) зависит не только от материалов гетероперехода, но и от формы переходной области. Связь полной волновой функции с огибающими функциями зоны проводимости дается выражением (18), в котором можно положить  $\Gamma(z) = \theta(z)$ .

#### 4.1.3. Граничные условия на огибающие

Из (22) несложно получить граничные условия, налагаемые на огибающие функции на гетерогранице; для этого необходимо свести (22) к дифференциальному уравнению второго порядка, пользуясь малостью вклада, пропорционального  $\mathbf{p}^4$  [28]. Приведем лишь результат:

$$\left. \begin{pmatrix} F_c(\mathbf{r}) \\ F'_c(\mathbf{r}) \end{pmatrix} \right|_{z=+0} = \begin{pmatrix} d_{11} & 0 \\ d_{21} & d_{22} \end{pmatrix} \left. \begin{pmatrix} F_c(\mathbf{r}) \\ F'_c(\mathbf{r}) \end{pmatrix} \right|_{z=-0},$$

где

$$d_{11} = 1 + 4m_1^2\alpha_0\Delta U_c + m_1\alpha \left( \frac{1}{m_2} - \frac{1}{m_1} \right),$$

$$d_{21} = \frac{2m_1}{\hbar^2} (d_1 + d_2 [\mathbf{pn}] \sigma),$$

и

$$d_{22} = \frac{m_2}{m_1} + 4m_1^2\alpha_0\Delta U_c - m_1\alpha \left( \frac{1}{m_2} - \frac{1}{m_1} \right).$$

В приближении стандартного метода эффективной массы мы имеем пространственно-независимую эффективную массу и скачкообразный потенциал. Поправки первого порядка по малому параметру задачи учитываются включением в стандартное уравнение  $\delta$ -функционального потенциала, пропорционального  $d_1$  (формально это поправка порядка  $\bar{k}_z d$ ). Полное уравнение (22) включает и все поправки порядка  $(\lambda \bar{k}_z)^2$ . Более малые вклады невозможно учесть в однозонном варианте метода огибающей функции из-за существования такого же порядка малости неизбежной ошибки, возникающей при переходе от многозонной системы интегральных уравнений к однозонному дифференциальному.

Несложно обобщить полученные результаты на случай произвольной гетероструктуры. Легко это сделать из соображений симметрии для симметричной квантовой ямы с двумя эквивалентными гетеропереходами с координатами  $z = 0$  и  $z = L$ , причем  $L = \mathcal{N}a/2$ , где  $\mathcal{N}$  — целое. Эффективный гамильтониан есть

$$H^{el} = \epsilon_{c0} + \Delta U_c \{ \theta(z) - \theta(z - L) \} + \frac{1}{2} m^\alpha(z) \mathbf{p} m^\beta(z) \mathbf{p} m^\alpha(z) + \alpha_0 \mathbf{p}^4 + \\ + \beta_0 \left( \mathbf{p}_x^2 p_z^2 + p_x^2 p_y^2 \right) + d_1 \{ \delta(z) + \delta(z - L) \} + d_2 \{ \delta(z) - \delta(z - L) \} [\mathbf{pn}] \sigma.$$

#### 4.2. Валентная зона

Отличие метода эффективной массы для валентной зоны от такового для зоны проводимости, в принципе, заключается лишь в необходимости рассматривать более громоздкие уравнения в первом случае. Основные особенности проблемы вывода уравнений для дырочных состояний с пространственно-зависящими параметрами эффективной массы были отмечены в [29]. Уже в уравнение для  $s$ -зоны входят три новых параметра, зависящие как от объемных свойств материалов гетероструктуры, так и от свойств гетерограницы. Для валентной же зоны подобных параметров будет существенно больше, что, возможно, приведет к малой практической ценности такого уравнения. Поэтому здесь мы ограничимся получением поправок лишь первого порядка по указанному выше параметру к стандартному уравнению.

В базе  $\{|J, j_z\rangle\}$  собственных функций полного момента  $J$  и его проекции  $j_z$ , являющихся комбинациями блоховских функций  $\mathcal{X}$ ,  $\mathcal{Y}$  и  $\mathcal{X}$  потолка валентной зоны

(преобразующиеся по представлению  $\Gamma_{15}$ ) и спина, в этом приближении матрица эффективного гамильтониана для валентной зоны  $\mathbf{H}$  представляет собой сумму матрицы  $6 \times 6$  стандартного оператора кинетической энергии  $\mathbf{T}$  (мы пренебрегаем малым вкладом объемных линейных по  $\mathbf{k}$  членов от спин-орбитального взаимодействия) и матрицы  $6 \times 6$  оператора потенциальной энергии  $\mathbf{V}$ :

$$\mathbf{H} = \mathbf{T} + \mathbf{V}.$$

Ясно, что в стандартном приближении в  $\mathbf{V}$  входят лишь диагональные скачкообразные потенциалы. В следующем приближении появляются дополнительные члены, представляющие собой диагональные и недиагональные  $\delta$ -функциональные потенциалы.

