

ФИЗИКА ПОЛУПРОВОДНИКОВ  
И ПОЛУПРОВОДНИКОВЫХ ПРИБОРОВ

---

Б. И. ШКЛОВСКИЙ, А. Л. ЭФРОС

# ЭЛЕКТРОННЫЕ СВОЙСТВА ЛЕГИРОВАННЫХ ПОЛУПРОВОДНИКОВ



МОСКВА «НАУКА»  
ГЛАВНАЯ РЕДАКЦИЯ  
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ

1979

**Электронные свойства легированных полупроводников.** Шкловский Б. И., Эфрос А. Л. Монография. — М.: Наука, Главная редакция физико-математической литературы, 1979, 416 стр.

В книге рассматриваются физические явления в легированных полупроводниках, для описания которых необходимо учитывать, что электроны находятся в хаотическом поле доноров и акцепторов. К этим явлениям относятся андерсоновская локализация электронов, прыжковая проводимость, переход от металлической проводимости к активационной при изменении степени легирования и компенсации, оптические явления, связанные с хвостами плотности состояний. Современный подход к указанным проблемам в значительной мере основан на теории протекания. В книге дается первый в монографической литературе обзор этой новой математической дисциплины. Детально обсуждается основанный на теории протекания метод вычисления электропроводности сильно неоднородных сред. Систематически излагается теория прыжковой проводимости, построенная с помощью этого метода. Большое внимание уделяется сопоставлению ее результатов с экспериментом. Обсуждаются нерешенные проблемы теории.

Табл. 13, библи. 358, рис. 85.

## ПРЕДИСЛОВИЕ

Полупроводники первого поколения нельзя назвать легированными. Они были просто очень грязными. Неконтролируемые примеси не позволяли выявить физические закономерности, что сбивало исследователей с толку и вызывало насмешки и пессимизм у гордых представителей бурно развивавшихся в то время «чистых» физических дисциплин. Когда удалось преодолеть эту «грязь», началась новая эра в развитии физики полупроводников. Эта эра прошла под девизом «чистота». Она принесла выдающиеся успехи пятидесятых годов, в результате которых возникла новая область техники, получившая название «полупроводниковая электроника». Эксперименты с чистыми кристаллами дали мощный стимул для развития теории полупроводников. В это время были разработаны и апробированы метод эффективной массы для сложных зон, теория примесных состояний, теория кинетических явлений.

По существу, эти достижения и составляют то, что сейчас принято называть физикой полупроводников. Однако в последние пятнадцать лет произошел ощутимый сдвиг в сторону грязных полупроводников. Дело в том, что именно примесные свойства ответственны за работу целого ряда важнейших полупроводниковых приборов, и потому технике нужны полупроводники с примесями. В отличие от полупроводников первого поколения их называют не грязными, а легированными, подчеркивая этим, что технологи в состоянии в определенных пределах регулировать примесный состав.

В теории электронных состояний легированных полупроводников возникли новые задачи. В них рассматривается электрон, находящийся не в периодическом поле атомов кристалла, а в хаотическом поле примесей,

причем потенциальную энергию этого поля нельзя считать малой. При низких температурах легированный кристаллический полупроводник оказывается неупорядоченной системой, напоминающей по своим общим свойствам аморфные системы. Это относится не только к сильно легированным, но и к слабо легированным полупроводникам. Чем слабее легирование, тем ниже температуры, при которых эти свойства проявляются.

Цель данной книги состоит в том, чтобы дать последовательное изложение теории электронных состояний и проводимости легированных полупроводников при низких температурах, т. е. в той области, где свойства электронных состояний сильно отличаются от свойств блоховских волн.

В зависимости от легирования электронные состояния проводника при нулевой температуре могут быть локализованными или делокализованными. Важным достижением теории неупорядоченных систем явилась так называемая теорема Андерсона, утверждающая, что при определенных условиях существуют строго локализованные состояния. С подробного обсуждения этого вопроса (гл. 2) фактически начато изложение свойств электронных состояний. Специфической чертой теории, отличающей ее от теории электронных свойств идеального кристалла, является необходимость учета электрон-электронного взаимодействия даже при очень малых концентрациях электронов. В связи с этим разработана теория нелинейного электронного экранирования (гл. 3, § 13), основанная на методе самосогласованного поля. В непосредственной окрестности уровня Ферми этот метод не работает и плотность состояний имеет интересные особенности (гл. 10).

Если уровень Ферми находится среди локализованных состояний, то электропроводность осуществляется путем электронных прыжков и экспоненциальным образом зависит от температуры и концентрации примесей. Явление прыжковой проводимости известно уже давно, однако в последнее десятилетие в этой области были достигнуты определенные успехи. Была построена теория, дающая хорошее количественное описание температурной, концентрационной и магнитопольевой зависимостей. Основой этих успехов стала молодая математиче-



ская дисциплина, называемая теорией протекания. Для изучения низкотемпературной проводимости метод протекания стал играть такую же роль, как метод кинетического уравнения для зонной проводимости. Термин «уровень протекания» употребляют в соответствующей литературе не реже, чем «время релаксации». Существуют хорошие обзоры по теории протекания, которые мы цитируем в гл. 5. Однако, во-первых, эти обзоры написаны сравнительно давно, а во-вторых, не «в том ключе», как это требуется для описания прыжковой проводимости. Поэтому мы сочли необходимым написать специальную главу (гл. 5), в которой изложены основные результаты теории протекания и дана библиография работ на эту тему.

Во всех главах, касающихся прыжковой проводимости, приведено подробное сопоставление теории с экспериментом. Результаты этого сопоставления кажутся нам в целом хорошими. Мы старались особо отметить те случаи, в которых возникают расхождения, а также теоретические проблемы, представляющиеся нам нерешенными.

Книга посвящена кристаллическим полупроводникам, однако излагаемые в ней идеи и методы столь близки к теории аморфных полупроводников, что «аморфные мотивы» неизбежно вплетаются в изложение. Иногда (см. гл. 9) для подтверждения тех или иных концепций мы пользуемся экспериментальным материалом, полученным на аморфных полупроводниках.

Объем книги не позволил включить в нее такие вопросы, как зависимость прыжковой проводимости от частоты и напряженности электрического поля. Читатель может ознакомиться с ними по обзорам Поллака [1] и Бётгера и Брыксина [2], а также по цитированному в книге обзорам Мотта.

Данная книга задумана не только как монография для специалистов, но и как продолжение обычного курса теории полупроводников, охватывающее новый круг вопросов. Связующими звеньями между этой книгой и каноническими курсами по теории «чистых» полупроводников должны служить глава 1, § 14, а также § 35. Книга рассчитана на широкий круг читателей: физиков теоретиков и экспериментаторов, аспирантов и инженеров, знающих основы физики твердого тела. Ее можно

прочитать в облегченном варианте, опустив §§ 3, 7—9, 19, 20, 28, 33, 34, 37, 40, 46. Важно также иметь в виду, что все основные вопросы, как правило, изложены дважды, сначала на качественном уровне, а затем на количественном. Читатель, которого не интересует математический аппарат, может продолжать чтение книги, ограничившись качественным объяснением. Места, где можно сделать пропуск, обычно указаны в тексте.

Как и большинство ученых, работающих в этой области, мы находимся под влиянием идей профессора Н. Ф. Мотта и отдаем ему этой книгой дань своего уважения. Его замечательные книги и обзоры пробудили наш интерес к теории неупорядоченных систем.

Важную стимулирующую роль в нашей работе сыграло постоянное общение с физиками-экспериментаторами полупроводниковых лабораторий ФТИ имени А. Ф. Иоффе и МГПИ имени В. И. Ленина. Мы благодарны также всем участникам теоретических семинаров ФТИ имени А. Ф. Иоффе и ИФП имени С. И. Вавилова, регулярно обсуждавшим наши работы. Мы благодарим всех советских и зарубежных коллег, работавших с нами и присылавших для обсуждения свои работы перед публикацией. Мы благодарны С. Д. Барановскому, А. Г. Забродскому, М. А. Кривоглазу, Нгуен Ван Лиену, Г. Е. Пикусу и И. И. Фельдману, взявшим на себя труд прочитать рукопись частично или полностью и сделавшим ряд ценных замечаний. С особой любовью мы благодарим своих родных М. М. Марголину, Д. Е. Шкловскую и Н. И. Эфрос за поддержку и самоотверженную помощь при подготовке рукописи к печати.

# ЧАСТЬ I

## СЛАБО ЛЕГИРОВАННЫЕ ПОЛУПРОВОДНИКИ

### ГЛАВА I

#### СТРУКТУРА ОДИНОЧНЫХ ПРИМЕСНЫХ СОСТОЯНИЙ

##### § 1. Мелкие примеси

Целью этой главы, которая носит вводный характер, является краткое изложение теории примесных центров в полупроводниках. Желающим ознакомиться с этим вопросом более детально можно рекомендовать обзор Кона [1], более новый обзор Бассани и др. [2], а также монографию Бира и Пикуса [3].

Как правило, электронные свойства полупроводников при низких температурах определяют примеси. Примеси принято делить на донорные и акцепторные. Донорные примеси обладают свойством сравнительно легко ионизоваться, находясь в кристалле. При этом примесный центр приобретает положительный заряд, а в зоне проводимости появляется электрон, который может участвовать в процессах переноса. С донорами связан электронный тип проводимости полупроводников. Если концентрация примесей не очень велика, то при достаточно низкой температуре электрон возвращается донору и донор становится нейтральным. Это явление иногда называют «вымораживанием электронов зоны проводимости».

Важнейшей характеристикой примеси является энергия связи, т. е. энергия, необходимая для перехода электрона с донорного уровня на дно зоны проводимости. Именно эта энергия в первую очередь определяет температуру вымораживания электронов. На энергетической схеме полупроводника энергии донорных уровней находятся в запрещенной зоне (рис. 1.1).

Донорная примесь называется мелкой, если ее уровень близок к дну зоны проводимости, т. е. энергия связи электрона мала по сравнению с шириной запрещенной зоны. Мелкие примеси играют особую роль в физике полупроводников. Как мы увидим ниже, они обладают особыми свойствами, допускающими универсальное описание. Именно о таких примесях в основном будет идти речь в этой главе и во всей книге.

Акцепторная примесь обладает свойством захватывать электрон, принадлежащий кристаллу. При этом примесный центр заряжается отрицательно, а в валентной зоне появляется дырка. С наличием акцепторов связана дырочная проводимость полупроводников. При низких температурах происходит вымораживание дырок. Оно связано с тем, что дырка локализуется вблизи акцептора. Температура вымораживания определяется главным образом энергией связи дырки. Акцепторные уровни также находятся в запрещенной зоне полупроводника

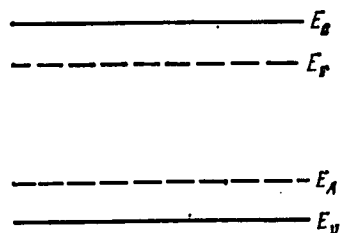


Рис. 1.1. Зонная схема полупроводника:  $E_c$ ,  $E_v$  — края зоны проводимости и валентной зоны,  $E_D$ ,  $E_A$  — энергии донорных и акцепторных уровней.

(см. рис. 1.1). Мелкими акцепторами называются такие акцепторы, уровни которых близки к вершине валентной зоны.

Окажется ли данная примесь донорной или акцепторной, чаще всего показывает просто ее положение в таблице Менделеева. Так, например, для полупроводников IV группы (Ge, Si) примеси, находящиеся в V группе (P, Sb, As), являются, как правило, донорами. Дело в том, что полупроводники четвертой группы обладают тетраэдральной атомной структурой. Химическая связь имеет ковалентную природу и осуществляется четырьмя  $sp^3$ -орбиталями. Каждый атом связан этими орбиталями с четырьмя ближайшими соседями, расположенными в углах тетраэдра. Элемент V группы имеет на один валентный электрон больше, и естественно, что, когда такой атом попадает в тетраэдральную структуру, пятый электрон легко отрывается, т. е. такая примесь является

донором. Элементу III группы, наоборот, не хватает одного электрона, и он легко его захватывает, образуя подвижную дырку. Поэтому элементы III группы (B, Al, Ga, In) оказываются в германии и кремнии акцепторами. Заметим, что характер примеси не всегда определяют только свойства примесного элемента. Например, алюминий в арсениде галлия может оказаться и донором и акцептором, в зависимости от того, замещает ли он атом галлия или атом мышьяка.

Существует ситуация, когда химическая природа примеси очень мало сказывается на структуре примесного состояния и на энергии примесного уровня. Такая ситуация возникает в широко распространенном случае мелких примесей. Будем для определенности говорить о донорах, хотя совершенно аналогичные рассуждения можно привести и для акцепторов. Близость донорного уровня ко дну зоны проводимости говорит о том, что избыточный электрон слабо связан с донорным центром и находится от него в среднем на большом расстоянии. Это означает, что атомная структура примесного центра слабо влияет на состояние избыточного электрона, а удерживает электрон вблизи центра лишь положительный заряд, который можно считать точечным. Поэтому для расчета энергии связи примем, что электрон движется в потенциальном поле

$$U(r) = -e^2/\kappa r, \quad (1.1)$$

где  $e$  — заряд электрона,  $r$  — расстояние от электрона до примесного центра, а  $\kappa$  — диэлектрическая проницаемость решетки, введение которой оправдано, если радиус примесного состояния велик по сравнению с постоянной решетки. В кристаллах низкой симметрии  $\kappa$  является тензором. Заметим, что, поскольку примесный центр неподвижен, в (1.1) следует подставлять статическую диэлектрическую проницаемость, даже если на частоте, соответствующей энергии связи примесного центра, частотная дисперсия  $\kappa(\omega)$  уже заметна (этим теория примесных центров отличается от теории экситонов, в которой положительный заряд также необходимо считать подвижным и потому решеточное экранирование может иметь динамический характер).

После введения потенциала (1.1) мы полностью исключили из описания примесного состояния индивидуальные черты примесного центра, связанные с его химической природой. Поэтому ясно, что в этом приближении уровни всех доноров оказываются одинаковыми. Разумеется, они сильно зависят от свойств полупроводника, в котором находится примесь, не только через диэлектрическую проницаемость  $\kappa$ , но и через закон дисперсии электронов.

Характерный размер области локализации электрона вблизи мелкого примесного центра (радиус состояния) гораздо больше постоянной решетки. Поэтому, если разложить волновую функцию электрона на плоские волны, то в разложении важную роль будут играть волновые векторы, малые по сравнению с вектором обратной решетки. Это означает, что динамика электрона определяется не всем спектром его энергии, а лишь областью с волновыми векторами, близкими к точке, соответствующей дну зоны проводимости. Математическое выражение этих идей составляет содержание метода эффективной массы, который является основой электронной теории полупроводников. Его изложению посвящена настоящая глава.

В § 2 рассмотрен случай невырожденной зоны и подробно описана структура донорных уровней в германии и кремнии; § 3 посвящен вырожденной зоне и структуре акцепторных уровней в типичных полупроводниках. Основной задачей этой главы является вывод асимптотических свойств волновых функций примесных состояний на больших расстояниях от центра. Эти свойства окажутся необходимыми при построении теории прыжковой проводимости. Асимптотике волновых функций мелких примесей посвящен § 4.

## § 2. Примесные уровни вблизи невырожденной зоны

Невырожденной зоной принято называть зону, имеющую вблизи экстремальной точки лишь двукратное вырождение по спину. Именно такой вид имеет зона проводимости типичных полупроводников (Ge, Si, соединения  $A_3B_5$ ).

Гамильтониан электрона в идеальной решетке имеет вид

$$H_0 = -\frac{\hbar^2}{2m_0} \Delta + V(r), \quad (2.1)$$

где  $m_0$  — масса свободного электрона, а  $V(r)$  — периодический, с периодом решетки, потенциал,  $\Delta$  — оператор Лапласа. Решением уравнения Шредингера являются блоховские функции

$$\varphi_{n, k}(r) = \frac{1}{V_0^{1/2}} u_{n, k}(r) \exp(ikr), \quad (2.2)$$

где  $u_{n, k}(r)$  — функция, периодическая с периодом решетки,  $V_0$  — нормировочный объем,  $k$  — волновой вектор,  $n$  включает номер зоны и спиновый индекс. Волновой вектор ограничим первой зоной Бриллюэна. Функция  $E_n(k)$ , представляющая собственные значения гамильтониана (2.1), будет считаться известной внутри этой зоны.

Задача состоит в решении уравнений Шредингера с гамильтонианом

$$H = H_0 + U(r), \quad (2.3)$$

где  $U(r)$  выражается формулой (1.1). Начнем рассмотрение с самого простого случая.

а) Экстремум в центре зоны Бриллюэна. В этом случае минимум функции  $E_n(k)$  достигается при  $k = 0$ , и мы будем считать, что вблизи этой точки функция изотропна, т. е. при малых  $k$

$$E_n(k) = \frac{\hbar^2 k^2}{2m}. \quad (2.4)$$

Величину  $m$  принято называть эффективной массой. Решение уравнения

$$(H_0 + U) \Psi = E \Psi \quad (2.5)$$

будем искать в виде

$$\Psi = \sum_{n', k'} B_{n'}(k') \varphi_{n', k'}(r). \quad (2.6)$$

Подставим (2.6) в (2.5), умножим (2.5) на  $\varphi_{n, k}^*(r)$  и проинтегрируем по  $r$ . Тогда с учетом условия нормировки

блеховских функций получим

$$(E_n(\mathbf{k}) - E) B_n(\mathbf{k}) + \sum_{n', \mathbf{k}'} U_{n', \mathbf{k}'}^{n\mathbf{k}} B_{n'}(\mathbf{k}') = 0, \quad (2.7)$$

где

$$U_{n', \mathbf{k}'}^{n\mathbf{k}} = \frac{1}{V_0} \int u_{n, \mathbf{k}}^* u_{n', \mathbf{k}'} \exp[i(\mathbf{k}' - \mathbf{k}) \mathbf{r}] U(r) dr. \quad (2.8)$$

Основываясь на качественных соображениях, изложенных в предыдущем параграфе, естественно предположить, что функция  $B_n(\mathbf{k})$  отлична от нуля лишь в малой области вблизи центра зоны Бриллюэна. Рассматривая матричный элемент (2.8) при малых  $\mathbf{k}$  и  $\mathbf{k}'$ , можно положить  $u_{n, \mathbf{k}}^* \approx u_{n, 0}^*$  и  $u_{n', \mathbf{k}'} \approx u_{n', 0}$ . Подынтегральное выражение в (2.8) содержит быстро меняющийся множитель  $u_{n, 0}^* u_{n', 0}$ , периодичный с периодом решетки, и плавно меняющийся множитель  $U(r) \exp[i(\mathbf{k}' - \mathbf{k}) \mathbf{r}]$ . Это позволяет записать интеграл в виде произведения интегралов:

$$U_{n', \mathbf{k}'}^{n\mathbf{k}} = \int_{\pi} u_{n, 0}^* u_{n', 0} dr \frac{1}{V_0} \int \exp[i(\mathbf{k}' - \mathbf{k}) \mathbf{r}] U(r) dr, \quad (2.9)$$

причем в первом интеграле интегрирование ведется по объему элементарной ячейки. Из условия нормировки и ортогональности блеховских функций при  $\mathbf{k} = 0$  следует, что

$$\int_{\pi} u_{n, 0}^* u_{n', 0} dr = \delta_{nn'}. \quad (2.10)$$

Таким образом, возникает замкнутое уравнение для функции  $B_n(\mathbf{k})$ :

$$(E_n(\mathbf{k}) - E) B_n(\mathbf{k}) + \sum_{\mathbf{k}'} U(\mathbf{k}\mathbf{k}') B_n(\mathbf{k}') = 0, \quad (2.11)$$

где

$$U(\mathbf{k}\mathbf{k}') = \frac{1}{V_0} \int \exp[i(\mathbf{k}' - \mathbf{k}) \mathbf{r}] U(r) dr = \\ = - \frac{4\pi e^2}{\kappa V_0 (\mathbf{k} - \mathbf{k}')^2}. \quad (2.12)$$

Подставляя (2.4) и (2.12) в (2.11), получим

$$\left( \frac{\hbar^2 \mathbf{k}^2}{2m} - E \right) B_n(\mathbf{k}) - \frac{4\pi e^2}{\kappa V_0} \sum_{\mathbf{k}'} \frac{1}{(\mathbf{k} - \mathbf{k}')^2} B_n(\mathbf{k}') = 0. \quad (2.13)$$



Строго говоря, суммирование в (2.13) должно происходить по векторам  $\mathbf{k}'$ , лежащим внутри первой зоны Бриллюэна. Фактически, однако, функция  $B_n(\mathbf{k})$ , удовлетворяющая (2.13), спадает в малой окрестности точки  $\mathbf{k} = 0$ . Поэтому с хорошей точностью можно считать, что суммирование в (2.13) распространяется до бесконечности. Это позволяет перейти к координатной волновой функции

$$F(\mathbf{r}) = \frac{1}{V_0^{1/2}} \sum_{\mathbf{k}} B_n(\mathbf{k}) \exp(i\mathbf{k}\mathbf{r}), \quad (2.14)$$

где суммирование происходит по всем  $\mathbf{k}$ . Умножая (2.13) на  $\exp(i\mathbf{k}\mathbf{r})$  и суммируя по всем  $\mathbf{k}$ , получим

$$\left(-\frac{\hbar^2}{2m} \Delta - \frac{e^2}{\kappa r}\right) F(\mathbf{r}) = EF(\mathbf{r}). \quad (2.15)$$

Уравнение (2.15) совпадает с уравнением Шредингера для атома водорода, в которое, однако, входят эффективная масса  $m$  и диэлектрическая проницаемость  $\kappa$ . Спектр связанных состояний определяется обычной формулой

$$E_p = -\frac{1}{t^2} \frac{e^4 m}{2\kappa^2 \hbar^2}, \quad t = 1, 2, \dots, \quad (2.16)$$

а волновая функция основного состояния имеет вид

$$F(\mathbf{r}) = \frac{1}{(\pi a^3)^{1/2}} \exp\left(-\frac{r}{a}\right), \quad (2.17)$$

где

$$a = \hbar^2 \kappa / m e^2 \quad (2.18)$$

есть эффективный борковский радиус, определяющий характерные размеры волновой функции\*).

Благодаря большим значениям диэлектрической проницаемости и малым значениям эффективной массы борковские радиусы оказываются очень большими. Например, в GaAs ( $\kappa = 12,6$ ,  $m = 0,066m_0$ )  $a = 100 \text{ \AA}$ . Это обстоятельство и позволяет применять метод эффективной

---

\*) Говоря о локализованных состояниях, мы будем использовать термины «характерный размер волновой функции» или «радиус состояния» для обозначения характерного размера области пространства, внутри которой волновая функция не мала экспоненциально. Если волновая функция спадает по закону (2.17), то указанные термины относятся к величине  $a$ .

массы. Определив из (2.14) и (2.17)  $B_n(\mathbf{k})$  для основного состояния, мы получим

$$B_n(\mathbf{k}) = \frac{8\pi^{1/2}}{V_0^{1/2} a^{5/2}} \frac{1}{(k^2 + a^{-2})^2}, \quad (2.19)$$

откуда видно, что функция  $B_n(\mathbf{k})$  действительно быстро убывает при  $k > a^{-1}$ . Для типичных полупроводников  $a^{-1}$  составляет сотые доли от модуля вектора обратной решетки, что и оправдывает сделанные выше приближения.

Полная волновая функция примесного центра определяется формулой (2.6), в которой можно положить  $B_{n'} = \delta_{nn'} B_n$ . Тогда

$$\begin{aligned} \Psi(\mathbf{r}) &= \sum_{\mathbf{k}} B_n(\mathbf{k}) \varphi_{n,\mathbf{k}}(\mathbf{r}) = \\ &= \frac{1}{V_0^{1/2}} \sum_{\mathbf{k}} B_n(\mathbf{k}) u_{n,\mathbf{k}} \exp(i\mathbf{k}\mathbf{r}). \end{aligned} \quad (2.20)$$

Пренебрегая, как и выше, зависимостью  $u_{n,\mathbf{k}}$  от  $\mathbf{k}$ , получим

$$\Psi(\mathbf{r}) = \frac{1}{V_0^{1/2}} u_{n,0}(\mathbf{r}) \sum_{\mathbf{k}} B_n(\mathbf{k}) \exp(i\mathbf{k}\mathbf{r}) = u_{n,0}(\mathbf{r}) F(\mathbf{r}). \quad (2.21)$$

Таким образом, волновая функция примесного состояния представляет собой блоховскую волновую функцию на дне зоны, промодулированную крупномасштабной водородоподобной функцией.

В заключение этого пункта приведем табл. 1.1, заимствованную из обзора [2], в которой сравниваются теоретические и экспериментальные значения донорных уровней. Некоторое различие между ними связано с тем, что метод эффективной массы плохо описывает потенциал, действующий на малых расстояниях от центра.

**б) Несколько эквивалентных экстремумов.** Зона проводимости таких полупроводников, как германий и кремний, имеет несколько экстремумов, расположенных в эквивалентных точках зоны Бриллюэна. Минимум зоны проводимости кремния расположен в направлении [100] на расстоянии  $k_z^0 = 0,85 (2\pi/a_0)$ , где  $a_0$  — постоянная решетки. Поверхности постоянной энергии представляют

Таблица 1.1

Диэлектрические проницаемости, эффективные массы, теоретические и экспериментальные значения энергии основного состояния для некоторых полупроводников

Вещество	$\kappa$	$m/m_0$	$E_{1s}, \text{ мэВ}$ теор.	$E_{1s}, \text{ мэВ}$ эксп.
GaAs	12,5	0,066	5,67	Ge: 6,1; Si: 5,8 Se: 5,9; S: 6,1 Sn: 5,9
InP	12,6	0,08	6,8	7,28
GdTe	10	0,1	13	13,8

собой эллипсоиды вращения вокруг направления [100]. Кубическая симметрия кристалла требует, чтобы такие же эллипсоиды существовали в каждом эквивалентном направлении. Таким образом, в кремнии имеется шесть эквивалентных минимумов. Выбирая ось  $z$  вдоль [100], получим следующее выражение для энергии вблизи экстремума:

$$E = \frac{\hbar^2}{2m_l} (k_z - k_z^0)^2 + \frac{\hbar^2}{2m_t} (k_x^2 + k_y^2). \quad (2.22)$$

Величины продольной  $m_l$  и поперечной  $m_t$  эффективных масс в кремнии следующие [2]:  $m_l = 0,916m_0$  и  $m_t = 0,190m_0$ .

Экстремумы зоны проводимости германия расположены точно на границах зоны Бриллюэна в направлениях телесных диагоналей куба. Каждый экстремум по этой причине принадлежит двум зонам, и число экстремумов вдвое меньше числа эквивалентных направлений, т. е. равно четырем. Поверхности постоянной энергии имеют вид эллипсоидов с осями вращения, направленными вдоль телесных диагоналей. Значения эффективных масс следующие:  $m_l = 1,58m_0$  и  $m_t = 0,08m_0$  [2].

Обобщение метода эффективной массы на случай нескольких экстремумов не представляет труда. Решение уравнения (2.5) вблизи экстремума, находящегося в точке  $\mathbf{k} = \mathbf{k}_j$ , в полной аналогии с (2.21) имеет вид

$$\chi_j(\mathbf{r}) = F_j(\mathbf{r}) \varphi_j(\mathbf{r}), \quad (2.23)$$

где  $\varphi_j$  — блоховская функция в точке  $k = k_j$ , нормированная на элементарной ячейке, а  $F_j$  — плавная огибающая функция, удовлетворяющая уравнению

$$\left(E_j(-i\nabla) - \frac{e^2}{\kappa r}\right) F_j(r) = E F_j(r), \quad (2.24)$$

где  $E_j(k)$  — энергия зоны проводимости вблизи точки  $k_j$ . Разумеется, собственные значения энергии  $E$  одинаковы для всех эквивалентных экстремумов. По этой причине в рамках метода эффективной массы примесные состояния оказываются дополнительно  $N$ -кратно вырожденными ( $N$  — число эквивалентных экстремумов). В германии и кремнии это вырождение частично снимается за счет поправок к методу эффективной массы. Эти поправки связаны с тем, что на малых расстояниях потенциал не описывается формулой (1.1), так как введение макроскопической диэлектрической проницаемости становится здесь бессмысленным. В результате следует вводить дополнительный к (1.1) короткодействующий потенциал. Матричные элементы этого потенциала, взятые на плоских волнах, не убывают с увеличением  $k$  подобно (2.12), а остаются при больших  $k$  постоянными. Поэтому они связывают волновые функции, принадлежащие разным экстремумам. Это приводит к расщеплению примесных состояний, вырожденных в рамках метода эффективной массы, и к сдвигу уровней, который принято называть химическим сдвигом. Характер расщепления можно предсказать, не вдаваясь в детальную структуру короткодействующего потенциала, а основываясь лишь на свойствах симметрии.

Мы обсудим лишь расщепление основного состояния. Правильные волновые функции являются суперпозицией:

$$\Psi^{(i)} = \sum_{j=1}^N \alpha_j^{(i)} \chi_j, \quad i = 1, \dots, N. \quad (2.25)$$

Характер расщепления и коэффициенты  $\alpha_j^{(i)}$  определяются для германия и кремния из свойств симметрии по отношению к преобразованиям группы тетраэдра  $T_d$ . В результате оказывается, что в кремнии шестикратно вырожденное основное состояние расщепляется на син-

глетное, двукратно вырожденное и трехкратно вырожденное. Наименьшая энергия — у синглетного состояния. Его волновая функция имеет вид

$$\Psi^{A_1} = \sum_{j=1}^6 \frac{1}{\sqrt{6}} \chi_j. \quad (2.26)$$

В германии четырехкратно вырожденное основное состояние расщепляется на синглетное и трехкратно вырожденное. Основное состояние — синглетное, его волновая функция имеет вид

$$\Psi^{A_1} = \sum_{j=1}^4 \frac{1}{2} \chi_j. \quad (2.27)$$

В рамках метода эффективной массы определение энергии примесного состояния сводится к решению уравнения (2.24). Подставляя в это уравнение эллипсоидальный спектр (2.22), получим

$$\left[ -\frac{\hbar^2}{2m_l} \frac{\partial^2}{\partial z^2} - \frac{\hbar^2}{2m_l} \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right) - \frac{e^2}{\kappa r} - E \right] F = 0. \quad (2.28)$$

Это уравнение зависит от одного безразмерного параметра, в качестве которого выбирают обычно отношение поперечной массы к продольной:  $\gamma = m_t/m_l$ .

Перепишем (2.28) в атомных единицах, соответствующих поперечной массе,  $r' = r/a_t$ ,  $E' = E/E_t$ ,  $a_t = \hbar^2 \kappa / e^2 m_l$ ,  $E_t = \hbar^2 / 2m_l a_t^2$ . При этом получим

$$\left[ \gamma \frac{\partial^2}{\partial z'^2} + \frac{\partial^2}{\partial x'^2} + \frac{\partial^2}{\partial y'^2} + 2 \left( \frac{1}{r'} + E' \right) \right] F = 0. \quad (2.29)$$

Это уравнение решается точно в двух предельных случаях:  $\gamma = 1$  и  $\gamma = 0$ . Первый случай соответствует обычной водородной задаче, а второй — двумерной кулоновской задаче. Первоочередной вопрос состоит, естественно, в том, чтобы найти энергию основного состояния уравнения (2.29) как функцию параметра  $\gamma$ . Эта задача была выполнена Коном и Латтинджером [4] с помощью вариационного метода, который оказался в этом случае

весьма удачным. Пробная функция была выбрана в виде

$$F = \frac{1}{(\pi a_{\perp}^2 a_{\parallel})^{1/2}} \exp \left[ - \left( \frac{x^2 + y^2}{a_{\perp}^2} + \frac{z^2}{a_{\parallel}^2} \right)^{1/2} \right], \quad (2.30)$$

а два параметра  $a_{\perp}$  и  $a_{\parallel}$  определены из условия минимума энергии. При  $\gamma = 1$  оказалось, что  $a_{\perp} = a_{\parallel}$ , пробная функция совпала с точной функцией основного состояния атома водорода и вариационная процедура дала в этом пределе точный результат. Менее тривиальный факт состоит в том, что при  $\gamma = 0$  результат вариационного расчета отличается от точного значения энергии всего на 7—8%.

Зависимость  $E'(\gamma)$ , полученная вариационным методом, показана на рис. 1.2. Используя приведенные выше

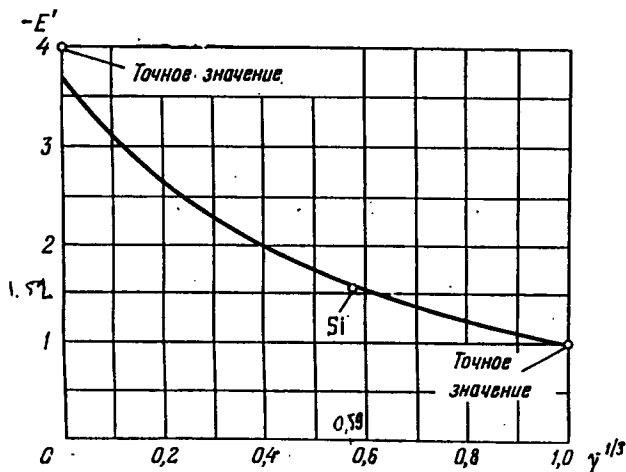


Рис. 1.2. Энергия основного состояния в единицах  $E_t$  как функция  $\gamma^{1/3}$ , вычисленная с пробной функцией (2.30) [4].

значения эффективных масс, а также значения диэлектрических проницаемостей: 11, 40 для кремния и 15,36 для германия, получим, что в рамках метода эффективной массы энергия основного состояния кремния равна 31,27 мэв, а германия — 9,81 мэв [2].

Аналогичным образом вычислены также энергии возбужденных состояний. Уравнение (2.29) имеет аксиальную симметрию, и потому квантовыми числами являются

магнитное квантовое число  $m$  и четность. Уровни, отличающиеся знаком  $m$ , без учета спина двукратно вырождены. Возбужденные состояния уравнения (2.29) принято классифицировать так же, как те водородные состояния, в которые они переходят при  $\gamma = 1$ . Однако при произвольных  $\gamma$  возникает дополнительное расщепление уровней. Например, вместо состояния  $2p$  атома водорода возникают состояние  $2p_0$  с  $m = 0$  и два вырожденных состояния  $2p_{\pm}$  с  $m = \pm 1$ .

Таблица 1.2  
Энергия связи донорных состояний в Ge и Si \*)

Вещество	$E_{1s}, \text{ эВ}$	$E_{2p_0}, \text{ эВ}$
Si (теория)	31,27	11,51
Si (P)	45,5; 33,9; 32,6	11,45
Si (As)	53,7; 32,6; 31,2	11,49
Si (Sb)	42,7; 32,9; 30,6	11,52
Ge (теория)	9,81	4,74
Ge (P)	12,9; 9,9	4,75
Ge (As)	14,17; 10,0	4,75
Ge (Sb)	10,32; 10,0	4,74

\*) Теоретические значения получены в рамках метода эффективной массы. Экспериментальные значения даны для примесей, химический символ которых указан в скобках.

В заключение этого параграфа мы приведем табл. 1.2, заимствованную из обзора [2], в которой даны теоретические и экспериментальные значения основных и возбужденных состояний в германии и кремнии.

Как видно из этой таблицы, метод эффективной массы далеко не блестяще описывает структуру основного состояния. Главная ошибка состоит в неправильном учете короткодействующей части потенциала. Этот потенциал создаст заметное увеличение энергии связи основного состояния и, как уже говорилось, расщепляет его. Значительно лучше метод эффективной массы описывает  $p$ -состояния. Это связано с тем, что волновые функции этих состояний обращаются в нуль вблизи примесного центра и потому мало чувствительны к виду потенциала на малых расстояниях.

### § 3. Примесные уровни вблизи точки вырождения зон

Допустим, что в некоторой точке  $k$ -пространства состояния, соответствующие гамильтониану (2.1), вырождены, т. е. существует  $t$  блоховских функций  $\varphi_{n, k}^j$ , удовлетворяющих уравнению

$$H_0 \varphi_{n, k}^j = E_n(k) \varphi_{n, k}^j, \quad j = 1, 2, \dots, t. \quad (3.1)$$

В дальнейшем мы будем рассматривать случай, когда вырожденной точкой является центр зоны Бриллюэна,

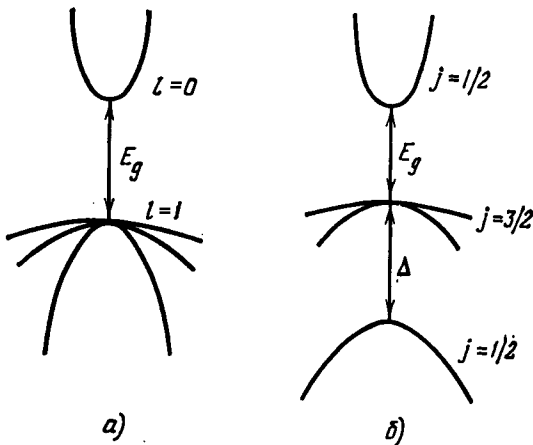


Рис. 1.3. Зонная схема типичных полупроводников: а) спин-орбитальное взаимодействие отсутствует, валентная зона шестикратно вырождена в точке  $k = 0$ ; б) спин-орбитальное взаимодействие отщепляет зону с полным моментом  $1/2$ .

т. е. точка  $k = 0$ . Именно так устроена валентная зона типичных полупроводников. Для простоты ее можно представить как результат расщепления атомного  $p$ -состояния. В точке  $k = 0$  зона вырождена с учетом спина шестикратно. При конечных значениях  $k$  это вырождение частично снимается (рис. 1.3).

Метод эффективной массы легко обобщается и на этот случай [5]. Решение уравнения Шредингера естественно искать в виде суперпозиции блоховских функций,



соответствующих точке вырождения:

$$\Psi = \sum_{j=1}^t F_j(\mathbf{r}) \varphi'_{n,0}(\mathbf{r}). \quad (3.2)$$

Коэффициенты  $F_j(\mathbf{r})$  оказываются плавными функциями координат, удовлетворяющими уравнению

$$\sum_{j=1}^t \left[ \sum_{\alpha, \beta=1}^3 H_{jj'}^{\alpha\beta} \hat{p}_\alpha \hat{p}_\beta + U(\mathbf{r}) \delta_{jj'} \right] F_{j'} = E F_j, \quad (3.3)$$

$$\hat{\mathbf{p}} = -i\hbar\nabla,$$

которое аналогично уравнению (2.15). Однако в этом случае огибающая функция представляет собой столбец, состоящий из  $t$  функций, а  $H_{jj'}^{\alpha\beta}$  является не только матрицей по отношению к декартовым индексам  $\alpha, \beta$ , но и матрицей, связывающей различные компоненты волновых функций  $F_j$ .

Такой результат представляется совершенно естественным обобщением метода эффективной массы, изложенного в предыдущем параграфе. Читателю, желающему в этом убедиться, мы рекомендуем оригинальную работу Латтинджера и Кона [5], которая содержит подробный вывод этого уравнения. В этой работе матрица  $H$  выражена через матричные элементы импульса, взятые на блоховских функциях. Возможен, однако, и другой подход к проблеме, состоящий в том, чтобы написать матрицу  $H$  просто из соображений симметрии. Этот путь избран для нашего изложения.

В предыдущем рассуждении мы не принимали во внимание эффекты, связанные со спин-орбитальным взаимодействием. Между тем к гамильтониану электрона в идеальной решетке  $H_0$  (см. (2.1)) необходимо добавить член

$$H_{c.o.} = \frac{1}{4m_0^2 c^2} [\sigma \nabla V] \hat{\mathbf{p}}, \quad (3.4)$$

который представляет собой релятивистскую поправку, следующую из уравнения Дирака. Здесь  $\sigma$  — оператор спина (матрицы Паули),  $V(\mathbf{r})$  — потенциальная энергия электрона в периодическом поле, фигурирующая в (2.1),  $c$  — скорость света.

Спин-орбитальное взаимодействие серьезным образом влияет на структуру валентной зоны. При наличии спин-орбитального взаимодействия спиновый и орбитальный моменты количества движения не являются более сохраняющимися величинами, но сохраняется полный момент количества движения. Рассмотрим шестикратно вырожденное  $p$ -состояние атома. Ему соответствуют три координатные волновые функции. Каждая из этих функций умножается на двухкомпонентный спинор, и возникает шесть вырожденных состояний. Спин-орбитальное взаимодействие разбивает эти шесть состояний на две группы, каждой из которых соответствует своя энергия. Одна группа содержит четыре функции и соответствует полному моменту, равному  $3/2$ , а вторая содержит две функции и соответствует моменту  $1/2$ .