Удобно сначала найти элементы матрицы потенциальной энергии  $\tilde{V}$  в базисе  $\{\mathcal{X}, \mathcal{Y}, \mathcal{X}\}$ , а затем составить из них необходимые линейные комбинации и перейти к  $\mathbf{V}$ :

$$\begin{aligned} \tilde{V}_{\mathcal{X}\mathcal{X}} &= \tilde{V}_{\mathcal{X}\mathcal{X}} = \epsilon_{\mathcal{X}0} + \delta U_{\mathcal{X}\mathcal{X}} G(z) + D_{0\mathcal{X}\mathcal{X}} \delta(z) \approx \\ &\approx \epsilon_{\mathcal{X}0} + \delta U_{\mathcal{X}\mathcal{X}} \theta(z) + (D_{0\mathcal{X}\mathcal{X}} + \rho_0 \delta U_{\mathcal{X}\mathcal{X}}) \delta(z), \end{aligned}$$

$$\begin{aligned} \tilde{V}_{\mathcal{X}\mathcal{Y}} &= \frac{1}{3i} (\Delta + \delta \Delta G(z)) \sigma_z + D_{0\mathcal{X}\mathcal{Y}} \delta(z) + S_{0\mathcal{X}\mathcal{Y}}^z \delta(z) \sigma_z \approx \\ &\approx \frac{1}{3i} (\Delta + \delta \Delta \theta(z)) \sigma_z + D_{0\mathcal{X}\mathcal{Y}} \delta(z) + \left( S_{0\mathcal{X}\mathcal{Y}}^z + \frac{\rho_0 \delta \Delta}{3i} \right) \delta(z) \sigma_z, \end{aligned}$$

где

$$\Delta = \frac{3\hbar i \langle \mathcal{X} | [\nabla U_1 \mathbf{p}]_z | \mathcal{Y} \rangle}{4m_0^2 c^2}, \quad \delta \Delta = \frac{3\hbar i \langle \mathcal{X} | [\nabla \delta U \mathbf{p}]_z | \mathcal{Y} \rangle}{4m_0^2 c^2}.$$

Аналогично,

$$\begin{aligned} \tilde{V}_{\mathcal{X}\mathcal{X}} &= \frac{i}{3} (\Delta + \delta \Delta G(z)) \sigma_y + S_{0\mathcal{X}\mathcal{X}}^y \delta(z) \sigma_y \approx \\ &\approx \frac{i}{3} (\Delta + \delta \Delta \theta(z)) \sigma_y + \left( S_{0\mathcal{X}\mathcal{X}}^y + \frac{i \rho_0 \delta \Delta}{3} \right) \delta(z) \sigma_y. \end{aligned}$$

Остальные элементы  $\tilde{V}$  можно получить из приведенных выше с помощью циклической перестановки индексов.

Вклад шестого и седьмого слагаемых из левой части (12), как видно, не вошел в  $\tilde{V}$ , поскольку он пренебрежимо мал по следующей причине:

$$\langle \mathcal{X} | \nabla_z U_1 | \mathcal{Y} \rangle = 0$$

(см. [27, § 21]), а значит, матричный элемент

$$\langle \mathcal{X} | \nabla_z \delta U | \mathcal{Y} \rangle = \langle \mathcal{X} | \nabla_z U_2 | \mathcal{Y} \rangle.$$

Теперь для блоховских функций  $n$ -ой зоны в  $\Gamma$ -точке правого кристалла в нерелятивистском пределе ( $\tilde{u}_{n0}$ ) имеем

$$\tilde{u}_{n0} = u_{n0} + \sum_{n'}' \frac{\delta U_{n'c} u_{n'0}}{\epsilon_{c0} - \epsilon_{n'0}}, \quad (23)$$

а для соответствующих функций края валентной зоны правого кристалла  $\tilde{\mathcal{X}}$  и  $\tilde{\mathcal{Y}}$  справедливо  $\langle \tilde{\mathcal{X}} | \nabla_z U_2 | \tilde{\mathcal{Y}} \rangle = 0$ . То есть, седьмое слагаемое из левой части (12) даст поправку лишь порядка  $(\lambda \bar{k}_z)^3$ , а не  $\lambda \bar{k}_z$ . В нашем же приближении разницей в блоховских функциях  $\tilde{u}_{n0}$  и  $u_{n0}$  нужно вообще пренебречь. Отсюда и из инвариантности уравнения относительно обращения времени следует, что вклад и шестого слагаемого левой части (12) пренебрежимо мал.

Несложно теперь получить элементы V. Фазы выберем так, как это сделано в [1] и обозначим

$$\begin{aligned} |1\rangle &\equiv \left| \frac{3}{2}, \frac{3}{2} \right\rangle, & |2\rangle &\equiv \left| \frac{3}{2}, -\frac{3}{2} \right\rangle, & |3\rangle &\equiv \left| \frac{3}{2}, \frac{1}{2} \right\rangle, \\ |4\rangle &\equiv \left| \frac{3}{2}, -\frac{1}{2} \right\rangle, & |5\rangle &\equiv \left| \frac{1}{2}, \frac{1}{2} \right\rangle, & |6\rangle &\equiv \left| \frac{1}{2}, -\frac{1}{2} \right\rangle \end{aligned}$$

(пары состояний тяжелых дырок, легких дырок и состояний спин-отщепленной зоны). Тогда искомая матрица потенциальной энергии принимает вид

$$V = \begin{pmatrix} V_{\Gamma_8 1} & V_0 \sigma_y & -i\sqrt{2} V_0 \sigma_y \\ V_0^\dagger \sigma_y & V_{\Gamma_8 1} & 0 \\ i\sqrt{2} V_0^\dagger \sigma_y & 0 & V_{\Gamma_7 1} \end{pmatrix}, \quad (24)$$

где

$$V_{\Gamma_8} = E_{\Gamma_8} + \Delta U_{\Gamma_8} \theta(z) + \chi_1 \delta(z),$$

$$V_{\Gamma_7} = E_{\Gamma_7} + \Delta U_{\Gamma_7} \theta(z) + \chi_2 \delta(z),$$

$$V_0 = \frac{D_0 \mathcal{X} \mathcal{Y}}{\sqrt{3}} \delta(z).$$

Здесь мы ввели следующие обозначения:

$$E_{\Gamma_8} = \epsilon_{\mathcal{X}0} + \frac{1}{3} \Delta, \quad \Delta U_{\Gamma_8} = \delta U_{\mathcal{X}\mathcal{X}} + \frac{1}{3} \delta \Delta,$$

$$E_{\Gamma_7} = \epsilon_{\mathcal{X}0} - \frac{2}{3} \Delta, \quad \Delta U_{\Gamma_7} = \delta U_{\mathcal{X}\mathcal{X}} - \frac{2}{3} \delta \Delta,$$

$$\chi_1 = D_0 \mathcal{X}\mathcal{X} + \rho_0 \delta U_{\mathcal{X}\mathcal{X}} + i S_0^z \mathcal{X}\mathcal{Y} + \frac{1}{3} \rho_0 \delta \Delta,$$

$$\chi_2 = D_0 \mathcal{X}\mathcal{X} + \rho_0 \delta U_{\mathcal{X}\mathcal{X}} - 2i S_0^z \mathcal{X}\mathcal{Y} - \frac{2}{3} \rho_0 \delta \Delta.$$

Выражения для  $D_0 \mathcal{X}\mathcal{X}$ ,  $D_0 \mathcal{X}\mathcal{Y}$  и  $S_0^z \mathcal{X}\mathcal{Y}$  имеют следующий вид:

$$\begin{aligned}
 D_{0\mathcal{X}\mathcal{X}} &= - \sum_{j \neq 0} \frac{\langle \mathcal{X} | \delta U \cos(K_j z) | \mathcal{X} \rangle}{K_j} \int_{-d}^d G'(z) \sin(K_j z) dz, \\
 D_{0\mathcal{X}\mathcal{Y}} &= \sum_{j \neq 0} \frac{\langle \mathcal{X} | \delta U \sin(K_j z) | \mathcal{Y} \rangle}{K_j} \int_{-d}^d G'(z) \cos(K_j z) dz, \\
 S_{0\mathcal{X}\mathcal{Y}}^z &= - \sum_{j \neq 0} \frac{\hbar \langle \mathcal{X} | [\nabla (\cos(K_j z) \delta U) \times \mathbf{p}_z] | \mathcal{Y} \rangle}{4K_j m_0^2 c^2} \int_{-d}^d G'(z) \sin(K_j z) dz.
 \end{aligned} \tag{25}$$

Таким образом, в рамках  $\mathbf{k}\mathbf{p}$ -метода мы доказали, что в (001) III-V гетероструктурах существует смешивание тяжелых ( $hh$ ) и легких ( $lh$ ) дырок в центре  $2D$ -зоны Бриллюэна (см. работу [33] и содержащиеся в ней ссылки), не связанное с объемными линейными по  $k_z$  членами от спин-орбитального взаимодействия. Это смешивание определяется параметром  $D_{0\mathcal{X}\mathcal{Y}}$ , оценка для которого была проведена в [33] с учетом экспериментальных данных для гетероструктур GaAs/AlAs:  $D_{0\mathcal{X}\mathcal{Y}} \simeq 500$  мэВ·Å.

В [29] был сделан вывод о том, что сила смешивания тяжелых и легких дырок в центре  $2D$ -зоны Бриллюэна больше для резких гетеропереходов, чем для гетеропереходов с плавно меняющимся химическим составом. Но это справедливо лишь в модели (1). Вообще говоря, можно сделать вывод лишь о зависимости силы этого смешивания от структуры переходной области гетеропереходов.

Для симметричной квантовой ямы с границами при  $z = 0$  и  $z = L$  элементы матрицы потенциальной энергии легко можно получить из соображений симметрии:

$$V_{\Gamma_1} = E_{\Gamma_1} + \Delta U_{\Gamma_1} (\theta(z) - \theta(z - L)) + \chi_1 (\delta(z) + \delta(z - L)),$$

$$V_{\Gamma_2} = E_{\Gamma_2} + \Delta U_{\Gamma_2} (\theta(z) - \theta(z - L)) + \chi_2 (\delta(z) + \delta(z - L)),$$

$$V_0 = \frac{D_{0\mathcal{X}\mathcal{Y}}}{\sqrt{3}} (\delta(z) - \delta(z - L)).$$