Такое же расщепление имеет место и в кристалле. При учете спин-орбитального взаимодействия валентная зона, шестикратно вырожденная в точке  $k = 0$ , расщепляется на четырехкратно и двукратно вырожденные зоны (см. рис. 1.3, б). При этом двукратно вырожденная зона опускается вниз на величину энергии спин-орбитального взаимодействия, которую принято обозначать буквой  $\Delta$ . (Значения  $\Delta$  для Ge и Si даны в табл. 1.3.)

Общий вид матрицы  $H$  при наличии спин-орбитального взаимодействия весьма сложен, и мы не будем его приводить. Все значительно упрощается в двух предельных случаях. Первый случай соответствует слабому спин-орбитальному взаимодействию, когда спиновые и координатные переменные не связаны и матрица  $H^{\alpha\beta}\hat{p}_\alpha\hat{p}_\beta$  фактически является матрицей  $3 \times 3$ . Второй случай соответствует сильному спин-орбитальному взаимодействию, когда зона с полным моментом  $1/2$  находится по энергии далеко и в окрестности верха валентной зоны ее можно не принимать во внимание. В этом случае в формуле (3.2) следует учитывать только четыре верхние зоны, а матрица  $H^{\alpha\beta}\hat{p}_\alpha\hat{p}_\beta$  должна быть матрицей  $4 \times 4$ .

Если мы интересуемся структурой примесного состояния, то первое приближение можно использовать при условии, что энергия спин-орбитального расщепления  $\Delta$  мала по сравнению с энергией примесного уровня, а второе приближение — в обратном предельном случае. Как видно из приведенных ниже данных, приближение силь-

ного спин-орбитального взаимодействия справедливо для германия и целого ряда других полупроводников. С этого приближения мы и начнем.

Итак, необходимо построить матрицу  $H^{\alpha\beta}\beta_\alpha\beta_\beta$  в окрестности четырехкратно вырожденной зоны. При этом гамильтониан должен удовлетворять требованиям симметрии кристалла относительно поворотов системы координат. Волновая функция представляет собой столбец из четырех координатных функций. С точки зрения квантовой механики это означает, что следует написать уравнение для частицы со спином  $3/2$ . Для этой цели надо ввести псевдовектор  $\mathcal{J}$ , который является оператором спинового момента. Его компоненты представляют собой матрицы четвертого ранга, эквивалентные по смыслу матрицам Паули для частицы со спином  $1/2$ . Явный вид матриц  $\mathcal{J}$  читатель найдет в книге Бира и Пикуса [3, стр. 335].

Если потребовать инвариантности гамильтониана по отношению к любым поворотам координатных осей, то найдется лишь два квадратичных по  $\hat{p}$  инварианта:  $\hat{p}^2$  и  $(\hat{p}\mathcal{J})^2$ . При этом матрица  $H$  зависит лишь от двух параметров и может быть представлена в виде

$$H = \frac{1}{m_0} \left[ \frac{\hat{p}^2}{2} \left( \gamma_1 + \frac{5}{2} \gamma \right) - \gamma (\hat{p}\mathcal{J})^2 \right]. \quad (3.5)$$

На самом деле кубическая симметрия не требует инвариантности по отношению к любым поворотам и в кубическом кристалле существует еще один инвариант  $\beta_x^2 \mathcal{J}_x^2 + \beta_y^2 \mathcal{J}_y^2 + \beta_z^2 \mathcal{J}_z^2$ . Поэтому полный гамильтониан может быть представлен в виде

$$H = \frac{1}{m_0} \left\{ \left( \gamma_1 + \frac{5}{2} \gamma_2 \right) \frac{\hat{p}^2}{2} - \gamma_3 (\hat{p}\mathcal{J})^2 + (\gamma_3 - \gamma_2) (\beta_x^2 \mathcal{J}_x^2 + \beta_y^2 \mathcal{J}_y^2 + \beta_z^2 \mathcal{J}_z^2) \right\}. \quad (3.6)$$

Этот гамильтониан был получен Латтинджером [6] и носит его имя. Параметры  $\gamma_1$ ,  $\gamma_2$ ,  $\gamma_3$  хорошо известны для многих полупроводников. Для Ge и Si они приведены в табл. 1.3.

Необходимо заметить, что выражение (3.6) для гамильтониана является полным, только если кристаллическая решетка обладает центром инверсии. В кристаллах

Таблица 1.3

Параметры гамильтониана Латтинджера, энергия спин-орбитального расщепления  $\Delta$  (эВ), диэлектрическая проницаемость  $\kappa$  для Si и Ge [2]

Вещество	$\gamma_1$	$\gamma_2$	$\gamma_3$	$\Delta$	$\kappa$
Si	4,22	0,39	1,44	0,04	11,4
Ge	13,35	4,25	5,69	0,29	15,4

без центра инверсии в гамильтониан могут входить не только квадратичные, но и линейные по  $\beta$  члены. Например, в кристаллах со структурой цинковой обманки (InSb, GaAs и т. д.) в гамильтониан может входить следующая величина:

$$\beta_x \{ \mathcal{G}_x, (\mathcal{G}_y^2 - \mathcal{G}_z^2) \} + \beta_y \{ \mathcal{G}_y, (\mathcal{G}_z^2 - \mathcal{G}_x^2) \} + \beta_z \{ \mathcal{G}_z, (\mathcal{G}_x^2 - \mathcal{G}_y^2) \}, \quad (3.7)$$

где фигурные скобки означают антикоммутатор:  $\{a, b\} = ab + ba$ .

Однако в хорошо изученных полупроводниках линейные члены очень малы и, как правило, ими можно пренебречь. По этой причине мы не будем их рассматривать.

Итак, получена система (3.3) четырех уравнений второго порядка, которую следует решить, чтобы найти энергию и волновые функции примесного центра. Прежде чем перейти к этой программе, решим систему уравнений (3.3) для свободных частиц, т. е. положим  $U(r) = 0$ . Гамильтониан (3.6) описывает движение частицы со спином  $3/2$ . Его первый член является оператором кинетической энергии, а остальные члены описывают своеобразное спин-орбитальное взаимодействие. Поскольку гамильтониан не зависит явно от координат, решение уравнения Шредингера следует искать в виде

$$F_j = A_j \left( \frac{\hbar}{|\mathbf{k}|} \right) \exp(i\mathbf{k}r), \quad (3.8)$$

где величины  $A_j$  — проекции четырехмерного собственного вектора матрицы  $H_{ij}$ , которая получена из (3.6)

заменой оператора  $\hat{p}$  на  $\hbar\mathbf{k}$ . По определению

$$\sum_j H_{ij} A_j \left( \frac{\hbar}{|\mathbf{k}|} \right) = E A_i \left( \frac{\hbar}{|\mathbf{k}|} \right). \quad (3.9)$$

Рассмотрим сначала гамильтониан (3.5). Поскольку он является скаляром, его собственные значения не зависят от направления вектора  $\mathbf{k}$  относительно кристаллографических осей. Воспользуемся тем, что матрица  $\mathcal{H}_z^2$  диагональна, и направим  $\mathbf{k}$  вдоль оси  $z$ . При этом система (3.9) распадается на четыре независимых уравнения с двумя различными собственными значениями энергии \*):

$$E_l = \frac{\gamma_l + 2\gamma}{2m_0} \hbar^2 k^2, \quad E_h = \frac{\gamma_l - 2\gamma}{2m_0} \hbar^2 k^2. \quad (3.10)$$

Если  $(\gamma_l + 2\gamma) > 0$  и  $(\gamma_l - 2\gamma) > 0$ , то полученные ветви спектра называются ветвями легких и тяжелых дырок. Массы дырок  $m_l$  и  $m_h$  связаны с параметрами  $\gamma_l$  и  $\gamma$  следующим образом:

$$m_l = \frac{m_0}{\gamma_l + 2\gamma}, \quad m_h = \frac{m_0}{\gamma_l - 2\gamma}. \quad (3.11)$$

(мы считаем  $\gamma > 0$ ). Валентная зона с положительными  $m_l$  и  $m_h$  изображена на рис. 1.3, б. При иных знаках  $\gamma_l$  и  $\gamma$  возможен случай, когда энергии  $E_l$  и  $E_h$ , определяемые (3.10), имеют разные знаки. Именно такая ситуация реализуется в полупроводниках с нулевой запрещенной зоной (серое олово, HgTe, HgSe). В этом случае одна из ветвей (3.10) соответствует пустой зоне проводимости, а вторая — заполненной валентной зоне. Обе зоны сходятся в точке  $\mathbf{k} = 0$ , т. е. запрещенная зона отсутствует. Теория примесных состояний в полупроводниках с нулевой запрещенной зоной весьма специфична, и мы не будем ее излагать, отсылая читателей к специальной литературе [7].

Рассмотрим теперь свойства гамильтониана (3.6). Его собственные значения зависят от направления вектора  $\mathbf{k}$  относительно кристаллографических осей. Спектр энергий сравнительно легко получить, используя уравнения (3.1) и явный вид матриц  $\mathcal{H}$ . Приведем лишь результат.

\*) В этом параграфе энергия отсчитывается вниз, в глубь валентной зоны.

Спектр имеет вид

$$E_{l, h} = \frac{\hbar^2}{2m_0} \{ \gamma_1 k^2 \pm [4\gamma_2^2 k^4 + 12(\gamma_3^2 - \gamma_2^2)(k_x^2 k_y^2 + k_y^2 k_z^2 + k_x^2 k_z^2)]^{1/2} \}. \quad (3.12)$$

Он содержит две двукратно вырожденные ветви, которые также принято называть ветвями легких и тяжелых дырок. Знак плюс в (3.12) соответствует легким, а знак минус — тяжелым дыркам. Поверхности постоянной энергии, описываемые уравнением (3.12), оказываются не сферически симметричными, а слегка гофрированными. В целом ряде полупроводников, в том числе и в германии, анизотропия изоэнергетических поверхностей оказывается довольно слабой. Поэтому достаточно хорошее приближение можно получить, ограничившись гамильтонианом (3.5). Это приближение иногда называют сферической моделью. Можно показать, что гамильтониан (3.5) аппроксимирует гамильтониан (3.6) наилучшим образом, если положить

$$\gamma = \frac{1}{5} (3\gamma_3 + 2\gamma_2). \quad (3.13)$$

Пользуясь (3.11), (3.13) и табл. 1.3, получим, что в германии  $m_l = 0,042m_0$ ,  $m_h = 0,32m_0$ .

Вернемся теперь к теории примесных состояний. Первоначально вычисление энергии связи акцепторных уровней проводилось вариационным методом с пробными функциями, определяемыми из соображений симметрии. При этом использовался точный гамильтониан (3.6) (см. обзор [1]). В последнее десятилетие значительные успехи достигнуты в рамках упомянутой выше сферической модели, т. е. с помощью гамильтониана (3.5) (см. обзор [2]). Как уже говорилось, для целого ряда полупроводников пренебрежение гофрированностью изоэнергетических поверхностей является хорошим приближением. В то же время оно позволяет значительно упростить вариационные расчеты и получить ряд точных результатов.

Дело в том, что сферически симметричный гамильтониан (3.5) коммутирует с оператором полного момента количества движения  $j$  (гамильтониан (3.6) не сохраняет

$j$ ). Оператор  $j$  можно представить в виде суммы,  $j = \mathcal{J} + L$ , спинного момента  $\mathcal{J}$  и орбитального момента  $L = \hbar^{-1}[r\hat{p}]$ , каждый из которых не коммутирует с гамильтонианом. Необходимо напомнить, что термин «спин» употребляется здесь лишь для того, чтобы подчеркнуть указанную выше математическую аналогию. На самом деле этот «спин» связан с орбитальным расщеплением атомных уровней, образующих валентную зону. Орбитальный же момент  $L$  характеризует лишь «макроскопическое» орбитальное движение, описываемое методом эффективной массы. Однако формальная аналогия позволяет использовать квантовомеханические правила сложения моментов и получить изложенные ниже результаты.

Как уже говорилось, спин частицы равен в нашем случае  $3/2$ . Отсюда следует, что все возможные собственные значения полного момента  $j$  суть полуцелые числа, не меньшие чем  $1/2$ . Из правила сложения моментов также следует, что при заданном значении  $j \geq 3/2$  орбитальное квантовое число  $l$  может принимать четыре значения, которые представляют собой целые числа, заключенные в интервале  $(j-3/2) \leq l \leq (j+3/2)$ . При  $j = 1/2$  возможны только два значения  $l$ , а именно 1 или 2. Уровни энергии должны быть вырожденными по проекции  $M$  вектора  $j$  на ось квантования.

Волновую функцию следует искать в виде [8]

$$F_{jM}(r, \theta, \varphi) = \sqrt{2j+1} \sum_l (-1)^{l-3/2+M} R_{j,l}(r) \times \\ \times \sum_{m,\mu} \begin{pmatrix} l & 3/2 & j \\ m & \mu & -M \end{pmatrix} Y_{lm}(\theta, \varphi) \chi_\mu. \quad (3.14)$$

Здесь  $Y_{lm}$  — шаровые функции;  $\chi_\mu$  — собственный вектор матрицы  $\mathcal{J}_z$ ;  $\mu$  — собственное значение, соответствующее этому вектору ( $\mathcal{J}_z \chi_\mu = \mu \chi_\mu$ ),  $\mu$  принимает значения  $-3/2, -1/2, 1/2, 3/2$ ;  $\begin{pmatrix} l & 3/2 & j \\ M & \mu & -M \end{pmatrix}$  —  $3j$ -символ Вигнера,  $M = m + \mu$ . Пределы суммирования по  $l$  при данном  $j$  рассмотрены выше. Теперь следует подставить (3.14) в матричное уравнение Шредингера

$$[H + U(r)] F_{jM} = E_j F_{jM}, \quad (3.15)$$

используя в качестве  $H$  выражение (3.5), а в качестве  $U$  выражение (1.1). Если  $j > 1/2$ , то возникает система обыкновенных дифференциальных уравнений, связывающая четыре радиальные функции  $R_{l, i}$ , соответствующие

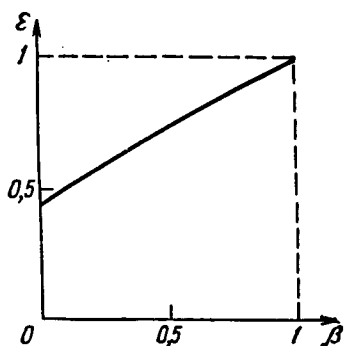


Рис. 1.4. Зависимость энергии основного состояния акцепторного уровня, отнесенной к боровской энергии тяжелой дырки, от отношения масс легкой и тяжелой дырок [9].

заданному значению  $j$ . Оказывается, что эта система распадается на две системы из двух уравнений, в одну из которых входят функции с  $l = j + 1/2$  и  $l = j - 3/2$ , а в другую — функции с  $l = j - 1/2$  и  $l = j + 3/2$ . Это важное свойство является следствием закона сохранения четности. Связанными оказываются лишь такие радиальные функции, которым соответствуют шаровые функции с одинаковой четностью. Основное состояние соответствует такому значению  $j$ , при котором в сумме (3.14) имеется член с  $l = 0$ .

В соответствии с изложенными выше правилами это означает, что основное состояние имеет полный момент  $j = 3/2$ . Ему соответствуют радиальные функции  $R_0 = R_{3/2, 0}$  и  $R_2 = R_{3/2, 2}$ , которые связаны уравнениями

$$(1 + \beta) \left( \frac{d}{dx} - \frac{1}{x} \right) \left( \frac{d}{dx} + \frac{3}{x} \right) R_2 + (1 - \beta) \left( \frac{d}{dx} - \frac{1}{x} \right) \frac{dR_0}{dx} + 2\beta \left( -\varepsilon + \frac{2}{x} \right) R_2 = 0, \quad (3.16)$$

$$(1 + \beta) \left( \frac{d}{dx} + \frac{2}{x} \right) \frac{dR_0}{dx} + (1 - \beta) \left( \frac{d}{dx} + \frac{2}{x} \right) \left( \frac{d}{dx} + \frac{3}{x} \right) R_2 + 2\beta \left( -\varepsilon + \frac{2}{x} \right) R_0 = 0,$$

где  $\beta = (\gamma_1 - 2\gamma) / (\gamma_1 + 2\gamma) = m_l / m_h$  и введены безразмерные атомные единицы  $x = r m_h e^2 / \kappa \hbar^2$ ,  $\varepsilon = -E 2 \kappa^2 \hbar^2 / e^4 m_h$ , соответствующие тяжелой дырке.

Итак, в рамках сферической модели вместо системы четырех связанных уравнений в частных производных



(3.3) возникает система из двух обыкновенных дифференциальных уравнений, которая содержит один безразмерный параметр  $\beta$ , представляющий собой отношение масс легкой и тяжелой дырок. Численное решение системы (3.16) при всех значениях  $\beta$  в интервале  $0 \leq \beta \leq 1$  было получено с помощью ЭВМ Гельмонтом и Дьяконовым [9]. Их результаты представлены на рис. 1.4.

По определению при  $\beta = 1$  имеем  $\epsilon = 1$ . Гельмонт и Дьяконов нашли, что при  $\beta \rightarrow 0$  энергия  $\epsilon = 4/9$ . Таким образом, во всем интервале возможных значений  $\beta$  энергия основного состояния изменяется примерно в два раза и по порядку величины совпадает с боровской энергией тяжелой дырки. Заметим, что в германии  $\beta = 0,13$ , а величина  $4m_{he}^4/18\hbar^2\kappa^2 = 8,1$  мэв, что дает неплохую оценку энергии основного состояния (см. табл. 1.4). В InSb  $\beta = 0,03$ , величина  $4m_{he}^4/18\hbar^2\kappa^2 = 8$  мэв, а экспериментальное значение энергии основного состояния для акцепторных примесей цинка и кадмия составляет 7,5 мэв. Таким образом, приближение, в котором отношение масс легкой и тяжелой дырок считается бесконечно малым, применимо к ряду полупроводников.

Таблица 1.4

Экспериментальные и теоретические значения энергий основного состояния акцепторов в Ge и Si \*) [2]

Вещество	$E_a$ , мэв эксп.	$E_a$ , мэв теор.
Ge { (B) (Al) (Ga) (In) (Te)	{ 10,47 10,80 10,97 11,61 13,10	9,73
Si { (B) (Al) (Ga) (In)	{ 44,5 68,5 72 155	31,56

\*) В скобках указаны химические символы примесей.

Итак, мы рассмотрели сферическую модель в случае, когда гамильтониан представляет матрицу четвертого ранга, что соответствует сильному спин-орбитальному взаимодействию. Можно ограничиться сферическим приближением и в обратном предельном случае малых  $\Delta$ , когда гамильтониан является матрицей третьего ранга: И в этом случае сферическое приближение приводит к разделению переменных и к значительному упрощению задачи.

В работе Балдерески и Липари [10] произведено вариационное вычисление энергий акцепторных уровней в обоих предельных случаях. Полученные ими энергии основного состояния для германия и кремния приведены в табл. 1.4.

В случае кремния Балдерески и Липари производили интерполяцию между случаями сильного и слабого спин-орбитального взаимодействия. Ими рассчитаны также основные и возбужденные состояния целого ряда других полупроводников.

#### § 4. Асимптотика волновых функций примесных уровней

При построении теории прыжковой проводимости слабо легированных полупроводников оказываются существенными свойства волновой функции электрона на большом расстоянии от примесного центра. Волновая функция локализованных состояний на больших расстояниях убывает экспоненциально, и для многих важных задач достаточно знать лишь показатель экспоненты.

В этом параграфе мы ограничимся задачей о вычислении показателя экспоненты волновой функции на больших расстояниях от центра. При этом используется метод эффективной массы, который хорошо работает на больших расстояниях, если характерная длина спада  $q^{-1}$  велика по сравнению с периодом кристаллической решетки.

Начнем со случая невырожденной зоны. Уравнение метода эффективной массы имеет вид

$$[E(\hat{p}) + U(r)] F = EF, \quad (4.1)$$

где  $E(\mathbf{p})$  — зависимость энергии от квазимпульса для рассматриваемой зоны, а  $\hat{\mathbf{p}} = -i\hbar\nabla$ .

При определении асимптотики волновой функции мы будем пользоваться квазиклассическим приближением, т. е. искать решение в виде

$$F = \exp\left[i \frac{S(\mathbf{r})}{\hbar}\right]. \quad (4.2)$$

Нас будет интересовать волновая функция глубоко под барьером, в классически недоступной области, где эта функция экспоненциально мала, т. е. велика мнимая часть действия  $S$ . Благодаря этому и выполняется условие применимости квазиклассического метода

$$\left| \frac{d}{dr} \frac{\hbar}{\left| \frac{dS}{dr} \right|} \right| \ll 1. \quad (4.3)$$

Некоторая необычность последующего изложения связана с тем, что квазиклассический метод использован при решении трехмерных задач с неразделяющимися переменными.

Подставим (4.2) в (4.1) и, следуя обычным правилам, сохраним наименьшие члены по постоянной Планка  $\hbar$ . Это значит, что следует оставить только первые производные от  $S$ , так что

$$E(\hat{\mathbf{p}}) F = E(\nabla S) F. \quad (4.4)$$

Тогда для функции  $S(\mathbf{r})$  получим уравнение Гамильтона — Якоби

$$E(\nabla S) + U = E. \quad (4.5)$$

Допустим, что нам известна волновая функция  $F$  на некоторой поверхности  $\sigma$  (рис. 1.5), находящейся значительно ближе к примесному центру, чем точка  $\mathbf{r}$ , в которой ищем функцию  $F(\mathbf{r})$ , однако удаленной от него так, что условие (4.3) на этой поверхности выполнено. Тогда в каждой точке поверхности  $\sigma$  задано действие  $S(\mathbf{r}_\sigma)$  и его градиент  $\nabla S$  (нормальная составляющая градиента определена уравнением (4.5)).

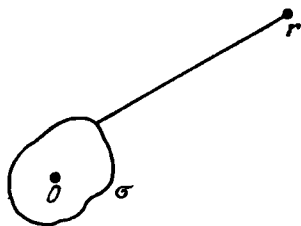


Рис. 1.5. Квазиклассическая траектория, проходящая через точку  $\mathbf{r}$ .

Проведем через каждую точку поверхности  $\sigma$  траектории (лучи), удовлетворяющие классическим уравнениям движения и имеющие на поверхности  $\sigma$  импульс  $\mathbf{p} = \nabla S$ . Тогда решение уравнения (4.5) в произвольной точке  $\mathbf{r}$  можно представить в виде интеграла по той из указанных траекторий, которая проходит через точку  $\mathbf{r}$ :

$$S(\mathbf{r}) = S(\mathbf{r}_\sigma) + \int_{\mathbf{r}_\sigma}^{\mathbf{r}} \mathbf{p} d\mathbf{r}', \quad (4.6)$$

где  $\mathbf{r}_\sigma$  — точка, в которой эта траектория пересекает поверхность  $\sigma$ . Уравнение траектории определяется уравнениями Гамильтона

$$\dot{\mathbf{r}} = \frac{\partial E}{\partial \mathbf{p}}, \quad \dot{\mathbf{p}} = - \frac{\partial U}{\partial \mathbf{r}}. \quad (4.7)$$

Выберем теперь поверхность  $\sigma$  на таком расстоянии от центра, чтобы значением потенциала  $U(\mathbf{r})$  можно было пренебречь. Тогда на траектории, фигурирующей в (4.6),  $\mathbf{p} = \mathbf{p}_0 = \text{const}$ , а сама траектория является прямой. Действие  $S(\mathbf{r}_\sigma)$ , по определению, мало по сравнению с  $S(\mathbf{r})$ , а вклад в интеграл (4.6) определяется верхним пределом. Таким образом, получим простой результат:

$$S(\mathbf{r}) = \mathbf{p}_0 \mathbf{r}. \quad (4.8)$$

Как видно из рис. 1.5, скорость на всякой траектории, проходящей через поверхность  $\sigma$  и далекую от нее точку  $\mathbf{r}$ , должна быть близкой по направлению к радиус-вектору  $\mathbf{r}$  (помещаем примесный центр в начало координат). Таким образом, вектор  $\mathbf{p}_0$  в (4.8) должен удовлетворять условиям

$$\left( \frac{\partial E}{\partial \mathbf{p}} \middle/ \left| \frac{\partial E}{\partial \mathbf{p}} \right| \right)_{\mathbf{p}=\mathbf{p}_0} = \mathbf{n}, \quad (4.9)$$

где  $\mathbf{n} = \mathbf{r}/r$ . К этим двум условиям следует добавить условие, следующее из (4.5):

$$E(\mathbf{p}_0) = E. \quad (4.10)$$

Итак, показатель экспоненты волновой функции на больших расстояниях определяется формулой

$$F = \exp\left(\frac{i\mathbf{p}_0 \mathbf{r}}{\hbar}\right), \quad (4.11)$$

где  $\mathbf{p}_0$  следует находить из уравнений (4.9), (4.10).

Для иллюстрации метода рассмотрим сначала тривиальный случай  $E(\mathbf{p}) = p^2/2m$ . В этом случае скорость  $\partial E/\partial \mathbf{p}$  совпадает по направлению с  $\mathbf{p}$ , а следовательно, в (4.11) следует считать, что вектор  $\mathbf{p}_0$  параллелен  $\mathbf{n}$ . Тогда

$$F(r) = \exp(-qr), \quad q = \frac{\sqrt{2m|E|}}{\hbar}. \quad (4.12)$$

Если подставить вместо  $E$  водородоподобный спектр (2.16), полученный в рамках метода эффективной массы, то, как и должно быть, показатель экспоненты в (4.12) совпадает с показателем экспоненты точных решений. Например, для основного состояния  $q = a^{-1}$ , где  $a = \frac{\hbar^2 \kappa}{me^2}$  — эффективный борковский радиус. Очень важно, однако, что выражение (4.12) имеет более общий характер, чем формула (2.16), и его область применимости в некотором смысле выходит за пределы области метода эффективной массы.

Как уже говорилось, истинные энергии примесных уровней отличаются от энергий, полученных методом эффективной массы, главным образом за счет того, что потенциал (1.1) сильно искажается на малых расстояниях от примеси. Это доказывается, например, тем, что метод эффективной массы дает несравненно лучшие результаты для энергий  $p$ -состояний, чем для  $s$ -состояний (см. табл. 1.2). С другой стороны, при выводе (4.12) вообще не было использовано уравнение Шредингера на малых расстояниях. Поэтому энергия  $E$  в (4.12) вовсе не должна быть энергией, полученной в приближении эффективной массы. В эти формулы имеет смысл подставлять экспериментально полученные энергии уровней, разные для разных примесей. Следует, конечно, иметь в виду, что эти формулы используют условие применимости метода эффективной массы на больших расстояниях, т. е. условие  $qa_0 \ll 1$ , где  $a_0$  — период кристаллической решетки.

Те же соображения относятся и к полученным ниже асимптотикам волновых функций при сложном спектре. Во всех случаях можно выразить показатель экспоненты через параметры спектра возбуждений и истинную энергию уровня.

Перейдем теперь к эллипсоидальному спектру

$$E(\rho) = \frac{\rho_x^2 + \rho_y^2}{2m_t} + \frac{\rho_z^2}{2m_l}. \quad (4.13)$$

Из условия (4.9) для компонент вектора  $\rho_0$  получим

$$\rho_{0x} = n_x d, \quad \rho_{0y} = n_y d, \quad \rho_{0z} = n_z d / \gamma, \quad (4.14)$$

где  $d = (\rho_{0x}^2 + \rho_{0y}^2 + \gamma^2 \rho_{0z}^2)^{1/2}$ ,  $\gamma = m_t / m_l$ . Подставляя (4.14) в (4.10), при отрицательной энергии  $E$  имеем

$$d^2 = - \frac{2m_t |E|}{n_x^2 + n_y^2 + n_z^2 / \gamma}. \quad (4.15)$$

Подставляя (4.15) и (4.14) в (4.11), получим

$$F(\mathbf{r}) = \exp[-q(\mathbf{n})r], \quad (4.16)$$

где

$$q(\mathbf{n}) = \hbar^{-1} [2m_t |E| (n_x^2 + n_y^2 + n_z^2 / \gamma)]^{1/2}. \quad (4.17)$$

Таким образом, поверхности постоянного значения показателя экспоненты волновой функции также являются эллипсоидами.

Произведем оценки для германия. Волновая функция, соответствующая одному из эллипсоидов, спадает в направлении оси вращения этого эллипсоида ( $n_x = n_y = 0$ ) с характерной длиной

$$q_l^{-1} = \hbar (2m_l |E|)^{-1/2} = 13,8 \text{ \AA},$$

а в поперечном направлении — с длиной

$$q_t^{-1} = \hbar (2m_t |E|)^{-1/2} = 61,4 \text{ \AA}.$$

(Использовано значение  $|E| = 12,9$  мэв, соответствующее основному состоянию примеси фосфора.)

Как указывалось в § 2, в полупроводниках с несколькими эквивалентными эллипсоидами (Ge, Si) волновая функция представляет линейную комбинацию волновых функций (см. (2.25) — (2.27)), соответствующих различным эллипсоидам. Каждому эллипсоиду соответствует своя функция  $q_l(\mathbf{n})$ , учитывающая ориентацию этого эллипсоида в  $\mathbf{k}$ -пространстве.

Однако на больших расстояниях от центра в каждом направлении координатного пространства некоторые модулирующие функции окажутся значительно больше других. Значение максимальной модулирующей функции в зависимости от направления определяется формулой (4.16) с

$$q(n) = \min \{q_j(n)\}. \quad (4.18)$$

Перейдем теперь к случаю вырожденной зоны и исследуем асимптотику волновых функций примесных центров в приближении сильного спин-орбитального взаимодействия. Начнем со сферического приближения и проанализируем систему уравнений (3.16), описывающую основное состояние акцепторного центра. Пренебрежем в (3.16) членами порядка  $x^{-1}$  по сравнению с  $\frac{d}{dx}$  и с  $2\beta|\varepsilon|$ . Тогда, складывая и вычитая уравнения, получим

$$\begin{aligned} \frac{d^2(R_2 + R_0)}{dx^2} &= \beta\varepsilon(R_2 + R_0), \\ \frac{d^2(R_2 - R_0)}{dx^2} &= \varepsilon(R_2 - R_0). \end{aligned} \quad (4.19)$$

При малом значении  $\beta$  ( $\beta = m_l/m_h$ ) сумма функций затухает значительно медленнее, чем разность. Поэтому при больших  $x$

$$R_2 \approx R_0 \approx \exp(-x \sqrt{\beta|\varepsilon|}) = \exp\left(-r \sqrt{\frac{2m_l|E|}{\hbar}}\right). \quad (4.20)$$

Таким образом, при сильном различии масс асимптотика волновых функций на больших расстояниях определяется легкой массой, в то время как энергия связи определяется в основном массой тяжелой.

Разумеется, этот важный вывод остается в силе и вне рамок сферической модели, т. е. при учете гофрированности изоэнергетических поверхностей. Для того чтобы рассмотреть этот случай, обобщим изложенный выше квазиклассический метод на случай вырожденных зон.

Решение уравнения (3.3) будем искать в виде

$$F_l = A_l \left(\frac{\hbar}{|k|}\right) \exp\left(\frac{iS(r)}{\hbar}\right), \quad (4.21)$$

$$\hbar k = \nabla S,$$

где  $A_l$  удовлетворяют уравнению (3.9), т. е. являются

компонентами собственного вектора матрицы, которая получается из гамильтониана (3.6) при замене оператора  $\hat{p}$   $s$ -числом. Подставим (4.21) в (3.3) и отбросим члены, содержащие вторые производные от  $S(\mathbf{r})$ . Тогда

$$\begin{aligned} & \sum_{j=1}^4 \sum_{\alpha, \beta=1}^3 H_{j\alpha\beta}^{\alpha\beta} \hat{p}_\alpha \hat{p}_\beta F_j = \\ & = \sum_{\alpha, \beta=1}^3 \exp\left(\frac{iS}{\hbar}\right) \frac{\partial S}{\partial x_\alpha} \frac{\partial S}{\partial x_\beta} \sum_{j=1}^4 H_{j\alpha\beta}^{\alpha\beta} A_j = E_{l, \hbar}(\nabla S) F_l, \quad (4.22) \end{aligned}$$

где функция  $E_{l, \hbar}(\nabla S)$  получается из (3.12) заменой  $\hbar k$  на  $\nabla S$ , а знак плюс или минус в (3.12) выбирается в зависимости от того, какой из ветвей спектра соответствует собственный вектор  $A_j$ .

Итак, вместо уравнения (3.3) мы получим уравнения Гамильтона — Якоби

$$E_{l, \hbar}(\nabla S) = E - U(\mathbf{r}). \quad (4.23)$$

Уравнения (4.23) определяют два разных действия для легких ( $S_l$ ) и тяжелых ( $S_h$ ) дырок. То обстоятельство, что в уравнение для  $S_l$  не входит  $S_h$  и, наоборот, в уравнение для  $S_h$  не входит  $S_l$ , означает, что в квазиклассическом приближении движения, соответствующие легкой и тяжелой дыркам, можно рассматривать независимо. На малых расстояниях от центра квазиклассическое приближение неприменимо. Эти движения не являются независимыми, и волновые функции связываются друг с другом. Поэтому на самом деле искомое решение на больших расстояниях является суперпозицией решений (4.21), соответствующих легким и тяжелым дыркам. Легко убедиться в том, что  $\text{Im } S_l < \text{Im } S_h$ , где  $S_{l, h}$  — решения уравнения (4.23), соответствующие легким и тяжелым дыркам. Поэтому асимптотика волновых функций  $F_j$  определяется действием  $S_l$ . Как и в случае простой зоны, результат выражается формулами (4.9) — (4.11), куда вместо  $E(\mathbf{p})$  следует подставлять  $E_l(\mathbf{p})$ .

Результат может быть представлен в виде (4.16), однако решить уравнения (4.9), (4.10) в произвольном направлении трудно. Поэтому ограничимся вычислением



$q(\mathbf{n})$  в двух направлениях: в направлении [100] ( $n_x = 1$ ,  $n_y = n_z = 0$ ) и в направлении [111] ( $n_x = n_y = n_z = 1/\sqrt{3}$ ). Легко убедиться, что в этих направлениях величины  $q(\mathbf{n})$  максимально отличаются друг от друга. Вычисление значительно упрощается тем, что в обоих направлениях групповая скорость  $\partial E_i / \partial \mathbf{p}$  параллельна импульсу  $\mathbf{p}$ . Поэтому направление вектора  $\mathbf{p}_0$  в (4.11) совпадает с  $\mathbf{n}$ . Согласно (3.12) в направлении [100] имеем

$$E_i = \frac{p^2}{2m_0} [\gamma_1 + 2\gamma_2]. \quad (4.24)$$

Из (4.10), (4.11) следует

$$q_{|100|} = \frac{(2m_0 |E|)^{1/2}}{\hbar (\gamma_1 + 2\gamma_2)^{1/2}}. \quad (4.25)$$

Аналогично получим, что

$$q_{|111|} = \frac{(2m_0 |E|)^{1/2}}{\hbar (\gamma_1 + 2\gamma_3)^{1/2}}. \quad (4.26)$$

Результат (4.20), полученный в рамках сферической модели, можно представить в аналогичном виде:

$$q_{\text{сф}} = \frac{(2m_0 |E|)^{1/2}}{\hbar \left[ \gamma_1 + \frac{2}{5} (3\gamma_3 + 2\gamma_2) \right]^{1/2}}. \quad (4.27)$$

Для германия, легированного галлием ( $|E| = 11$  мэв), получим  $q_{|100|}^{-1} = 87 \text{ \AA}$ ,  $q_{|111|}^{-1} = 92 \text{ \AA}$ ,  $q_{\text{сф}}^{-1} = 90 \text{ \AA}$ . Таким образом, в этом случае сферическая модель дает хорошее приближение и для асимптотики волновых функций.

## ГЛАВА 2

### ЛОКАЛИЗАЦИЯ ЭЛЕКТРОННЫХ СОСТОЯНИЙ

В предыдущей главе рассмотрена структура электронных состояний, возникающих в окрестности одного примесного центра. Теперь мы переходим к одному из самых важных вопросов в теории легированных полупроводников — к вопросу о том, как влияют друг на друга примесные состояния, принадлежащие разным центрам, и как выглядит энергетический спектр кристалла, обладающего конечной концентрацией примесей.

Принято говорить, что при конечной концентрации примесей одного сорта вместо примесного уровня возникает примесная зона, имеющая конечную энергетическую ширину. В этой книге мы также будем пользоваться термином «примесная зона», хотя сразу же предупреждаем читателя, что применять этот термин следует с большими оговорками. В частности, примесная зона слабо легированного полупроводника не обладает самым важным свойством кристаллических зон: электрон, локализованный вблизи одного из примесных центров, не расплывается по всем центрам, из которых составлена зона. Его волновая функция остается локализованной. Тем не менее мы будем пользоваться термином «примесная зона» для обозначения совокупности энергетических уровней, возникающих благодаря примесным центрам.

Важнейшей характеристикой примесной зоны является плотность состояний. Она определяется как число уровней, попадающих в малый интервал энергий, отнесенное к этому интервалу и к объему системы. Следует иметь в виду, что в макроскопической системе плотность состояний является непрерывной функцией энергии в некотором интервале, даже если речь идет о примесной зоне, которая представляет собой набор дискретных уровней. Таким образом, плотность состояний не содер-

жит информации, позволяющей отличить истинную зону от набора дискретных уровней, не связанных друг с другом и случайно разбросанных в энергетическом пространстве.

Настоящая глава посвящена самым общим вопросам, связанным с локализацией электронных состояний. Начнем с изучения квантового уширения уровней. Для этой цели будет использован метод сильной связи и обсуждены свойства узких зон, образованных периодически расположенными примесными центрами, находящимися на больших расстояниях друг от друга. Зонная структура, возникающая при таком стандартном рассмотрении, неустойчива. Электрон-электронное взаимодействие приводит к расщеплению зоны, в результате чего заполненная и пустая зоны оказываются разделенными щелью (переход Мотта). Флуктуации потенциалов отдельных центров создают упомянутую выше локализацию электронных состояний вблизи центров, ставших теперь неэквивалентными (переход Андерсона). К этому же явлению приводит неупорядоченность в расположении центров (модель Лифшица). Все три названных выше фактора играют важную роль в примесных полупроводниках.

## § 5. Узкие зоны и переход Мотта

Этот параграф мы начнем со вспомогательной задачи, которая будет важна в дальнейшем. Допустим, что примеси в кристалле не распределены хаотически, а образуют правильную кристаллическую решетку, имеющую гораздо больший период, чем основная решетка кристалла (подрешетка примесей). Используя метод сильной связи [1], получим энергетический спектр и волновые функции электронов в примесной зоне. Эта модель позволяет выработать терминологию, используемую в следующих параграфах. Кроме того, мы обсудим здесь важный вопрос о роли электрон-электронного взаимодействия в условиях узкой разрешенной зоны.

Потенциал, создаваемый примесной подрешеткой, имеет вид

$$V(\mathbf{r}) = \sum_I U(\mathbf{r} - \mathbf{r}_I), \quad (5.1)$$

где суммирование ведется по всем узлам примесной подрешетки,  $U(\mathbf{r})$  — потенциал одного примесного центра. Задачу будем решать в рамках метода эффективной массы. Допустим, что нам известны волновые функции  $\varphi_n$  и энергетические уровни  $E_n$ , соответствующие решению уравнения Шредингера с потенциалом одного примесного центра

$$\left[ -\frac{\hbar^2}{2m} \Delta + U(\mathbf{r}) \right] \varphi_n = E_n \varphi_n, \quad (5.2)$$

где  $m$  — эффективная масса.

Для простоты ограничимся в этой главе случаем стандартной зоны, т. е. невырожденной зоны с изотропным квадратичным спектром. Мы не будем рассматривать переходы между состояниями, соответствующими разным зонам основного кристалла. Поэтому можно опустить блоховские множители в волновых функциях и интересоваться только уравнениями для огибающих.

В дальнейшем будем считать, что ширина примесной зоны гораздо меньше расстояний между уровнями  $E_n$ , и рассматривать окрестность одного из этих уровней  $E_0$ .

Известно, что волновую функцию, соответствующую потенциалу (5.1), следует конструировать в виде суперпозиции функций  $\varphi(\mathbf{r} - \mathbf{r}_j)$ , удовлетворяющих (5.2):

$$\Psi = \sum_j a_j \varphi(\mathbf{r} - \mathbf{r}_j), \quad (5.3)$$

$$\sum_j |a_j|^2 = 1. \quad (5.4)$$

Это приближение применимо, если характерный размер  $a$  функции  $\varphi(\mathbf{r})$  мал по сравнению с периодом подрешетки примесей  $b_0$ . Действительно, основной вклад в энергию дают те области пространства, в которых волновая функция  $\Psi(\mathbf{r})$  велика, а эти точки, как видно из (5.3), находятся очень близко от примесных центров, т. е. в сфере действия одной примеси, где справедливо уравнение (5.2). Таким образом, выражение (5.3) должно быть близко к решению уравнения Шредингера с потенциалом (5.1). Коэффициенты  $a_j$  следует найти из условия минимума энергии.