#### 4.3. Уравнение для огибающих функций для узкой ямы в зоне проводимости

Проблеме электронных состояний в узких квантовых ямах мы уделяем отдельное внимание, потому что в настоящее время существует твердое убеждение о неприменимости метода эффективной массы в таких случаях. Здесь мы рассмотрим лишь состояния  $s$ -зоны в (001) III-V гетероструктуре, составленной из родственных полупроводников, типа одиночной квантовой ямы (или узкого барьера), для случая, когда ее ширина  $L \lesssim \lambda$ . Дырочные состояния легко рассмотреть аналогичным образом. Теперь вклады резкости потенциала начинают играть существенно большую роль, чем в случае широких квантовых ям. В самом деле, оценка  $\bar{k}_z^2/2m \sim \delta U_{cc}$  справедлива лишь тогда, когда ширина квантовой ямы больше или порядка характерной длины локализации состояний. Для состояний же в узкой квантовой яме, естественно,  $\bar{k}_z^2/2m \sim \delta U_{cc} \bar{k}_z L$ , а вклад

в собственное значение энергии членов, связанных с резкостью потенциала, в первом порядке можно оценить как  $\delta U_{cc} \bar{k}_z a$ , что означает необходимость их учета даже в нулевом приближении.

Итак, потенциал рассматриваемой гетероструктуры можно записать как

$$U = U_1 + P[U_2 - U_1] \equiv U_1 + P(z)\delta U,$$

где  $U_1$  и  $U_2$  есть периодические кристаллические потенциалы «барьерного» (базисного) и «ямного» полупроводников соответственно,  $P(z)$  — формфактор гетероструктуры. Определим  $L$  так, что  $P(z < -L/2) = P(z > L/2) = 0$ . Естественно рассматривать  $P(z)$  как локальную функцию на масштабе изменения огибающей функции зоны проводимости, поскольку  $L\bar{k}_z \ll 1$ . Однозонное уравнение будет иметь следующий вид:

$$\left[ \epsilon_{c0} - \epsilon + \frac{p^2}{2m} + b_1\delta(z) + b_2\delta'(z) + b_3\delta(z)[\mathbf{pn}]\sigma \right] F_c(\mathbf{r}) = 0, \quad (26)$$

где  $m$  является пространственно-независимой эффективной массой, а для трех параметров  $b_i$  ( $i = 1, 2, 3$ ) мы имеем следующие выражения:

$$\begin{aligned} b_1 &= \sum_j \langle c | \delta U \cos(K_j z) | c \rangle \int_{-L/2}^{L/2} P(z) \cos(K_j z) dz, \\ b_2 &= \sum_{n;j} \frac{\hbar \langle c | p_z | n \rangle \langle n | \delta U \sin(K_j z) | c \rangle}{im_0(\epsilon_{c0} - \epsilon_{n0})} \int_{-L/2}^{L/2} P(z) \sin(K_j z) dz - \\ &\quad - \sum_j \langle c | \delta U \cos(K_j z) | c \rangle \int_{-L/2}^{L/2} P(z) \cos(K_j z) z dz, \\ b_3 &= \sum_j \left( \sum_n \frac{\hbar \langle c | [\nabla (\sin(K_j z) \delta U) \times \mathbf{p}]_x | n \rangle \langle n | p_y c \rangle}{2m_0^3 c^2 (\epsilon_{c0} - \epsilon_{n0})} - \right. \\ &\quad \left. - \frac{\hbar \langle c | \nabla_z (\sin(K_j z) \delta U) | c \rangle}{4m_0^2 c^2} \right) \int_{-L/2}^{L/2} P(z) \sin(K_j z) dz, \end{aligned} \quad (27)$$

в которых индекс суммирования  $j$  включает и нуль (слагаемые с  $j = 0$  представляют вклад плавной части потенциала). Член в потенциальной энергии, пропорциональный  $b_1$ , дает основной вклад, а два остальных — поправки порядка  $\lambda \bar{k}_z$ . Для симметричной структуры,  $P(z) = P(-z)$ , уравнение упрощается:

$$b_2 = b_3 = 0.$$

Строго говоря, (26) некорректно с математической точки зрения (оно не имеет строгого нетривиального решения), поэтому ему следует придать другую форму, пользуясь малостью члена, пропорционального  $b_2$ , и отбрасывая поправки порядка  $(\lambda \bar{k}_z)^2$ . Нужно

воспользоваться приближенным равенством

$$b_1 \delta(z) + b_2 \delta'(z) \approx b_1 \delta \left( z + \frac{b_2}{b_1} \right), \quad (28)$$

а в члене, пропорциональном  $b_3$ ,  $\delta(z)$  можно для простоты заменить на  $\delta(z + b_2/b_1)$ . Мы получим корректное уравнение:

$$\left[ \epsilon_{c0} - \epsilon + \frac{p^2}{2m} + b_1 \delta \left( z + \frac{b_2}{b_1} \right) + b_3 \delta \left( z + \frac{b_2}{b_1} \right) [\text{pn}]\sigma \right] F_c(\mathbf{r}) = 0. \quad (29)$$

Заметим, что для состояний одной зоны (если не интересоваться, например, межзонными переходами) в гетероструктуре с одним узким слоем значение  $b_2$  не играет роли: в (26) можно сделать сдвиг начала координат  $z' = z + b_2/b_1$ .

Предположим, что мы имеем дело со структурой, содержащей два узких слоя вблизи друг друга (так что расстояние между ними порядка  $\lambda$ ). В этом случае оценка сверху для ошибки, возникающей из-за перехода к однозонному дифференциальному уравнению, дает величину  $\lambda \bar{k}_z$  (это, в частности, справедливо в случае, когда константа  $b_1$ , описывающая потенциал первого слоя, равна взятой с противоположным знаком такой же константе, описывающей потенциал второго слоя). Тогда мы должны изъять из рассмотрения члены, содержащие  $b_2$  и  $b_3$  в (26). Можно, казалось бы, рассмотреть оба слоя как одно локальное возмущение, тем самым уменьшив допустимую ошибку, и получить уравнение (26) с одним набором параметров локального возмущения. В ситуации, для которой мы получили оценку ошибки сверху, нельзя тогда гарантировать малость параметра  $\bar{k}_z b_2/b_1$ , на чем основано (28). Если он не мал, то полученному уравнению нельзя придать математически корректную форму, что означает неприменимость такого подхода.

Таким образом, с учетом изложенной выше модификации метод эффективной массы можно применять и при рассмотрении электронных состояний в гетероструктурах со сверхтонкими слоями. В связи с этим возможна ситуация, когда потенциал толстого слоя некоторого полупроводника выступает в роли барьера для электронных состояний, а в то же время тонкий слой того же материала будет связывать состояния и наоборот, в зависимости от знака параметра  $b_1$ . Знак этого параметра, как видно из (27), может отличаться от знака параметра  $\langle c|\delta U|c \rangle$ , который определяет разрыв зоны проводимости на гетеропереходе.

## 5. ИЕРАРХИЯ УРАВНЕНИЙ ЭФФЕКТИВНОЙ МАССЫ И ОБСУЖДЕНИЕ РЕЗУЛЬТАТОВ

Мы вывели многозонную  $k\mathbf{r}$ -систему интегральных уравнений (12), применимую для описания электронных состояний в гетероструктурах с атомно резким изменением химического состава. В системе содержатся в виде сходящихся рядов по степеням  $(k_z - k'_z)$  вклады, обусловленные неплывностью гетероперехода. Например, в [14, 20, 19] такие члены были отброшены, так что эффектами резкого изменения кристаллического потенциала структуры вблизи гетероперехода фактически пренебрегалось. Именно наличие этих членов отличает выведенную здесь многозонную систему  $k\mathbf{r}$ -уравнений от системы, полученной Лейблером [3] для гетероструктур с плавными гетеропереходами. Во всех известных нам работах, использующих для описания электронных состояний

в гетероструктурах однозонные или многозонные схемы метода огибающих функций, применяется как раз  $k\mathbf{p}$ -система Лейблера. Очень часто также пренебрегается отличием блоховских функций для составляющих гетероструктуру полупроводников, что, в частности, приводит, к  $\delta U_{nn'} = 0$  при  $n \neq n'$ . С учетом членов, обусловленных неплавностью гетероперехода,  $k\mathbf{p}$ -метод можно применить и к описанию междолинного смешивания состояний в гетероструктурах, в том числе к проблеме  $\Gamma - X_z$ -смешивания состояний в (001)-гетероструктурах.

Самым существенным ограничением на точность развиваемого метода огибающих функций, использующего дифференциальные уравнения, является процедура перехода от  $k$ - к  $\mathbf{k}$ -представлению. Однозонное дифференциальное уравнение (четвертого порядка) с пространственно-зависящей эффективной массой оказывается корректным для структур с характерной шириной слоев, (много) большей  $\lambda$ , где, напомним, длина  $\lambda$  определена в разд. 2.

Выше мы подробно рассмотрели, как учет вкладов резкости интерфейсного потенциала модифицирует уравнение для состояний зоны проводимости в (001)-гетероструктурах из родственных согласованных по постоянной решетки III-V полупроводников, выведенное в [28]. Формально полученное уравнение для резкого гетероперехода отличается от такового для плавного гетероперехода лишь перенормировкой входящих в него параметров. В случае же валентной зоны учет резкости гетероперехода приводит к качественно новым эффектам (смешивание тяжелых и легких дырок при  $k_{\parallel} = 0$ ).

Для гетероструктур с широкими слоями можно построить иерархию приближений однозонного метода огибающих функций по параметру  $\lambda \bar{k}_z$ , где  $\bar{k}_z$  есть характерное значение квазиимпульса состояния. Например, для изолированного гетероперехода мы имеем следующее.

0) Нулевой уровень иерархии для электронов.

В приближении эффективной массы, в котором пренебрежено малыми по этому параметру поправками, мы имеем обычное уравнение с пространственно-независящей эффективной массы и скачкообразным потенциалом:

$$\left( \epsilon_{c0} + \theta(z)\Delta U_c + \frac{\mathbf{p}^2}{2m} \right) F_c(\mathbf{r}) = \epsilon F_c(\mathbf{r}). \quad (30)$$

1) Первый уровень иерархии.

Поправки первого порядка (здесь, на самом деле, играет роль малый параметр  $\bar{k}_z d$ , где  $2d$  есть ширина переходной области гетероперехода) учитываются включением в уравнение (30)  $\delta$ -функционального потенциала, локализованного на гетерогранице:

$$\left( \epsilon_{c0} + \theta(z)\Delta U_c + d_1 \delta(z) + \frac{\mathbf{p}^2}{2m} \right) F_c(\mathbf{r}) = \epsilon F_c(\mathbf{r}),$$

где  $d_1$  определяется громоздким выражением в разд. 4.1.

2) Второй уровень иерархии.