Так как волновые функции  $\varphi(\mathbf{r} - \mathbf{r}_j)$ , соответствующие разным узлам, не ортогональны, математическое

ожидание энергии не является квадратичной формой коэффициентов  $a_j$ . Однако, если  $a \ll b_0$ , перекрытие между соседними состояниями мало. В этом случае можно ограничиться первым приближением в разложении математического ожидания энергии по  $a_j^* a_{j+m}$ . Для зависящей от  $a_j^* a_{j+m}$  части энергии получим

$$E = \sum_{j,m} a_j^* a_{j+m} I(m). \quad (5.5)$$

Величину  $I(m)$  называют энергетическим интегралом перекрытия (или просто интегралом перекрытия). Мы не будем в этой главе приводить для нее явных выражений, заметим только, что в интересующих нас условиях она очень мала, так как содержит множитель  $\exp(-\beta b_0/a)$  (для ближайших соседей), где  $\beta$  — численный коэффициент.

Как известно, набор  $a_j$ , удовлетворяющий теореме Блоха

$$\Psi(r + r_j) = \Psi(r) \exp(ikr_j), \quad (5.6)$$

имеет вид  $a_j = (\mathcal{N})^{-1/2} \exp(ikr_j)$ , где  $\mathcal{N}$  — полное число узлов примесной решетки. Подставив это выражение в (5.3) и (5.5), получим

$$\Psi = \frac{1}{\sqrt{\mathcal{N}}} \sum_j \varphi(r - r_j) \exp(ikr_j), \quad (5.7)$$

$$E = \sum_{m \neq 0} I(m) \exp(ikm), \quad (5.8)$$

где  $m$  — вектор, соединяющий заданный узел с другими узлами решетки. Так как величина  $I(m)$  убывает с ростом  $|m|$  экспоненциально, то в (5.8) имеет смысл ограничиться ближайшими соседями. Конкретные выражения зависят от вида решетки. Например, для простой кубической решетки получаем

$$E = 2I(b_0) [\cos k_x b_0 + \cos k_y b_0 + \cos k_z b_0]. \quad (5.9)$$

Для простой квадратной решетки

$$E = 2I(b_0) [\cos k_x b_0 + \cos k_y b_0]. \quad (5.10)$$

Итак, вместо отдельных энергетических уровней мы получили зону разрешенных значений энергии. Для простой квадратной и кубической решеток первая зона

Бриллюэна соответствует набору  $k$  в интервале  $-\frac{\pi}{b_0} < k_\alpha < \frac{\pi}{b_0}$  (здесь  $\alpha$  — декартов индекс).

Энергетическая ширина разрешенной зоны  $V_b$  равна  $12|I(b_0)|$  для простой кубической и  $8|I(b_0)|$  для простой квадратной решеток. В этих случаях выполняется правило  $V_b = 2Z|I(b_0)|$ , где  $Z$  — число ближайших соседей. Заметим, что распространенное мнение о том, что это правило справедливо для всех простых решеток, является заблуждением. Например, для треугольной двумерной решетки с числом ближайших соседей  $Z = 6$  оказывается, что  $V_b = 9|I(b_0)|$ . При малых  $k$  из (5.9) получим

$$\varepsilon(k) = -I(b_0)k^2b_0^2, \quad (5.11)$$

где по определению

$$\varepsilon(k) = E - 6I(b_0). \quad (5.12)$$

Величина  $\hbar^2/2Ib_0^2$ , как ясно из (5.11), играет роль массы электрона в образовавшейся зоне.

С увеличением расстояния между ближайшими соседями ширина разрешенной зоны экспоненциально уменьшается ( $I(b_0) \propto \exp(-\beta b_0/a)$ , где  $\beta$  — численный коэффициент). При этом масса становится экспоненциально большой. Это наводит на мысль, что в интересующем нас случае малой концентрации примесей, т. е. больших значений  $b_0$ , зонные свойства в известном смысле являются фиктивными. Тем не менее пока мы видим, что при сколь угодно узкой зоне волновая функция электрона является модулированной плоской волной и электрон может двигаться без рассеяния.

Зона, образованная примесями, заполнена не более чем на половину, поскольку каждая примесь дает (или забирает) один электрон, а зона двукратно вырождена по спину. Таким образом, получается, что если примеси действительно можно считать расположенными периодически, то проводимость примесных электронов должна иметь металлический характер при сколь угодно малой концентрации примесей.

Это утверждение является неправильным даже с оговоркой относительно периодического расположения примесей. Дело в том, что в предшествующем выводе использовано одноэлектронное приближение. Это прибли-

жение, которое удовлетворительно работает при расчете широких разрешенных зон металлов, оказывается недопустимым в случае узких зон.

Как видно из (5.7), волновая функция электрона вблизи каждого узла мало отличается от узельной функции  $\varphi(r)$ . Оценим теперь энергию взаимодействия двух электронов с разными спинами, находящихся на одном узле. Она окажется порядка  $U_0 \approx e^2/a$ . Если величина  $U_0$  мала по сравнению с шириной разрешенной зоны  $V_b$ , то перестройка волновых функций, связанная с взаимодействием электронов, будет незначительной. Так обстоит дело в хороших металлах. Однако в интересующем нас случае величина  $V_b$  экспоненциально мала и значительно меньше  $U_0$ . Пусть постоянная примесной подрешетки  $b_0$  бесконечно велика. Тогда на каждом узле имеется два электронных уровня. Энергия электрона равна  $E_0$  или  $E_0 + U_0$  в зависимости от того, есть ли на этом узле еще один электрон.

При конечном значении  $b_0$  оба уровня расплываются в зоны с шириной порядка  $|I(b_0)|$  (рис. 2.1). Число мест в каждой из таких зон вдвое меньше, чем в зоне (5.8), и равно числу узлов (в нижней зоне не может быть узла с двумя электронами). Нижняя зона окажется заполненной, а верхняя — пустой. Таким образом, если  $|I(b_0)| \ll U_0$ , наше вещество будет диэлектриком. При уменьшении  $b_0$  в некоторой точке  $A$  ширина запрещенной зоны обращается в нуль и происходит переход диэлектрик — металл, который принято называть переходом Мотта. Вопрос о характере этого перехода и о поведении электропроводности в точке перехода до сих пор остается открытым.

Для количественного изучения перехода Мотта пользуются, как правило, моделью Хаббарда. В этой

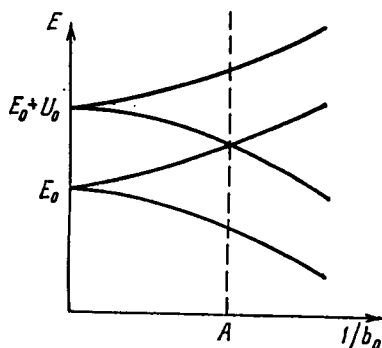


Рис. 2.1. Электронные зоны в зависимости от периода примесной подрешетки  $b_0$ . Слева от точки  $A$  — диэлектрик, справа — металл.

максимально упрощенной модели предполагается, что электроны отталкиваются, только если они находятся на одном узле. Гамильтониан Хаббарда имеет вид

$$H = \sum_{j, m \neq 0} I(m) a_{j, \sigma}^{\dagger} a_{j+m, \sigma} + \frac{U_0}{2} \sum_{j, \sigma} n_{j, \sigma} n_{j, -\sigma}, \quad (5.13)$$

где  $n_{j, \sigma} = a_{j, \sigma}^{\dagger} a_{j, \sigma}$  — оператор чисел заполнения состояния на узле  $j$  со спином  $\sigma$ . Этот гамильтониан получается из (5.5) добавлением члена, описывающего отталкивание электронов, имеющих разный спин и находящихся на одном узле.

Модель Хаббарда допускает точное решение только в одномерном случае [2]. Результат противоречит приведенным выше качественным рассуждениям. В одномерном случае щель в спектре сохраняется при любых значениях  $I(b_0)/U_0$ , т. е. одномерная модель всегда является изолятором. Принято, однако, считать, что это — исключительное свойство одномерных систем и что указанное обстоятельство не бросает тени на изложенную выше качественную картину, коль скоро она относится к двух- и трехмерным системам.

В нашу задачу не входит обзор теории, развитой в связи с переходом Мотта. Изложение этих вопросов можно найти в книге Мотта [3] или в обзоре Хомского [4]. Мы же, возвращаясь после этой экскурсии назад к узким зонам, хотим подчеркнуть, что, говоря в дальнейшем о примесной зоне при малой концентрации примесей, будем всегда иметь в виду нижнюю хаббардовскую зону (см. рис. 2.1). Уширение этой зоны, которое может быть связано и с квантовым перекрытием и с разбросом уровней вследствие флуктуирующего классического поля, будем предполагать малым по сравнению с расстоянием до второй зоны  $U_0$ . Поэтому, изучая структуру примесной зоны при нулевой температуре, будем считать, что на каждом узле может находиться только один электрон. Заметим, что это следует не из принципа Паули, допускающего наличие двух электронов в одном орбитальном состоянии, а из зарядового ограничения, которое оказывается в этом случае более жестким.



## § 6. Переход Андерсона

В этом параграфе мы возвращаемся снова к одно-электронному приближению и продолжаем изучать структуру примесной зоны. Не будем, однако, более считать, что примеси образуют идеальную решетку. Нарушить идеальность примесной подрешетки можно несколькими способами. Самый простой способ состоит в том, чтобы считать расположение примесей случайным. Мы вернемся к нему позднее, а сейчас рассмотрим другой, более разработанный и тоже практически важный вариант разупорядочения.

Допустим, что примеси находятся в узлах правильной решетки, но уровень электрона на всех узлах разный. Таким образом, рассматривается система периодически расположенных потенциальных ям разной глубины (рис. 2.2).



Рис. 2.2. Потенциальные ямы в модели Андерсона.

Обозначим через  $\varepsilon_j$  отклонение уровня электрона на узле  $j$  от среднего значения. Гамильтониан системы в узельном представлении получается из (5.5) добавлением члена, представляющего сумму энергий электронов, которые находятся на изолированных узлах:

$$H = \sum_j \varepsilon_j a_j^\dagger a_j + \sum_{l, m \neq 0} I(m) a_j^\dagger a_{j+m}. \quad (6.1)$$

Энергии  $\varepsilon_j$  считаются случайными величинами, между которыми нет корреляции. Иными словами, вероятность того, что некоторый узел имеет заданную энергию, не зависит от энергий прочих узлов. Распределение энергий будем считать равномерным в некотором интервале  $W$ . Функция распределения  $P(\varepsilon)$  имеет вид

$$P(\varepsilon) = \begin{cases} 1/W, & |\varepsilon| < W/2, \\ 0, & |\varepsilon| > W/2. \end{cases} \quad (6.2)$$

Сформулированная модель предложена Андерсоном [5] и носит его имя. К сожалению, модель Андерсона не допускает точного решения. Несмотря на огромное количество работ, посвященных локализации Андерсона, хорошего математического описания этого явления до сих пор не существует. Тем не менее качественное его понимание, основанное на сравнительно простых физических концепциях, несомненно внушает доверие. Как мы увидим, оно подтверждается также экспериментами на ЭВМ. На качественной картине явления мы остановимся в первую очередь.

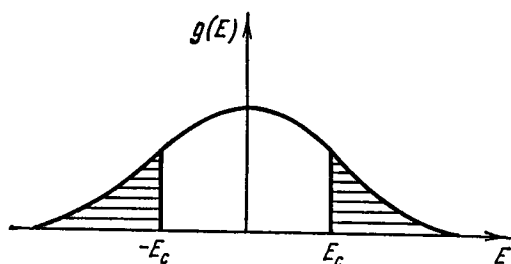


Рис. 2.3. Плотность состояний в модели Андерсона. Локализованные состояния заштрихованы. Энергии  $E_c$  и  $-E_c$ , отделяющие области локализованных и делокализованных состояний, являются порогами подвижности.

Основной вопрос, который ставится при рассмотрении модели Андерсона, состоит в том, чтобы определить, являются ли волновые функции электрона локализованными в окрестности некоторого узла или они распространяются на всю систему. В обоих случаях волновая функция вблизи каждой примеси похожа на узельную, коль скоро перекрытие мало. Важно понять, образуется ли когерентное состояние, являющееся суперпозицией бесконечного числа узельных функций, входящих с примерно одинаковым весом, которое простирается на макроскопическое расстояние. Альтернативный вариант состоит в том, что узельные функции входят в суперпозицию с весом, экспоненциально убывающим по мере удаления от некоторого узла. Такое состояние называют локализованным вблизи этого узла. Если все состояния локализованы, то проводимость системы при нулевой температуре равна нулю.

Вопрос о том, как сформулировать условие локализации математическим языком, далеко не тривиален. Используют несколько различных критериев локализации. Андерсон [5] использовал следующий критерий. Допустим, что в нулевой момент времени волновая функция электрона совпадала с узельной функцией, соответствующей узлу  $i$ . Эта функция не является собственной функцией гамильтониана (6.1), и она будет изменяться со временем. Решив уравнение Шредингера, найдем квадрат модуля волновой функции  $|\Psi_i(t)|^2$  на узле  $i$  при больших временах  $t$ .

Если состояния не локализованы, начальный волновой пакет расплывается по всей системе, тогда в бесконечной системе  $\lim_{t \rightarrow \infty} |\Psi_i(t)|^2 = 0$ . Если разброс уровней

отсутствует и имеется идеальная зонная структура, рассмотренная в предыдущем параграфе, то расплывание волнового пакета произойдет за время порядка  $\hbar/V_b$ . Если же истинные волновые функции локализованы, то расплывания вообще не произойдет и волновая функция приобретет со временем «хвосты» на соседних узлах с экспоненциально малой амплитудой и будет сосредоточена примерно в той же области пространства, что и в начальный момент. Поэтому  $\lim_{t \rightarrow \infty} |\Psi_i(t)|^2$  окажется ко-

нечной величиной. В этом и состоит критерий Андерсона.

Модель Андерсона содержит один безразмерный параметр  $W/I$ , где  $I$  — модуль интеграла перекрытия для соседних узлов. Результат Андерсона, подтвержденный большим числом более поздних исследований, состоит в следующем. При достаточно больших значениях  $W/I$  все состояния являются локализованными. Существует критическое значение  $W_c/I$ , при котором в центре зоны впервые появляются делокализованные состояния. При дальнейшем уменьшении  $W/I$  область делокализованных состояний увеличивается, захватывая почти всю зону (рис. 2.3).

Любопытно, что все сказанное не относится к одномерной системе, где состояния всегда являются локализованными [6—11]. Ситуация здесь аналогична ситуации с переходом Мотта, упомянутой в предыдущем параграфе. Таким образом, учет сколь угодно слабой

флуктуации энергий узлов, так же как и учет электрон-электронного взаимодействия, неизбежно превращает одномерный проводник в диэлектрик.

И переход Андерсона, и переход Мотта являются разновидностями перехода металл — диэлектрик. Как мы видели, основная разница между ними состоит в том, что переход Андерсона связан с неупорядоченностью системы, а переход Мотта происходит и в идеальной периодической структуре. Кроме того, надо иметь в виду, что переход Мотта возникает вследствие электрон-электронного взаимодействия, а в теории перехода Андерсона рассматривается движение одного электрона.

Для того чтобы осмыслить результат Андерсона, рассмотрим вспомогательную задачу о двух ямах с разной глубиной, расположенных на большом расстоянии друг от друга. Такая квантовомеханическая задача легко решается. Пусть энергии электронов в каждой из ям, без учета второй ямы, равны  $\varepsilon_1$  и  $\varepsilon_2$  соответственно, а волновые функции равны  $\varphi_1$  и  $\varphi_2$ . Если ямы одинаковы, то  $\varepsilon_1 = \varepsilon_2$  и возникают два состояния с волновыми функциями

$$\Psi_I = \frac{1}{\sqrt{2}}(\varphi_1 + \varphi_2), \quad \Psi_{II} = \frac{1}{\sqrt{2}}(\varphi_1 - \varphi_2). \quad (6.3)$$

Разность энергий этих состояний равна  $E_I - E_{II} = 2I$ . Важно отметить, что, как бы далеко ни отстояли друг от друга эти ямы и как бы мал ни был интеграл перекрытия, в обоих состояниях электрон в равной степени принадлежит каждой из ям, т. е. его можно найти в любой из них с одинаковой вероятностью. Характер решения мало меняется, если  $|\varepsilon_1 - \varepsilon_2| \ll I$ . В обратном предельном случае  $|\varepsilon_1 - \varepsilon_2| \gg I$  картина совершенно иная. По-прежнему есть два состояния:

$$\Psi_I = c_1\varphi_1 + c_2\varphi_2, \quad \Psi_{II} = c_2\varphi_1 - c_1\varphi_2. \quad (6.4)$$

Однако в состоянии I энергия  $E_I$  близка к энергии  $\varepsilon_1$ , а волновая функция близка к  $\varphi_1$ . Отношение  $c_2/c_1 = I/|\varepsilon_1 - \varepsilon_2|$  и экспоненциально убывает с увеличением расстояния между ямами. Наоборот, в состоянии II энергия  $E_{II}$  близка к энергии  $\varepsilon_2$ , а волновая функция  $\Psi_{II}$  близка к  $\varphi_2$ . Таким образом, в этом случае обобщения электрона не происходит. Каждое состояние

фактически представляет состояние электрона в одной из ям.

Этот пример поможет понять суть перехода Андерсона. (Идеи последующих рассуждений принадлежат Таулессу [12].) Рассмотрим полосу энергий, находящихся в интервале  $-\frac{\Delta}{2} < \varepsilon < \frac{\Delta}{2}$ , причем ширина полосы  $\Delta$  порядка интеграла перекрытия  $I$ . Будем называть узлы, энергия которых попадает в эту полосу, резонансными, а узлы с энергией вне полосы — нерезонансными. Суть этого определения в том, что электронное состояние обобществляется между двумя резонансными узлами, если эти узлы являются ближайшими соседями.

Назовем, далее, два резонансных узла связанными друг с другом, если они являются ближайшими соседями или если есть третий резонансный узел, который является ближайшим соседом их обоих. Два резонансных узла также связаны друг с другом, когда они соединены цепочкой из связанных друг с другом резонансных узлов. Совокупность связанных между собой узлов назовем кластером. Этим кластерам будут соответствовать электронные состояния, у которых квадрат модуля волновой функции одного порядка на всех узлах, принадлежащих кластеру, и мал везде вне кластера.

Следующий шаг состоит в том, что мы вообще выбросим из рассмотрения нерезонансные узлы. Этот шаг сомнителен. Гамильтониан Андерсона (6.1) связывает друг с другом только соседние узлы. Тем не менее два резонансных узла могут иметь общее электронное состояние и в том случае, когда между ними находится нерезонансный узел. Однако эффективный интеграл перекрытия будет при этом не порядка  $I$ , а порядка  $I^2/W$ . Если отношение  $I/W$  мало, то связь через нерезонансный узел приведет к обобществлению состояний в значительно более узкой полосе, чем рассматриваемая, и этим явлением можно будет пренебречь. Итак, выбросив нерезонансные узлы, получим, что характерный размер волновой функции определяется размером кластеров из связанных резонансных узлов.

Распределение энергий  $\varepsilon_i$  в модели Андерсона считается равномерным в интервале  $W$ . По этой причине доля резонансных узлов должна быть порядка  $I/W$ . При

малых значениях этого параметра резонансных узлов мало и они расположены в основном поодиночке. Однако при некотором критическом его значении возникает бесконечный кластер из связанных резонансных узлов, т. е. образуются пути, уходящие на бесконечность. По этим путям расплываются волновые функции электронных состояний. В этом и состоит переход Андерсона.

Теория протекания (см. гл. 5) позволяет найти точное значение величины  $\Delta/W_c$ , при котором образуется бесконечный кластер. Для этого следует приравнять долю резонансных узлов  $\Delta/W_c$  к порогу протекания  $x_c(s)$  для задачи узлов, т. е.

$$\Delta/W_c = x_c(s). \quad (6.5)$$

Значения  $x_c(s)$  для различных решеток даны в табл. 5.1.

Чтобы оценить критическое значение  $W_c/I$ , нужно найти связь между шириной резонансной полосы  $\Delta$  и интегралом перекрытия  $I$ . В этом заключается наиболее сложный и (спорный) момент предлагаемой оценки. Дело в том, что само определение резонансных узлов не является точным, так как при уменьшении разности  $|e_i - e_j|$  обобществление электронных состояний между двумя узлами происходит постепенно, а не скачком.

Тем не менее переход Андерсона происходит при вполне определенном значении  $W_c/I$ . Действительно, если волновая функция экспоненциально убывает от узла к узлу, то она не может принадлежать делокализованному состоянию. Поэтому при больших  $W/I$  все состояния локализованы. В бесконечной системе возникновение делокализованного состояния при уменьшении  $W/I$  есть четко выраженное критическое явление, которому должно соответствовать определенное пороговое значение  $W_c/I$ . Поэтому отмеченная выше неоднозначность определения резонансных узлов свидетельствует лишь о приближенности описания, использующего понятие резонансной полосы.

Можно, однако, предложить некоторый способ оценки ширины  $\Delta$  и критического значения  $W_c/I$ , основанный на топологических свойствах бесконечного кластера (подробно описанных в гл. 5). Вблизи порога протекания бесконечный кластер представляет собой сетку, составленную практически одномерными цепочками резонанс-

ных узлов. Если пренебречь различием энергий резонансных узлов, то окажется, что вследствие перекрытия волновых функций вдоль цепочки образуется зона шириной  $4I$ . Естественно поэтому считать, что ширина резонансной полосы  $\Delta \approx 4I$ . Подставляя это значение в (6.5), получим следующую оценку:

$$\frac{W_c}{I} = \frac{4}{x_c(s)}. \quad (6.6)$$

Для ряда решеток значения  $W_c/I$  известны из численных экспериментов на ЭВМ. Эдвардс и Таулесс [13], а также Лиссиарделло и Таулесс [14] провели эксперименты для двумерных решеток, а Веер и Сривастава [15] — для трехмерных и гиперкубических решеток. Результаты этих работ собраны в табл. 2.1. Для сравнения приведены также результаты оценки (6.6). Нужные для оценки пороги протекания заимствованы из табл. 5.1, а для гиперрешеток — из работы Киркпатрика [28 к гл. 5]. Разумеется, предположения, сделанные при выводе (6.6), являются очень грубыми и могут вызвать возражения. Данные, приведенные в табл. 2.1, однако,

Таблица 2.1  
Пороги протекания и критические значения  $W_c/I$   
для некоторых решеток

Решетки	$x_c(s)$	$4/x_c(s)$	$W_c/I$
Двумерные:			
шестиугольная	0,70	5,7	4,3
квадратная	0,59	6,8	6,1
треугольная	0,50	8,0	9,4
Трехмерные:			
алмаза	0,43	9,3	8,0
простая кубическая	0,31	12,9	14,4
Гиперкубические:			
четырёхмерная	0,20	20	25
пятимерная	0,14	28,6	36

показывают, что оценка хорошо согласуется с результатами численных экспериментов.

Итак, при  $W < W_c$  в электронном спектре возникает полоса делокализованных состояний. Естественно спросить, является ли эта полоса вблизи перехода бесконечно узкой, или она сразу имеет конечную ширину? Обоснованного ответа на этот вопрос не существует. Обычно считают (см., например, [14]), что полоса делокализованных состояний «раскрывается» постепенно, хотя серьезных доводов в пользу такой точки зрения, на наш взгляд, нет. Более того, нам представляется, что существование сколь угодно узких делокализованных полос трудно совместить с концепцией минимальной металлической проводимости (см. § 7). Вряд ли сколь угодно узкая полоса может обеспечить конечную проводимость.

Выше под переходом Андерсона мы подразумевали возникновение полосы делокализованных состояний. Однако в этот термин вкладывают и другой смысл. Допустим, что в рассматриваемой зоне уже существует резкая граница между локализованными и делокализованными состояниями. Если каким-либо образом изменять заполнение зоны электронами, то будет меняться и положение уровня Ферми. При этом изменении уровень Ферми может пересечь границу области локализации, что приведет к существенным изменениям электронных свойств системы. Это явление также часто называют переходом Андерсона, и рассматриваемые ниже примеры будут относиться в основном к этому случаю.

### § 7. Примеры переходов Андерсона. Электропроводность и волновые функции вблизи перехода

На представлениях о переходе Андерсона основана важнейшая концепция теории неупорядоченных систем — концепция порога подвижности. Особую роль эта концепция играет в теории аморфных полупроводников. Согласно современной теории зонная структура аморфного полупроводника имеет вид, изображенный на рис. 2.4.

Наличие ближнего порядка приводит к тому, что сохраняется понятие о разрешенных и запрещенных зонах.



Однако многочисленные структурные нарушения создают «хвост» плотности состояний в запрещенной зоне, так что границы зон (в смысле обращения в нуль плотности состояний) оказываются размытыми. Уменьшение плотности состояний ведет к тому, что состояния в запрещенной зоне оказываются локализованными, и в некотором смысле пороги локализации и играют роль зонных границ. Уровень Ферми находится в области локализованных состояний, и поэтому перенос заряда осуществляется путем теплового заброса носителей в область делокализованных состояний или путем активированных прыжков по локализованным состояниям (см. книгу Мотта и Дэвиса [16]). И тот и другой способ приводит при низких температурах к экспоненциальному росту электропроводности с температурой. Утверждение о строгой локализации электронных состояний играет при этом очень важную роль.

Если бы существовала возможность расплывания волнового пакета, то было бы возможно безактивационное движение электрона в электрическом поле. При достаточно низкой температуре такой механизм переноса

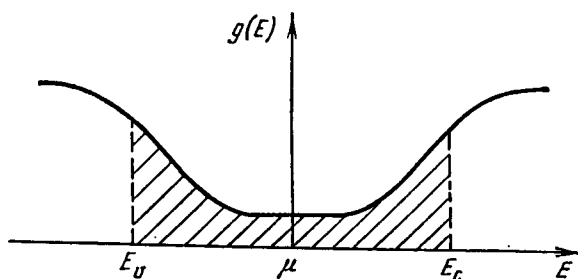


Рис. 2.4. Плотность состояний в аморфном полупроводнике. Локализованные состояния заштрихованы.

имел бы несомненное преимущество и электропроводность перестала бы зависеть от температуры. Как будет показано в следующих главах этой книги, ситуация, когда уровень Ферми оказывается в области локализованных состояний, весьма часто осуществляется в легированных кристаллических полупроводниках, и потому в их теории вопрос о локализации также играет первостепенную роль.

Весьма важным оказался вопрос об андерсоновской локализации в теории МДП-структуры. Эта структура (металл, диэлектрик, полупроводник) чаще всего представляет собой пленку двуокиси кремния (диэлектрик), находящуюся между подложкой из кремния (полупроводник) и плоским металлическим электродом. Приложенный к металлу потенциал изгибает зоны и перерас-

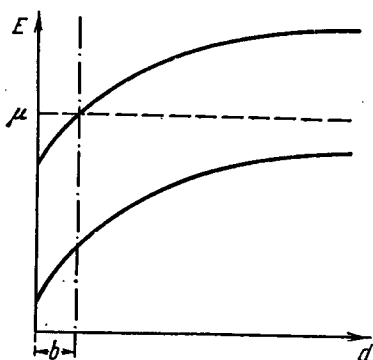


Рис. 2.5. Зонная схема вблизи поверхности полупроводника в МДП-структуре:  $d$  — расстояние от поверхности,  $b$  — ширина инверсионного слоя. Штриховая линия — уровень Ферми.

пределяет заряд таким образом, что на поверхности полупроводника возникает узкий инверсионный слой, т. е. слой с носителями тока, другого знака, чем в объеме (рис. 2.5).

Изменяя приложенный потенциал, можно менять концентрацию носителей в инверсионном слое, а следовательно и поверхностную проводимость слоя, в весьма широких пределах. На этом основано действие полевого транзистора, представляющего собой такую структуру. Инверсионный слой заполнен двумерным

электронным газом, находящимся в случайном потенциале, который создается главным образом за счет неоднородностей диэлектрической пленки. Согласно идеям Мотта [17] и Штерна [18], в этом потенциале и происходит андерсоновская локализация носителей. При малых концентрациях носителей проводимость инверсионного слоя носит активационный характер, причем при низких температурах энергия активации уменьшается с температурой. Это означает, что в окрестности уровня Ферми находятся локализованные состояния, по которым происходят перескоки. При увеличении концентрации носителей уровень Ферми оказывается в области делокализованных состояний и проводимость носит металлический характер (подробное обсуждение обширного экспериментального материала см. в обзоре Мотта и др. [19]). Простой способ изменения концентрации с по-

мощью приложенного напряжения позволяет сравнительно легко исследовать электропроводность вблизи порога локализации. При этом возникает еще одна интересная теоретическая проблема, к обсуждению которой мы переходим.

Допустим, что уровень Ферми приближается к границе делокализованных состояний. Спрашивается, обращается ли электропроводность при нулевой температуре в нуль скачком или непрерывным образом в момент пересечения уровнем Ферми границы делокализованной области (рис. 2.6)?

Строго обоснованного ответа на этот вопрос нет. Ниже будет изложена близкая нам точка зрения Мотта [20, 21], которая, по-видимому, занимает господствующее место в литературе, хотя неоднократно подвергалась критике [22].

Допустим, что уровень Ферми находится посередине разрешенной зоны. Если беспорядок мал и электроны почти свободные, то проводимость выражается обычной формулой Друде:

$$\sigma = \frac{e^2 k_F^2 l}{3\pi^2 \hbar}, \quad (7.1)$$

где  $l$  — длина свободного пробега,  $k_F$  — модуль волнового вектора на поверхности Ферми. Точка зрения Мотта основана на утверждении, высказанном пионерами физики неупорядоченных систем А. Ф. Иоффе и А. Р. Регелем [23], которое состоит в том, что длина свободного пробега по самому ее физическому смыслу не может быть меньше длины волны электрона. В случае зоны, заполненной наполовину,  $k_F \approx 1/a_0$  ( $a_0$  — постоянная решетки), утверждение  $k_F l = l/a_0 > 1$  является совершенно прозрачным. Из него следует, что электропроводность

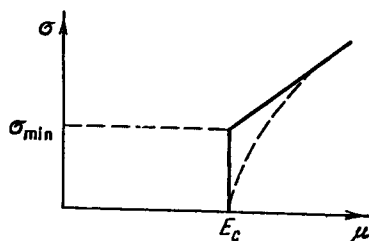


Рис. 2.6. Электропроводность при нулевой температуре как функция положения уровня Ферми  $\mu$ . В точке, где уровень Ферми пересекает порог локализации  $E_c$ , электропроводность должна обратиться в нуль. Но происходит ли это скачком (сплошная кривая) или плавно (штриховая)?

не может быть меньше величины

$$\sigma_{\min} = \frac{c_3 e^2}{\hbar a_0}, \quad (7.2)$$

где  $c_3$  — численный коэффициент.

Концепция Мотта состоит, по существу, в том, что при нулевой температуре электронный газ не может обладать никакой другой проводимостью, кроме той, которая описана в курсах общей физики и по порядку величины выражается формулой (7.1). Но такая проводимость имеет минимально допустимое значение (7.2). Отсюда Мотт делает вывод, что при увеличении беспорядка (например, при увеличении величины  $W$  в модели Андерсона) проводимость уменьшается, достигая величины (7.2), после чего скачком обращается в нуль (см. рис. 2.6). В литературе эта точка зрения получила название «концепции минимальной металлической проводимости». Разумеется, если исходить из микроскопических позиций, то причина, по которой проводимость обращается в нуль, состоит в андерсоновской локализации, и точка, в которой это происходит, соответствует андерсоновскому порогу.

В модели Андерсона константа  $c_3$  в (7.2), по оценке Мотта [19], лежит в пределах 0,025—0,1 (для различных решеток). Поэтому  $\sigma_{\min}$  заключена в пределах 250—1000  $\text{ом}^{-1} \cdot \text{см}^{-1}$ .

Особый интерес приобретает концепция минимальной металлической проводимости в двумерном случае. При этом

$$\sigma_{\min} = \frac{c_2 e^2}{\hbar} \quad (7.3)$$

и в качестве размерных величин в  $\sigma_{\min}$  входят только универсальные константы. Поэтому можно надеяться, что константа  $c_2$  также является универсальной и не зависит от рассматриваемой модели. Это подтверждается изящными аргументами, приведенными Лиссиарделло и Таулессом [24].

Относительно концепции минимальной металлической проводимости необходимо иметь в виду следующее. Все описанные выше рассуждения и расчеты велись в одноэлектронном приближении. Очень важный момент этих рассуждений состоял в том, что плотность состоя-

ний на уровне Ферми считалась отличной от нуля заданной величиной. Мы увидим в § 33, что электрон-электронное взаимодействие приводит к обращению в нуль плотности состояний на уровне Ферми. Поэтому вопрос о том, существует ли минимальная металлическая проводимость, в связи с этим обстоятельством требует дополнительного обсуждения.

Мы переходим теперь к другому важному аспекту теории перехода Андерсона, а именно к вопросу о поведении волновых функций при энергиях, близких к критической энергии  $E_c$ , разделяющей локализованные и делокализованные состояния. Из общих принципов квантовой механики следует, что поведение волновых функций не может измениться при малом изменении энергии скачком. Локализованные состояния, отделенные от делокализованных состояний очень малой энергией, должны иметь большой радиус локализации, так как именно энергия  $E_c - E$  определяет затухание локализованных состояний в классически запрещенных областях.

Рассмотрим, например, потенциальную яму и допустим, что вследствие уменьшения ее глубины какое-либо связанное состояние «выталкивается» в непрерывный спектр. При этом неизбежно радиус локализации этого состояния обращается в бесконечность, как  $E^{-1/2}$ , где  $E$  — энергия, отсчитанная от границы непрерывного спектра. Поэтому следует принять, что радиус локализованных состояний  $a(E)$  ведет себя вблизи порога по закону

$$a(E) = \frac{a_0}{|E_c - E|^s}, \quad (7.4)$$

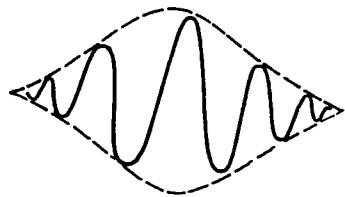


Рис. 2.7. Волновая функция вблизи перехода Андерсона.

где  $s$  — неизвестный критический индекс. Следует иметь в виду, что соответствующие волновые функции осциллируют в масштабах порядка постоянной решетки и лишь их огибающая ведет себя, как  $\exp[-r/a(E)]$  (рис. 2.7).

Вопрос о величине критического индекса  $s$  до сих пор окончательно не решен. Если бы речь шла о простой потенциальной яме, где считается, что потенциал на достаточно большом расстоянии от ямы есть константа, то

естественный ответ состоял бы в том, что  $s = 1/2$ . Именно это и предположили в первую очередь Мотт и Дэвис [16, стр. 49]. Однако серьезных оснований для такого предположения нет. Андерсон [25], используя результаты теории самоизбегающих путей, показал, что наиболее вероятная функция Грина убывает с расстоянием, как  $\exp[-r/a(E)]$ , где  $a(E)$  выражается формулой (7.4) с  $s = 0,6$ .

Следует, однако, заметить, что аргументация Андерсона, с помощью которой задача об асимптотике наиболее вероятной функции Грина сводится к задаче о самоизбегающих путях, содержит ряд не вполне очевидных предположений. Любопытно тем не менее, что Абрам и Эдвардс [26] получили тот же результат из совершенно других соображений. Фрид [27] получил близкий, но не тождественный результат  $s = 2/3$ . Близкий результат получил также Садовский [28], который, используя аргументацию Андерсона и теорию самоизбегающих путей, развитую Де Женом [29], нашел, что индекс  $s$  совпадает с индексом радиуса корреляции обычной теории критических явлений.

Вполне естественно считать, что индекс  $s$  не зависит от конкретной модели неупорядоченной системы, но зависит от числа измерений пространства. Приведенные выше значения индексов относились к трехмерному случаю. Абрам [30] нашел, что в двумерном случае  $s = 0,75$ . Наконец, Ласт и Таулесс [31] предположили, что в трехмерном случае непосредственно вблизи энергии  $E_c$  существует область энергий, где волновые функции убывают с расстоянием степенным образом. В этой же работе сделан численный эксперимент, свидетельствующий в пользу этого предположения.

## § 8. Некоторые аспекты теории перехода Андерсона

В предыдущих параграфах рассматривались, по существу, лишь качественные соображения, относящиеся к переходу Андерсона. В этом параграфе мы кратко остановимся на аналитических подходах к проблеме. Идея аналитического описания перехода содержалась в первой работе Андерсона [5].

Андерсон рассматривал одночастичную функцию Грина в узельном представлении  $G_{ik}$  и изучил собственно энергетическую часть  $\Sigma_i$  диагонального по индексу узла матричного элемента  $G_{ii}$ . Переходу должно соответствовать обращение в нуль  $\text{Im}\Sigma_i$  при некотором значении энергии, лежащем на вещественной оси. Фактически это означает, что обращается в нуль вероятность ухода с данного узла на любой другой узел. Прежде всего Андерсон показал, что такая вероятность во втором порядке теории возмущений равна нулю при любых значениях параметра  $I/W$ .

На первый взгляд это выглядит очень странно. Действительно, во втором порядке теории возмущений для вероятности ухода с узла  $i$  на любой из соседних узлов  $g$  имеем

$$V_i = \frac{2\pi}{\hbar} \sum_g |I(g)|^2 \delta(\epsilon_i - \epsilon_g). \quad (8.1)$$

Усредним это выражение по всевозможным значениям энергии на узле  $g$ :

$$\langle V_i \rangle = \frac{2\pi}{\hbar} \sum_g |I(g)|^2 \int P(E) \delta(\epsilon_i - E) dE, \quad (8.2)$$

где  $P(E)$  — функция распределения энергий на узле. Воспользовавшись формулой (6.2), получим

$$\langle V_i \rangle = \frac{2\pi}{\hbar} \frac{\sum_g |I(g)|^2}{W}. \quad (8.3)$$

В противоположность сказанному, эта величина не обращается в нуль ни при каких конечных значениях  $I/W$ . Более того, такая ситуация имеет место и во всех порядках теории возмущений. Однако сейчас мы увидим, что это ничего не означает. Заменим в (8.1)  $\delta$ -функцию на  $\frac{1}{\pi} \frac{s}{s^2 + (\epsilon_i - \epsilon_g)^2}$  и будем считать  $s$  очень малой величиной. Пусть

$$V_i = \lim_{s \rightarrow 0} V_i(s), \quad V_i(s) = sX_i(s), \quad (8.4)$$

$$X_i(s) = \frac{2}{\hbar} \sum_g |I(g)|^2 \frac{1}{s^2 + (\epsilon_i - \epsilon_g)^2}. \quad (8.5)$$

Для простоты считается, что величина  $I(g)$  отлична от нуля только для ближайших соседей.

Теперь легко интерпретировать проделанные выкладки. Введя конечную ширину уровня  $s$ , мы, естественно, получили, что возможны переходы между узлами, энергии которых различаются на величину порядка  $s$  (или еще меньше). При сколь угодно малых значениях  $s$  в системе существуют резонансные пары, составленные из узлов, являющихся ближайшими соседями и имеющих очень близкие значения энергий ( $|\epsilon_g - \epsilon_l| \leq s$ ). Доля узлов, принадлежащих к резонансным парам, пропорциональна  $s/W$ , т. е. исчезающе мала, однако при усреднении (8.4), (8.5) по всем значениям  $\epsilon_g$  конфигурациям, содержащим резонансные пары, приписывается очень большой вес (порядка  $1/s$ ). Вследствие этого  $\langle V_i \rangle$  оказывается конечной величиной, отличной от нуля.

Резонансные пары очень редки и располагаются на больших расстояниях друг от друга. Переходы между различными парами невозможны, и потому существование таких пар не может обеспечить расплывание волнового пакета на макроскопическое расстояние. Таким образом, величина  $\langle V_i \rangle$  не имеет непосредственного отношения к вопросу о локализации.

Используя (8.5), Андерсон вычислил функцию распределения случайной величины  $X_i$  при фиксированном  $s$ . Оказалось, что эта функция имеет вполне определенный предел при  $s \rightarrow 0$ . В этом пределе функция нормируема и, следовательно, вероятность того, что  $X$  принимает конечное значение, равна единице. Вероятность же отличного от нуля значения случайной величины  $V_i$  согласно (8.4) равна нулю\*).