Уравнение (22) включает все поправки порядка  $(\lambda \bar{k}_z)^2$ . Более малые вклады, третьего порядка и выше, невозможно корректно учесть в однозонном варианте метода.

Для дырочных состояний мы получили, см. (24), первые поправки к стандартному уравнению (получили первую ступень иерархии уравнений эффективной массы для дырок) и показали, что для (001)-гетероструктур действительно существует смешивание

тяжелых и легких дырок в центре  $2D$ -зоны Бриллюэна, причем именно вклады резкости потенциала гетероперехода определяют силу этого смешивания, так что она зависит от микроскопического строения гетерограниц. В [20] утверждается, в частности, что такое смешивание тяжелых и легких дырок вызвано различием блоховских функций для составляющих гетероструктуру полупроводников и оно отсутствует, если пренебречь таким различием, или, что должно быть эквивалентно, если один и тот же полный набор функций является набором функций Блоха составляющих структуру объемных полупроводников. Если бы это имело место, то (даже не принимая во внимание соображения симметрии) вклад таких членов имел бы малость лишь  $(\lambda \bar{k}_z)^2$ , что видно из (23). На самом деле разница в блоховских функциях не играет существенной роли. Для доказательства рассмотрим идеализированную ситуацию резкого (001)-гомоперехода — задачу о дырочных состояниях в слабом, но не плавном внешнем потенциале, скажем,  $W(z) = G(z)W_0$ , где  $W_0$  является постоянной, определяющей скачок потенциала, малой по сравнению с шириной запрещенной зоны. И в этом случае точечная симметрия структуры ( $C_{2v}$ ) допускает существование смешивания тяжелых и легких дырок в центре  $2D$ -зоны Бриллюэна [33], а для определяющего это смешивание коэффициента  $D_{0\mathcal{X}\mathcal{Y}}$  вместо (25) мы будем иметь

$$D_{0\mathcal{X}\mathcal{Y}} = \sum_{j \neq 0} \frac{W_0 \langle \mathcal{X} | \sin(K_j z) | \mathcal{Y} \rangle}{K_j} \int_{-d}^d G'(z) \cos(K_j z) dz.$$

Это прямое доказательство нашего утверждения.

Введенные в данной работе независимые параметры  $\alpha$ ,  $d_1$  и  $d_2$ , входящие в (22), а также  $\chi_1$ ,  $\chi_2$  и  $D_{0\mathcal{X}\mathcal{Y}}$ , входящие в оператор эффективной потенциальной энергии для состояний валентной зоны, зависят не только от объемных свойств материалов гетероперехода, но и от микроскопической структуры последнего. Все эти параметры определяют гетероинтерфейсный вклад в потенциальную энергию. Между тем известно [34], что вследствие возможного возникновения на границе раздела двух материалов электрического дипольного момента и сама величина скачка потенциала на переходе может зависеть от микроскопического строения границы. Это не описывается в нашей модели гетероперехода из-за неучета эффекта такого диполя. Включение соответствующего резкого электростатического потенциала в (1) и обработка его по изложенной выше схеме как раз и обеспечит нужный эффект.

Электронные состояния в гетероструктурах, составленных из тонких слоев, толщина которых меньше или порядка  $\lambda$ , можно рассматривать лишь в квадратичном по оператору импульса приближении. При этом учет членов, обусловленных резкостью потенциала, становится необходим уже в нулевом приближении. Это видно из (27). В связи с этим возможна следующая ситуация: по мере уменьшения ширины квантовой ямы связанное состояние может исчезнуть, или, наоборот, тонкий номинально «барьерный» слой некоторого полупроводника может создать притягивающий потенциал и образовать связанное состояние. Возможно, именно такая ситуация наблюдалась в [35], а затем моделировалась в [36].

Авторы благодарны Е. Л. Ивченко и А. А. Горбачевичу за полезное обсуждение ряда результатов работы. Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (грант № 99-02-17592) и ИНТАС-РФФИ (грант № 97-11475), а также Министерства науки РФ в рамках программ «Физика твердо-

тельных наноструктур» (грант № 99-1124) и «Поверхностные атомные структуры» (грант № 3.1.99).

## ПРИЛОЖЕНИЕ 1

### Уравнение Дирака с переменной в пространстве щелью

Рассмотрим модельное уравнение Дирака с переменной в пространстве щелью  $2m(\mathbf{r})c^2$ :

$$\begin{bmatrix} m(\mathbf{r})c^2 & c\boldsymbol{\sigma}\mathbf{p} \\ c\boldsymbol{\sigma}\mathbf{p} & -m(\mathbf{r})c^2 \end{bmatrix} \begin{pmatrix} \varphi_e \\ \varphi_p \end{pmatrix} = \epsilon \begin{pmatrix} \varphi_e \\ \varphi_p \end{pmatrix},$$

где  $\varphi_e$  и  $\varphi_p$  — электронная и позитронная компоненты волновой функции соответственно. Пусть  $m(\mathbf{r})$  слабо меняется в пространстве, т.е.  $m(\mathbf{r}) = \bar{m} + \delta m(\mathbf{r})$ , так что  $\delta m(\mathbf{r})/\bar{m} \ll 1$ . С помощью унитарного преобразования типа Фолди—Вотхойзена достаточно просто, следуя, например, схеме, изложенной в [24, гл. XX, § 33], получить однозонное уравнение, описывающее состояния электрона. Так, уравнение, в котором отброшены все малые члены, представляет собой обычное уравнение Шредингера:

$$\left( m(\mathbf{r})c^2 + \frac{\mathbf{p}^2}{2\bar{m}} \right) \bar{\varphi}_e = \epsilon \bar{\varphi}_e,$$