Таким образом, во втором порядке теории возмущений делокализации электронных состояний не возникает. Однако Андерсон показал, что ряд теории возмущений для величины  $X_i$  расходится с вероятностью, равной единице, при больших значениях  $1/W$ . Точка, где он начинает расходиться, и есть порог делокализации.

---

\* При больших значениях  $X$  функция распределения убывает лишь как  $X^{-3/2}$ , так что интеграл, представляющий среднее значение  $X$ , расходится. С этим и связано отличное от нуля выражение (8.3) для  $\langle V_i \rangle$ .



Таковы основные идеи работы Андерсона. В последние годы эти идеи интенсивно развивались. Обзор работ на эту тему читатель найдет в статьях Таулесса [12] и Лиссиарделло и Иконому [32]. Интересно отметить, что полученные в работе [32] пороговые значения  $W_c/I$  оказались очень близкими к вычисленным с помощью ЭВМ.

Любопытно, что одночастичная функция Грина, усредненная по всем реализациям случайных энергий и содержащая полную информацию относительно плотности состояний, по-видимому, «не чувствует» границу области делокализованных состояний (см. [33]). Так, по крайней мере, дело обстоит в модели Ллойда [34]. Эта модель отличается от модели Андерсона всего лишь тем, что  $P(E)$  выражается не формулой (6.2), а формулой

$$P(E) = \gamma/\pi (\gamma^2 + E^2). \quad (8.6)$$

Все качественные соображения относительно перехода Андерсона, приведенные в предыдущих параграфах, в равной мере относятся и к модели Ллойда. Однако для этой модели Ллойду удалось найти усредненную функцию Грина. Оказалось, что она не имеет никаких особенностей, связанных с делокализацией. Единственной моделью, в рамках которой удалось математически точно описать делокализацию, является модель Андерсона на решетке Бете. Метод, с помощью которого решена эта задача, положен в основу приближенных вычислений для обычных решеток [35, 36].

Нам представляется, что теория перехода Андерсона должна быть построена в духе общей теории критических явлений. Вероятно, этому должна предшествовать феноменологическая теория, вводящая параметр порядка и критические индексы, подобно тому, как это сделано в теории протекания (см. § 19). Вероятно, индексом радиуса корреляции и является индекс  $s$  в формуле (7.4). Однако, насколько нам известно, работ такого типа до сих пор не было.

## § 9. Локализация в модели Лифшица

В модели Андерсона потенциальные ямы разной глубины располагались в узлах правильной кристаллической решетки. Теперь мы рассмотрим в некотором смысле

альтернативную модель: одинаковые потенциальные ямы, случайно расположенные в пространстве. В литературе эту модель иногда называют «моделью со структурным беспорядком». Будем считать потенциал ямы короткодействующим, а среднее расстояние между ямами большим не только по сравнению с радиусом действия потенциала, но и по сравнению с радиусом волновой функции, соответствующей одной яме. Впервые эта модель была подробно изучена И. М. Лифшицем [37].

Основной вопрос, который мы хотим выяснить, состоит в том, будут ли электронные состояния в этом случае локализованными вблизи отдельных ям или они расплывутся по всей системе. На первый взгляд может показаться, что поскольку все ямы одинаковые, то, по терминологии § 6, они являются резонансными и, следовательно, должно произойти расплывание волнового пакета. Однако перекрытие волновых функций соседних ям приведет к сдвигу энергетических уровней, а поскольку у разных ям ближайшие соседи находятся на разных расстояниях, то и сдвиги окажутся разными. Модель Лифшица существенно отличается от модели Андерсона. В последней считалось, что разброс уровней и перекрытие имеют разную природу и могут быть заданы в произвольном отношении. Если разброс значительно больше перекрытия, то состояния локализованы, а если значительно меньше, то они делокализованы. В модели Лифшица перекрытие и разброс имеют одну природу и в среднем одинаковый порядок величины. Единственный малый параметр задачи представляет отношение радиуса волновой функции к среднему расстоянию между ямами, однако сразу не ясно, каким образом этим параметром можно пользоваться.

Мы не будем полностью излагать математический аппарат, разработанный Лифшицем для этого случая, а рассмотрим задачи о двух и трех потенциальных ямах. Это прояснит проблему настолько, что окажется возможным объяснить основные результаты без дальнейшей математики. Они состоят в следующем. Электронные состояния оказываются либо локализованными вблизи отдельных потенциальных ям, либо принадлежат сразу паре близлежащих ям, образуя суперпозицию вида (6.3).

И те и другие содержат экспоненциально малые примеси прочих состояний. Состояния, в которые равноправным образом входят волновые функции трех и более ям, не реализуются, так как это может быть лишь при строго симметричном расположении этих ям. Таким образом, состояния в модели Лифшица являются локализованными. Этих сведений достаточно, чтобы перейти к следующим параграфам, не изучая предлагаемого ниже вывода. Рассмотрим гамильтониан

$$\hat{H} = \frac{\hat{p}^2}{2m} + \sum_j U(\mathbf{r} - \mathbf{r}_j). \quad (9.1)$$

Допустим, что потенциал отдельной ямы  $U(\mathbf{r})$  создает локализованное состояние  $\Psi_0(\mathbf{r})$ , имеющее энергию  $E_0$ , и что прочие состояния, образованные этой ямой, отделены от  $E_0$  на энергию, значительно превосходящую смещение уровня под действием других потенциальных ям. Будем считать, что  $U(\mathbf{r})$  спадает гораздо быстрее, чем  $\Psi_0(\mathbf{r})$ , так что волновая функция в сфере действия потенциала может считаться константой. Поэтому в уравнении Шредингера член  $\sum_j U(\mathbf{r} - \mathbf{r}_j) \Psi(\mathbf{r})$  заменяется членом  $\sum_j U(\mathbf{r} - \mathbf{r}_j) \tau_j$ , где  $\tau_j = \Psi(\mathbf{r}_j)$ . Тогда

$$\left(\frac{\hat{p}^2}{2m} - E\right) \Psi(\mathbf{r}) + \sum_j U(\mathbf{r} - \mathbf{r}_j) \tau_j = 0, \quad (9.2)$$

где  $E$  — искомая энергия, или

$$\Psi(\mathbf{r}) = \left(E - \frac{\hat{p}^2}{2m}\right)^{-1} \sum_j U(\mathbf{r} - \mathbf{r}_j) \tau_j. \quad (9.3)$$

Разложив правую часть в ряд Фурье, найдем

$$\Psi(\mathbf{r}) = \sum_j F(E, \mathbf{r} - \mathbf{r}_j) \tau_j, \quad (9.4)$$

где

$$F(E, \mathbf{r}) = \int \frac{c(\mathbf{k}) \exp(i\mathbf{k}\mathbf{r}) d\mathbf{k}}{E - \hbar^2 k^2 / 2m}, \quad (9.5)$$

$$U(\mathbf{r}) = \int c(\mathbf{k}) \exp(i\mathbf{k}\mathbf{r}) d\mathbf{k}.$$

Положив  $r = r_i$ , получим систему линейных однородных уравнений

$$\tau_i - \sum_j F(E, r_{ij}) \tau_j = 0 \quad (9.6)$$

( $r_{ij} = r_i - r_j$ ), условие разрешимости которой и определяет энергию  $E$ .

Если есть всего одна яма, то энергия уровня  $E_0$  определяется уравнением

$$1 - F_0(E_0) = 0, \quad (9.7)$$

где

$$F_0(E) = \int \frac{c(k) dk}{E - \hbar^2 k^2 / 2m}. \quad (9.8)$$

В одномерном случае можно заменить  $c(k)$  на  $c(0)$ . При этом интеграл (9.8) переписывается в виде

$$F_0(E_0) = 2\pi c(0) \int \frac{g_0(E') dE'}{E_0 - E'}, \quad (9.9)$$

где

$$g_0(E) = (2\pi)^{-1} \int dk \delta(E - \hbar^2 k^2 / 2m)$$

есть плотность состояний свободных электронов. Так как  $g_0(E) \propto 1/\sqrt{E}$ , интеграл (9.9) расходится при  $E_0 = 0$ . Поэтому в одномерном случае уровень  $E_0$  существует при сколь угодно малых отрицательных значениях  $c(0)$ . В трехмерном случае ситуация иная. Замена  $c(k)$  на  $c(0)$  недопустима, так как при этом интеграл (9.8) расходится при больших  $k$ . Это связано с тем, что трехмерная яма дельта-функционной формы имеет бесконечно большую мощность и для того, чтобы получить конечную энергию уровня  $E_0$ , необходимо ограничить потенциал на малых расстояниях.

Волновая функция электрона на одной яме  $\Psi_0(r)$  определяется из (9.4):

$$\Psi_0(r) = \Psi_0(0) F(E_0, r). \quad (9.10)$$

При больших  $r$  в интеграле (9.5) можно заменить  $c(k)$  на  $c(0)$ , и мы получим

$$F(E_0, r) = - \frac{A \exp(-\alpha r)}{r}, \quad (9.11)$$

где

$$A = (2\pi)^2 \frac{mc(0)}{\hbar^2}, \quad \alpha^2 = -\frac{2mE_0}{\hbar^2},$$

что и определяет асимптотику волновой функции.

Теперь напишем условие разрешимости системы (9.6), определяющее энергию  $E$  в потенциале большого числа ям. Мы предполагаем, что величина  $\varepsilon = E_0 - E$  мала по сравнению с расстоянием до следующего уровня одной ямы, и произведем разложение в (9.8). Для диагональных членов (9.6), воспользовавшись (9.7), получим

$$1 - F_0(E) = B\varepsilon, \quad (9.12)$$

где

$$B = \int \frac{dk c(k)}{(E(\hbar k) - E_0)^2}.$$

Для недиагональных членов мы воспользуемся формулой (9.11), так как среднее расстояние между ямами  $N^{-1/2} \gg \alpha^{-1}$ . В результате получим следующее уравнение для энергии  $\varepsilon$ :

$$\text{Det} \left( \varepsilon \delta_{ij} + \lambda \frac{\exp(-\alpha r_{ij})}{r_{ij}} (1 - \delta_{ij}) \right) = 0, \quad (9.13)$$

где  $\lambda = A/B$ . Для случая двух ям, находящихся на расстоянии  $r$ , оно легко решается:

$$\varepsilon_{1,2} = \pm \lambda \frac{\exp(-\alpha r)}{r}. \quad (9.14)$$

Волновые функции имеют вид

$$\Psi_{1,2} = \frac{1}{\sqrt{2}} (\Psi_0(r - r_1) \pm \Psi_0(r - r_2)), \quad (9.15)$$

где  $\Psi_0(r)$  определяется (9.10). Таким образом, электронные состояния в равной мере принадлежат обеим ямам, а сдвиг уровней экспоненциально падает с расстоянием между ямами.

Теперь рассмотрим три ямы, находящиеся на одной прямой, как это показано на рис. 2.8. Расстояние  $r_{12}$  будем считать фиксированным и посмотрим, как меняются энергии уровней в зависимости от  $x_{23} = r_{23}\alpha$ . Каждый

из трех корней уравнения (9.13) представим в виде

$$\varepsilon = \varepsilon_0 \exp(-s), \quad (9.16)$$

где  $\varepsilon_0$  — степенная функция от  $r_{12}$ ,  $r_{23}$ , которой мы не будем интересоваться, сосредоточив внимание на экспоненциальной зависимости. Зависимость  $s$  от  $x_{23}$  для каждого корня уравнения изображена на рисунке. Если  $r_{23} > r_{12}$ , то состояния классифицируются следующим образом.

Два состояния обобществлены между ямами 1 и 2 и имеют волновые функции (9.15). Их энергии определяются (9.14) и  $s = x_{12}$ . Третья яма на эти состояния практически не влияет. Она имеет собственное состояние с волновой функцией  $\Psi_0(r-r_3)$  и с  $s = 2x_{23}$ . Следует обратить внимание на тот факт, что сдвиг уровня в резонансной ситуации, когда электрон обобществляется между двумя ямами, пропорционален интегралу перекрытия ( $s = x_{12}$ ), а в нерезонансной ситуации — его квадрату ( $s = 2x_{23}$ ).

Совершенно аналогично при  $r_{23} < r_{12}$  возникает дублет, обобществленный ямами 2 и 3

с  $s = x_{23}$ , и состояние в яме 1 с  $s = 2x_{12}$ . Суть дела состоит в том, что если, например, ямы 1 и 2 близки друг к другу, то их состояния обобществляются, а два возникающих при этом уровня так сильно «отталкиваются» друг от друга, что выходят из резонанса с третьим уровнем.

Состояние, которое в равной мере принадлежит всем трем ямам, возникает, только если  $r_{12} = r_{23}$ . Важно понять, что максимально допустимое отклонение от этого равенства, при котором состояния все еще обобществлены, равно радиусу состояния  $\alpha^{-1}$ . Это и является ключом к пониманию ситуации в целом. Дело в том, что среднее расстояние между примесями гораздо больше

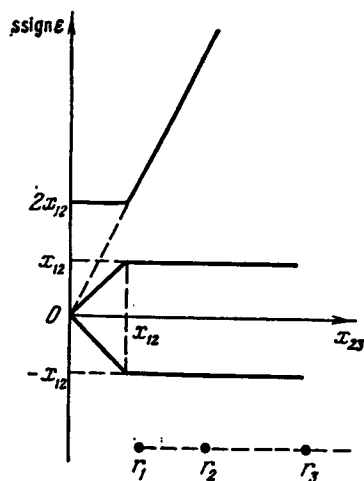


Рис. 2.8. Уровни энергии для трех потенциальных ям как функции расстояния  $r_{23} = x_{23}/a$ .

$\alpha^{-1}$ . По этой причине вероятность упорядоченных конфигураций ничтожно мала и в системе из любого числа ям состояния будут представлять собой синглеты и дублеты с волновыми функциями (9.10) и (9.15). Таким образом, возможность расплывания волнового пакета по зоне совершенно исключена.

Легко понять, как выглядит плотность состояний в модели Лифшица. Очевидно, что характерная энергия разброса уровней (ширина примесной зоны, если потенциальные ямы связаны с примесями) порядка

$$\Delta = \lambda N^{1/3} \exp[-\alpha N^{-1/3}]. \quad (9.17)$$

Эта оценка получается, если подставить в (9.14) вместо  $r$  среднее расстояние между ямами ( $N$  — концентрация ям). Легко также получить асимптотику плотности состояний  $g(\epsilon)$  при  $|\epsilon| \ll \Delta$  и  $|\epsilon| \gg \Delta$ .

Начнем со случая  $|\epsilon| \gg \Delta$ . Состояния с такой энергией возникают за счет пар ям, находящихся на anomalously малом расстоянии ( $r \ll N^{-1/2}$ ) друг от друга. Вероятность того, что в сферическом слое между  $r$  и  $r+dr$  вокруг фиксированной ямы имеется другая яма, равна  $4\pi r^2 N dr$ . Число таких конфигураций в единице объема получится умножением на концентрацию ям  $N$  и делением на 2. (Последнее необходимо, чтобы не учитывать каждую конфигурацию дважды.) Энергия  $\epsilon$  связана с расстоянием  $r$  формулой (9.14). Поэтому для плотности состояний имеем

$$g(\epsilon) = 2\pi r^2 N^2 \frac{dr}{d\epsilon} \approx \frac{2\pi N^2}{\alpha^3} \frac{[\ln(|\epsilon|/\lambda\alpha)]^2}{|\epsilon|}. \quad (9.18)$$

(Мы приближенно решили уравнение (9.14) относительно  $r$ , положив  $r \approx \alpha^{-1} \ln(|\epsilon|/\lambda\alpha)$ .) Формула (9.18) описывает спад плотности состояний при  $|\epsilon| \gg \Delta$  (однако при  $|\epsilon| \ll |E_0|$ ).

Уровни с очень малыми значениями  $|\epsilon|$  возникают в разрежениях, т.е. в таких ямах, вокруг которых отсутствуют другие ямы на расстояниях, значительно превышающих  $N^{-1/2}$ . При этом с подавляющей вероятностью ближайшая соседка рассматриваемой ямы будет иметь ближайшей соседкой другую яму, так что сдвиг уровня будет нерезонансным и определится формулой  $\epsilon = \epsilon_0 \exp(-2\alpha r)$ . Поэтому для существования уровня

с энергией  $\varepsilon$  необходимо отсутствие ям в сфере с радиусом  $r_s \approx (2\alpha)^{-1} \ln(\varepsilon_0/|\varepsilon|)$ . Согласно распределению Пуассона вероятность того, что в объеме  $4\pi r_s^3/3$ , где в среднем имеется  $4\pi N r_s^3/3$  ям, нет ни одной ямы, равна  $\exp(-4\pi N r_s^3/3)$ . Поэтому

$$g(\varepsilon) \propto \exp\left[-\frac{\pi N}{6\alpha^3} (\ln(\varepsilon_0/|\varepsilon|))^3\right] \quad (9.19)$$

и стремится к нулю при  $\varepsilon \rightarrow 0$ . Энергия  $\varepsilon_0$  порядка  $\lambda N^{-1/2}$ . Это видно из того, что при  $\varepsilon$  порядка  $\Delta$  экспоненциальная малость должна исчезнуть.

Таким образом, плотность состояний имеет два пика по разные стороны от энергии  $E_0$ . Заметим также, что плотность состояний, строго говоря, не обращается в нуль при  $\varepsilon = 0$ . Действительно, если радиус  $r_s$  столь велик, что в сферическом слое с объемом  $4\pi r_s^2 \alpha^{-1}$  находится много ям, то сдвиг уровня определяется не одной ямой, а всеми ямами слоя и формула (9.19) не работает.

В заключение еще раз подчеркнем, что вывод о локализации электронных состояний в модели Лифшица является следствием предположения о том, что среднее расстояние между ямами велико по сравнению с радиусом волновой функции,  $\alpha N^{-1/2} \gg 1$ . При  $\alpha N^{-1/2} \approx 1$  делокализованные состояния возникают и в модели Лифшица. В двумерном случае переход от локализованных состояний к делокализованным был прослежен с помощью ЭВМ [38, 39], однако четко определить критическую концентрацию не удалось.



## ГЛАВА 3

# СТРУКТУРА ПРИМЕСНОЙ ЗОНЫ СЛАБО ЛЕГИРОВАННЫХ ПОЛУПРОВОДНИКОВ

### § 10. Общие замечания

Полупроводник называется слабо легированным, если перекрытие электронных состояний, принадлежащих разным примесям, мало. Это означает, что среднее расстояние между примесями значительно превышает характерный размер волновой функции  $a$  или что  $Na^3 \ll 1$ , где  $N$  — концентрация примесей.

Как мы видели в гл. 1, радиусы электронных состояний мелких примесей сильно зависят от того, вблизи какой зоны находится состояние и о каком полупроводнике идет речь. (Перекрытие между глубокими примесями практически всегда мало.) Поэтому диапазон концентраций, соответствующих слабому легированию, сильно меняется от вещества к веществу и даже в одном веществе в зависимости от того, является примесь донором или акцептором.

Имеет смысл определить четкую границу сильного и слабого легирования. Для этого естественно использовать статическую проводимость при низких температурах. Дело в том, что в сильно легированных полупроводниках эта проводимость носит металлический характер (слабо зависит от температуры), а в слабо легированных полупроводниках — активационный характер (экспоненциально обращается в нуль при  $T \rightarrow 0$ ). Концентрацию примесей  $N_m$ , при которой в слабо компенсированном полупроводнике впервые появляется отличная от нуля проводимость при нулевой температуре, разумно назвать границей сильного и слабого легирования.

Например, в  $p$ -Ge  $N_M \approx 10^{17} \text{ см}^{-3}$ . Указанный переход называют переходом „металл — диэлектрик” или, более конкретно, переходом Мотта. Его физическая природа близка к природе перехода, описанного в § 5, однако отличается тем, что расположение примесей в легированном полупроводнике неупорядочено. Как мы увидим ниже, электронные состояния в диэлектрической фазе являются по этой причине локализованными. В этом и состоит основное отличие рассматриваемого перехода от перехода, описанного в § 5.

Заметим, что слабая компенсация упомянута в определении не случайно. При сильной компенсации переход металл — диэлектрик происходит при значительно больших концентрациях примесей и является переходом Андерсона. Этому вопросу посвящена гл. 13. В первой части книги речь пойдет только о полупроводниках с малой концентрацией примесей ( $N \ll N_M$ ).

Рассмотрим сначала случай, когда имеются примеси лишь одного типа. Для определенности будем говорить в этой главе только о полупроводниках  $n$ -типа, хотя все сказанное относится и к полупроводникам  $p$ -типа. Переход от одного случая к другому в рассуждениях и формулах осуществляется простой заменой символов. Итак, если есть только доноры, то при нулевой температуре на каждом из них будет находиться электрон. Отсутствие электрона на одном из доноров будет представлять собой элементарное возбуждение.

Естественно поставить вопрос о спектре возбуждений и о плотности состояний. При этом необходимо учитывать, что возбуждение может находиться на любом из доноров. Разброс энергетических уровней связан с двумя обстоятельствами. Во-первых, когда с одного из доноров удаляется электрон, оставшееся положительно заряженное ядро поляризует нейтральные доноры, расположенные в его окрестности. Этот механизм дает вклад в энергию ионизации донора, вклад, который меняется в зависимости от конфигурации соседних нейтральных доноров. Во-вторых, имеется квантовое перекрытие между волновыми функциями, соответствующими возбуждениям, находящимся на разных донорах. Если пренебречь первой причиной, то мы приходим к модели И. М. Лифшица, рассмотренной в § 9. Ширина возникающей при

этом примесной зоны дается формулой (9.17). Весьма существенно, что возбуждения оказываются строго локализованными. При нулевой температуре дырка, находящаяся на одном из доноров, не может перейти на другие доноры. Электропроводность при нулевой температуре строго равна нулю.

Следует, однако, иметь в виду, что в полупроводниках область применимости модели Лифшица ограничена. Пренебрежение поляризацией нейтральных доноров может быть оправдано, если концентрация примесей не очень мала. Однако в пределе малых концентраций это пренебрежение недопустимо, так как ширина зоны, связанная с перекрытием, как видно из (9.17), обращается в нуль экспоненциально. В то же время разброс уровней, связанный с нейтральными донорами, убывает с концентрацией по степенному закону.

В полупроводниках с компенсирующими примесями главной причиной разброса уровней являются не нейтральные, а заряженные примеси. Допустим, что, кроме доноров, есть акцепторы. При низкой температуре каждый акцептор забирает электрон у донора и заряжается отрицательно. Таким образом, кроме нейтральных доноров, возникают положительно заряженные доноры и равное им количество отрицательно заряженных акцепторов. Эти случайно расположенные заряды создают флуктуирующий потенциал и разброс уровней, значительно больший, чем экспоненциально малый разброс, связанный с перекрытием, и значительно больший, чем разброс, создаваемый наведенными дипольными моментами. Обсудим теперь основные свойства примесной зоны компенсированного полупроводника.

В предыдущей главе мы видели, что разброс уровней всегда способствует локализации электронных состояний. Поэтому сделанное выше заключение остается в силе: электронные состояния в примесной зоне слабо легированного полупроводника являются строго локализованными.

Энергию электронного состояния на доноре  $i$  мы будем обозначать  $\epsilon_i$  и отсчитывать от энергии изолированного донора  $-E_0$ . Величина  $\epsilon_i$  представляет собой потенциальную энергию электрона в поле всех прочих

заряженных доноров и акцепторов:

$$\varepsilon_i = \frac{e^2}{\kappa} \left[ \sum_t^{\text{акц}} \frac{1}{|r_t - r_i|} - \sum_{k \neq i}^{\text{дон}} \frac{(1 - n_k)}{|r_i - r_k|} \right]. \quad (10.1)$$

В первом члене суммирование ведется по координатам акцепторов, а во втором — по координатам доноров. Для заполненных доноров  $n_i=1$ , а для пустых доноров  $n_i=0$ . При нулевой температуре набор чисел заполнения  $\{n_i\}$  определяется минимумом полной электростатической энергии.

Так же, как в теории металлов, имеет смысл ввести понятие об энергии (уровне) Ферми  $\mu$ . В основном состоянии все доноры с энергиями  $\varepsilon_i < \mu$  должны быть заполненными, а с энергиями  $\varepsilon_i > \mu$  — пустыми.

Обсудим теперь вопрос о положении уровня Ферми в примесной зоне при нулевой температуре. Энергия Ферми  $\mu$  как функция степени компенсации  $K$  представлена на рис. 3.1. Она также отсчитана

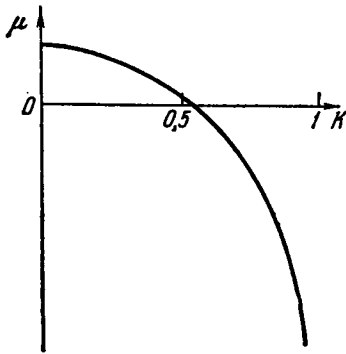


Рис. 3.1. Положение уровня Ферми относительно уровня изолированной примеси  $-E_0$  как функция степени компенсации  $K$ .

от уровня изолированной примеси. Степенью компенсации принято называть отношение концентраций неосновных и основных примесей. В нашем случае  $K = N_A/N_D$ , где  $N_D$  — концентрация доноров, а  $N_A$  — концентрация акцепторов. По определению  $0 \leq K \leq 1$ .

При малых значениях  $K$  почти все доноры имеют электрон и, следовательно, уровень Ферми находится в верхней части примесной зоны (см. рис. 3.3). По этой причине энергия Ферми, отсчитанная от  $-E_0$ , положительна. Как будет показано в § 11, при  $K \rightarrow 0$  она стремится к конечному пределу, определяемому концентрацией доноров (формула (11.11)). При  $K \rightarrow 1$  электроны имеются лишь на самых глубоких донорах (см. рис. 3.4). При этом энергия Ферми отрицательна и, как будет показано в § 13, ведет себя, как  $-(1-K)^{-1/4}$ .

Детальному изучению структуры примесной зоны посвящены последующие параграфы этой главы. Здесь мы сделаем лишь несколько замечаний относительно изложенной ниже теории.

Как уже говорилось, разброс уровней в примесной зоне связан с потенциалом заряженных доноров и акцепторов. С другой стороны, вопрос о том, какой из доноров пуст и, следовательно, заряжен, определяется потенциалом, который создают на нем другие доноры. Отсюда должно быть ясно, что рассматриваемая задача не может быть одноэлектронной. В последующих параграфах показано, что при малых и больших степенях компенсации потенциал определяется только ближайшим окружением примесного центра и крупномасштабными флуктуациями плотности примесного заряда. (Промежуточные масштабы практически не играют роли.)

Крупномасштабные флуктуации возникают потому, что концентрация носителей (дырок при малой и электронов при большой степени компенсации) мала, вследствие чего велик радиус экранирования. Потенциал ближайшего окружения учитывается с помощью простых комбинаторных вычислений, а крупномасштабный потенциал исследуется методом самосогласованного поля, который применим в этих случаях всюду, за исключением малой области энергий, непосредственно примыкающей к энергии Ферми.

В случае промежуточной компенсации метод самосогласованного поля неприменим. В гл. 10 мы увидим, что в этом случае электрон-электронное взаимодействие создает «мягкую щель» в плотности состояний вблизи уровня Ферми, приводящую к ряду интересных явлений. Щель возникает и в описанных выше предельных случаях ( $K \rightarrow 0$ ,  $K \rightarrow 1$ ), однако там ее энергетическая ширина мала по сравнению с полной шириной примесной зоны.

В заключение этого параграфа остановимся на вопросе о характере распределения примесей в полупроводнике. Это — один из наиболее сложных и неизученных вопросов теории легированных полупроводников. Простейшее предположение, из которого мы будем исходить в первую очередь, состоит в том, что это распределение

является пуассоновским, т. е. корреляция в распределении примесей полностью отсутствует. Ясно, что это предположение не может быть правильным во всех случаях. Например, если образец получен вытягиванием из расплава, то корреляция в расположении примесей устанавливается при таких температурах, когда они в состоянии диффундировать, а затем фиксируется при закалке. Отсюда видно, что корреляция, вообще говоря, существует и существенно зависит от способа приготовления образца. Точный учет этой корреляции весьма сложен. Поэтому, как правило, мы будем исходить из пуассоновского распределения и специально отметим те результаты, которые в существенной мере основаны на предположении об отсутствии корреляции.

### § 11. Примесная зона в случае малой степени компенсации

При малой компенсации ( $K = N_A/N_D \ll 1$ ) большая часть доноров заполнена электронами и лишь небольшая часть свободна и имеет положительный заряд. Свободный донор называется дыркой. Концентрация заряженных доноров (дырок) в точности равна концентрации акцепторов, каждый из которых имеет отрицательный заряд. Электроны могут перемещаться по донорам, вследствие чего будет двигаться дырка. (Следует четко понимать, что доноры и акцепторы в нашей схеме считаются неподвижными, и когда мы говорим о движении дырки, то под этим понимается перемещение электронов, при котором остается свободным то тот, то другой донор.) При нулевой температуре распределение дырок будет соответствовать минимуму энергии при заданной конфигурации доноров и акцепторов. Благодаря малому параметру  $K \ll 1$  оказалось возможным построить количественную теорию, описывающую это состояние [1].

В основном состоянии положительно заряженная дырка должна находиться как можно ближе к отрицательному акцептору. (Иными словами, отрицательные акцепторы «сдувают» электроны с соседних доноров.) Таким образом, дырка стремится занять ближайший к акцептору донор. Акцепторов столько же, сколько дырок, и гораздо меньше, чем доноров, по которым могут

странствовать дырки. По этой причине в основном состоянии все дырки окажутся связанными вблизи акцепторов. Поскольку акцепторов мало, в первом приближении можно считать, что каждый из них погружен в бесконечное море доноров. Для переноса дырки с ближайшего к акцептору донора на бесконечно удаленный от него донор следует совершить работу, которая в среднем окажется порядка  $e^2/\kappa r_D$ , где  $r_D$  — среднее расстояние между донорами. Эта работа по порядку величины и определяет энергию связи дырок вблизи акцепторов.

Приведенные выше соображения были высказаны Моттом [2] в 1956 г. Однако, как мы увидим ниже, их недостаточно для построения количественной теории. Изложенная схема требует некоторого дополнения. Оказывается, что в основном состоянии вблизи одних акцепторов вообще нет дырок, зато существуют акцепторы, удерживающие сразу две дырки. Вспомним, что при пуассоновском распределении некоторые акцепторы не имеют близко расположенных доноров, так что приписанные к таким акцепторам дырки были бы связаны с ними очень слабо. Из условия минимума полной энергии следует, что каждая дырка должна предпочитать такому акцептору другой акцептор, имеющий более выгодную конфигурацию соседних доноров, и стать второй дыркой в его окружении. При этом она будет испытывать отталкивание со стороны первой дырки, однако при удачном взаимном расположении притяжение к акцептору окажется более сильным.

Действительно, вычислим энергию связи дырки в конфигурации, изображенной на рис. 3.2, а. Она равна работе, необходимой для переноса дырки с одного из доноров на большое расстояние от акцептора, где мы считаем электростатический потенциал равным нулю. Это эквивалентно тому, чтобы снять электрон с одного из удаленных доноров и посадить его на один из рассматриваемых доноров. Такая работа равна сдвигу уровня того электрона, который окажется на месте одной из дырок, в поле, создаваемом акцептором и второй дыркой. Так как радиус электронного состояния мы считаем малым по сравнению с расстояниями между примесями, сдвиг уровня электрона равен потенциалу, создаваемому

акцептором и второй дыркой в месте нахождения первой дырки, помноженному на заряд электрона.

Итак, энергия связи равна  $\frac{e^2}{\kappa r} - \frac{e^2}{2\kappa r} = \frac{e^2}{2\kappa r}$  ( $\kappa$  — диэлектрическая проницаемость). Мы видим, что при малых  $r$  она может быть очень большой. Таким образом, рассмотренная конфигурация, несомненно, может конкурировать с конфигурацией, в которой вблизи акцептора есть только одна дырка. В частности, место второй

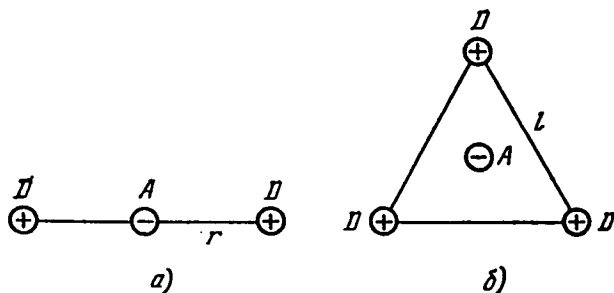


Рис. 3.2. а) Акцептор с двумя заряженными донорами; б) конфигурация трех заряженных доноров вокруг акцептора, оптимальная с точки зрения близости доноров к акцептору и удаленности друг от друга.

дырки в такой конфигурации с точки зрения минимума энергии выгоднее, чем место единственной дырки вблизи акцептора, у которого ближайший донор находится на расстоянии, большем чем  $2r$ .

Далее, естественно, возникает вопрос: существуют ли акцепторы, обладающие еще большим количеством дырок? Ответ на него отрицательный. Более двух дырок акцептор удерживать не может. Рассмотрим случай трех дырок. Экстремальной в смысле близости всех трех дырок к акцептору и удаленности их друг от друга является конфигурация равностороннего треугольника, в углах которого находятся дырки, а в центре — акцептор (рис. 3.2, б). Для того чтобы исследовать эту конфигурацию, нужно подсчитать работу, необходимую для переноса одной из дырок на бесконечность. Для этого нужно вычислить потенциал, создаваемый на одной из дырок двумя другими дырками и акцептором. Энергия притяжения дырки к акцептору равна  $\sqrt{3} e^2/\kappa l$ , где  $l$  — сторона треугольника. Энергия отталкивания от двух



других дырок равна  $2e^2/kl$ . Таким образом, отталкивание сильнее притяжения и такая конфигурация должна развалиться.

Итак, один акцептор не может связать три дырки. Очевидно, что он не может связать и большее число дырок. Это обстоятельство и позволяет построить количественную теорию. Действительно, мы получили простую классификацию возможных конфигураций (рис. 3.3).

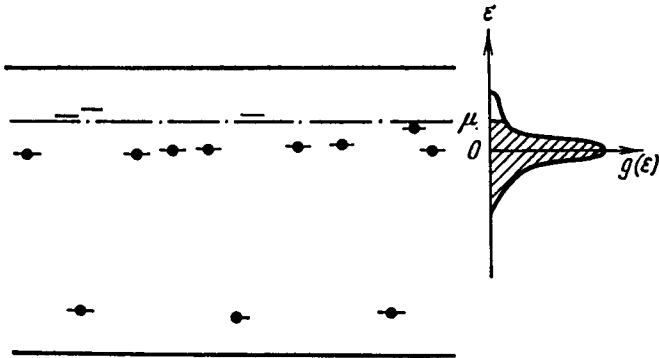


Рис. 3.3. Энергетическая схема слабо компенсированного полупроводника. Сплошные линии изображают дно зоны проводимости и потолок валентной зоны. Черными кружками изображены электроны, занимающие донорные и акцепторные уровни. Штрих-пунктирная линия показывает положение уровня Ферми. Первый слева акцептор образует 2-комплекс, второй — 1-комплекс, третий — 0-комплекс. Справа показана плотность состояний на донорных уровнях. Занятые состояния заштрихованы.

В зависимости от расположения доноров вблизи данного акцептора реализуются три ситуации. Во-первых, может оказаться, что около заданного акцептора нет ни одного ионизованного донора (0-комплекс). Это может быть в том случае, если вокруг акцептора доноры случайно образовали разрежение, так что добавка к энергии электрона  $\epsilon$ , создаваемая отрицательным акцептором на ближайшем к нему доноре, не превосходит энергии Ферми  $\mu$ . Во-вторых, вблизи акцептора может быть ионизован один донор (1-комплекс). Такая ситуация возможна, когда добавка к энергии, создаваемая акцептором на ближайшем доноре, больше  $\mu$ , а добавки, создаваемые совместным действием акцептора и ионизованного донора на всех остальных донорах, меньше  $\mu$ .

В-третьих, вблизи акцептора могут быть два ионизованных донора (2-комплекс), причем они расположены так, что энергия уровня на каждом из них поднята выше  $\mu$  за счет того, что потенциал, создаваемый акцептором, значительно превосходит потенциал, создаваемый вторым ионизованным донором. Такой комплекс может возникнуть, если два донора расположены близко от акцептора, но по разные стороны от него.

Из сказанного ясно, что 0-комплексы отрицательно заряжены, 1-комплексы нейтральны, а 2-комплексы обладают положительным зарядом. Поэтому условие нейтральности, которое формулировалось как «один ионизованный донор на каждый акцептор», может быть сформулировано и как условие равенства концентраций 0- и 2-комплексов:

$$N_0(\mu) = N_2(\mu). \quad (11.1)$$

С помощью этого условия мы и будем находить уровень Ферми  $\mu$ . Концентрация  $N_0(\mu)$  легко вычисляется. Действительно, 0-комплекс существует, когда на расстоянии  $r_\mu = e^2/\kappa\mu$  от фиксированного акцептора нет ни одного донора. Вероятность этого события  $\exp\left(-\frac{4\pi}{3} N_D r_\mu^3\right)$ . Таким образом, для  $N_0(\mu)$  имеем

$$N_0(\mu) = N_A \exp\left(-\frac{4\pi}{3} \frac{e^6 N_D}{\kappa^3 \mu^3}\right). \quad (11.2)$$

Точное вычисление  $N_2(\mu)$  затруднительно. Поэтому вместо концентрации  $N_2(\mu)$  будем использовать две близкие к ней величины  $N_2^>(\mu)$  и  $N_2^<(\mu)$ , берущие ее в «вилку»:

$$N_2^>(\mu) > N_2(\mu) > N_2^<(\mu).$$

Мы покажем, что эта «вилка» настолько узка, что соответствующая неопределенность в значении  $\mu$  меньше 1%. Величина  $N_2^>(\mu)$  представляет собой полную концентрацию пар доноров, у которых в ионизованном состоянии обе добавки к энергиям уровней  $\varepsilon_1$  и  $\varepsilon_2$  превышают  $\mu$  (будем называть такую пару  $\mu$ -парой). Почти всегда наличие  $\mu$ -пары вблизи акцептора означает существование 2-комплекса. Исключение составляет случай, когда, кроме  $\mu$ -пары, около акцептора имеется еще один

очень близкий донор, так что энергетически оказывается выгоднее 1-комплекс с этим донором, чем 2-комплекс с парой. Другая причина, по которой величина  $N_2^>(\mu)$  несколько превышает  $N_2(\mu)$ , заключается в том, что около акцептора, вообще говоря, может оказаться несколько  $\mu$ -пар доноров. Величина  $N_2^>(\mu)$  включает все такие пары. Даже пары, имеющие один общий донор, считаются двумя разными парами. На самом же деле вблизи акцептора будет ионизована только одна наиболее энергетически выгодная  $\mu$ -пара, которая и образует 2-комплекс.

Вычислим  $N_2^>(\mu)$ . Поместим акцептор в начале координат и допустим, что в точке  $r_1$  находится донор. Вычислим среднее число доноров, образующих  $\mu$ -пару с этим донором и находящихся в любых точках  $r_2$  таких, что  $r_2 > r_1$ . Среднее число доноров в элементе  $dr_2$  есть  $N_D dr_2$ . Поэтому искомая величина равна

$$N_D \int_{r_2 > r_1} dr_2 \theta \{ \varepsilon_1(r_1, r_2) - \mu \} \theta \{ \varepsilon_2(r_1, r_2) - \mu \}.$$

Здесь  $\varepsilon_1$  и  $\varepsilon_2$  — добавки к энергиям электронов на донорах за счет потенциала акцептора и второго донора:

$$\varepsilon_1 = \frac{e^2}{\kappa |r_1|} - \frac{e^2}{\kappa |r_1 - r_2|}, \quad \varepsilon_2 = \frac{e^2}{\kappa |r_2|} - \frac{e^2}{\kappa |r_1 - r_2|}. \quad (11.3)$$

Как уже говорилось, мы считаем, что характерные значения  $|r_1|$ ,  $|r_2|$  и  $|r_1 - r_2|$  велики по сравнению с борновским радиусом электрона на доноре, так что сдвиг уровня можно считать классическим. Функция  $\theta(x) = 1$  при  $x > 0$  и  $\theta(x) = 0$  при  $x < 0$ . Умножив найденное среднее число доноров на величину  $N_D dr_1$ , представляющую собой вероятность нахождения донора в элементе  $dr_1$ , получим среднее число  $\mu$ -пар с одним из доноров в элементе  $dr_1$ . Проинтегрировав по всему объему, получим среднее число  $\mu$ -пар в окрестности данного акцептора. Наконец, умножив получившуюся величину на  $N_A$ , получим для концентрации  $\mu$ -пар

$$\begin{aligned} N_2^>(\mu) &= \\ &= N_A N_D^2 \int dr_1 \int_{r_2 > r_1} dr_2 \theta \{ \varepsilon_1(r_1, r_2) - \mu \} \theta \{ \varepsilon_2(r_1, r_2) - \mu \}. \end{aligned} \quad (11.4)$$

Вторая введенная нами величина,  $N_2^<(\mu)$ , представляет собой концентрацию  $\mu$ -пар, образованных двумя ближайшими к своему акцептору донорами. Докажем, что  $N_2^<(\mu) < N_2(\mu)$ . Прежде всего заметим, что  $\mu$ -пара из ближайших доноров обязательно является 2-комплексом. Однако наличие такой  $\mu$ -пары есть лишь достаточное, но не необходимое условие существования 2-комплекса: 2-комплекс может существовать, например, если два ближайших донора вообще не составляют  $\mu$ -пару, но  $\mu$ -пара образована первым и третьим донором. Эта ситуация возможна, когда первый и второй ближайшие доноры находятся по одну сторону от акцептора, а третий — по другую. Таким образом,  $N_2^<(\mu) < N_2(\mu)$ .

Вероятность того, что два ближайших к акцептору донора находятся в интервалах  $dr_1$  и  $dr_2$ , равна

$$N_D^2 \exp \left\{ -\frac{4\pi}{3} N_D r_1^3 - \frac{4\pi}{3} N_D (r_2^3 - r_1^3) \right\} dr_1 dr_2.$$

Отсюда для  $N_2^<(\mu)$  мы имеем

$$N_2^<(\mu) = N_A N_D^2 \int dr_1 \int_{r_1 < r_2} dr_2 \exp \left( -\frac{4\pi}{3} N_D r_2^3 \right) \times \\ \times \theta \{ \epsilon_1(r_1, r_2) - \mu \} \theta \{ \epsilon_2(r_1, r_2) - \mu \}, \quad (11.5)$$

где интегралы берутся так же, как в (11.4), по окрестности одного акцептора. Вычисление (11.4) и (11.5) удобно проводить, вводя сферические системы координат: для  $r_1$  — с полярной осью, направленной по  $r_2$ , а для  $r_2$  — с произвольной осью. Тогда подынтегральные выражения (11.4) и (11.5) не зависят от переменных  $\theta_2$ ,  $\varphi_2$ ,  $\varphi_1$ . Для дальнейшего вычисления удобно ввести безразмерные переменные

$$z = \left( \frac{4\pi}{3} \right)^{1/3} \frac{e^2 N_D^{1/3}}{\kappa \mu}, \quad (11.6) \\ x_1 = \left( \frac{4\pi}{3} \right)^{1/3} N_D^{1/3} r_1, \quad x_2 = \left( \frac{4\pi}{3} \right)^{1/3} N_D^{1/3} r_2.$$

В этих переменных  $\theta$ -функции в (11.4) и (11.5) приводят к следующим пределам интегрирования по  $\theta_1$ ,  $x_1$ ,  $x_2$

(геометрическим требованиям к паре):

$$\frac{(x_1^2 + x_2^2)(z - x_2)^2 - z^2 x_2^2}{2x_1 x_2 (z - x_2)^2} > \cos \theta_1 > -1, \quad (11.7)$$

$$x_2 > x_1 > \frac{x_2^2}{z - x_2}, \quad \frac{z}{2} > x_2 > 0.$$

Вычисление трехкратного интеграла (11.4) в пределах (11.7) оказывается очень громоздким, но весьма простым и доводится до конца:

$$N_2^>(z) = 7,14 \cdot 10^{-4} z^6 N_A. \quad (11.8)$$

При интегрировании (11.5) удается взять лишь два интеграла:

$$N_2^<(z) = 3z^6 N_A \int_0^{0.5} y^5 \exp(-y^3 z^3) \left[ 1 + \frac{8y - 5}{16(1 - y)^4} \right] dy. \quad (11.9)$$

В переменных (11.6) выражение (11.2) для концентрации имеет вид

$$N_0(z) = N_A \exp(-z^3). \quad (11.10)$$

Приравнивая (11.10) по очереди (11.8) и (11.9), получим два значения  $z$ . Вычисление интеграла (11.9) и решение трансцендентных уравнений производились на ЭВМ [1]. Оказалось, что первое значение  $z_1 = 1,63$ , второе  $-z_2 = 1,65$ . Поскольку  $N_2^>(z) > N_2(z) > N_2^<(z)$ , правильное значение  $z$  лежит между  $z_1$  и  $z_2$ . Переходя к размерным величинам, получаем, что при  $K \rightarrow 0$  с точностью, превышающей 1%,

$$\mu = 0,61 \varepsilon_D = 0,99 e^2 N_D^{1/3} / \kappa, \quad (11.11)$$

где

$$\varepsilon_D = e^2 / \kappa r_D \quad (11.12)$$

— энергия кулоновского взаимодействия на среднем расстоянии между донорами,  $r_D = \left( \frac{4\pi}{3} N_D \right)^{-1/3}$ . Малое отличие величин  $z$ , найденных двумя способами, связано с тем, что при значении  $\mu$  (11.11) те ситуации, за счет которых  $N_2^>(\mu)$  и  $N_2^<(\mu)$  отличаются от  $N_2(\mu)$ , очень редки.

Подставляя (11.11) в (11.2), получим, что концентрация 0-комплексов реально составляет всего 1,3% от полной концентрации акцепторов. Столько же имеется 2-комплексов. Остальные 97,4% акцепторов образуют нейтральные 1-комплексы.

Подобные цифры дают основание думать, что в первом приближении 0- и 2-комплексы вообще можно было не учитывать. Однако это неверно. Если пренебречь ими и считать, что каждый акцептор образует 1-комплекс, то среди этих 1-комплексов встретится и такой, в котором доноры образуют разрежение вокруг акцептора, а добавка к энергии  $\epsilon$  ближайшего к акцептору донора близка к нулю. Но поскольку этот донор ионизован (1-комплекс!), это означает, что  $\mu = 0$ , а вовсе не  $0,61\epsilon_D$ . Таким образом, пренебрежение 0- и 2-комплексами при вычислении уровня Ферми недопустимо.

Задолго до появления излагаемых здесь работ [1] положение уровня Ферми при  $K \ll 1$  исследовалось Миллером и Абрахамсом в связи с предложенной ими теорией прыжковой проводимости. Их результат сильно отличается от (11.11), и нам представляется важным обсудить причины этого отличия, Миллер и Абрахамс рассуждали следующим образом. Вероятность того, что ближайший к донору акцептор расположен на расстоянии от  $r$  до  $r + dr$ , равна

$$p(r) dr = \frac{3r^2}{r_A^3} \exp \left[ - \left( \frac{r}{r_A} \right)^3 \right] dr, \quad (11.13)$$

где  $r_A = (3/4\pi N_A)^{1/3}$  — среднее расстояние между акцепторами. Акцептор сдвигает энергию электрона на доноре на величину  $\epsilon = e^2/\kappa r$ . Поэтому Миллер и Абрахамс считали, что распределение доноров по энергиям имеет вид

$$p(\epsilon) d\epsilon = \frac{3\epsilon_A^3}{\epsilon^4} \exp \left[ - \left( \frac{\epsilon_A}{\epsilon} \right)^3 \right] d\epsilon, \quad (11.14)$$

где  $\epsilon_A = e^2/\kappa r_A$ . Уровень Ферми они получили, приравняв два выражения для вероятности того, что произвольный донор ионизован:

$$\int_{\mu}^{\infty} p(\epsilon) d\epsilon = K. \quad (11.15)$$

Этот метод представляется нам непоследовательным. При выводе (11.14) предполагалось, что на любой донор из окружения акцептора действует кулоновский потенциал. Это неверно, поскольку один из этих доноров, как правило, ионизован и на все остальные действует лишь потенциал диполя.

Итак, мы определили положение уровня Ферми. Из сказанного ясен и общий характер плотности состояний  $g(\epsilon)$  (см. рис. 3.3). Большинство доноров находится далеко от редко расположенных акцепторов, и их уровни мало сдвинуты и образуют острый пик вблизи  $-E_0$ . Вверх от уровня  $-E_0$  простирается хвост  $g(\epsilon)$ . Можно оценить величину  $g(\epsilon)$  вблизи уровня Ферми. Над уровнем Ферми находится  $N_A$  состояний, а характерной энергией спада в хвосте  $g(\epsilon)$  является величина  $\epsilon_D$ . Таким образом, вблизи уровня Ферми

$$g(\epsilon) \approx \frac{N_A}{\epsilon_D}. \quad (11.16)$$

Перейдем теперь к более подробному описанию структуры основного пика вблизи энергии уровня  $-E_0$ .

## § 12. Крупномасштабный потенциал при малой степени компенсации

В предыдущих рассуждениях мы полностью пренебрегли электростатическим взаимодействием между различными комплексами. Основанием было то, что расстояния между комплексами ( $N_i^{-1/3}$ ) значительно больше размеров комплексов ( $N_D^{-1/3}$ ). Тем не менее электростатические поля заряженных 0- и 2-комплексов приводят к разбросу энергий в основном пике вблизи  $-E_0$ . При достаточно слабом легировании этот разброс всегда оказывается значительно больше разброса, связанного с перекрытием состояний, вследствие чего структура основного пика, как правило, не описывается моделью Лифшица (см. § 10).

Если учитывать взаимодействие между комплексами, то вопрос о том, какой комплекс реализуется около данного акцептора, решается только характером случайной конфигурации доноров вокруг него. Поэтому расположение

заряженных комплексов в пространстве оказывается абсолютно случайным, т. е. напоминает идеальную плазму. Плотность заряда комплексов имеет флуктуации всех масштабов, начиная от среднего расстояния между комплексами. Легко понять, что флуктуации такого масштаба создают в пространстве потенциал порядка  $e^2 N_0^{1/3} / \kappa$ .

Сейчас мы покажем, что крупномасштабные гауссовы флуктуации создают значительно больший потенциал. Напомним, что флуктуации называются гауссовыми, если избыточное число частиц в рассматриваемом объеме значительно меньше среднего числа частиц в этом объеме. По этой причине масштаб гауссовых флуктуаций должен быть непременно больше, чем среднее расстояние между частицами (в нашем случае между комплексами), а флуктуирующую концентрацию можно считать плавной функцией координат. Пусть  $N_0(\mathbf{r})$  и  $N_2(\mathbf{r})$  — соответственно концентрации 0- и 2-комплексов в точке  $\mathbf{r}$ . Введем величины

$$\xi_0(\mathbf{r}) = N_0(\mathbf{r}) - \langle N_0(\mathbf{r}) \rangle, \quad \xi_2(\mathbf{r}) = N_2(\mathbf{r}) - \langle N_2(\mathbf{r}) \rangle, \quad (12.1)$$

где символ  $\langle \dots \rangle$  означает усреднение,  $\langle N_0(\mathbf{r}) \rangle = \langle N_2(\mathbf{r}_2) \rangle = N_0$ . Математическим выражением того, что 0- и 2-комплексы флуктуируют независимо, являются равенства [4]

$$\begin{aligned} \langle \xi_0(\mathbf{r}) \xi_0(\mathbf{r}') \rangle &= N_0 \delta(\mathbf{r} - \mathbf{r}'), \\ \langle \xi_2(\mathbf{r}) \xi_2(\mathbf{r}') \rangle &= N_0 \delta(\mathbf{r} - \mathbf{r}'), \\ \langle \xi_0(\mathbf{r}) \xi_2(\mathbf{r}') \rangle &= 0, \end{aligned} \quad (12.2)$$

причем более высокие корреляторы распадаются на парные, так как речь идет о гауссовых флуктуациях. Нас интересуют флуктуации заряда

$$\rho(\mathbf{r}) = e(N_2(\mathbf{r}) - N_0(\mathbf{r})) = e(\xi_2(\mathbf{r}) - \xi_0(\mathbf{r})). \quad (12.3)$$

С помощью (12.2) получаем

$$\langle \rho(\mathbf{r}) \rho(\mathbf{r}') \rangle = 2e^2 N_0 \delta(\mathbf{r} - \mathbf{r}'), \quad (12.4)$$

а более высокие корреляторы также распадаются на парные.

Однако сказанного выше еще недостаточно для того, чтобы описать флуктуации потенциала. Рассмотрим шар радиуса  $R$ . В таком шаре в среднем находится  $N_0 R^3$



комплексов (численными коэффициентами мы пренебрегаем). Среднеквадратичная флуктуация числа комплексов в шаре порядка  $(N_0 R^3)^{1/2}$ . Следовательно, типичный заряд шара порядка  $e(N_0 R^3)^{1/2}$ , а потенциал, который в нем возникает, порядка  $e(N_0 R^3)^{1/2}/R$ . Этот потенциал расходится при  $R \rightarrow \infty$ . Итак, в нашем приближении потенциал крупномасштабных флуктуаций оказывается бесконечно большим. Эта расходимость исчезает, если учесть экранирование.

Ниже мы увидим, что результирующий потенциал мало меняется на среднем расстоянии между комплексами. Это позволяет воспользоваться приближением самосогласованного поля. Мысленно разобьем систему на отдельные блоки, каждый из которых является макроскопическим в том смысле, что содержит много комплексов, однако настолько малым, что потенциал крупномасштабных флуктуаций  $\varphi(r)$  можно считать внутри него постоянным. Внутри каждого блока можно определить плотность состояний, которую будем теперь называть локальной. Она совпадает с вычисленной в предыдущем параграфе, но сдвинута по оси энергий на величину  $-e\varphi(r)$  и, таким образом, зависит от координаты. (Под точками  $r$  можно понимать, например, точки, являющиеся геометрическими центрами блоков.) Блоки находятся в термодинамическом равновесии друг с другом, так что энергия Ферми всюду одинакова, но зато флуктуируют концентрации дырок.

Концентрации 0- и 2-комплексов в точке  $r$  соответственно равны  $N_0(\mu + e\varphi(r))$  и  $N_2(\mu + e\varphi(r))$ . Теперь вспомним, что равенство  $N_0(\mu) = N_2(\mu)$ , с помощью которого мы находили  $\mu$ , выполняется не при всех значениях аргументов, а лишь при выполнении (11.11). Легко убедиться, что в местах с положительным значением  $\varphi$  оказывается больше отрицательных 0-комплексов и, наоборот, в местах с отрицательными  $\varphi$  — больше 2-комплексов. Таким образом и происходит экранирование потенциала  $\varphi$ . Математическим выражением сказанного выше является уравнение Пуассона для флуктуирующего потенциала  $\varphi(r)$ :

$$\Delta\varphi = -\frac{4\pi}{\kappa} \rho(r) - \frac{4\pi e}{\kappa} [N_2(\mu + e\varphi) - N_0(\mu + e\varphi)]. \quad (12.5)$$

Здесь  $\rho(r)$  — случайная гауссова функция с коррелятором (12.4), а функции  $N_0$  и  $N_2$  выражаются формулами (11.2) и (11.8) или (11.9).

В этом месте может возникнуть следующий вопрос. Почему мы сначала считали, что 0- и 2-комплексы флуктуируют независимо друг от друга, а затем ввели их корреляцию, которая и представляет собой экранирование? Почему вообще эти два явления можно считать независимыми? Ведь, в конце концов, и флуктуируют и экранируют те же самые комплексы! Ответ состоит в следующем. Такая процедура правильна, если для характерного значения  $\varphi$  справедливо неравенство

$$|\epsilon\varphi(r)| \ll \mu. \quad (12.6)$$

При этом экранирование осуществляется комплексами, находящимися в узкой полосе энергий, близкой к уровню Ферми, а флуктуации заряда  $\rho(r)$  связаны со всеми комплексами и эти флуктуации можно считать некоррелированными.

Неравенство (12.6) позволяет линеаризовать уравнение (12.5):

$$\Delta\varphi = -\frac{4\pi\rho}{\kappa} + \frac{\varphi}{r_0^2}. \quad (12.7)$$

Его решение имеет вид

$$\epsilon\varphi(r) = \int \rho(r') K(r-r') dr', \quad (12.8)$$

$$K(r) = \frac{e \exp(-r/r_0)}{r\kappa},$$

где радиус экранирования  $r_0$  определяется формулой

$$r_0^{-2} = \frac{4\pi e}{\kappa} \frac{d}{d\varphi} [N_0(\mu + \epsilon\varphi) - N_2(\mu + \epsilon\varphi)]_{\varphi=0}. \quad (12.9)$$

Производные в (12.9) вычисляются с помощью (11.2) и (11.8) или (11.9). При этом разница в значениях  $r_0$ , даваемая (11.8) и (11.9), оказывается меньше 2%. Взяв среднее значение, получим

$$r_0 = 0,58 N_A^{-1/2} N_D^{1/6}. \quad (12.10)$$

Легко получить функцию распределения потенциала. Как видно из (12.8), потенциал связан линейно со слу-

чайной гауссовой величиной  $\rho(r)$ . Следовательно,  $\varphi(r)$  также является случайной гауссовой величиной с функцией распределения

$$F(e\varphi) = \frac{1}{\gamma\sqrt{\pi}} \exp\left(-\frac{(e\varphi)^2}{\gamma^2}\right), \quad (12.11)$$

где  $\gamma^2$  есть, по определению, удвоенный средний квадрат потенциальной энергии:

$$\gamma^2 = 2e^2 \langle \varphi^2(r) \rangle. \quad (12.12)$$

Подставляя (12.8) в (12.12) и пользуясь (12.4), найдем

$$\begin{aligned} \gamma^2 &= 2 \int dr' \int dr'' \langle \rho(r') \rho(r'') \rangle K(r-r') K(r-r'') = \\ &= 4e^2 N_0 \int K^2(r) dr, \\ \gamma &= 2 \sqrt{2\pi} \frac{e^2}{\pi r_0} (N_0 r_0^3)^{1/2} = 0,26 e_D (N_A/N_D)^{1/4}. \end{aligned} \quad (12.13)$$

Мы получили в результате, что энергия уровней доноров модулирована медленно изменяющимся потенциалом с распределением (12.11).

Таким образом, плотность состояний вблизи центра пика имеет вид

$$g(e) = \frac{N_D}{\gamma\sqrt{\pi}} \exp\left(-\frac{e^2}{\gamma^2}\right). \quad (12.14)$$

Оценим теперь приближения, сделанные в этом параграфе. Согласно (12.13) и (11.11) условие  $\gamma \ll \mu$  хорошо выполняется при  $K \ll 1$ . Предположение о гауссовом характере флуктуаций справедливо, если радиус экранирования  $r_0$  превышает среднее расстояние между комплексами, т. е.  $4/3\pi r_0^3 N_0 \gg 1$ . С помощью (12.10) это неравенство переписывается в виде  $5 \cdot 10^{-2} K^{-1/2} \gg 1$ , откуда видно, что оно выполняется лишь при очень малых  $K$ . Таким образом, результаты, полученные в этом параграфе, к реальным ситуациям применимы лишь качественно.

Изложенные в §§ 11, 12 представления о комплексах используются в теориях прыжковой проводимости (см. гл. 8), оптических явлений [5] и фотопроводимости [6].

### § 13. Примесная зона при большой степени компенсации

В случае сильной компенсации ситуация оказывается несколько более сложной, чем при слабой компенсации. Однако наличие малого параметра  $1 - K \ll 1$  позволяет разобраться в структуре примесной зоны, и, хотя не все конечные результаты могут быть «доведены до чисел»,

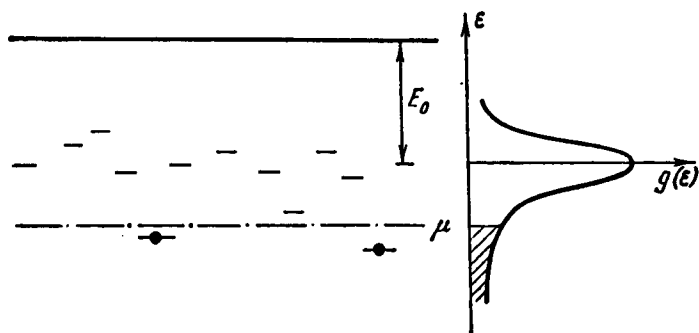


Рис. 3.4. Энергетическая схема сильно компенсированного полупроводника в пренебрежении крупномасштабным потенциалом. Сплошная прямая — зона проводимости, штрих-пунктир — уровень Ферми. Короткие черточки изображают уровни доноров, а кружки — занимающие их электроны. Справа изображена плотность состояний на донорных уровнях. Заполненные состояния заштрихованы.

асимптотические зависимости ширины примесной зоны и энергии Ферми от концентрации примесей и степени компенсации в пределе  $K \rightarrow 1$  можно считать четко установленными, по крайней мере, в модели пуассоновского распределения примесей [7].

При сильной компенсации ( $1 - K \ll 1$ ) концентрация электронов на донорах  $n = N_D - N_A$  намного меньше концентрации доноров  $N_D$ . Поэтому при низких температурах всем электронам удастся разместиться на донорах, энергетические уровни которых существенно понижены потенциалом соседних заряженных примесей. В результате уровень Ферми при сильной компенсации располагается ниже уровня изолированной примеси (рис. 3.4). Чем меньше концентрация электронов  $n$ , тем глубже в запрещенную зону проваливается уровень Ферми.

Если не учитывать гауссовых флуктуаций концентраций заряженных доноров и акцепторов с характерными размерами, много большими среднего расстояния между примесями (о них речь пойдет ниже), то основная причина возникновения донорных уровней с пониженными энергиями состоит в следующем. Вблизи некоторого донора на расстоянии  $r \ll N_D^{-1/3}$  может оказаться другой положительно заряженный донор. Если  $r \gg a$ , поправка  $\epsilon$  к энергии уровня на первом доноре определяется просто кулоновским потенциалом второго донора, т. е.  $\epsilon = -e^2/\kappa r$ . (Мы по-прежнему отсчитываем величины  $\epsilon$  и  $\mu$  вверх от уровня изолированной примеси.) При этом имеется в виду, что второй донор пуст. Энергия связи второго электрона на паре доноров близка к  $E_0$ . Поэтому при сильной компенсации, когда уровень Ферми отрицателен и велик по абсолютной величине, на паре близких доноров может находиться лишь один электрон.

Для нахождения  $g(\epsilon)$  необходимо подсчитать концентрацию пар, имеющих уровень с энергией в интервале  $\epsilon, \epsilon + d\epsilon$ . Вероятность того, что около некоторого донора на расстоянии от  $r$  до  $r + dr$  имеется другой донор, есть  $4\pi N_D r^2 dr$ , а вероятность того, что сдвиг уровня лежит в интервале от  $\epsilon$  до  $\epsilon + d\epsilon$ , есть  $4\pi N_D r^2 \frac{dr}{d\epsilon} d\epsilon$ , где  $r(\epsilon) = e^2/\kappa\epsilon$ . Находя концентрацию пар, надо умножить указанную вероятность на  $N_D$  и затем разделить на 2, чтобы не учитывать каждый донор дважды. Это дает

$$g(\epsilon) = \frac{3}{2} \frac{e_D^3}{\epsilon^4} N_D, \quad \epsilon_D = \left( \frac{4\pi}{3} N_D \right)^{1/3} \frac{e^2}{\kappa}. \quad (13.1)$$

Теперь нетрудно найти уровень Ферми при  $T = 0$  из условия, что

$$\int_{-\infty}^{\mu} g(\epsilon) d\epsilon = n. \quad (13.2)$$

В результате получим

$$\mu = - \left( \frac{2\pi}{3} \right)^{1/3} \frac{e^2 N_D^{2/3}}{\kappa n^{1/3}} = - \frac{\epsilon_D}{2^{1/3} (1 - K)^{1/3}}. \quad (13.3)$$

Формула (13.3) справедлива, если существенные расстояния в парах  $r \approx r_\mu = e^2/\kappa\mu$  велики по сравнению

с  $a$ , но малы по сравнению со средним расстоянием между примесями  $N_D^{-1/3}$ . Это накладывает следующие ограничения на  $|\mu|$ :

$$\frac{e^2 N_D^{1/3}}{\kappa} < |\mu| < E_0. \quad (13.4)$$

Существование «окна» между  $e^2 N_D^{1/3}/\kappa$  и  $E_0$  гарантируется условием слабого легирования,  $N_D a^3 \ll 1$ . Когда  $r_\mu$  сравниваются с  $a$ , классическая формула  $\epsilon = -e^2/\kappa r$  для сдвига уровня становится непригодной. Ясно, что с учетом квантовых эффектов даже при  $r \rightarrow 0$  два донора, образующие двухзарядное ядро водородоподобного атома, могут создать уровень только с энергией  $-4E_0$ , т. е. обеспечить сдвиг  $\epsilon = -3E_0$  относительно уровня изолированного донора. Образование более глубокого локального хвоста плотности состояний, как будет показано в § 39, связано с флуктуационными скоплениями нескольких доноров, образующих ядро многозарядного атома. При этом

$$g(\epsilon) \propto \exp \left\{ -\sqrt{\frac{|\epsilon|}{E_0}} \ln \left[ \left( \frac{\epsilon}{E_0} \right)^2 (N_D a^3)^{-1} \right] \right\}, \quad |\epsilon| \gg E_0. \quad (13.5)$$

Видно, что  $g(\epsilon)$  спадает очень быстро с ростом  $\epsilon$ . Поэтому при очень больших степенях компенсации, когда  $|\mu|$  доходит до  $3E_0$ , дальнейшее опускание уровня Ферми без учета крупномасштабного потенциала должно происходить медленно ( $|\mu| \sim |\ln n|$ ).

Обратимся теперь к рассмотрению крупномасштабных флуктуаций концентрации заряженных примесей. В отличие от случая  $K \ll 1$ , при  $1 - K \ll 1$  крупномасштабный потенциал велик и играет столь же важную, а иногда более существенную роль в погружении уровня Ферми, чем локальные эффекты. Причина состоит в том, что при  $1 - K \ll 1$  велика концентрация заряженных частиц, создающих флуктуационные поля, и слабо электронное экранирование. Крупномасштабный потенциал оказывается столь большим, что приводит к глубокой модуляции плотности электронов в пространстве. При этом обычная теория экранирования, основанная на лиnearизации уравнения Пуассона, эквивалентной предпо-

ложению о слабой неоднородности электронной плотности, оказывается неприменимой. В этом и состоит основная трудность в описании крупномасштабного потенциала при  $1 - K \ll 1$ .

На первый взгляд может показаться, что если заряженных примесей очень много, а электронов мало, то потенциальный рельеф создается только примесями, а потенциал электронов можно вообще не учитывать. Однако это неправильно. Как мы видели в § 12, флуктуации потенциала случайно расположенных кулоновских центров бесконечно велики. Среднеквадратичный потенциал таких флуктуаций, равный  $\left[ N_t \int \varphi^2(r) dr \right]^{1/2}$ , где  $\varphi = e^2/\kappa r$ , расходится на больших расстояниях. Эту расходимость легко понять. Допустим, что возникла флуктуация, имеющая объем  $R^3$ . Среднеквадратичная флуктуация числа примесей в таком объеме порядка  $(N_t R^3)^{1/2}$ , где  $N_t$  — суммарная концентрация доноров и акцепторов. Потенциальная энергия пробного электрона, создаваемая такой флуктуацией, равна

$$\gamma(R) = \frac{e^2}{\kappa R} (N_t R^3)^{1/2}. \quad (13.6)$$

Отсюда видно, что при  $R \rightarrow \infty$  потенциал флуктуации  $\gamma(R) \rightarrow \infty$ . Таким образом, если примеси расположены случайно, электронное экранирование нужно учитывать при сколь угодно малых концентрациях электронов. Как же происходит это экранирование?

Для ответа на этот вопрос рассмотрим перераспределение электронов, вызываемое флуктуациями заряженных примесей с размером порядка  $R$ . Избыточная плотность заряженных примесей в таких флуктуациях порядка

$$\Delta N = (N_t R^3)^{1/2} / R^3. \quad (13.7)$$

Важно иметь в виду, что мы рассматриваем не редко встречающиеся, а типичные флуктуации. Таким образом, если разбить кристалл на области объемом порядка  $R^3$ , то половина из них будет обладать положительным зарядом, а половина — отрицательным, с плотностью порядка  $e \Delta N$ ,

При  $T = 0$  электроны стремятся занять состояния с минимальной энергией. Поэтому они переходят в области с положительным примесным зарядом. Если избыточная примесная плотность  $\Delta N$  меньше, чем средняя электронная плотность  $n$ , то флуктуации будут полностью нейтрализованы небольшим изменением электронной плотности. Если же  $\Delta N \gg n$ , то флуктуации вообще не могут нейтрализоваться электронами. Действительно, если все электроны перейдут из отрицательных областей в положительные, то электронная плотность в положительных областях увеличится, грубо говоря, в два раза и останется по-прежнему малой по сравнению с зарядом флуктуаций.

Итак, для ответа на вопрос, экранируются ли флуктуации данного масштаба, необходимо сравнить избыточную плотность заряженных примесей в таких флуктуациях (формула (13.7)) со средней концентрацией электронов. Согласно (13.7) величина  $\Delta N$  стремится к нулю при  $R \rightarrow \infty$ . Поэтому достаточно большие флуктуации окажутся полностью нейтрализованными, как бы ни была мала средняя концентрация электронов. Благодаря этому и ликвидируется расходимость среднеквадратичного потенциала примесей.

Введем длину  $r_s$ , определив ее равенством

$$n = \frac{(N_t r_s^3)^{1/2}}{r_s^3}$$

или

$$r_s = \frac{N_t^{1/3}}{n^{2/3}} = \frac{2^{2/3} N_t^{-1/3}}{(1 - K)^{2/3}} \quad (N_D \approx N_A). \quad (13.8)$$

Из сказанного ясно, что флуктуации с размерами  $R > r_s$  нейтрализуются электронами, а флуктуации с  $R < r_s$  остаются неэкранированными. Поэтому амплитуда потенциала определяется флуктуациями размера  $r_s$ , имеет порядок величины

$$\gamma(r_s) = \frac{e^2 N_t^{2/3}}{\kappa n^{1/3}} \quad (13.9)$$

и возрастает с уменьшением  $n$ .



Мы говорили сейчас только о флуктуациях, близких к среднеквадратичным. Но существуют и маловероятные флуктуации, обладающие зарядом, большим чем  $e(N_i R^3)^{1/2}$ , которые, казалось, должны быть более «привлекательными» для электронов. Фактически наши рассуждения основывались на том, что число электронных мест в притягивающей флуктуации лимитируется ее зарядом. В маловероятной флуктуации с избыточным числом примесей  $Z$  число мест больше в  $Z/(N_i R^3)^{1/2}$  раз. Однако число таких флуктуаций меньше в  $\exp(-Z^2/N_i R^3)$  раз. Таким образом, полное число мест в маловероятных флуктуациях гораздо меньше, чем в типичных флуктуациях того же масштаба. Масштаб  $r_s$  выбран так, что в типичных флуктуациях этого масштаба число мест близко к полному числу электронов. Это означает, что разместить все электроны в маловероятных флуктуациях масштаба  $r_s$  невозможно, в них поместится лишь малая часть электронов.

Как сказывается крупномасштабный потенциал на положении уровня Ферми? Поскольку электроны переходят в ямы крупномасштабного потенциала, уровень Ферми понижается на величину, близкую к амплитуде флуктуаций  $\gamma(r_s)$ . Интересно отметить, что  $\gamma(r_s)$  имеет тот же порядок величины, что и опускание уровня Ферми (13.3) за счет локальной плотности состояний донорных пар. Таким образом, при  $|\mu| < 3E_0$ , когда пары существенны, вклады локальной плотности состояний и крупномасштабного потенциала в понижение уровня Ферми одинаковы, а при  $|\mu| > 3E_0$  главную роль играет крупномасштабный потенциал. В результате, несмотря на то, что при  $|\mu| > 3E_0$  пары не играют важной роли,  $|\mu|$  продолжает возрастать по закону  $(1-K)^{-1/3}$ . Вытекающая из изложенных качественных соображений энергетическая схема сильно компенсированного полупроводника изображена на рис. 3.5.

Описанное выше нелинейное экранирование осуществляется, если  $\gamma(r_s)$  меньше ширины запрещенной зоны  $E_g$ . В противном случае в экранирование включаются также дырки и все характерные масштабы несколько изменяются. Мы рассмотрим этот случай подробно в § 46.



Предыдущее изложение основано на результатах работы Шкловского и Эфроса [7]. Позднее аналогичные результаты были получены Штерном [8]. Штерн считал распределение потенциала гауссовым. Ширину гауссова распределения  $\gamma$  он связывал с радиусом экранирования с помощью формулы (12.13), где вместо  $N_0$  стояла величина  $N_D$ . В свою очередь радиус экранирования он связывал с плотностью состояний на уровне Ферми с помощью обычной формулы, следующей из теории линейного экранирования:  $r_0 = (4\pi e^2 g(\mu)/\kappa)^{-1/2}$ . Это давало возможность определить радиус экранирования и уровень Ферми.

Как показано выше, экранирование при сильной компенсации является нелинейным, поэтому все три аппроксимации Штерна являются, строго говоря, неправильными (например, ниже мы увидим, что асимптотика функции распределения потенциала отличается от гауссовой). Тем не менее по порядку величин результаты Штерна совпадают с (13.8) и (13.9). Более того, возможно, что предложенный им метод гауссовой аппроксимации является перспективным методом приближенных вычислений.

Интересно отметить, что в спектре оптического поглощения компенсированных кристаллов  $n$ -InSb наблюдался пик, который может быть интерпретирован как результат фотоионизации электронов, находящихся на парах близких друг к другу доноров [9]. Ширина линии примесного оптического поглощения в сильно компенсированных полупроводниках вычислялась в рамках рассмотренной в этом параграфе модели в работе [10].

Мы изложили основные результаты, относящиеся к случаю сильной компенсации. Оставшаяся часть этого параграфа посвящена количественной теории для случая  $|\mu| < 3E_0$  и модели пуассоновского распределения примесей. Цель этой теории состоит в подтверждении полученных выше результатов и в выводе уравнений, с помощью которых на ЭВМ можно получить численные коэффициенты в выражении для уровня Ферми. Кроме того, мы приведем асимптотический вид функции распределения потенциала флуктуаций при больших по абсолютной величине значениях этого потенциала.

Читатель, которого удовлетворили приведенные качественные соображения, может опустить эту часть и перейти к следующим параграфам книги.

Как видно из сказанного выше, в области  $|\mu| < 3E_0$  основную роль в распределении электронов играют близкие пары доноров и крупномасштабный потенциал. Количественная теория может быть построена благодаря тому, что характерный размер  $r_s$  крупномасштабного потенциала гораздо больше, чем среднее расстояние между электронами. Поэтому в основу математического описания может быть положено самосогласованное уравнение, аналогичное (12.5). Так же, как и при выводе (12.5), мысленно разобьем систему на отдельные блоки, каждый из которых является макроскопическим в том смысле, что содержит много электронов, сидящих на донорных парах, однако настолько малым, что потенциал крупномасштабных флуктуаций  $\varphi(r)$  можно считать внутри него постоянным. Пусть  $N_D(r)$ ,  $N_A(r)$  и  $n(r)$  — средние по таким блокам концентрации доноров, акцепторов и электронов, находящихся на донорных парах. Напишем макроскопическое уравнение Пуассона:

$$\Delta\varphi = -\frac{4\pi e}{\kappa} [N_D(r) - N_A(r) - n(r)]. \quad (13.10)$$

Концентрации  $N_D(r)$  и  $N_A(r)$  флуктуируют независимо в соответствии с гауссовой статистикой, а концентрация  $n(r)$  определяется числом пар, опускающих энергию электрона ниже уровня Ферми, причем последний не меняется вдоль системы. Для вычисления  $n(r)$  следует воспользоваться локальной плотностью состояний  $g(e, r)$ , которая получается из (13.1) заменой  $e$  на  $e + e\varphi$ . Концентрацию  $N_D$  в (13.1) можно считать не зависящей от  $r$ , так как при гауссовых флуктуациях она мало отличается от своего среднего значения. Тогда

$$\begin{aligned} n(r) &= \int_{-\infty}^{\mu} g(e, r) de = \frac{3}{2} e_D^3 N_D \int_{-\infty}^{\mu} \frac{de}{(e + e\varphi)^4} = \\ &= -\frac{1}{2} \frac{N_D e_D^3}{(\mu + e\varphi)^3}. \end{aligned} \quad (13.11)$$

Введем безразмерные переменные:

$$\chi = -\frac{2^{2/3} e\varphi}{\gamma(r_s)}, \quad \eta = -\frac{2^{2/3} \mu}{\gamma(r_s)}, \quad (13.12)$$

$$x = \frac{2^{1/3} r}{r_s}, \quad (13.13)$$

$$f = \frac{\xi_D(r) - \xi_A(r)}{N_D - N_A}, \quad (13.14)$$

$$\xi_D(r) = N_D(r) - N_D, \quad \xi_A(r) = N_A(r) - N_A.$$

Подставляя (13.11)—(13.14) в (13.10), получим уравнение

$$\Delta_x \chi = 4\pi \left( f + 1 - \frac{2\pi}{3} \frac{1}{(\chi + \eta)^3} \right), \quad (13.15)$$

справедливое при  $\chi > -\eta$ . Так как мы считаем распределение примесей некоррелированным, из определения (13.14) и соотношений

$$\begin{aligned} \langle \xi_D(r) \xi_A(r') \rangle &= 0, \\ \langle \xi_D(r) \xi_D(r') \rangle &= N_D \delta(r - r'), \\ \langle \xi_A(r) \xi_A(r') \rangle &= N_A \delta(r - r') \end{aligned} \quad (13.16)$$

следует

$$\langle f(x) f(x') \rangle = 2\delta(x - x'). \quad (13.17)$$

Предполагая флуктуации гауссовыми, будем считать, что более высокие корреляторы  $f(x)$  распадаются на произведения парных. Для определения величины  $\eta$  следует задать функцию  $f$  в области с линейным размером  $X \gg 1$  и потребовать, чтобы уравнение (13.15) имело решение с равным нулю полным зарядом области (интегралом от правой части (13.15)) и потенциалом, стремящимся к нулю на бесконечности. Ясно, что при  $X \gg 1$  химический потенциал  $\eta$  не зависит от реализации случайной функции  $f(x)$ . Соответствующие численные расчеты могут быть выполнены на ЭВМ. Однако из того, что уравнения (13.15) и (13.17) не содержат параметров, следует, что  $\eta$  порядка единицы. Таким образом, в согласии с результатами качественного рассмотрения энергия Ферми, отсчитанная от уровня изолированной

примеси, равна

$$|\mu| = \eta \frac{e^2 N_D^{2/3}}{\chi n^{1/3}}. \quad (13.18)$$

Обсудим теперь распределение потенциала  $\chi$ . Введем функцию  $F(\chi)$ , представляющую собой плотность вероятности потенциала  $\chi$ . Эта функция должна удовлетворять условиям нормировки и нейтральности:

$$\int_{-\eta}^{\infty} F(\chi) d\chi = 1, \quad (13.19)$$

$$\frac{2\pi}{3} \int_{-\eta}^{\infty} \frac{F(\chi) d\chi}{(\chi + \eta)^3} = 1. \quad (13.20)$$

Разумеется,  $F(\chi)$  нельзя вычислить без точного решения уравнения (13.15), однако можно исследовать асимптотические свойства этой функции. При больших положительных значениях  $\chi$  это исследование можно произвести с помощью метода оптимальной флуктуации, изложенного в гл. 12. В § 43 показано, что при построении теории сильно легированного и сильно компенсированного полупроводника мы приходим к уравнению (43.11), отличающемуся от (13.15) только структурой члена, описывающего электронную плотность. Исследование (43.11) с помощью метода оптимальной флуктуации, приведенное в § 43, показывает, что при  $\chi \gg 1$

$$F(\chi) \propto \exp\left(-\frac{\sqrt{2}}{3\sqrt{\pi}} \chi^{3/2}\right). \quad (13.21)$$

Легко убедиться, что этот вывод и формула (13.21) применимы и для уравнения (13.15). Дело в том, что области с  $\chi \gg 1$  являются горбами потенциальной энергии и электроны в этих областях отсутствуют. Поэтому вид электронной плотности не играет никакой роли при выводе (13.21).