где  $\bar{\varphi}_e$  — преобразованная электронная волновая функция. Уравнение же, в котором пренебрежено всеми членами более высокого порядка малости, чем  $\delta m(\mathbf{r})/\bar{m}$  (следующее приближение), имеет вид

$$H\bar{\varphi}_e = \epsilon\bar{\varphi}_e,$$

$$H = m(\mathbf{r})c^2 + \frac{\mathbf{p}^2}{2\bar{m}} - \frac{\mathbf{p}\delta m(\mathbf{r})\mathbf{p}}{2\bar{m}^2} + \frac{\hbar^2\nabla^2\delta m(\mathbf{r})}{8\bar{m}^2} - \frac{\mathbf{p}^4}{8\bar{m}^3c^2} - \frac{\hbar[\nabla\delta m(\mathbf{r})\mathbf{p}]\boldsymbol{\sigma}}{4\bar{m}^2}. \quad (\text{П.1})$$

Все слагаемые (П.1), за исключением третьего, описывающего пространственно-зависящую массу, можно считать «обычными». Второе, третье и четвертое (член Дарвина) слагаемые можно либо записать в следующей форме:

$$T_2 = \frac{1}{2} \mathbf{p} \frac{1}{m(\mathbf{r})} \mathbf{p} + \frac{\hbar^2\nabla^2\delta m(\mathbf{r})}{8\bar{m}^2},$$

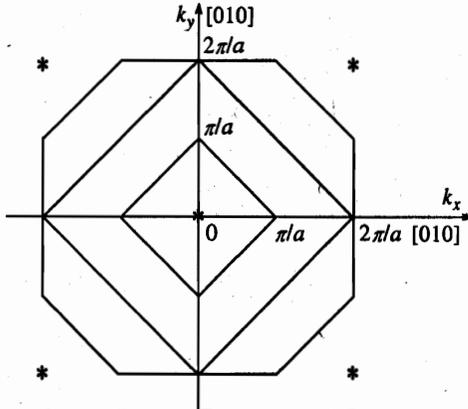
либо объединить в один (квадратичный по импульсу) оператор кинетической энергии:

$$T_2 = \frac{1}{2} \frac{1}{\sqrt{m(\mathbf{r})}} \mathbf{p} \frac{1}{\sqrt{m(\mathbf{r})}} \mathbf{p} \frac{1}{\sqrt{m(\mathbf{r})}},$$

либо использовать любую другую эквивалентную форму записи. Например, в [21] приведена следующая форма для  $T_2$ :

$$T_2 = \frac{1}{4} \left[ \mathbf{p} \frac{1}{\sqrt{m(\mathbf{r})}} \mathbf{p} \frac{1}{\sqrt{m(\mathbf{r})}} + \frac{1}{\sqrt{m(\mathbf{r})}} \mathbf{p} \frac{1}{\sqrt{m(\mathbf{r})}} \mathbf{p} \right].$$

Таким образом, в модельном уравнении Дирака с переменной в пространстве щелью концепция пространственно-зависящей эффективной массы проявляется лишь за рамками нерелятивистского (квадратичного) приближения.



Проекция объемной зоны Бриллюэна (область, ограниченная восьмиугольником) и узлов обратной решетки (обозначены звездочками) на плоскость (001). Квадрат с диагональю  $2\pi/a$  ограничивает область, где нет  $2D$ -процессов переброса

## ПРИЛОЖЕНИЕ 2

### О процессах переброса в двумерной зоне Бриллюэна для (001)-гетероструктуры

Рассмотрим вторую сумму в (4), описывающую процессы переброса в двумерной зоне Бриллюэна, и докажем, что она не дает вклада в интересующем нас случае состояний вблизи ее центра в (001)-гетероструктурах. Поскольку функция  $\mathcal{F}(q)$  отлична от нуля при любых  $q$ , существуют такие отличные от нулевых векторы обратной решетки  $\mathbf{K}_j$ , для которых  $\mathbf{k}'_{\parallel} - \mathbf{k}_{\parallel} = \mathbf{K}_{\parallel j}$ , а  $\mathbf{k}'_{\parallel}$  и  $\mathbf{k}_{\parallel}$  — компоненты векторов, принадлежащих зоне Бриллюэна. Существует конечное число векторов  $\mathbf{K}_j$ , обладающих этим свойством. Поэтому в общем случае мы должны сохранить и вторую сумму в выражении (4). Рассмотрим теперь интересующий нас случай (001)-гетероструктуры. Восьмиугольник на рисунке представляет проекцию объемной зоны Бриллюэна на плоскость (001), а проекции узлов обратной  $3D$ -решетки обозначены звездочками. Функция  $\mathcal{F}'_n(k'_z, \mathbf{k}'_{\parallel})$ , на которую действует оператор  $\mathcal{M}_{nn'}(\mathbf{k}, \mathbf{k}')$  в уравнении (3), определена лишь при  $\mathbf{k}'_{\parallel}$ , принадлежащих проекции объемной зоны Бриллюэна на плоскость (001). А поскольку все узлы  $\mathbf{K}_{\parallel j}$  при  $j \neq 0$  лежат вне этой проекции, существует область  $\mathbf{k}_{\parallel}$ , для которой  $|\mathbf{k}'_{\parallel} - \mathbf{k}_{\parallel}| < \mathbf{K}_{\parallel j}$ ,  $j \neq 0$ . Эта область определяется выражением  $|k_x| + |k_y| < \pi/a$  и на рисунке указана квадратом с диагональю  $2\pi/a$ . Площадь этой области в 4 раза меньше площади первой  $2D$ -зоны Бриллюэна (квадрат с диагональю  $4\pi/a$ ). Именно в области  $|k_x| + |k_y| < \pi/a$  вторая сумма в (4) не дает вклада в уравнения для огибающих функций. Ее достаточно большие размеры обеспечивают выполнение условий применимости выводимых уравнений для огибающих функций, описывающих состояния вблизи  $\Gamma$ -точки.