Мы изложили теорию для случая, когда компенсация такова, что  $|\mu|$  не превосходит  $3E_0$ . При более высоких степенях компенсации опускание уровня Ферми связано в основном с крупномасштабным потенциалом. Детальный вид локальной плотности состояний здесь не играет

роли. Важно лишь, что плотность состояний убывает быстрее, чем плотность состояний пар (13.1), выдерживающая конкуренцию с крупномасштабным потенциалом. Этому условию удовлетворяет закон убывания (13.5). При этом можно считать, что потенциальная энергия электронов  $-e\varphi$  нигде не опускается ниже уровня Ферми, а в областях, где электронная плотность  $n(\mathbf{r})$  отлична от нуля,  $e\varphi = -\mu$ . Так как потенциал в этих областях не зависит от  $\mathbf{r}$ , в них должно выполняться условие нейтральности  $n(\mathbf{r}) = N_D(\mathbf{r}) - N_A(\mathbf{r})$ .

Возникающая при этом математическая задача в точности совпадает с задачей, рассмотренной нами в § 43 в связи с теорией сильно легированных полупроводников. Соответствующая система уравнений не содержит параметров, если перейти к переменным (13.12) — (13.14). Поэтому выражение для глубины опускания уровня Ферми (13.18) справедливо и при  $|\mu| > 3E_0$ , однако значение численного коэффициента  $\eta$  должно определяться в этом случае из решения другой задачи и может быть иным.

## ГЛАВА 4

# ОБЩЕЕ ПРЕДСТАВЛЕНИЕ О ПРЫЖКОВОЙ ЭЛЕКТРОПРОВОДНОСТИ СЛАБО ЛЕГИРОВАННЫХ ПОЛУПРОВОДНИКОВ

### § 14. Основные экспериментальные факты

При высоких температурах полупроводники обладают собственной электропроводностью, связанной с термической активацией электронов из валентной зоны в зону проводимости. Концентрации собственных носителей, электронов ( $n$ ) и дырок ( $p$ ), — экспоненциальные функции температуры [1]:

$$n = p = \frac{(2\pi \sqrt{m_e m_h} kT)^{3/2}}{4\pi^3 \hbar^3} \exp\left(-\frac{E_g}{2kT}\right). \quad (14.1)$$

Здесь  $m_e$  и  $m_h$  — массы электронов и дырок,  $E_g$  — ширина запрещенной зоны.

Ввиду большой величины энергии активации  $E_g/2$  при понижении температуры концентрация собственных носителей очень быстро убывает до тех пор, пока их не станет меньше, чем носителей, поставляемых в зоны примесями. При дальнейшем понижении температуры проводимость целиком определяется концентрацией и свойствами примесей и поэтому называется примесной.

На рис. 4.1 мы схематически изобразили зависимость логарифма удельного сопротивления слабо легированного полупроводника от обратной температуры. Собственной проводимости соответствует участок  $A$  с почти вертикальным наклоном. Участки  $B$  —  $\Gamma$  соответствуют примесной проводимости. Если речь идет о мелких примесях с энергией ионизации, много меньшей  $E_g$ , то суще-



ствуует область температур  $B$ , называемая областью истощения, в которой все примеси ионизованы и соответственно концентрация носителей в зоне не зависит от температуры. В этой области температурная зависимость сопротивления полностью определяется температурной зависимостью подвижности. Например, убывание сопротивления при понижении температуры связано с ослаблением рассеяния на фонах.

Дальнейшее понижение температуры (область  $B$ ) приводит к постепенному вымораживанию примесных электронов, т. е. к возвращению электронов на доноры (для определенности мы будем говорить о полупроводнике  $n$ -типа). Температурная зависимость электропроводности здесь практически целиком определяется быстрым убыванием концентрации свободных электронов. Теоретический расчет температурной зависимости концен-

трации в области  $B$  не содержит трудностей, если пренебречь разбросом примесных уровней (шириной примесной зоны) и считать, что все доноры обладают одинаковой энергией ионизации  $E_0$ . Подробности этого расчета можно найти в книге Блекмора [2]. Результаты расчета существенно зависят от соотношения числа пустых мест  $KN_D$ , создаваемых на донорах компенсацией, и числа пустых мест  $n(T)$ , создаваемых термическим забросом электронов в зону проводимости.

При  $KN_D \ll n(T)$  уровень Ферми находится практически посередине между уровнем донора и дном зоны проводимости и

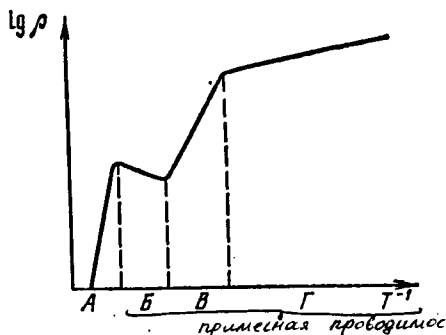


Рис. 4.1. Схематическая температурная зависимость сопротивления слабо легированного полупроводника:  $A$  — область собственной проводимости,  $B$  — область истощения примесной проводимости,  $B$  — область вымораживания,  $\Gamma$  — область прыжковой проводимости.

$$n(T) = \left( \frac{N_D N_c}{2} \right)^{1/2} \exp\left( - \frac{E_0}{2kT} \right), \quad (14.2)$$

где

$$N_c = \frac{(2\pi m_e kT)^{3/2}}{4\pi^3 \hbar^3}. \quad (14.2a)$$

При  $n(T) \ll KN_D$  уровень Ферми близок к уровню изолированного донора и

$$n(T) = \frac{N_c N_D}{2N_A} \exp\left(-\frac{E_0}{kT}\right). \quad (14.3)$$

Зависимость (14.2) имеет место в области, где температура не слишком мала и не очень высока ( $KN_D \ll \ll n(T) \ll N_D$ ). Поэтому область применимости (14.2), как правило, невелика и существует лишь при весьма малых степенях компенсации. Часто экспериментальная температурная зависимость удельного сопротивления в области  $B$  имеет одну энергию активации  $e_1$ :

$$\rho(T) = \rho_1 \exp\left(\frac{e_1}{kT}\right), \quad (14.4)$$

где  $e_1$  — величина, близкая к энергии ионизации изолированного донора  $E_0$ , определенной, например, по оптическим данным. Лишь в очень слабо компенсированных полупроводниках в высокотемпературной части области  $B$  проявляется небольшой участок с энергией активации порядка  $E_0/2$ .

Постепенное вымораживание электронов в зоне проводимости при понижении температуры приводит к тому, что наступает момент, когда наибольшую роль в электропроводности начинают играть прыжки электронов непосредственно по примесям, без активации в зону. Такая проводимость называется прыжковой. Электрон прыгает с занятого донора на свободный. Поэтому необходимым условием прыжковой проводимости является наличие свободных мест на донорах. При низких температурах оно может быть обеспечено только компенсацией.

Прыжковому механизму электропроводности соответствует очень малая подвижность, так как прыжки электронов связаны со слабым перекрытием хвостов волновых функций соседних доноров. Несмотря на это, прыжковая проводимость выигрывает в конкуренции с зонной, так как в прыжковой могут принимать участие все

находящиеся на донорах электроны, а в зонной — только экспоненциально малое число свободных электронов. На рис. 4.1 прыжковой проводимости соответствует область  $\Gamma$ . В германии с концентрацией мелких доноров  $N_D \approx 10^{15} \text{ см}^{-3}$  переход к прыжковой проводимости происходит при  $T \approx 7^\circ\text{K}$ . Области  $B$  соответствуют температуры  $7 < T < 50^\circ\text{K}$ , области  $B$  —  $50 < T < 400^\circ\text{K}$  и области  $A$  —  $T > 400^\circ\text{K}$ .

Возможность прыжковой проводимости была теоретически предсказана в работе Гуддена и Шоттки [3]. Первые экспериментальные данные о прыжковой проводимости были получены Бушем и Лабхартом [4] на карбиде кремния и Хангом и Глиссманом [5] на германии и кремнии. В дальнейшем прыжковая электропроводность многократно исследовалась в германии, кремнии, сурьмянистом индии, арсениде галлия, сурьмянистом галлии, фосфиде индия, сульфиде кадмия, алмазе и других полупроводниках. Ниже мы обсудим основные экспериментальные факты, используя в качестве примера данные для германия. Начнем с температурной зависимости прыжковой проводимости.

На рис. 4.2 изображены полученные в работе Фрицше и Кейваса [6] зависимости удельного сопротивления от обратной температуры для германия  $p$ -типа со степенью компенсации  $K = 0,4$ . Каждой кривой соответствует своя концентрация акцепторов ( $G_a$ ). Все кривые с хорошей точностью могут быть аппроксимированы с помощью формулы

$$\rho^{-1}(T) = \rho_1^{-1} \exp\left(-\frac{e_1}{kT}\right) + \rho_3^{-1} \exp\left(-\frac{e_3}{kT}\right). \quad (14.5)$$

Первое слагаемое соответствует зонной проводимости. Оно оказывается практически одинаковым для всех концентраций акцепторов. Действительно, как видно из формулы (14.3), концентрация свободных носителей зависит лишь от степени компенсации, которая, как уже говорилось, для всех образцов была одинаковой. Слабая зависимость подвижности от концентрации примесей в масштабе рис. 4.2 не заметна.

Второе слагаемое соответствует прыжковой проводимости (почему ему «без очереди» присвоен номер 3, мы объясним ниже). Видно, что прыжковая проводимость

обладает хотя и малой по сравнению с  $\epsilon_1$ , но все же заметной энергией активации. Ее происхождение связано с разбросом уровней примесей. Прыгая по донорам,

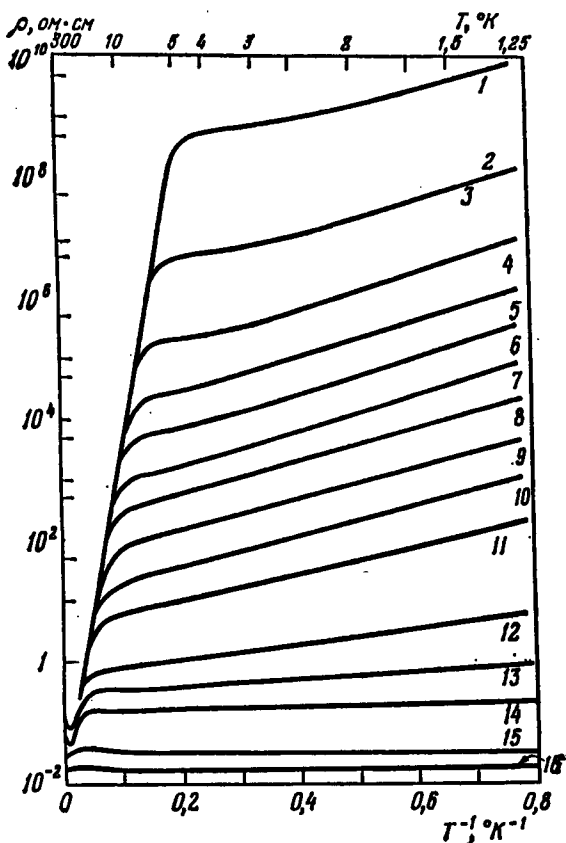


Рис. 4.2. Удельное сопротивление Ge p-типа со степенью компенсации  $K = 0,4$  [6]. Кривые 1—16 соответствуют следующим концентрациям акцепторов (в  $\text{см}^{-3}$ ): 1 —  $7,5 \cdot 10^{14}$ ; 2 —  $1,4 \cdot 10^{15}$ ; 3 —  $1,5 \cdot 10^{15}$ ; 4 —  $2,7 \cdot 10^{15}$ ; 5 —  $3,6 \cdot 10^{15}$ ; 6 —  $4,9 \cdot 10^{15}$ ; 7 —  $7,2 \cdot 10^{15}$ ; 8 —  $9 \cdot 10^{15}$ ; 9 —  $1,4 \cdot 10^{16}$ ; 10 —  $2,4 \cdot 10^{16}$ ; 11 —  $3,5 \cdot 10^{16}$ ; 12 —  $7,3 \cdot 10^{16}$ ; 13 —  $1,0 \cdot 10^{17}$ ; 14 —  $1,5 \cdot 10^{17}$ ; 15 —  $5,3 \cdot 10^{17}$ ; 16 —  $1,35 \cdot 10^{16}$ .

электрон поглощает и излучает фононы. В результате и возникает экспоненциальная зависимость электропроводности от температуры.

При увеличении концентрации примесей энергия активации поначалу несколько возрастает. Это связано с

увеличением кулоновского случайного потенциала, создаваемого заряженными примесями. Однако дальнейшее увеличение концентрации приводит к усилению перекрытия волновых функций соседних центров и уменьшению  $\epsilon_3$ . В результате при концентрации  $N_A = 10^{17} \text{ см}^{-3}$  энергия активации  $\epsilon_3$  обращается в нуль. При этой концентрации происходит переход от активационной проводимости к металлической, описанный в гл. 3. Область концентраций, близких к переходу диэлектрик — металл, ввиду отсутствия малых параметров плохо поддается количественной интерпретации. Поэтому мы все время будем говорить о прыжковой проводимости в области значительно меньших концентраций.

На рис. 4.2 видна еще одна характерная черта прыжковой проводимости: исключительно сильная зависимость величины  $\rho_3$  от концентрации примесей. Значение  $\rho_3$  можно определить с помощью рис. 4.2, если экстраполировать низкотемпературный прямолинейный участок к  $T^{-1} = 0$ , т. е. найти точку пересечения этой прямой с осью ординат. При возрастании концентрации примесей в 30 раз  $\rho_3$  уменьшилось в  $10^7$  раз. Такую резкую зависимость естественно считать экспоненциальной и  $\rho_3$  записывать в виде

$$\rho_3 = \rho_{03} \exp [f(N_D)], \quad (14.6)$$

где  $\rho_{03}$  и  $f(N_D)$  — степенные функции концентрации примесей. Физическая причина экспоненциальной зависимости (14.6) ясна. Она состоит в том, что вероятность прыжка между двумя примесями определяется перекрытием волновых функций. Нас интересуют расстояния между примесями, много большие борковского радиуса. На таких расстояниях, как мы видели в гл. 1, волновые функции спадают экспоненциально. Поэтому интегралы перекрытия между примесями экспоненциально убывают с ростом расстояния между ними. При уменьшении концентрации растет среднее расстояние между примесями, экспоненциально убывают вероятности прыжков и, следовательно, электропроводность. Наличие экспоненциальной зависимости электропроводности от концентрации примесей является обычно главным экспериментальным доводом в пользу прыжкового механизма проводимости.

Величины  $\varepsilon_3$  и  $\rho_3$  зависят от степени компенсации. На рис. 4.3 приведены экспериментальные значения энергии активации  $\varepsilon_3$  при различных степенях компенсации в образцах  $p$ -типа с одинаковой концентрацией основных примесей  $N_A = 2,66 \cdot 10^{15} \text{ см}^{-3}$  [7]. Видно, что с увеличением степени компенсации  $\varepsilon_3$  сначала резко убывает, где-то вблизи  $K = 0,4$  достигает минимума, а затем при  $K \rightarrow 1$  быстро растет.

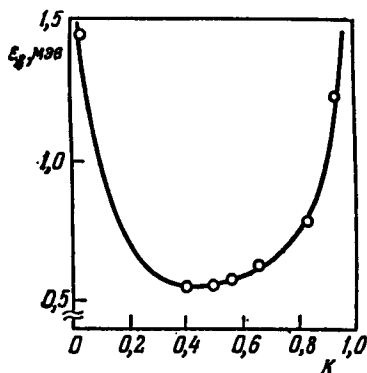


Рис. 4.3. Экспериментальная зависимость энергии активации  $\varepsilon_3$  от степени компенсации  $K$  [7].

Относительно зависимости  $\rho_3(K)$  имеются противоречивые данные. В обзоре Мотта и Туза [7] утверждается, что  $\rho_3$  монотонно возрастает с ростом  $K$ . С другой стороны, Доброго и Ермолаев [8] в близких по концентрации примесей образцах германия наблюдали убывание  $\rho_3$  при увеличении компенсации от 0 до 0,2. Такое же поведение при  $K < 0,2$  наблюдали ранее Дэвис и

Комптон [9]. При  $K \approx 0,2$  сопротивление  $\rho_3$  имело минимум, а затем росло с увеличением  $K$  [9].

В слабо компенсированных полупроводниках ( $K < 0,2$ ) в ограниченном интервале концентраций вблизи перехода Мотта, кроме зонной и прыжковой проводимостей, в температурной зависимости проводимости проявляется еще один активационный механизм, который создает в формуле (14.5) дополнительное слагаемое

$\rho_2^{-1} \exp\left(-\frac{\varepsilon_2}{kT}\right)$ . Этот механизм работает в промежуточном между зонной и прыжковой проводимостями интервале температур, поскольку  $\rho_1 \ll \rho_2 \ll \rho_3$  и  $\varepsilon_1 > \varepsilon_2 > \varepsilon_3$ .

Механизму  $\varepsilon_2$ -проводимости посвящено много работ (см., например, [10, 11]). Большинство авторов считает, что  $\varepsilon_2$ -проводимость связана с передвижением электронов по однократно заполненным нейтральным донорам. Гершензон и др. [12] показали, что в германии на нейтральном доноре существует состояние второго электрона ( $D^-$ -состояние) с энергией связи порядка  $0,1E_0$ .

$D$ -состояния обладают большим радиусом и поэтому при промежуточных концентрациях примесей сильно перекрываются. В результате должна образовываться широкая зона (рис. 4.4).

Эта зона представляет собой аналог верхней зоны Хаббарда, о которой говорилось в гл. 2, с тем, однако, отличием, что зона образовалась в неупорядоченной системе. Неупорядоченность же может привести к полной или частичной локализации состояний в  $D$ -зоне. Однако не вызывает сомнения, что даже в случае заметной локализации благодаря большому радиусу  $D$ -состояния (большой ширине зоны)  $D$ -зоне отвечает существенно большая подвижность, чем прыжковой проводимости по пустым состояниям. Поэтому можно себе представить, что несмотря на экспоненциально малое число электронов в небольшой области температур  $\varepsilon_2$ -проводимость может быть больше  $\varepsilon_3$ -проводимости. С другой стороны, благодаря большой ширине  $D$ -зоны ее нижний порог подвижности может оказаться существенно ближе к уровню Ферми, чем дно зоны проводимости (см. рис. 4.4). В результате  $D$ -зона может выиграть конкуренцию и у зоны проводимости, которой, конечно, соответствует бóльшая подвижность, но существенно меньшая концентрация электронов.

В пользу концепции  $D$ -зоны для  $\varepsilon_2$ -проводимости говорит то, что  $\varepsilon_2$ -проводимость наблюдается лишь при значительной концентрации примесей и при малой компенсации. Действительно, уменьшение легирования должно приводить к сильному сужению  $D$ -зоны и приближению ее энергии к энергии изолированного иона  $D$ . Эта энергия, как мы говорили, имеет значение порядка  $-0,1E_0$ . Таким образом, энергия активации  $\varepsilon_2$  в пределе малых концентраций должна стремиться к величине порядка  $0,9E_0$ , т. е. должна стать очень близкой к  $\varepsilon_1$ .

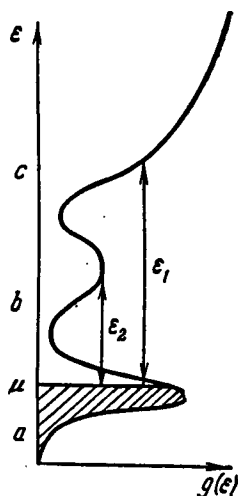


Рис. 4.4. Плотность состояний электронов в полупроводнике  $n$ -типа [13]. Показано расщепление примесной зоны на зону  $\varepsilon_3$  ( $a$ ) и зону  $\varepsilon_2$  ( $b$ );  $c$  — зона проводимости.

Естественно, что в этом случае  $\varepsilon_2$ -проводимость не может оказаться конкурентоспособной с зонной проводимостью.

С другой стороны, ясно, что при полном отсутствии компенсации, когда  $\varepsilon_3$ -проводимость равна нулю, для  $\varepsilon_2$ -проводимости имеются самые благоприятные условия (наибольшая концентрация нейтральных доноров). Увеличение компенсации существенно облегчает условия  $\varepsilon_3$ -проводимости и ухудшает условия  $\varepsilon_2$ -проводимости. В результате исчезает область температур, где первая уступает второй.

Несмотря на вполне разумную качественную интерпретацию экспериментальных фактов с помощью модели  $D$ -зоны, существует важное, на наш взгляд, возражение против этой модели. Из того, что в германии наблюдаемые значения  $\varepsilon_2$  имеют порядок величины  $1/3 E_0$ , следует, что  $D$ -зона за счет перекрытия волновых функций опустилась от энергии  $-0,1 E_0$  до энергии  $-2/3 E_0$ . Казалось бы, нет никаких оснований считать, что при этом ширина  $D$ -зоны осталась небольшой, как это изображено на рис. 4.4. Напротив, более логичным было бы предположение о том, что  $D$ -зона расширилась на величину порядка  $E_0$ , перекрылась с зоной проводимости и потеряла всякую структуру. Но тогда не ясно, чем достигается существование определенной энергии активации  $\varepsilon_2$ .

Хотелось бы, конечно, разобраться во всех этих вопросах с помощью количественной теории. Однако построение количественной теории  $\varepsilon_2$ -проводимости сталкивается с очень большими трудностями. Основная трудность состоит в том, что необходимо работать в области концентраций, где перекрытие волновых функций играет роль, сравнимую с кулоновским взаимодействием электронов с примесями и друг с другом, причем ситуация усугубляется существенной ролью беспорядка в распределении примесей. Ниже мы больше не будем возвращаться к вопросу о  $\varepsilon_2$ -проводимости. Под термином «прыжковая проводимость» мы всегда будем понимать именно  $\varepsilon_3$ -проводимость.

Хорошо известно, что в областях  $A$ ,  $B$ ,  $V$ , где электропроводность определяется электронами зоны проводимости, совместное исследование электропроводности и эф-



факта Холла позволяет найти в отдельности концентрацию и подвижность носителей и установить механизм рассеяния электронов. Посмотрим, что дает исследование эффекта Холла при переходе к прыжковой проводимости.

На рис. 4.5 даны зависимости коэффициента Холла  $R$  для германия  $p$ -типа от  $T^{-1}$  для некоторых образцов, удельное сопротивление которых приведено на рис. 4.2. Видно, что в области  $B$  коэффициент Холла растет по экспоненциальному закону благодаря вымораживанию концентрации свободных электронов  $n(T)$ . Подвижность же остается большой.

Крутой спад коэффициента Холла в области  $\Gamma$  можно объяснить, если воспользоваться так называемой двухзонной моделью. Согласно этой модели все электроны, участвующие в электропроводности, делятся на две группы: электроны зоны проводимости, обладающие электропроводностью  $\sigma_c = n(T)e\mu_c$  и коэффициентом Холла  $R_c = [n(T)ec]^{-1}$ , где  $n(T)$  — концентрация электронов в зоне, и электроны примесной зоны с прыжковой электропроводностью  $\sigma_h$  и коэффициентом Холла  $R_h$ . Тогда простой феноменологический расчет для наблюдаемого коэффициента Холла дает

$$R = \frac{R_c \sigma_c^2 + R_h \sigma_h^2}{(\sigma_c + \sigma_h)^2}. \quad (14.7)$$

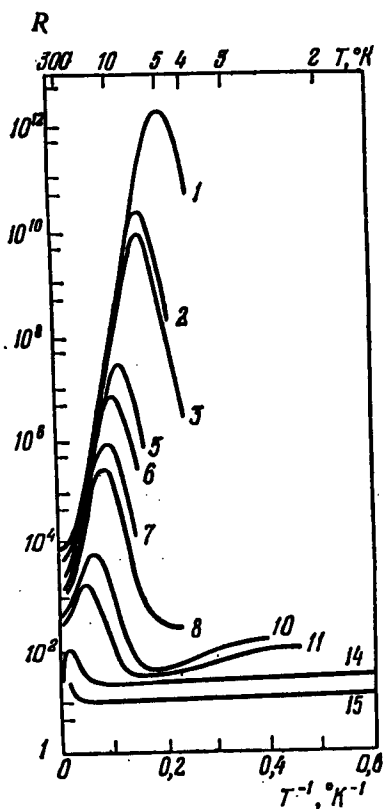


Рис. 4.5. Зависимость коэффициента Холла  $R$  (в  $\text{см}^3 \cdot \text{К}^{-1}$ ) для Ge  $p$ -типа от  $T^{-1}$  для некоторых из образцов, данные для которых приведены на рис. 4.2 [6].

Формула (14.7) объясняет температурную зависимость  $R$ , если предположить, что прыжковая холловская подвижность  $\mu_h = cR_h\sigma_h$  очень мала по сравнению с зонной  $\mu_c = cR_c\sigma_c$ . В этом случае при не слишком низких температурах вторым слагаемым в числителе (14.7) можно пренебречь и формула (14.7) дает резкий максимум при  $\sigma_c \approx \sigma_h$ , т. е. при переходе от зонной электропроводности к прыжковой. Левее максимума, при  $\sigma_c \gg \sigma_h$ ,

$$R = R_c = [n(T)ec]^{-1} \propto \exp\left(\frac{e_1}{kT}\right). \quad (14.8)$$

Правее максимума, т. е. при  $\sigma_c \ll \sigma_h$ ,

$$R = \frac{R_c\sigma_c^2}{\sigma_h^2} = \frac{\mu_c^2 e}{c\sigma_h^2} n(T) \propto \exp\left(\frac{2e_3 - e_1}{kT}\right). \quad (14.9)$$

Если пренебречь  $e_3$  по сравнению с  $e_1$ , то окажется, что справа от максимума значение  $R$  должно убывать пропорционально  $\exp(-e_1/kT)$ . Это объясняет тот факт, что углы наклона прямых слева и справа от максимума близки.

Видно, что измерение коэффициента Холла в области действия (14.9) никак не позволяет выделить холловскую подвижность и концентрацию из  $\sigma_h$ . Такая возможность возникла бы, если бы удалось продвинуться в область более низких температур, где существенно второе слагаемое в числителе (14.7). Но это обычно не удается сделать из-за высокоомности образцов. По этой причине в настоящее время про холловскую подвижность твердо известно лишь то, что она очень мала. Ссылки на работы, посвященные теории прыжкового эффекта Холла, можно найти в обзоре Бетгера и Брыксина (см. [2] к предисловию).

Дрейфовую подвижность можно оценить, если считать, что в прыжковой проводимости принимают участие практически все носители, находящиеся на примесях. Например, для образца 4 р-германия при 2,5°K (см. рис. 4.2) мы получим

$$\mu = [\rho e(N_A - N_D)]^{-1} \approx 6 \cdot 10^{-3} \text{ в}^{-1} \cdot \text{сек}^{-1} \cdot \text{см}^2. \quad (14.10)$$

Такое малое значение подвижности полностью исключает какую бы то ни было возможность интерпретиро-

вать прыжковую проводимость, исходя из представлений об эпизодическом рассеянии квазисвободных электронов (например, на фононах), т. е. использовать для ее вычисления кинетическое уравнение Больцмана. Действительно, если допустить такую мысль и записать  $\mu$  в виде  $e\tau/m$ , где  $\tau$  — время релаксации, то, используя значение (14.10) и массу свободного электрона, мы получим для  $\tau$  величину  $4 \cdot 10^{-18}$  сек. Неопределенность в энергии  $\hbar/\tau$ , соответствующая такому значению  $\tau$ , составляет сотни электронвольт, т. е. существенно превышает характерные атомные энергии. Это, конечно, противоречит представлению о свободных электронах и делает его неприменимым.

Для построения теории прыжковой проводимости нужно, таким образом, использовать иные исходные представления. Ясно, что главным среди них должно быть представление о локализованных электронных состояниях. Взаимодействие с фононами и перекрытие волновых функций локализованных состояний должны приводить к редким прыжкам с одного состояния на другое.

В следующем параграфе мы сформулируем аппарат, основанный на таких представлениях и являющийся основой теории прыжковой проводимости.

### § 15. Модель сетки сопротивлений Миллера и Абрахамса

Миллер и Абрахамс [14] предложили следующий способ рассуждений. Исходя из локализованных на отдельных донорах волновых функций электронов, вычислим вероятность перехода электрона между двумя донорами,  $i$  и  $j$ , с поглощением или излучением фонона. Подсчитаем затем число переходов  $i \rightarrow j$  в единицу времени. В отсутствие электрического поля такое же количество электронов совершает обратный переход, т. е. имеет место детальный баланс. В слабом электрическом поле прямые и обратные переходы окажутся несбалансированными, т. е. возникнет ток, пропорциональный электрическому полю. Вычислив ток, мы найдем сопротивление данного перехода  $R_{ij}$  и сведем всю задачу к вычислению электропроводности эквивалентной сетки случайных

сопротивлений. Ниже мы подробно проследим, как реализуется эта программа.

Рассмотрим два донора  $i$  и  $j$  с координатами  $r_i$  и  $r_j$ , на которых находится один электрон. Пусть  $\Psi(r - r_i) \equiv \Psi_i(r)$  — волновая функция основного состояния изолированного донора, удовлетворяющая уравнению (2.5) с  $U = -\frac{e^2}{\kappa|r - r_i|}$ . Будем считать, что расстояние  $r_{ij} = |r_i - r_j|$  много больше характерного размера волновой функции, так что перекрытие функций  $\Psi_i$  и  $\Psi_j$  слабое. За счет взаимодействия электрона с обоими донорами вырожденное состояние расщепляется. В рамках метода линейных комбинаций атомных орбиталей (ЛКАО) волновые функции расщепившихся состояний представляют собой симметричную и антисимметричную комбинации атомных функций:

$$\Psi_{1,2} = 2^{-1/2} \left( 1 \pm \int \Psi_i^* \Psi_j dr \right)^{-1/2} (\Psi_i \pm \Psi_j). \quad (15.1)$$

Вычисление энергий состояний 1, 2 с гамильтонианом

$$H = H_0 - \frac{e^2}{\kappa|r - r_i|} - \frac{e^2}{\kappa|r - r_j|} \quad (15.2)$$

даёт

$$E_{1,2} = -E_0 - \frac{e^2}{\kappa r_{ij}} \mp I_{ij}. \quad (15.3)$$

Здесь  $H_0$  — гамильтониан электрона в периодическом кристалле,  $-E_0$  — уровень энергии изолированного донора, а энергетический интеграл перекрытия  $I_{ij}$  равен

$$I_{ij} = \int \Psi_i^* \Psi_j \frac{e^2}{\kappa|r - r_j|} dr - \int \Psi_i^* \Psi_j dr' \int \frac{e^2 |\Psi_i|^2}{\kappa|r - r_j|} dr. \quad (15.4)$$

Рассмотрим простой случай, когда состояние на доноре связано с одним экстремумом в центре зоны, т. е.  $\Psi_i(r)$  имеет вид (2.21). Подставим (2.21) в (15.4) и разобьем каждый интеграл на сумму интегралов по элементарным ячейкам кристалла. Затем, воспользовавшись тем, что огибающие функции  $F(r)$  практически не меняются на длине постоянной решетки  $a_0$ , вынесем их за знак интеграла по ячейке. То же самое можно проделать и с множителями  $|r - r_j|^{-1}$  в существенной для интегрирования области  $|r - r_j| \gg a_0$ . Интегралы по ячейке

от  $|u_{n,0}|^2$  в соответствии с (2.10) равны единице. Переходя обратно от сумм по ячейкам к интегралам, получим

$$I_{ij} = \int F_i(r) F_j(r) \frac{e^2}{\kappa |r - r_j|} dr - \int F_i(r') F_j(r') dr' \int \frac{e^2 F_i^2(r)}{\kappa |r - r_j|} dr. \quad (15.5)$$

Для водородной функции (2.17) вычисление (15.5) приводит к простой формуле:

$$I_{ij} = \frac{2}{3} \left( \frac{e^2}{\kappa a} \right) \left( \frac{r_{ij}}{a} \right) \exp \left( - \frac{r_{ij}}{a} \right). \quad (15.6)$$

Заметим, что существует более точный расчет энергетического расщепления (15.3), выходящий за рамки метода ЛКАО и учитывающий деформацию волновой функции посередине между двумя центрами, где их потенциалы сравнимы. Этот расчет приводит к замене множителя  $2/3$  в (15.6) на  $2/e$  (см. [8 (гл. 1), стр. 363]).

Выше мы уже говорили, что в слабо легированном полупроводнике рассмотренная резонансная ситуация практически не реализуется из-за сильного влияния энергии взаимодействия  $W(r)$  электрона с окружающими доноры  $i, j$  заряженными примесями. Почти для всех пар доноров выполняется неравенство

$$\Delta_i^j \gg I_{ij}, \quad \text{где} \quad \Delta_i^j \equiv W(r_j) - W(r_i). \quad (15.7)$$

Включив потенциальную энергию  $W(r)$  в гамильтониан и выполнив вариационный расчет с функцией

$$\Psi = A_i \Psi_i + A_j \Psi_j, \quad (15.8)$$

можно показать [14], что при условии (15.7) волновые функции двух низких состояний имеют вид

$$\Psi_i' = \Psi_i + \frac{I_{ij}}{\Delta_i^j} \Psi_j, \quad (15.9)$$

$$\Psi_j' = \Psi_j - \frac{I_{ij}}{\Delta_i^j} \Psi_i. \quad (15.10)$$

Они мало отличаются от функций изолированных доноров, что неудивительно, так как речь идет о ситуации,

когда, согласно (15.7), имеется полная андерсоновская локализация. Поэтому, в отличие от (15.1), в (15.9) и (15.10) имеет смысл сохранять индексы  $i$  и  $j$  у функций  $\Psi'$ . Ясно, что переход электрона из состояния  $\Psi'_i$  в состояние  $\Psi'_j$  означает перенос заряда  $-e$  на расстояние  $r_{ij}$ . С такими переходами и связано возникновение тока.

Разность энергий состояний (15.9) и (15.10) при условии (15.7) с точностью до  $I_{ij}^2/\Delta_i'$  совпадает с величиной  $\Delta_i'$ . Поэтому энергия фонона, поглощаемого при переходе  $i \rightarrow j$ , равна  $\Delta_i'$ . Поскольку величина  $\Delta_i'$  обычно составляет несколько миллиэлектронвольт, переход  $i \rightarrow j$  может происходить только с участием длинноволнового акустического фонона. Для простоты мы вслед за Миллером и Абрахамсом будем считать, что электрон взаимодействует только с одной акустической ветвью, спектр которой изотропен.

Вероятность перехода  $i \rightarrow j$  с поглощением фонона равна

$$\gamma_{ij} = \frac{2\pi}{\hbar} \frac{V_0}{(2\pi)^3} \int |M_q|^2 \delta(\hbar s q - \Delta_i') dq, \quad (15.11)$$

где  $V_0$  — объем кристалла,  $s$  — скорость звука,  $q$  — волновой вектор фонона,

$$M_q = iE_1 \left( \frac{\hbar q N_q}{2V_0 s d} \right)^{1/2} \int \Psi'_j e^{iqr} \Psi'_i dr \quad (15.12)$$

— матричный элемент электрон-фононного взаимодействия,  $E_1$  — константа деформационного потенциала,  $d$  — плотность кристалла,  $N_q$  — число фононов с импульсом  $q$ . Подставим в (15.12) волновые функции (15.9), (15.10) и, используя (2.21), так же как при переходе от (15.4) к (15.5), исключим из рассмотрения периодические функции  $u_{n,0}$ . В результате получим

$$M_q = iE_1 \left( \frac{\hbar q N_q}{2V_0 s d} \right)^{1/2} \left\{ \frac{I_{ij}}{\Delta_i'} \left[ \int F_j^2 e^{iqr} dr - \int F_i^2 e^{iqr} dr \right] + \right. \\ \left. + \left[ 1 - \left( \frac{I_{ij}}{\Delta_i'} \right)^2 \right] \int F_i F_j e^{iqr} dr \right\}. \quad (15.13)$$

Для дальнейшего важно оценить величину безразмерных параметров

$$qr_{ij} = \frac{\Delta_i^j r_{ij}}{\hbar s} \quad \text{и} \quad qa = \frac{a\Delta_i^j}{\hbar s}.$$

Учитывая, что  $\Delta_i^j$  имеет порядок величины  $e^2 N_D^{1/3} \kappa^{-1}$ , а  $r_{ij} \approx N_D^{-1/3}$ , легко увидеть, что параметр  $qr_{ij}$  оказывается большим ( $\approx 20-30$ ), а параметр  $qa$  порядка единицы. Величина  $F_i F_j$  вблизи прямой, соединяющей центры, имеет характерный масштаб изменения  $\sqrt{r_{ij} a}$ . Поэтому в последнем интеграле (15.13) за счет осциллирующего множителя происходят сильные сокращения, позволяющие им пренебречь. Учитывая, что  $F_j(r - r_{ij}) = F_j(r)$ , выражение в первых квадратных скобках можно переписать в виде

$$(e^{iqr_{ij}} - 1) \int F_i^2 e^{iqr} dr. \quad (15.14)$$

Для водородной функции (2.17)

$$\int F_i^2 e^{iqr} dr = \left[ 1 + \left( \frac{qa}{2} \right)^2 \right]^{-2}. \quad (15.15)$$

Тогда

$$\begin{aligned} |M_q|^2 &= \\ &= \left( \frac{\hbar q E_1^2}{2 d V_0 s} \right) \left( \frac{I_{ij}}{\Delta_i^j} \right)^2 N_q (1 - \cos qr_{ij}) \left[ 1 + \left( \frac{qa}{2} \right)^2 \right]^{-4}. \end{aligned} \quad (15.16)$$

При подстановке (15.16) в (15.11) интегралом, содержащим  $\cos qr_{ij}$ , при интегрировании по направлениям  $q$  можно пренебречь благодаря большой величине параметра  $qr_{ij}$ . Кроме того, ниже мы всегда будем считать, что электрическое поле достаточно слабое, чтобы не выводить фононы из равновесия. Поэтому  $N_q$  — равновесная планковская функция распределения. Тогда из (15.11) мы получаем для вероятности перехода

$$\gamma_{ij} = \gamma_{ij}^0 \exp\left(-\frac{2r_{ij}}{a}\right) N(\Delta_i^j), \quad (15.17)$$

где

$$\gamma_{ij}^0 = \frac{E_1^2 \Delta_i^j}{\pi ds^5 \hbar^4} \left( \frac{2e^2}{3\kappa a} \right)^2 \frac{r_{ij}^2}{a^2} \left[ 1 + \left( \frac{\Delta_i^j a}{2\hbar s} \right)^2 \right]^{-4}, \quad (15.18)$$

$$N(\Delta_i^j) = \left[ \exp\left(\frac{\Delta_i^j}{kT}\right) - 1 \right]^{-1}. \quad (15.19)$$

Пусть  $n_i = (0; 1)$  — флуктуирующее во времени число заполнения донора  $i$ . Переход  $i \rightarrow j$  возможен только в те моменты времени, когда  $n_i = 1$ ,  $n_j = 0$ . Поэтому число электронов, совершающих переход в единицу времени, равно

$$\Gamma_{ij} = \langle \gamma_{ij} n_i (1 - n_j) \rangle, \quad (15.20)$$

где усреднение происходит по времени. Величина  $\gamma_{ij}$  также флуктуирует во времени. Это происходит из-за того, что числа заполнения соседних с  $i$  и  $j$  доноров, которые своим потенциалом создают разность энергий  $\Delta_i^j$ , меняются во времени. За счет этого возникают флуктуации  $\Delta_i^j$  и, следовательно, флуктуации  $\gamma_{ij}$ .

Сделаем теперь очень важное упрощающее предположение. Допустим, что числа заполнения и энергии узлов не флуктуируют во времени и равны своим средним значениям. Другими словами, будем описывать систему в рамках приближения самосогласованного поля, аналогичного приближению Хартри. Это приближение заключается в следующем. 1) Каждый донор будем характеризовать средним числом заполнения  $\langle n_i \rangle \equiv f_i$ . 2) Каждому донору сопоставим среднюю по времени энергию электронного уровня в поле всех остальных примесей и электронов

$$e_i = \sum_l^{\text{акч}} \frac{e^2}{\kappa |r_l - r_i|} - \sum_{k \neq i}^{\text{дон}} \frac{e^2 (1 - f_k)}{\kappa |r_l - r_k|}. \quad (15.21)$$

Здесь первое суммирование происходит по всем акцепторам, а второе — по всем донорам, за исключением донора  $i$ ; величина  $(1 - f_k)e$  имеет смысл среднего заряда донора  $k$ . 3) Энергию фонона, поглощаемого при переходе  $i \rightarrow j$ , запишем в виде  $\Delta_i^j = e_j - e_i$ . В таком



приближении

$$\Gamma_{ij} = \gamma_{ij}^0 \exp\left(-\frac{2r_{ij}}{a}\right) N(\epsilon_i - \epsilon_j) f_i (1 - f_j). \quad (15.22)$$

Переход от (15.20) к (15.22) нетривиален, и в общем случае его нельзя обосновать. Анализ применимости этого приближения будет дан в гл. 10. Мы увидим, что многие результаты, полученные с его помощью, правильны. Например, экспоненциальная зависимость  $\rho_3$  от концентрации вообще не чувствительна к этому приближению. Это видно уже сейчас, так как экспоненциальный множитель  $\exp\left(-\frac{2r_{ij}}{a}\right)$ , создающий концентрационную зависимость, в силу независимости от времени вообще, не «замешан» в усреднении (15.20). Как мы увидим в гл. 10, в рамках этого приближения правильно вычисляется и энергия активации  $\epsilon_3$  при очень слабой и очень сильной компенсациях. Однако в случае промежуточной компенсации или в случае электропроводности с переменной длиной прыжка (см. гл. 9) указанное приближение не удастся оправдать, и это приводит, как нам кажется, к неизвестному численному множителю в показателе экспоненты температурной зависимости электропроводности.