## Литература

1. J. M. Luttinger and W. Kohn, Phys. Rev. **97**, 869 (1955).
2. J. M. Luttinger, Phys. Rev. **102**, 1030 (1956).

3. L. Leibler, Phys. Rev. B **12**, 4443 (1975).
4. G. Bastard, Phys. Rev. B **24**, 5693 (1981).
5. Q. G. Zhu and H. Kroemer, Phys. Rev. B **27**, 3519 (1983).
6. R. A. Morrow and K. R. Brownstein, Phys. Rev. B **30**, 678 (1984).
7. D. L. Smith and C. Mailhot, Phys. Rev. B **33**, 8345 (1986).
8. W. Trzeciakowski, Phys. Rev. B **38**, 12493 (1988).
9. T. Ando, S. Wakahara, and H. Akera, Phys. Rev. B **40**, 11609 (1989).
10. K. Young, Phys. Rev. B **39**, 13434 (1989).
11. B. Laikhtman, Phys. Rev. B **46**, 4769 (1992).
12. J. P. Cuypers and W. van Haeringen, J. Phys: Condens. Matter **4**, 2587 (1992).
13. B. A. Foreman, Phys. Rev. B **48**, 4964 (1993).
14. M. G. Burt, J. Phys.: Condens. Matter. **4**, 6651 (1994).
15. G. T. Einevoll and L. J. Sham, Phys. Rev. B **49**, 10533 (1994).
16. T. Yamanaka, H. Kamada, Y. Yoshikuni, W. W. Lui, S. Seki, and K. Yokoyama, J. Appl. Phys. **76**, 2347 (1994).
17. W. Trzeciakowski, Semicond. Sci. Technol. **10**, 768 (1995).
18. A. Brezini, M. Sebbani, and S. Marouf, Phys. Stat. Sol. (b) **189**, 389 (1995).
19. B. A. Foreman, Phys. Rev. B **52**, 12241 (1995).
20. B. A. Foreman, Phys. Rev. B **54**, 1909 (1996).
21. F. S. A. Cavalkante, R. N. Costa Filho, J. Ribeiro Filho, C. A. S. de Almeida, and V. N. Freire, Phys. Rev. B **55**, 1326 (1997).
22. M. V. Kisin, B. L. Gelmont, and S. Luryi, Phys. Rev. B **58**, 4605 (1998).
23. В. Келдыш, ЖЭТФ **45**, 364 (1963).
24. А. Мессиа, *Квантовая механика*, Наука, Москва (1979), т. 2.
25. В. Б. Берестецкий, Е. М. Лифшиц, Л. П. Питаевский, *Квантовая электродинамика*, Наука, Москва (1989), § 33.
26. Р. А. Сурис, ФТП **20**, 2008 (1986).
27. Г. Л. Бир, Г. Е. Пикус, *Симметрия и деформационные эффекты в полупроводниках*, Наука, Москва (1972).
28. E. E. Takhtamirov and V. A. Volkov, Semicond. Sci. Technol. **12**, 77 (1997).
29. В. А. Волков, Э. Е. Тахтамиров, УФН **167**, 1123 (1997).
30. E. E. Takhtamirov and V. A. Volkov, in *Proc. of Int. Symp. «Nanostructures: Physics and Technology»*, Ioffe Institute, St. Petersburg (1997), p. 134.
31. E. E. Takhtamirov and V. A. Volkov, in *Proc. of Int. Symp. «Nanostructures: Physics and Technology»*, Ioffe Institute, St. Petersburg (1998), p. 332.
32. Ф. Т. Васько, Письма в ЖЭТФ **30**, 574 (1979).
33. E. L. Ivchenko, A. Yu. Kaminski, and U. M. Rössler, Phys. Rev. B **54**, 5852 (1996).
34. R. G. Dandrea, C. B. Duke, and A. Zunger, J. Vac. Sci. Technol. B **10**, 1744 (1992).
35. R. Schwabe, F. Pietag, W. Faulkner, and S. Lassen, J. Appl. Phys. **77**, 6295 (1995).
36. M. Di Ventra and A. Baldereschi, in *Proc. 23<sup>d</sup> Int. Conf. on the Physics of Semiconductors*, Berlin, Germany (1996), World Scientific (1996), p. 1719.