По аналогии с (15.22), для обратного процесса,  $j \rightarrow i$ , с излучением фонона имеем

$$\Gamma_{ji} = \gamma_{ji}^0 \exp\left(-\frac{2r_{ij}}{a}\right) [N(\epsilon_i - \epsilon_j) + 1] f_j (1 - f_i). \quad (15.23)$$

С помощью  $\Gamma_{ij}$  и  $\Gamma_{ji}$  можно записать ток между донорами  $i$  и  $j$ :

$$\mathcal{J}_{ij} = -e(\Gamma_{ij} - \Gamma_{ji}). \quad (15.24)$$

Если электрическое поле  $E$  отсутствует, то функции  $f_i$  имеют равновесный вид:

$$f_i = f_i^0 = \left[ \frac{1}{2} \exp\left(\frac{\epsilon_i^0 - \mu}{kT}\right) + 1 \right]^{-1}, \quad (15.25)$$

где  $\epsilon_i^0$  — среднее значение энергии на узле  $i$  при  $E = 0$ , а множитель  $1/2$  связан с наличием двух спиновых состояний занятого узла. Величины  $\epsilon_i^0$  и  $f_i^0$  самосогласо-

ваны с помощью (15.21). При интересующих нас низких температурах энергии  $\epsilon_i^0$  очень слабо отличаются от энергий, соответствующих нулевой температуре, распределение которых подробно изучено в гл. 3.

Легко убедиться, что в равновесии имеет место детальный баланс между переходами  $i \rightarrow j$  и  $j \rightarrow i$ . Действительно, используя (15.19) и (15.23)—(15.25), видим, что при  $E = 0$   $\Gamma_{ij} = \Gamma_{ji} = \Gamma_{ij}^0$  и, следовательно, ток  $\mathcal{I}_{ij} = 0$ . Электрическое поле, конечно, нарушает этот баланс. Происходит это по следующим причинам. Во-первых, поле перераспределяет электроны по донорам, т. е. создает добавки  $\delta f_i$  к функциям (15.25). Запишем эти добавки, введя малые величины  $\delta\mu_i$ :

$$f_i(E) = f_i^0 + \delta f_i = \left[ 1 + \frac{1}{2} \exp\left(\frac{\epsilon_i^0 - \delta\mu_i - \mu}{kT}\right) \right]^{-1}. \quad (15.26)$$

Во-вторых, под действием поля происходит изменение энергий уровней доноров  $i$  и  $j$ :

$$\epsilon_i = \epsilon_i^0 + \delta\epsilon_i,$$

$$\delta\epsilon_i \equiv eEr_i + \frac{e^2}{\kappa} \sum_{k \neq i}^{\text{дон}} \frac{\delta f_k}{|r_i - r_k|}. \quad (15.27)$$

Первое слагаемое величины  $\delta\epsilon_i$  описывает непосредственное воздействие внешнего поля  $E$ , а второе — изменение кулоновского потенциала за счет перераспределения электронов (см. (15.21)). В результате изменяется энергия поглощаемого фонона  $\epsilon_j - \epsilon_i$ , входящая в аргумент функции Планка.

Если внешняя цепь разомкнута (образец находится в конденсаторе), то в электрическом поле возникает новое состояние равновесия, в котором  $\delta\mu_i = -\delta\epsilon_i$ . Для вычисления электропроводности следует рассмотреть случай замкнутой цепи. В этом случае равновесие нарушается и  $\delta\mu_i \neq -\delta\epsilon_i$ . Если электрическое поле настолько мало, что добавки  $\delta\mu_i$  и  $\delta\epsilon_i$  малы по сравнению с  $kT$ , то в выражениях для  $\Gamma_{ij}$  и  $\Gamma_{ji}$  функции  $f_i$ ,  $f_j$ ,  $N(\epsilon_j - \epsilon_i)$  можно разложить в ряд по этим добавкам. Тогда для тока  $\mathcal{I}_{ij}$  в линейном по внешнему полю приближении с помощью простых выкладок из (15.22)—(15.25)

получаем

$$\mathcal{J}_{ij} = \frac{e\Gamma_{ij}^0}{kT} [\delta\mu_i + \delta e_j - (\delta\mu_i + \delta e_i)], \quad (15.28)$$

где  $\Gamma_{ij}^0$  — частота переходов  $i \rightarrow j$  и  $j \rightarrow i$  в равновесии. Выражение для тока  $\mathcal{J}_{ij}$  можно переписать в виде, напоминающем закон Ома:

$$\mathcal{J}_{ij} = R_{ij}^{-1} (U_i - U_j), \quad (15.29)$$

где

$$R_{ij} = \frac{kT}{e^2 \Gamma_{ij}^0}, \quad (15.30)$$

$$-eU_i = \delta e_i + \delta\mu_i = eEr_i + \delta\mu_i + \frac{e^2}{\kappa} \sum_{k \neq i}^{\text{дон}} \frac{\delta f_k}{|r_i - r_k|}. \quad (15.31)$$

Величину  $-eU_i$  можно рассматривать как отсчитанное от  $\mu$  локальное значение электрохимического потенциала электрона на доноре  $i$ . Тогда  $U_i - U_j$  есть напряжение, приходящееся на переход  $i \rightarrow j$ , а величина  $R_{ij}$  имеет смысл сопротивления перехода  $i \rightarrow j$ . Переходя от двух доноров к реальной системе из большого числа доноров, мы видим, что если каким-либо образом нам удастся найти величины  $U_i$ , то можно вычислить ток между любыми двумя донорами. Полный ток, проходящий через образец, определяется суммой токов, пересекающих сечение образца. Знание отдельных слагаемых (15.31), как мы видим, не обязательно для вычисления тока.

Схема определения напряжений  $U_i$  в принципе состоит в следующем. Для доноров в непосредственной близости металлических контактов напряжения  $U_i$  равны потенциалу контакта. Если образец имеет длину  $\mathcal{L}$ , то по условиям задачи между контактами есть разность потенциалов  $eE\mathcal{L}$ . Именно через эти граничные условия в задачу входит внешнее поле  $E$ . Все остальные напряжения определяются из условия равенства втекающего и вытекающего токов для каждого донора, т. е. из условия стационарности функций  $f_i$ .

Таким образом, электропроводность образца полностью определяется сопротивлениями  $R_{ij}$ . Фактически задача о вычислении прыжковой проводимости свелась к вычислению проводимости случайной сетки, каждый

узел которой совпадает с одним из доноров, а между каждой парой узлов включено сопротивление (15.30) (рис. 4.6).

Для дальнейшего нам удобно в выражении для  $R_{ij}$  (15.30) выделить множитель, экспоненциально зависящий от  $r_{ij}$  и энергий доноров, и множитель, содержащий более слабую степенную зависимость от этих параметров. Подставляя в выражение для  $\Gamma_{ij}$  равновесные функции  $N$ ,  $f_i^0$  и  $f_j^0$ , легко проверить, что в случае низких температур,  $kT \ll |e_i^0 - e_j^0|$ ,  $|e_i^0 - \mu|$ ,  $|e_j^0 - \mu|$ , и при произвольном расположении  $e_i^0$  и  $e_j^0$  относительно уровня Ферми величину  $\Gamma_{ij}^0$  можно записать в виде

$$\Gamma_{ij}^0 = \gamma_{ij}^0 \exp\left(-\frac{2r_{ij}}{a}\right) \exp\left(-\frac{\epsilon_{ij}}{kT}\right), \quad (15.32)$$

где

$$\epsilon_{ij} = \frac{1}{2}(|e_i - e_j| + |e_i - \mu| + |e_j - \mu|). \quad (15.33)$$

Тогда согласно (15.30)

$$R_{ij} = R_{ij}^0 \exp \xi_{ij}, \quad (15.34)$$

где

$$\xi_{ij} = \frac{2r_{ij}}{a} + \frac{\epsilon_{ij}}{kT}, \quad (15.35)$$

$$R_{ij}^0 = \frac{kT}{e^2 \gamma_{ij}^0}. \quad (15.36)$$

В формуле (15.33) и везде ниже мы для упрощения записи опускаем индекс 0 в обозначении равновесных энергий, самосогласованных с помощью (15.21) и (15.25). Подчеркнем еще раз, что, поскольку нас интересуют температуры, при которых разброс энергий примесей много больше, чем  $kT$ , энергии, фигурирующие в (15.33), имеют тот же смысл, что и энергии донорных уровней в основном состоянии всей системы, распределение которых исследовалось в гл. 3.

Важной особенностью сетки сопротивлений (15.34) является чрезвычайно широкий спектр значений  $R_{ij}$ . В типичной для исследования прыжковой проводимости в слабо легированных полупроводниках ситуации сред-

нее расстояние между донорами  $r_D = \left(\frac{4\pi}{3} N_D\right)^{-1/3}$  составляет  $6 \div 12$  боровских радиусов  $a$ . При этом сопротивления пары с  $r_{ij} = 2r_D$  и пары с  $r_{ij} = r_D$  отличаются в  $e^{12} \div e^{24}$  раз. К сильному разбросу сопротивлений при достаточно низких температурах приводит и энергетическое слагаемое (15.35).

Нахождение электропроводности случайной сетки с экспоненциально широким спектром сопротивлений представляет собой наиболее сложную задачу на пути вычисления прыжковой проводимости. Общий метод ее решения будет сформулирован в §§ 21, 22. Здесь же мы рассмотрим некоторые неверные, но поучительные подходы к ее решению и, разобрав их слабые места, лучше почувствуем суть задачи.

На первый взгляд, неплохое приближение можно было бы получить, если пренебречь вторым и третьим слагаемыми в (15.31). Тогда все токи оказались бы известными:

$$j_{ij} = R_{ij}^{-1} r_{ij} E \cos \theta_{ij}. \quad (15.37)$$

Здесь  $\theta_{ij}$  — угол между  $r_{ij}$  и  $E$ . Для вычисления плотности тока можно воспользоваться следующим приемом. Пусть  $P$  — поляризация (дипольный момент единицы объема). Тогда плотность тока  $j$  можно записать в виде

$$j = \frac{dP}{dt} = V_0 \sum_I \left(-e \frac{df_I}{dt}\right) r_I, \quad (15.38)$$

где  $V_0$  — объем системы, а суммирование идет по всем донорам. Учитывая, что  $-e \frac{df_I}{dt} = \sum_j j_{ij}$  и  $j_{ij} = -j_{ji}$ ,

получаем

$$j = \frac{1}{V_0} \sum_{i \neq j} j_{ij} r_i = -\frac{1}{V_0} \sum_{i \neq j} j_{ij} r_j = \frac{1}{2V_0} \sum_{i \neq j} j_{ij} r_{ij}. \quad (15.39)$$

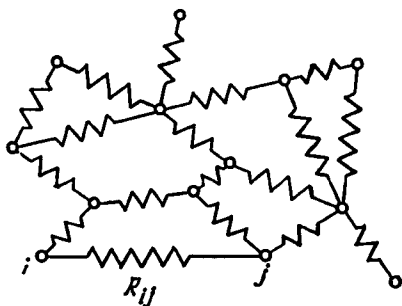


Рис. 4.6. Случайная сетка сопротивлений Миллера и Абрахамса [14].

Взяв проекцию (15.39) на направление поля, используя (15.37), находим

$$\rho^{-1} = \frac{1}{2V_0} \sum_{i \neq j} R_{ij}^{-1} r_{ij}^2 \cos^2 \theta_{ij}. \quad (15.40)$$

Ввиду экспоненциальной малости подавляющего большинства обратных сопротивлений  $R_{ij}^{-1}$  в сумме (15.40) будут играть определяющую роль слагаемые с  $r_{ij} \approx a$ ,  $v_{ij} \approx kT$ . В результате окажется, что  $\rho$  невелико и неэкспоненциально зависит от концентрации и температуры. Ясно, что этот результат неверен. Ведь при интересующих нас значениях  $r_D \gg a$  пары с  $r_{ij} \approx a$  встречаются очень редко. Они напоминают маленькие изолированные металлические островки в сравнительно плохо проводящем море и, очевидно, не могут определять электропроводность системы. Ясна и причина ошибки. В действительности на «металлических» парах падения напряжения  $U_j - U_i$  очень малы по сравнению с  $eEr_{ij}$ , т. е. именно для них наше исходное предположение ошибочно.

Выше мы практически имели дело с усреднением локальных значений электропроводности. Миллер и Абрахамс [14] использовали противоположный подход к задаче, основанный на усреднении сопротивлений  $R_{ij}$ . Мы дадим простую интерпретацию их расчета для случая сравнительно высоких температур, когда температурный экспоненциальный множитель в (15.32) порядка единицы и

$$R_{ij} = R_{ij}^0 \exp\left(\frac{2r_{ij}}{a}\right). \quad (15.41)$$

Фактически мы вычислим удельное сопротивление  $\rho_3$  в формуле (14.5). Миллер и Абрахамс предположили, что случайная сетка эффективно эквивалентна набору независимых цепочек сопротивлений, причем каждая цепочка проходит через весь образец от контакта до контакта. Сопротивления одной цепочки включены последовательно, и поэтому сопротивление всей цепочки равно сумме отдельных сопротивлений  $R_{ij}$ . Далее, они обратили внимание на то, что вследствие случайного распределения примесей в кристалле существуют области раз-

мера  $r$ , большего, чем  $N_D^{-1/3}$ , в которых вообще нет ни одного донора (рис. 4.7).

Вероятность возникновения разрежения в виде сферы с радиусом  $r$  дается распределением Пуассона и равна  $\exp\left(-\frac{4\pi}{3} N_D r^3\right)$ . Миллер и Абрахамс сочли, что любая цепочка доноров, соединяющая контакты, пересекает эти

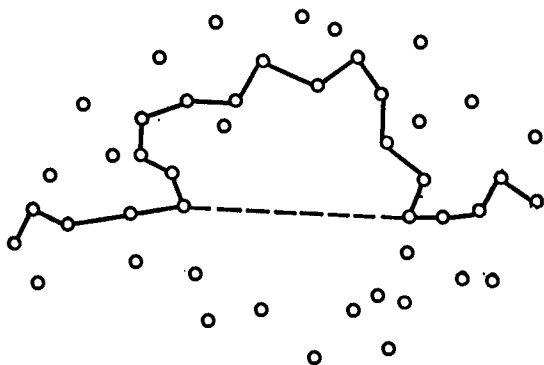


Рис. 4.7. К выводу формулы (15.44). Кружками изображены доноры. Ломаная линия представляет собой цепочку со сравнительно малым сопротивлением, обходящую случайно возникшее разрежение. Пунктиром показано, как, согласно Миллеру и Абрахамсу, цепочки пересекают разрежение.

полости. Перескок электрона через такую полость очень маловероятен, т. е. полость имеет очень большое сопротивление, порядка  $R_{ij}^0 \exp(2r/a)$ . Вклад от полостей доминирует в сопротивлении цепочки и равен

$$\int R_{ij}^0 \exp\left(\frac{2r}{a}\right) \exp\left(-\frac{4\pi}{3} N_D r^3\right) dr. \quad (15.42)$$

Произведение экспонент в (15.42) имеет острый максимум при  $r = r_m$ , где

$$r_m = \sqrt{\frac{2}{3}} r_D \left(\frac{r_D}{a}\right)^{1/2} \gg r_D, \quad r_D = \left(\frac{4\pi}{3} N_D\right)^{-1/3}. \quad (15.43)$$

Таким образом, основной вклад в сопротивление цепочки вносят элементы с длиной, близкой к  $r_m$ . Подставляя (15.43) в (15.41), мы найдем сопротивление типичной цепочки. Считая, что все цепочки имеют одинаковое

сопротивление, для удельного сопротивления  $\rho_3$  получаем

$$\rho_3 \propto \exp \left\{ 1,09 \left( \frac{r_D}{a} \right)^{3/2} \right\}. \quad (15.44)$$

Ясно, в чем недостаток рассуждений, приводящих к (15.44). Действительно, мы получили, что сопротивление определяется полостями с радиусом  $r_m$ , которые встречаются с экспоненциально малой вероятностью  $\exp \left( -\frac{4\pi}{3} N_D r_m^3 \right)$ . Такие полости занимают ничтожную часть объема. Поэтому, кроме цепочек, пересекающих полости, существует огромное число цепочек, обходящих их стороной (см. рис. 4.7). Эти цепочки обладают гораздо меньшим сопротивлением. Мы имеем здесь как бы островки диэлектрика в море проводящей фазы, и ясно, что эти островки не могут играть определяющую роль в эффективной электропроводности. Недостаток предыдущих рассуждений фактически заключается в том, что считалось, что большие полости всегда включены последовательно, т. е. через них течет ток того же порядка, что и через типичные сопротивления длиной  $r_D$ . На самом же деле эти полости включены параллельно к основным токопроводящим путям и ток через них исчезающе мал.

Мы рассмотрели два противоположных, можно сказать, экстремистских подхода к вычислению эффективной электропроводности сетки: усреднение проводимостей и усреднение сопротивлений. У этих подходов есть общий недостаток: и тот и другой приводят к выделенной роли аномально редко встречающихся сопротивлений. Положительное значение рассмотренных расчетов состоит в том, что они дают оценки электропроводности сверху и снизу, правда, очень слабые. Еще один метод вычисления эффективной электропроводности сетки Миллера и Абрахамса был предложен Пиппардом и реализован Тузом (см. [7]). Метод Пиппарда—Туза является более уравновешенным, в том смысле, что в нем удается избежать указанных выше крайностей. Он представляет собой вариант метода эффективной среды, широко используемого для усреднения характеристик неоднородных сред [15]. Рассматривалась задача о вычисле-



нии удельного сопротивления  $\rho_3$  случайной сетки с сопротивлениями вида (15.41). Сначала отдельный элемент с сопротивлением  $R_{ij}$  заменялся шаром радиуса  $r_{ij}$  и проводимостью  $(R_{ij}r_{ij})^{-1}$ . Окружение этого элемента заменялось сплошной средой с искомой макроскопической проводимостью  $\rho_3^{-1}$ . После этого находилась плотность тока внутри сферы как функция  $r_{ij}$  и  $\rho_3^{-1}$ . Найденная плотность тока усреднялась по распределению длин  $r_{ij}$ . Выразив, таким образом, среднюю плотность тока через среднюю проводимость  $\rho_3^{-1}$ , можно получить интегральное уравнение для  $\rho_3^{-1}$ . Его решение приводит к результату

$$\rho_3 \propto \exp\left(\frac{0,9}{N_D^{1/3}a}\right). \quad (15.45)$$

Удельное сопротивление (15.45) имеет промежуточное значение между результатами первых двух методов. Согласно (15.45) сопротивление определяется прыжками с эффективной длиной порядка среднего расстояния  $r_D$ , что кажется вполне разумным. Тем не менее численный коэффициент в показателе не надежен. Ведь теория эффективной среды дает точные результаты лишь при слабой относительной неоднородности и по своему смыслу не претендует на хорошие результаты при очень сильной экспоненциальной неоднородности. В гл. 6 мы получим численный коэффициент в показателе экспоненты  $\rho_3$ , почти в два раза больший, чем в (15.45). При интересующих нас малых значениях параметра  $N_D^{1/3}a \approx 0,1$  такое различие оказывается очень существенным.

Новый метод вычисления электропроводности случайной сетки, основанный на математической теории протекания, был независимо предложен Амбегаокаром, Гальпериным и Ланжером [16], Поллаком [17] и Шкловским и Эфросом [18]. Этот метод позволил построить количественную теорию прыжковой проводимости. Подробное изложение этой теории и ее результатов авторы считают одной из основных задач книги. Ввиду того, что теория протекания не вошла еще в круг хорошо известных математических дисциплин, мы дадим в следующей главе довольно подробное введение в эту теорию.

## ГЛАВА 5

### ТЕОРИЯ ПРОТЕКАНИЯ

#### § 16. Решеточные задачи

Слово «протекание» — вольный перевод английского термина *percolation* (просачивание), впервые употребленного в 1957 г. Бродбентом и Хаммерсли [1] в связи с ими же введенным новым классом математических задач. Задачи эти естественно возникают при рассмотрении протекания жидкости по случайному лабиринту. Это приложение и дало название всей математической теории.

Рассмотрим постановку простейших задач теории протекания. Представим себе бесконечную пространственную или плоскую решетку. Назовем связями отрезки между ближайшими узлами решетки. Пусть по каждой связи в обе стороны может протекать некоторая жидкость, так что каждый смоченный узел мгновенно смачивает все соседние узлы. Если в такой решетке смочить один узел, то все остальные узлы также оказываются смоченными. Различным образом вводя в условия задачи случайные элементы, можно получить разные задачи теории протекания.

Начнем с так называемой «задачи связей» (*bond problem*). Предположим, что каждая связь может находиться в двух состояниях. Она может быть разорванной (и тогда она не пропускает жидкость) или целой. Пусть вероятность того, что произвольная связь цела, есть  $x$  и не зависит от состояния других связей. Тогда мы имеем «идеальный раствор» целых и разорванных связей. Концентрация (относительная доля) первых есть  $x$ , а концентрация вторых  $1 - x$ . Картина распределения целых и разорванных связей фиксирована и во времени не изменяется.

Теперь после смачивания одного узла решетки могут возникнуть две разные ситуации. Исходный узел может смочить либо конечное, либо бесконечное число узлов. Какая из этих двух возможностей реализуется, зависит, конечно, от доли целых связей решетки. Однако в силу случайного расположения целых и разорванных связей важное значение имеет и выбор исходного узла. Так, на рис. 5.1 видно, что узел *A* смачивает большое (и, возможно, бесконечное) число узлов, а узел *B* не смачивает ни одного другого узла.

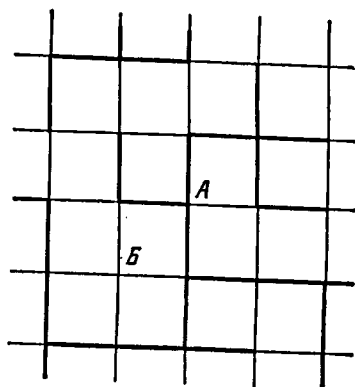


Рис. 5.1. Задача связей на квадратной решетке. Толстыми линиями изображены целые связи, тонкими — разорванные связи.

Для характеристики всей системы в целом удобно говорить не об одном конкретном исходном узле, а о вероятности того, что произвольный исходный узел смачивает бесконечное число узлов. Очень важно, что в бесконечной решетке эта вероятность не зависит от конкретной реализации, т. е. от того, каким образом легли целые и разорванные связи. Для заданной решетки она зависит только от  $x$ , и мы будем обозначать ее  $P^{(b)}(x)$  (индекс  $b$  указывает на то, что величина относится к задаче связей).

Графики функций  $P^{(b)}(x)$ , полученные с помощью численных расчетов, приведены для различных решеток на рис. 5.2. Кривые 1, 2, 4 относятся к трехмерным решеткам, а остальные три кривые — к двумерным. При малых  $x$  величина  $P^{(b)}(x) \equiv 0$ , так как разорванные связи не позволяют жидкости отойти далеко от исходного узла. При  $x$ , близких к единице, вероятность  $P^{(b)}(x)$  также близка к единице. Важнейшее значение имеет введенное Бродбентом и Хаммерсли [1] понятие порога протекания (критической вероятности). Порогом протекания  $x_c$  называют верхнюю границу тех значений  $x$ , для которых  $P^{(b)}(x) = 0$  (см. рис. 5.2). Общепринятая точка зрения, подтверждаемая численными расчетами, состоит в том, что за порогом протекания функция  $P^{(b)}(x)$

непрерывно нарастает до единицы. При  $x - x_c \ll 1$

$$P^{(b)}(x) \propto (x - x_c)^\beta, \quad (16.1)$$

где  $\beta$  — критический индекс, численные значения которого будут приведены ниже.

Таким образом, поведение  $P^{(b)}(x)$  с увеличением  $x$  напоминает поведение параметра порядка фазового перехода второго рода при понижении температуры. Например, так ведет себя спонтанная намагниченность в

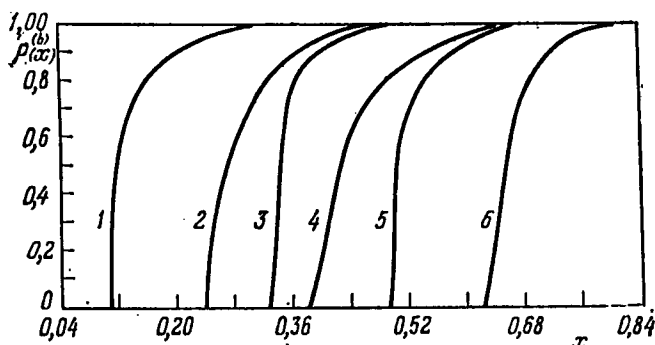


Рис. 5.2. Вероятность  $P^{(b)}(x)$  (плотность бесконечного кластера) для задачи связей на различных решетках [2]. Решетки: 1 — гранецентрированная кубическая, 2 — простая кубическая, 3 — треугольная, 4 — тетраэдрическая, 5 — квадратная, 6 — шестиугольная.

ферромагнетике или поляризация в сегнетоэлектрике. Имея в виду пороговое поведение  $P^{(b)}(x)$ , часто говорят, что при  $x < x_c$  протекания нет, а при  $x = x_c$  протекание возникает. Существование определенного значения порога протекания  $x_c$ , так же как и самой функции  $P^{(b)}(x)$ , связано, конечно, с тем, что речь идет о бесконечной решетке, для которой все случайные реализации разорванных связей с заданным значением  $x$  с точки зрения протекания эквивалентны. В этом же причина неаналитичности  $P^{(b)}(x)$  при  $x = x_c$ . Аналогичным образом, благодаря бесконечным размерам системы, возникает неаналитичность статистической суммы и свободной энергии в точке фазового перехода второго рода.

Наряду с вероятностью смочить бесконечное число узлов  $P^{(b)}(x)$ , можно говорить о вероятности  $P_N^{(b)}(x)$  того,

что данный узел смачивает по крайней мере  $N$  узлов, где  $N$  — большое, но конечное число. Вероятность  $P_N^{(b)}(x)$ , разумеется, отлична от нуля при всех  $0 < x < 1$ , хотя при  $N$  много больших 1 и  $x < x_c$  она очень мала. Вероятность  $P^{(b)}(x)$  получается из  $P_N^{(b)}(x)$  предельным переходом к  $N = \infty$ :

$$P^{(b)}(x) = \lim_{N \rightarrow \infty} P_N^{(b)}(x). \quad (16.2)$$

Рассмотрим пример, в котором возникает задача связей [1, 3]. Пусть проектируется фруктовый сад, представляющий собой квадратную решетку, в узлах которой растут деревья. Известно, что заболевшее дерево заражает другое дерево, находящееся от него на расстоянии  $r$ , с вероятностью  $f(r)$ , где  $f(r)$  — очень быстро убывающая функция  $r$ . Требуется найти минимальный период решетки, при котором одно заболевшее дерево способно заразить лишь конечное число деревьев, т. е. отсутствует опасность эпидемии. Из условия задачи очевидно, что контактами ближайших соседей можно пренебречь, а доля контактов между ближайшими деревьями, приводящих к заражению, равна  $f(h)$ , где  $h$  — период решетки. Если таким контактам сопоставить целые связи решетки, а остальным — разорванные, то мы приходим к задаче связей, причем  $x = f(h)$ . Поэтому искомым минимальным периодом  $h_{\min}$  определяется равенством

$$f(h_{\min}) = x_c. \quad (16.3)$$

Перейдем теперь ко второй основной задаче теории протекания — задаче узлов (*site problem*). В этой задаче все связи считаются целыми, а портятся узлы. Узлы могут быть перекрытыми и неперекрытыми. Перекрытые узлы не пропускают жидкость ни в какую сторону. Они не могут быть смочены и не смачивают другие узлы. Обозначим через  $x$  долю неперекрытых узлов и через  $P^{(s)}(x)$  — вероятность того, что произвольный узел смачивает бесконечное число узлов (индекс  $s$  означает, что соответствующая величина относится к задаче узлов).

Графики  $P^{(s)}(x)$  для трех кубических решеток приведены на рис. 5.3. Здесь так же, как и в задаче связей, существует порог протекания  $x_c(s)$  — верхняя граница значений  $x$ , для которых  $P^{(s)}(x) = 0$ .

Задачи узлов и связей чаще всего формулируют на языке статистики кластеров, не используя представления о растекании жидкости из одного узла. Сделаем это на примере задачи узлов. Пусть случайным образом долю  $x$  узлов решетки выкрасили в черный цвет, а оставшиеся узлы — в белый. Будем называть связанными любые

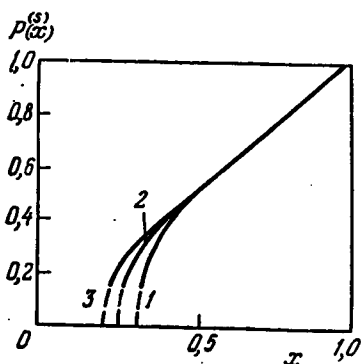


Рис. 5.3. Плотность бесконечного кластера  $P^{(s)}(x)$  для трех кубических решеток из работы Киркпатрика [4]. Решетки: 1 — простая, 2 — объемноцентрированная, 3 — границентрированная.

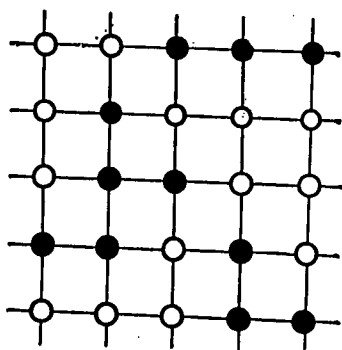


Рис. 5.4. Задача узлов на квадратной решетке.

два черных узла, являющиеся ближайшими соседями. Назовем кластером совокупность черных узлов, связанных друг с другом как непосредственно, так и посредством цепочек связанных черных узлов. Например, на рис. 5.4 имеется один кластер из пяти черных узлов и два кластера по три черных узла.

На языке кластеров динамика возникновения протекания при увеличении  $x$  выглядит следующим образом. При малых  $x$  все кластеры невелики. Однако по мере приближения к порогу протекания отдельные кластеры сливаются и средний размер кластеров возрастает. В точке  $x = x_c$  впервые возникает бесконечный черный кластер. Он напоминает собой случайную сетку и пронизывает все пространство. В «порах» бесконечного кластера размещаются конечные изолированные кластеры (рис. 5.5).

Понятие бесконечного кластера позволяет дать несколько иную трактовку вероятности  $P^{(s)}(x)$ . Так как только узлы бесконечного кластера смачивают бесконечное число узлов, вероятность  $P^{(s)}(x)$  равна отношению числа узлов, принадлежащих бесконечному кластеру, к полному числу узлов решетки. Другими словами,  $P^{(s)}(x)$  есть плотность бесконечного кластера. Возрастание

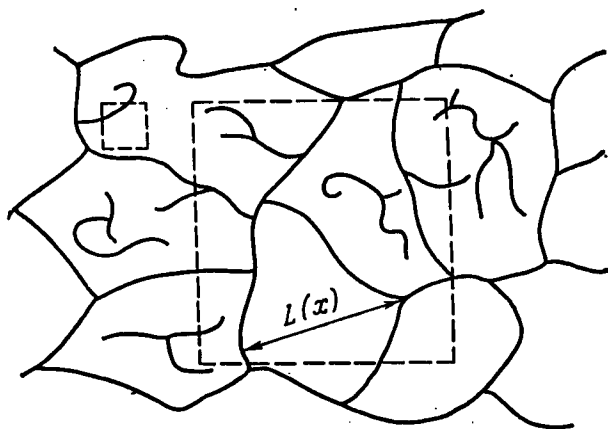


Рис. 5.5. Схематическое изображение бесконечного и больших (критических) конечных кластеров при  $0 < x - x_c \ll 1$ . Линии представляют собой цепочки черных узлов, изображенные со столь малым разрешением, что детали с размером порядка периода исходной решетки не заметны.  $L(x)$  — радиус корреляции. Штриховыми линиями изображены два квадрата со сторонами  $l > L(x)$  и  $l < L(x)$ .

$P^{(s)}(x)$  при удалении от порога протекания в сторону больших  $x$  означает, что бесконечный кластер, постепенно присоединяя конечные кластеры, из очень редкого, бестелесного становится все более густым. Средний размер его «пор» постепенно убывает. Соответственно убывает среднее число частиц тех конечных кластеров, которые остаются изолированными.

Выше мы говорили о задаче узлов. В задаче связей таким же образом можно говорить о конечных и бесконечном кластерах из целых связей и связанных ими узлов. Величина  $P^{(b)}(x)$  опять имеет смысл отношения числа узлов, принадлежащих бесконечному кластеру, к полному числу узлов решетки. Следует, однако, иметь в виду, что существует и другое определение  $P^{(b)}(x)$ . За

плотность бесконечного кластера  $P^{(b)}(x)$  принимают отношение числа его связей к полному числу связей решетки. Графики так определенных функций  $P^{(b)}(x)$  приведены в обзоре Шанте и Киркпатрика [5].

Принято считать, что в системе не может существовать несколько бесконечных черных кластеров, не связанных друг с другом. Это утверждение кажется нам правильным, хотя оно и не является строго доказанным. Мы основываемся на следующих соображениях, являющихся некоторым развитием аргументов Кикучи [6]. Предположим противное, т. е. допустим, что в некоторой области значений  $x_c < x < x_1$  существуют два бесконечных кластера, пронизывающих все пространство. Возьмем значение  $x = x_c + \Delta/2$ , где  $\Delta < x_1 - x_c$ . Соответствующие этому  $x$  два бесконечных кластера обладают конечной плотностью и везде находятся на конечном расстоянии друг от друга. Увеличим  $x$  до значения  $x_c + \Delta$ , вводя дополнительные черные узлы. Вероятность застраивания этими узлами любой фиксированной перемычки между двумя бесконечными кластерами окажется малой, но конечной. В то же время число возможных перемычек между бесконечными кластерами бесконечно. Поэтому с достоверностью при  $x = x_c + \Delta$  оба кластера окажутся связанными в один, что противоречит исходному предположению, поскольку  $x_c + \Delta < x_1$ .

Теперь мы естественным образом подошли к важному физическому применению задачи узлов [7]. Представим себе кристаллический раствор ферромагнитного вещества в немагнитном. Пусть доля ферромагнитных атомов равна  $x$ . Допустим, что обменное взаимодействие между ферромагнитными атомами, приводящее к параллельной ориентации их магнитных моментов, столь резко убывает с расстоянием, что его можно считать отличным от нуля лишь тогда, когда эти атомы являются ближайшими соседями в решетке. При нулевой температуре любая пара ближайших соседей ориентирует свои спины одинаково. Ясно, что при малых  $x$  это приведет к образованию небольших, изолированных друг от друга магнитных кластеров, внутри которых все спины параллельны. Однако взаимная ориентация моментов этих кластеров будет произвольной. Правила для объединения в магнитный кластер оказываются в точности



такими же, как и правила объединения черных узлов в задаче узлов. Поэтому, пока концентрация магнитных атомов меньше  $x_c(s)$ , макроскопическая намагниченность не возникает даже при  $T=0$ . С другой стороны, при  $x > x_c(s)$  конечная доля магнитных атомов оказывается объединенной в бесконечный кластер, в результате чего возникает макроскопический магнитный момент.

Из этих рассуждений следует два вывода. Во-первых, магнитный фазовый переход в разбавленном ферромагнетике возможен лишь при  $x > x_c(s)$ . Во-вторых, намагниченность насыщения  $M(T=0)$  при  $x > x_c(s)$  выражается через плотность бесконечного кластера:

$$M(T=0) = \mu P^{(s)}(x), \quad (16.4)$$

где  $\mu$  — магнитный момент атома. Большое внимание привлекает зависимость температуры ферромагнитного перехода от концентрации магнитных атомов  $T_c(x)$  [8—10]. Закон, по которому  $T_c(x)$  стремится к нулю при  $x \rightarrow x_c + 0$ , найден в работах [11, 12].

Во второй главе мы уже встретились с другим важным применением задачи узлов. Эта задача возникла, когда мы получали оценки критической точки перехода Андерсона. Черным узлам соответствовали резонансные атомы, а белым — нерезонансные. Образованию бесконечного кластера черных узлов соответствовало возникновение нелокализованных состояний.

Используя язык статистики кластеров, дадим теперь формальные определения  $P(x)$  и других величин теории протекания на примере задачи узлов. Пусть  $n_s$  — число кластеров из  $s$  узлов в бесконечной решетке, приходящееся на один узел решетки. Тогда сумма  $\sum_s sn_s$  есть доля узлов решетки, принадлежащих конечным кластерам. Поскольку доля всех черных узлов есть  $x$ , доля узлов решетки, принадлежащих бесконечному кластеру  $P(x)$ , равна

$$P(x) = x - \sum_s sn_s. \quad (16.5)$$

Второй важной величиной теории протекания является среднее число узлов конечного кластера (в английской

литературе *mean cluster size*):

$$S(x) = \frac{\sum_s s^2 n_s}{\sum_s s n_s}. \quad (16.6)$$

Здесь, как и в (16.5), суммирование происходит по всем конечным  $s$ . Согласно численным расчетам, при  $x \rightarrow x_c + 0$  или  $x \rightarrow x_c - 0$  величина  $S(x)$  обращается в бесконечность:

$$S(x) \propto |x - x_c|^{-\gamma} \quad (|x - x_c| \ll 1), \quad (16.7)$$

где  $\gamma$  (как и  $\beta$  в (16.1)) — критический индекс теории протекания.

Нетрудно видеть, что в (16.6) речь идет о взвешенном среднем. Усредняется по узлам число узлов в кластере, к которому принадлежит данный узел. Поэтому обращение в бесконечность  $S(x)$  при  $|x - x_c| \rightarrow 0$  не означает ни того, что больших кластеров больше, чем маленьких, ни даже того, что большим кластерам принадлежит подавляющее число черных узлов. Действительно, как легко увидеть из (16.5),  $\lim_{x \rightarrow x_c} \sum_s s n_s = x$  и, следовательно, простое среднее  $\sum_s s n_s / \sum_s n_s$  остается конечным при  $x = x_c$ . Таким образом, основная часть черных узлов и при  $x = x_c$  находится в кластерах с  $s \sim 1$ .

Согласно численным расчетам, при  $|x - x_c| \ll 1$  величина  $n_s$  как функция  $s$  ведет себя следующим образом. Существует некоторое критическое значение числа узлов в кластере  $s_c$ , которое растет при  $x \rightarrow x_c + 0$  и  $x \rightarrow x_c - 0$ :

$$s_c \propto |x - x_c|^{-\Delta}, \quad \Delta > 0. \quad (16.8)$$

При  $s \ll s_c$  величина  $n_s$  убывает с ростом  $s$  по слабому степенному закону, а при  $s \gg s_c$  — по более резкому экспоненциальному закону. Числитель (16.6) определяется кластерами с  $s \approx s_c$ , которые мы будем называть критическими. Обращение в бесконечность величины  $S(x)$  при  $x \rightarrow x_c$  является отражением роста числа узлов в критических кластерах, т. е. удаления границы

экспоненциального спада  $n_s$ . Ниже, употребляя выражение «среднее число узлов в кластере», мы всегда будем иметь в виду величину (16.6).

Третьей важной характеристикой теории протекания является корреляционная функция. Чтобы ввести ее, определим функцию  $g(r_i, r_j)$ , положив ее равной единице, если узлы  $i$  и  $j$  черные и принадлежат одному конечному кластеру, и равной нулю во всех остальных случаях. После этого введем парную корреляционную функцию, усреднив  $g(r_i, r_j)$  по всем узлам решетки:

$$G(r, x) = G(r_i - r_j, x) \equiv \langle g(r_i, r_j) \rangle. \quad (16.9)$$

При  $r \rightarrow \infty$  функция  $G(r, x)$  стремится к нулю, так как число кластеров  $n_s$  размера  $s$  убывает с ростом  $s$ . На расстояниях  $r$ , меньших среднего размера критических кластеров, который мы обозначим  $L(x)$ , определяющую роль в создании корреляции играют кластеры с  $s \ll s_c(x)$ . Следовательно, в этой области значений  $r$  функция  $G(r, x)$  убывает степенным образом с ростом  $r$ . При  $r \gg L(x)$  главную роль играют экспоненциально редкие кластеры с  $s \gg s_c(x)$ , и поэтому  $G(r, x)$  убывает экспоненциально с ростом  $r$ .

Естественно считать, что средний размер критического кластера  $L(x)$  является единственной характерной длиной функции  $G(r, x)$ , и назвать эту длину радиусом корреляции. Поскольку число узлов критического кластера при  $x \rightarrow x_c \pm 0$  стремится к бесконечности, радиус корреляции  $L$  должен возрастать при приближении к порогу протекания:

$$L(x) \approx |x - x_c|^{-\nu} \quad (|x - x_c| \ll 1), \quad (16.10)$$

где  $\nu$  — критический индекс радиуса корреляции. Обычно считают, что критические индексы  $\gamma$ ,  $\Delta$ ,  $\nu$  одинаковы при  $x > x_c$  и  $x < x_c$ . При  $x > x_c$  малая вероятность возникновения конечных кластеров с размером больше  $L(x)$  означает одновременно малую вероятность пор бесконечного кластера такого размера (см. рис. 5.5). Поэтому при  $x > x_c$  радиус корреляции можно интерпретировать как характерный размер пор сетки бесконечного кластера. Можно сказать, что радиус корреляции дает представление о среднем расстоянии между узлами сетки, или о ее «периоде».

Сформулируем теперь задачу узлов (и связей) третьим эквивалентным способом. Речь пойдет о протекании сквозь конечную решетку. Рассмотрим решетку конечного размера, например квадрат, на каждой стороне которого помещается  $l$  узлов (см. рис. 5.4). Пусть сначала все узлы решетки белые. Будем затем перекрашивать в черный цвет случайно выбранные узлы, постепенно увеличивая долю перекрашенных узлов  $x$ . При некотором значении  $x = x_{cl}$  впервые возникнет путь по связанным черным узлам с левой стороны квадрата на правую. Назовем  $x_{cl}$  порогом протекания для конечной решетки. Если повторить эту процедуру, то черные узлы окажутся в других позициях и порог протекания также будет, вообще говоря, иным. Таким образом,  $x_{cl}$  есть случайная величина. Много раз повторяя описанную процедуру, мы можем вычислить среднее значение  $\langle x_{cl} \rangle$ . Порог протекания, соответствующий бесконечной системе, очевидно, определяется выражением

$$x_c = \lim_{l \rightarrow \infty} \langle x_{cl} \rangle. \quad (16.11)$$

В трехмерном случае аналогичным образом можно говорить о протекании с какой-либо грани куба на противоположную. Другой вариант постановки задачи состоит в том, чтобы задавать  $x$  не как долю черных узлов, а как априорную вероятность того, что данный узел черный. Для бесконечной решетки между этими двумя определениями  $x$  нет разницы. Однако в конечном объеме благодаря флуктуациям эти два определения  $x$  приводят к разным величинам  $x_{cl}$  и разным значениям  $\langle x_{cl} \rangle$ . Можно сказать, что в первом случае, вычисляя  $\langle x_{cl} \rangle$ , мы имели дело с микроканоническим ансамблем, а в новом варианте рассматривается канонический ансамбль. Разумеется, несмотря на разницу в значениях  $x_{cl}$ , предельный переход (16.11) дает для обоих ансамблей один и тот же результат.

На рис. 5.6, заимствованном из работы Левинштейна и др. [13], представлены функции распределения  $f_l$  величин  $x_{cl}$  в каноническом ансамбле для задачи узлов на квадратной решетке. Они были получены с помощью непосредственного моделирования протекания сквозь конечную решетку методом Монте-Карло на ЭВМ.

Видно, что форма функций распределения близка к гауссовой. Более детальный анализ кривых показал [13], что в рамках погрешности расчета нельзя обнаружить отклонение  $f_l(x_{cl})$  от гауссовых функций. Как и следовало ожидать, чем больше сторона квадрата  $l$ , тем меньше дисперсия  $W_l^2 = \langle (x_{cl} - \langle x_{cl} \rangle)^2 \rangle$ . Однако совсем

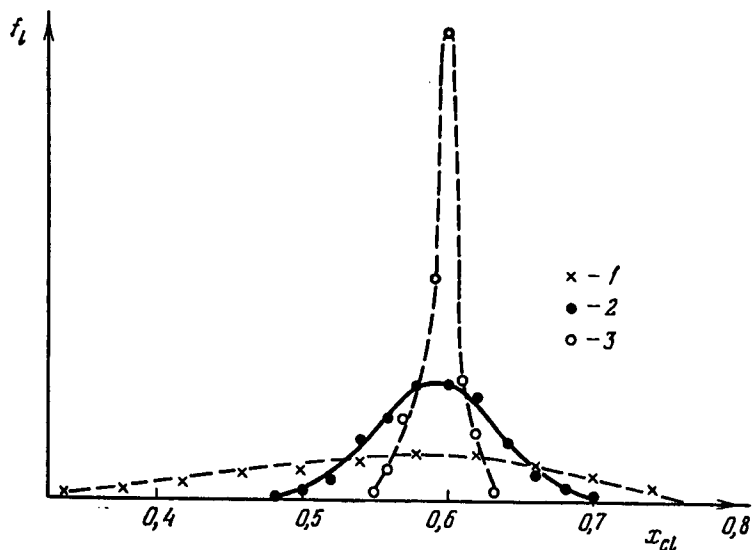


Рис. 5.6. Функции распределения значений порога протекания через квадратичную решетку  $l \times l$  для задачи узлов [13]. Значения  $l$  в единицах постоянной решетки: 1—8, 2—32, 3—128.

нетривиально хорошо заметное на рис. 5.6 превышение среднеквадратичной флуктуации  $W_l$  над сдвигом точки максимума  $f_l(x_{cl})$  (или среднего значения  $\langle x_{cl} \rangle$ ). Аналогичное поведение имеет место и в трехмерном случае. В работе [13] показано, что убывание  $W_l$  можно описать степенным законом

$$W_l = Bl^{-1/\nu}. \quad (16.12)$$

Для нахождения порога протекания в [13] также использовалась степенная экстраполяция  $\langle x_{cl} \rangle$ :

$$\langle x_{cl} \rangle = x_c + Al^{-1/\lambda}. \quad (16.13)$$

В трехмерном случае было найдено, что  $\nu \approx 0,9$  и  $\lambda \approx 1$ .

Покажем теперь, что определенный с помощью (16.12) индекс  $\nu$  равен индексу радиуса корреляции [14]. Для этого определим величину

$$\mathcal{P}_l(x) = \int_0^x f_l(x') dx', \quad (16.14)$$

смысл которой — вероятность того, что при заданном значении  $x$  имеется протекание через решетку. Из геометрической интерпретации радиуса корреляции при  $x > x_c$  следует, что при  $l \gg L(x)$  протекание практически достоверно, т. е.

$$1 - \mathcal{P}_l(x) \ll 1. \quad (16.15)$$

Это можно понять, мысленно вырезав конечный квадрат со стороной  $l$  из бесконечной решетки (см. рис. 5.5). Тем же способом можно понять, что при  $l \ll L(x)$  достоверность протекания квадрата исчезает, т. е. неравенство (16.15) больше не выполняется. Рассмотрим теперь случай  $x < x_c$ . В этом случае при  $l \gg L(x)$  почти все конечные кластеры меньше  $l$  и, следовательно, практически достоверно, что протекание отсутствует, т. е.

$$\mathcal{P}_l(x) \ll 1. \quad (16.16)$$

В обратном предельном случае  $l \ll L(x)$  неравенство (16.16) не выполняется. Используя эти свойства радиуса корреляции, найдем его с помощью расчетных данных о  $f_l(x)$ . Зафиксируем, например, значение  $x - x_c > 0$  и, постепенно уменьшая  $l$ , найдем радиус корреляции  $L$  как то значение  $l$ , при котором неравенство (16.15) нарушается. Это произойдет, очевидно, когда полуширина  $W_l$  функции  $f_l(x)$  приблизится к  $x - x_c$ . Используя (16.12), получаем

$$BL(x)^{-1/\nu} \approx x - x_c, \quad \text{т. е.} \quad L(x) \approx (x - x_c)^{-\nu}. \quad (16.17)$$

Таким образом, индекс  $\nu$  в (16.12) действительно является индексом радиуса корреляции.

В качестве примера протекания с грани на грань для задачи связей рассмотрим электропроводность большого куба, в котором вместо целых связей стоят одинаковые сопротивления, соединяющие соседние узлы, а разорванным связям соответствует бесконечное сопротивление и

напряжение подводится к плоским металлическим контактам, покрывающим две противоположные грани куба. Рис. 5.7 иллюстрирует такую постановку задачи для двумерного случая.

Если куб достаточно велик, то его эффективная электропроводность  $\sigma(x)$  отлична от нуля только при  $x > x_c(b)$ . Согласно численным расчетам, при  $0 < x - x_c \ll 1$

$$\sigma(x) \propto (x - x_c)^t, \quad (16.18)$$

где  $t$  — новый критический индекс. Ниже мы увидим, что поведение  $\sigma(x)$  при  $x > x_c(b)$  представляет интерес для применений, связанных с электропроводностью неупорядоченных систем. Очевидно, что  $\sigma(x)$ , вообще говоря, не может быть вычислена на основе информации о  $P^{(b)}(x)$ . Для нахождения  $\sigma(x)$  нужно знать не только плотность, но и топологическую организацию бесконечного кластера.

Выше мы рассмотрели три основные постановки задач узлов и связей: растекание «жидкости» из смоченного узла на бесконечность, статистику кластеров и протекание сквозь конечную решетку. Ясно, что во всех трех случаях речь идет об одном и том же критическом явлении. Численные расчеты порогов протекания в пределах своей точности подтверждают эквивалентность этих трех постановок задачи в пределе бесконечной системы.

Из сказанного выше ясен характер вопросов, на которые должна отвечать теория протекания. Вот, например, некоторые из них. Каковы пороговые величины  $x_c$ ? Чему равны критические индексы  $\beta$ ,  $\gamma$ ,  $\nu$  и  $t$ ? Как пороги и индексы зависят от геометрии решетки и размерности пространства? Какова топология конечных и бесконечного кластеров?

Значения порогов протекания задач связей и узлов для наиболее часто исследуемых решеток собраны нами

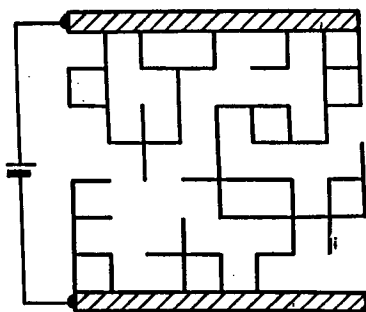


Рис. 5.7. Случайная сетка со противлений, включенная между двумя электродами.

Таблица 5.1

Пороги протекания  $x_c$  для различных решеток \*)

Решетка	$x_c(b)$	$x_c(s)$
Шестиуголь- ная	<u>0,6527</u> [15, 16], <u>0,640±0,018</u> [17—19]	0,700±0,010 [20], 0,688±0,017 [17—19], 0,697±0,001 [21, 22], 0,698±0,003 [23]
Квадратная	0,5000 [15, 16], <u>0,493±0,013</u> [17—19], <u>0,498±0,017</u> [21, 22, 24]	0,590±0,010 [20], 0,581±0,015 [17—19], 0,591±0,001 [21, 22], 0,593±0,002 [23], 0,594±0,002 [13], 0,5898±0,0008 [25]
Треугольная	<u>0,3473</u> [15, 16], <u>0,341±0,011</u> [17—19], <u>0,349±0,010</u> [21, 22]	<u>0,5000</u> [15, 16], <u>0,493±0,018</u> [17—19], <u>0,500±0,001</u> [21, 22]
Алмаза	<u>0,388±0,005</u> [20], <u>0,390±0,011</u> [17—19]	0,425±0,015 [20], 0,436±0,012 [21, 22]
Простая ку- бическая	<u>0,247±0,005</u> [20], <u>0,254±0,013</u> [17—19]	0,307±0,010 [20], 0,325±0,023 [17—19], 0,320±0,001 [21, 22], 0,312±0,002 [4], 0,312±0,003 [26], 0,312±0,001 [28]
Объемноцен- трированная кубическая	<u>0,178±0,005</u> [20]	0,243±0,010 [20], 0,254±0,001 [21, 22], 0,248±0,003 [4], 0,248±0,003 [26]
Гранецентри- рованная кубическая	<u>0,119±0,002</u> [20], <u>0,125±0,005</u> [17—19], <u>0,1190±0,0005</u> [27]	0,195±0,005 [20], 0,199±0,008 [17—19], 0,208±0,001 [21, 22], 0,200±0,002 [4], 0,198±0,003 [29]

\*) Точные значения подчеркнуты. Остальные величины получены численными методами.



в табл. 5.1. В табл. 5.2 мы приводим значения наиболее важных критических индексов  $\beta$ ,  $\gamma$ ,  $\nu$  и  $t$ . Для некоторых двумерных решеток Сайкс и Эссам [15, 16] нашли точные значения порога  $x_c$  (в табл. 5.1 эти значения подчеркнуты). В теории протекания существуют и другие точные результаты. Значительное место среди них занимают неравенства, связывающие величины  $x_c$  для различных решеток и различных задач [42]. Кроме того, большое количество точных результатов удается получить для решеток Бете. В этом случае найдены не только пороги протекания, но и критические индексы  $\beta$ ,  $\gamma$ ,  $\nu$  и  $t$  [43—46]. Подробное изложение точных результатов теории протекания, полученных до 1972 г., можно найти в превосходных обзорах Эссама [47] и Шанте и Киркпатрика [5].

подавляющее большинство результатов, приведенных в табл. 5.1 и 5.2, получены с помощью численных методов или в редких случаях с помощью модельных экспериментов. О модельных экспериментах мы расскажем в § 20, а сейчас кратко остановимся на численных методах.

В 1959 г. Домб [48] предложил использовать для отыскания  $x_c$  метод рядов. Первые расчеты произвели Эллиот и др. [7] и Домб и Сайкс [49]. Идея метода состоит в том, что при малых  $x$  среднее число узлов конечного кластера  $S(x)$  записывают в виде степенного ряда:

$$S(x) = \sum_n a_n x^n. \quad (16.19)$$

Поскольку  $S(x)$  обращается в бесконечность при  $x \rightarrow x_c$ , порог протекания  $x_c$  ищут как точку ближайшей особенности суммы ряда на положительной половине вещественной оси. Аналогичным образом в теории ферромагнетизма находят температуру фазового перехода, исследуя сходимость ряда по обратным степеням температуры для магнитной восприимчивости [50].

Сайкс и Эссам [20], изучая методом рядов поведение  $S(x)$  вблизи  $x = x_c$ , впервые вычислили критический индекс  $\gamma$  для задачи узлов и связей. Оказалось, что на всех исследованных ими плоских решетках — квадратной, треугольной и шестиугольной — значения  $\gamma$  близки. Аналогичное совпадение было зафиксировано в трех-

Таблица 5.2

Критические индексы:  $\beta$  — плотности бесконечного кластера,  $\gamma$  — среднего числа частиц в кластере,  $\nu$  — радиуса корреляции и  $t$  — электропроводности  $\sigma(x)$  по результатам численных расчетов ( $d$  — размерность пространства)

Критический индекс	$d=2$	$d=3$
$\beta$	$0,14 \pm 0,03$ [30] $0,138 \pm 0,007$ [31] $0,148 \pm 0,004$ [32]	$0,35 \pm 0,05$ [4] $0,39 \pm 0,02$ [28] $0,42 \pm 0,06$ [29] $0,47 \pm 0,02$ [33]
$\gamma$	$1,85 \pm 0,20$ [32] $2,38 \pm 0,03$ [20] $2,38 \pm 0,02$ [27] $2,43 \pm 0,03$ [23] $2,34 \pm 0,3$ [34]	$1,69 \pm 0,05$ [20] $1,70 \pm 0,11$ [27] $1,80 \pm 0,05$ [28]
$\nu$	$1,34 \pm 0,02$ [27] $1,33 \pm 0,04$ [13] $1,36 \pm 0,04$ [26] $\leq 1,5$ [25]	$0,94 \pm 0,05$ [13, 26] $0,82 \pm 0,05$ [27] $0,83 \pm 0,13$ [35]
$t$	$1,38 \pm 0,12$ [36] $1,1 \pm 0,1$ [28] $1,15 \pm 0,15$ [37] $1,1 \pm 1,25$ [38]	$2$ [39] $1,725 \pm 0,005$ [40] $1,6 \pm 0,1$ [28] $1,75 \pm 0,10$ [38] $1,6 \pm 0,1$ [41]

мерном случае для простой, объемноцентрированной и гранецентрированной кубических решеток и решетки алмаза. Основываясь на этих данных, Сайкс и Эссам предположили, что индекс  $\gamma$  универсален, т. е. для фиксированной размерности пространства он не зависит ни от характера задачи, ни от вида решетки. Имеющиеся теперь данные не противоречат гипотезе о том, что все индексы теории протекания универсальны. Для того чтобы указать размерность пространства, к которой относится данный критический показатель, ниже мы будем использовать дополнительный цифровой индекс (напри-

мер,  $\gamma_2, \beta_3$ ). В последние годы метод рядов используют и при  $x > x_c$ . Например, строят ряды для величины  $P(x)$  по степеням величины  $(1 - x_c)$ . Исследуя эти ряды с помощью аппроксимант Паде (см. [50]), находят критический индекс  $\beta$  [29—33]. Результаты расчетов приведены в табл. 5.2.

Другим основным методом численного решения задач теории протекания является метод случайных испытаний, или, как его называют, метод Монте-Карло. Он включает в себя практически все расчеты на ЭВМ, где существенно используются датчики случайных чисел. Впервые метод Монте-Карло был применен к задачам протекания в работах Хаммерсли и группы американских авторов [2, 17—19]. Впоследствии Дин [21] разработал новый вариант метода, с помощью которого Дин и Берд получили очень точные результаты для  $x_c$  [22]. Третий вариант метода Монте-Карло использовался в работах [51—53, 13, 26]. Он состоял в непосредственной численной реализации протекания через конечную решетку. Характер результатов, получаемых этим способом, мы уже обсуждали выше.

Обсудим теперь численные значения порогов  $x_c$ . Из табл. 5.1 видно, что величина  $x_c$  сильно зависит от решетки и характера задачи. Для задачи связей существует приближенное эмпирическое правило [49, 54], по которому  $x_c(b)$  выражается через простые характеристики решеток — размерность  $d$  и число ближайших соседей  $Z$ :

$$B_c \equiv Zx_c(b) = \frac{d}{d-1}, \quad (16.20)$$

где  $B = Zx$  — среднее число узлов, связанных с данным узлом решетки. С помощью табл. 5.1 легко можно убедиться в том, что соотношение (16.20) выполняется с точностью до нескольких процентов при изменении величин  $Z$  и  $x_c(b)$  в 3 раза.

Для задачи узлов величина  $Zx_c$  не является даже приближенным инвариантом. Однако в этом случае существует другой приближенный инвариант, замеченный Шером и Залленом [55].

Построим вокруг каждого узла решетки окружность (сферу) с радиусом, равным половине расстояния до

ближайшего соседа (рис. 5.8). Пусть  $f$  есть доля площади (объема), попадающая внутрь этих окружностей (сфер). Величину  $f$  называют плотностью упаковки. Ее легко вычислить для каждой решетки. Шер и Заллен подсчитали долю площади (объема), которая при  $x=x_c$

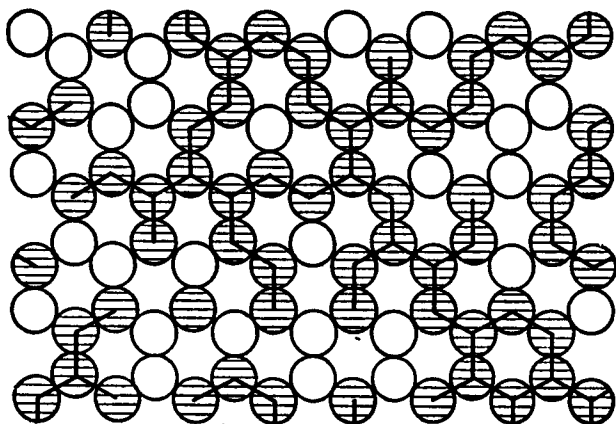


Рис. 5.8. Построение Шера и Заллена для шестиугольной решетки [55]. Заштрихованы окружности, соответствующие черным узлам.

находится внутри окружностей (сфер), построенных вокруг черных узлов. Для этого надо вычислить произведение  $\phi_c = f x_c(s)$ . Оно оказывается приблизительно одинаковым для всех решеток одной размерности:

$$\phi_c = \begin{cases} 0,15 \pm 0,01 & \text{при } d=3, \\ 0,45 \pm 0,02 & \text{при } d=2. \end{cases} \quad (16.21)$$

Задачи узлов и связей называют решеточными. Исследование решеточных задач началось раньше других, и им посвящена подавляющая часть оригинальных работ. Именно на этих задачах отработаны методы и проверены основные идеи теории протекания. Поэтому мы и начали главу с подробного знакомства с решеточными задачами. Кроме них существуют еще два больших класса задач теории протекания: континуальные задачи и задачи на случайных узлах. Именно эти задачи непосредственно используются в теории примесной проводимости; им посвящены §§ 17 и 18. В § 19 излагается теория по-

добия, позволяющая связать друг с другом критические индексы. § 20 посвящен электропроводности  $\sigma(x)$  и связанным с ней вопросам о топологии бесконечного кластера. Очень важную роль в нашей книге играет последний в этой главе § 21. Здесь показано, как с помощью теории протекания можно вычислять электропроводность сильно неоднородных сред.

## § 17. Континуальные задачи

Континуальные задачи формулируются следующим образом. Допустим, что во всем пространстве задана случайная непрерывная функция  $V(r)$ , определенная корреляционными соотношениями. Не нарушая общности, будем считать, что  $\langle V(r) \rangle = 0$ . ( $\langle \dots \rangle$  означает среднее значение.) Зададим вещественное число  $V$  и покрасим мысленно области пространства, где  $V(r) \leq V$ , черной краской, а остальные области — белой.

При изменении  $V$  от  $-\infty$  до  $+\infty$  объем черных областей меняется от нуля до объема всего пространства (рис. 5.9). Изолированные вначале черные «озера» сливаются друг с другом и при больших  $V$  образуют «океан». Требуется найти нижнюю границу  $V_c$  тех значений  $V$ , для которых имеется возможность, стартуя из некоторых черных точек, уходить на бесконечное расстояние, двигаясь только по черным областям.

Такая задача возникает, например, когда нужно найти минимальную энергию, которую должен иметь электрон, чтобы пройти бесконечное расстояние в пространстве с профилем потенциальной энергии  $V(r)$ , не выходя из классически разрешенных областей [54, 56—60]. Величина  $V_c$  называется в этом случае уровнем протекания. Очевидно, что, кроме приведенной выше постановки задачи, аналогичной первой формулировке решеточных задач (растекание), можно сформулировать задачу и на языке, аналогичном кластерному. Возникновение протекания соответствует на этом языке образованию связанной черной области, обладающей бесконечным объемом. Естественно, обобщается и формулировка о протекании через конечный объем. Здесь нужно говорить о возникновении путей протекания «по черному» между противоположными сторонами квадрата (гранями куба).

Легко понять, что континуальная задача близка к сформулированным ранее решеточным задачам. Вообразим решетку со столь малым периодом, что функция  $V(\mathbf{r})$  практически на нем не меняется. Зафиксируем  $V$  и будем считать, что узлы решетки, находящиеся в черных областях, — черные, а в белых — белые. При увеличении  $V$  увеличивается доля черных узлов

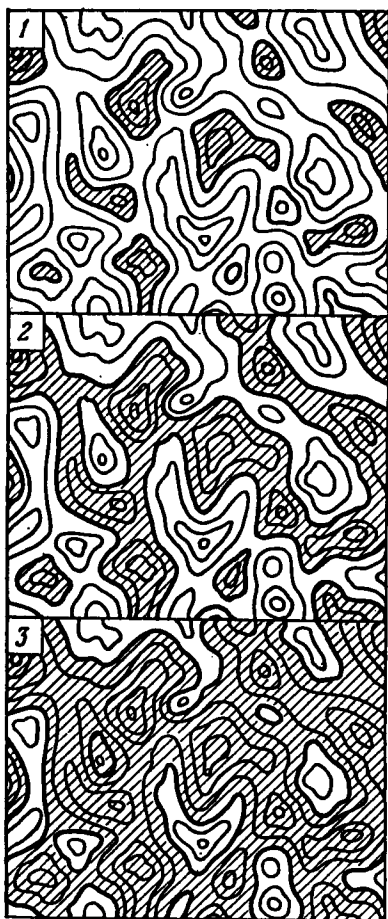


Рис. 5.9. Протекание в двумерном потенциале [60]. Контурные линии отвечают поверхностям  $V(\mathbf{r}) = \text{const}$ . Черные ( $V(\mathbf{r}) < V$ ) области заштрихованы. Три «раскраски» 1—3 соответствуют трем различным значениям  $V$  ( $V_1 < V_2 < V_3$ ).

равна порогу протекания  $x_c$  для задачи узлов, отличающейся от сформулированной ранее тем, что распределение черных и белых узлов коррелировано в соответствии с видом  $V(\mathbf{r})$ , т. е. черные и белые узлы сгруппированы в кластеры. Это утверждение и служит основой для вычисления  $\phi_c$  (см. ниже).

$$x = \phi(V) = \int_{-\infty}^V F(V') dV', \quad (17.1)$$

где  $\phi(V)$  — доля пространства, в которой  $V(\mathbf{r}) < V$ , а  $F(V)$  — функция распределения потенциала  $V$ . Уровень протекания  $V_c$ , таким образом, может быть найден из того условия, что доля пространства, занятого черными областями на пороге протекания:

$$\phi_c = \int_{-\infty}^{V_c} F(V) dV, \quad (17.2)$$

равна порогу протекания  $x_c$  для задачи узлов, отличающейся от сформулированной ранее тем, что распределение черных и белых узлов коррелировано в соответствии с видом  $V(\mathbf{r})$ , т. е. черные и белые узлы сгруппированы в кластеры. Это утверждение и служит основой для вычисления  $\phi_c$  (см. ниже).

Наряду с протеканием по черному можно говорить и о протекании по белому. При малых  $V$  белое заполняет почти все пространство, образуя «континент», окружающий своими «берегами» черные «озера». При больших  $V$  «континент» разбивается на изолированные «острова» и протекания по белому нет. Назовем уровнем протекания по белому значение  $V = V'_c$ , при котором протекание по белому исчезает. Если говорить о случайных потенциалах  $V(\mathbf{r})$ , обладающих конечным радиусом корреляции, то в двумерном случае существует простая связь между  $V_c$  и  $V'_c$ .

Действительно, рассмотрим протекание по черному и по белому между противоположными сторонами квадрата с размерами, много большими радиуса корреляции потенциала. В этом случае ясно, что отсутствие протекания по черному слева направо означает существование протекания по белому сверху вниз. С другой стороны, если имеется протекание по белому, например, слева направо, то протекание по черному сверху вниз отсутствует.

Поскольку в очень больших квадратах протекание по одному из цветов в обоих направлениях возникает при одинаковом значении  $V$ , то в предыдущих двух фразах можно опустить слова «слева направо» и «сверху вниз», и мы приходим к следующему заключению. Протекание по черному исчезает при том же значении  $V$ , при котором возникает протекание по белому, т. е.

$$V_c = V'_c. \quad (17.3)$$

С помощью (17.3) можно определить  $V_c$  для потенциала, статистические свойства которого симметричны относительно значения  $V=0$ . Для таких потенциалов очевидно, что

$$V_c = -V'_c, \quad (17.4)$$

и с учетом (17.3)

$$V_c = V'_c = 0 \quad \text{и} \quad \theta_c = 0,5, \quad (17.5)$$

т. е. на пороге протекания площадь поровну распределена между белым и черным цветом. Результат (17.5) был получен в работах [60—62].

Для трехмерного потенциала, как и для двумерного, не может отсутствовать протекание по обоим цветам. Здесь, однако, протекания по черному и по белому могут сосуществовать, так как каналы, по которым осуществляется протекание по белому и по черному, не обязаны пересекаться. Поэтому в трехмерном случае имеет место лишь неравенство

$$V_c \leq V'_c, \quad (17.6)$$

т.е. с точки зрения протекания возможны три различные области значений  $V$ :

1)  $V < V_c$  — есть протекание по белому, но нет по черному;

2)  $V_c \leq V < V'_c$  — есть протекание и по белому, и по черному;

3)  $V \geq V'_c$  — есть протекание по черному, но нет по белому.

Для статистически симметричного потенциала неравенство (17.6) совместно с (17.4) дает

$$V_c \leq 0, \quad V'_c \geq 0 \quad \text{и} \quad \phi_c \leq 0,5. \quad (17.7)$$

Заллен и Шер предположили [60], что величина  $\phi_c$  значительно менее чувствительна к виду потенциала, чем  $V_c$ . Для оценки  $\phi_c$  они предложили использовать инвариант  $f_{x_c}(s)$ . Для плоского случая согласно (16.21)  $f_{x_c}(s) = 0,45$ , что близко к значению  $\phi_c = 0,5$  для симметричного потенциала. Точного совпадения не следует ожидать, так как в построении Заллена и Шера отсутствует симметрия между точками, внутренними и внешними по отношению к окружностям или сферам.

Скал и др. [51, 52] решали на ЭВМ континуальные задачи методом Монте-Карло. Рассматривался гауссов потенциал, определяемый соотношением

$$V(\mathbf{r}) = \int K(\mathbf{r} - \mathbf{r}') f(\mathbf{r}') d\mathbf{r}', \quad (17.8)$$

где  $f(\mathbf{r})$  — случайная гауссова функция с коррелятором

$$\langle f(\mathbf{r}) f(\mathbf{r}') \rangle = \delta(\mathbf{r} - \mathbf{r}'), \quad (17.9)$$

а ядро  $K(\mathbf{r})$  достаточно быстро убывает с  $r$  за пределами радиуса корреляции  $r_0$ . Из (17.8) и (17.9) следует,



что функция распределения потенциала  $V$  имеет вид

$$F(V) = \frac{1}{\sqrt{\pi v}} \exp\left(-\frac{V^2}{v^2}\right), \quad v^2 = 2 \int K^2(r) dr. \quad (17.10)$$

Для моделирования континуальной задачи в узлах решетки с периодом, меньшим  $r_0$ , задавались случайные значения  $f_i$ , а значения  $V_i$  образовывались с помощью суммирования:  $V_i = \sum_j K(r_i - r_j) f_j$ . После этого задавалось значение потенциала  $V$  и образовывался массив чисел  $\{A_i\}$ , состоящий из единиц и нулей:  $A_i = 1$ , если  $V_i < V$ , и  $A_i = 0$ , если  $V_i \geq V$ . Затем проверялось протекание по единицам через конечную решетку со стороны, много большей  $r_0$ .

В двумерном случае в качестве  $K(r)$  использовалась функция  $\exp(-r/r_0)$ . Если принять период решетки за единицу, то при переходе от  $r_0 < 1$  к  $r_0 \gg 1$  пороговое значение концентрации единиц вследствие постепенного увеличения корреляции в расположении единиц должно изменяться от порога задачи узлов  $x_c(s) = 0,59$  до порога континуальной задачи  $\phi_c = 0,50$ . Такой переход действительно наблюдался [51]. Аналогичное явление исследовано в работах [63, 64].

В работе [52] было определено значение  $\phi_c$  для трехмерных гауссовых потенциалов с пятью различными функциями  $K(r)$ , которые характеризовались радиусом корреляции  $r_0$ . Затем производилась экстраполяция к большим значениям  $r_0$  и тем самым находилось значение  $\phi_c$  континуальной задачи. Это значение в пределах погрешностей оказалось одинаковым для всех пяти функций. Оно равно

$$\phi_c = 0,17 \pm 0,01. \quad (17.11)$$

Результат (17.11), в согласии с предположением Заллена и Шера, близок к значению их инварианта (16.21).

В работе [65] методом Монте-Карло решалась задача узлов на простой кубической решетке с термодинамической корреляцией узлов. Рассматривался неразбавленный изинговский ферромагнетик при  $T < T_c$  и исследовалось протекание по перевернутым спинам, т.е. по спинам, направленным противоположно спонтанной намагниченности. Такие спины благодаря взаимодействию

друг с другом расположены коррелированным образом. Оказалось, что даже довольно умеренной корреляции, которая имеет место, когда температура и энергия взаимодействия — величины одного порядка, достаточно для того, чтобы понизить  $x_c$  от 0,31 до 0,19. В задаче связей постепенный переход к континууму за счет увеличивающейся положительной корреляции соседних звеньев численно исследовался Вебманом и др. [41]. Они пришли к выводу, что при увеличении корреляции  $x_c$  стремится к величине  $\vartheta_c = 0,15 \pm 0,02$ .

### § 18. Задачи на случайных узлах

В теории прыжковой проводимости, как мы увидим ниже, наиболее важную роль играют задачи о протекании по случайным узлам. Случайные узлы — это точки, хаотически распределенные в пространстве. Среднее число узлов в единице объема задано и равно  $N$ . Пусть  $\xi_{ij}$  — некоторая функция вектора  $r_{ij}$ , соединяющего два узла  $i$  и  $j$ . Зададим некоторое число  $\xi$  и будем считать, что узлы  $i$  и  $j$  связаны, если выполнено неравенство

$$\xi_{ij} \leq \xi, \quad (18.1)$$

которое в дальнейшем мы будем называть условием связности. Если два узла связаны непосредственно или с помощью попарно связанных узлов, то говорят, что они принадлежат одному кластеру. Требуется найти порог протекания  $\xi_c$ , т. е. нижнюю границу значений параметра  $\xi$ , для которых существует бесконечный кластер.

Простейшей из этого класса задач является задача с условием связности

$$r_{ij} \leq r. \quad (18.2)$$

Условие связности (18.2) выполняется, если узел  $j$  находится внутри сферы радиуса  $r$ , описанной вокруг узла  $i$ . Поэтому задача имеет следующую геометрическую интерпретацию. Пусть вокруг каждого узла построены сферы одинакового радиуса  $r$ . Требуется найти наименьшее значение  $r = r_c$ , при котором существуют бесконечные цепочки узлов, где каждый последующий узел лежит внутри сферы, построенной на предыдущем (рис. 5.10, а).

Мы будем называть порог протекания  $r_c$  перколяционным радиусом. Иногда предпочитают строить сферы с радиусом  $r/2$  и искать значение  $r$ , при котором бесконечный кластер образуют перекрывающиеся сферы (рис. 5.10, б). Пайк и Сигер [66] предложили называть первое построение охватывающими сферами, а второе — перекрывающимися сферами. Ниже мы увидим, что для

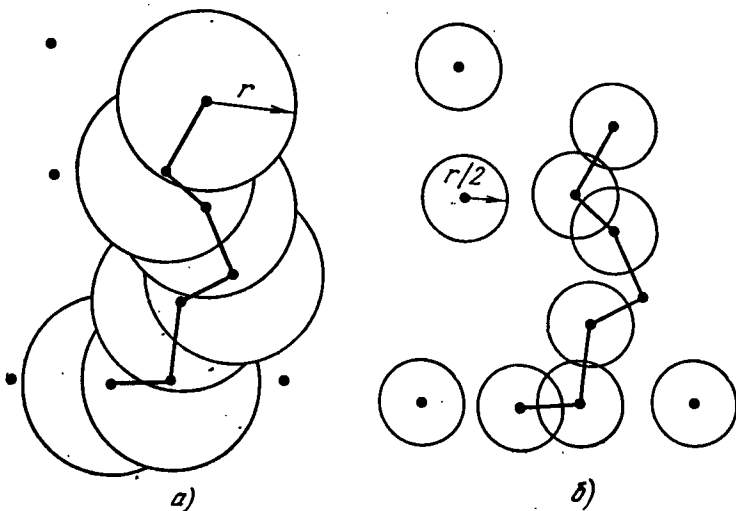


Рис. 5.10. Построения охватывающих (а) и перекрывающихся (б) окружностей. Связанные узлы соединены прямыми линиями.

использования в теории прыжковой проводимости удобнее первое построение. Поэтому мы будем уделять ему больше внимания.

Перколяционный радиус  $r_c$  может зависеть только от концентрации узлов  $N$  и по соображениям размерности пропорционален  $N^{-1/3}$ . Часто бывает удобным вместо  $r_c$  говорить о пороговом значении

$$B_c = 4/3\pi N r_c^3 \quad (18.3)$$

безразмерного параметра  $B = 4/3\pi N r^3$ , имеющего смысл среднего числа связей на один узел. Задачу об определении  $r_c$  и  $B_c$  мы будем называть задачей сфер. Аналогичную двумерную задачу можно назвать задачей окружностей. В этом случае удобно пользоваться

величиной

$$B_c = \pi \tilde{N} \tilde{r}_c^2, \quad (18.4)$$

где  $\tilde{N}$  — число случайных узлов на единицу площади, а  $\tilde{r}_c$  — двумерный перколяционный радиус.

Задача сфер естественным образом возникает при рассмотрении ферромагнетизма кристаллических разбавленных растворов магнитных и немагнитных атомов, когда обменные силы, действующие между магнитными атомами, моделируются взаимодействием с конечным, но превышающим постоянную решетки радиусом  $r_{вз}$ . В пределе очень большого  $r_{вз}$ , когда в сферу взаимодействия попадает очень много узлов решетки и дискретность положения частиц больше несущественна, возникает задача сфер, с той только разницей, что радиус  $r_{вз}$ , играющий роль правой части (18.2), теперь задан, а ищется концентрация частиц (случайных узлов). Если обозначить через  $Z$  число узлов решетки, попадающих в сферу взаимодействия, то  $Zx$  есть среднее число черных узлов, связанных с данным узлом, и, очевидно, для всех решеток одной размерности мы должны иметь

$$\lim_{Z \rightarrow \infty} (Zx_c) = B_c, \quad (18.5)$$

где  $B_c$  в зависимости от размерности определяется (18.3) или (18.4).

В задачах на случайных узлах можно ввести критические величины, близкие к тем, о которых шла речь для решеточной задачи узлов. Пусть  $n_s$  — число кластеров из  $s$  узлов в некотором объеме, отнесенное к полному числу узлов объема. Тогда

$$P(B) = 1 - \sum_s s n_s, \quad S(B) = \sum_s s^2 n_s / \sum_s s n_s. \quad (18.6)$$

Вблизи порога  $B_c$

$$P(B) \propto (B - B_c)^\beta, \quad S(B) \propto |B - B_c|^{-\nu}. \quad (18.7)$$

По аналогии с решеточными задачами можно ввести корреляционную функцию и радиус корреляции  $L$ . Вблизи порога

$$L(B) \approx N^{-1/3} |B - B_c|^{-\nu}, \quad (18.8)$$

где среднее расстояние между узлами  $N^{-1/3}$  фигурирует в качестве естественного масштаба длины. Радиус корреляции по-прежнему имеет смысл характерного размера критических кластеров, т.е. самых больших из тех кластеров, для которых величины  $n_s$  еще не экспоненциально малы. Формулы (18.7) и (18.8) можно записать, введя вместо разности  $B - B_c$  безразмерную переменную  $(r - r_c)/r_c$  или, в более общем случае,  $(\xi - \xi_c)/\xi_c$ . Например,

$$L \approx N^{-1/3} \left| \frac{r - r_c}{r_c} \right|^{-\nu} \quad \text{или} \quad L \approx N^{-1/3} \left| \frac{\xi - \xi_c}{\xi_c} \right|^{-\nu}. \quad (18.9)$$

Поскольку критические индексы не зависят от геометрии решетки, естественно считать, что они не изменяются и при переходе от решеток к случайному расположению узлов. Мы увидим, что существующие данные не противоречат этой гипотезе.

Ввиду особой важности задачи сфер для теории прыжковой проводимости мы дадим краткий обзор работ, посвященных вычислению  $B_c$  и критических индексов в этом случае. Домб и др. [67, 68] рассчитали методом рядов пороги протекания  $x_c(s)$  для задачи узлов на различных решетках с учетом связывания ближайших соседей. Экстраполируя для каждой решетки произведение  $Zx_c(s)$  к  $Z = \infty$ , они получили значения  $B_c$ , приведенные в первой строке табл. 5.3. Большое число работ [70—77, 66, 35] посвящено вычислению  $B_c$  методом Монте-Карло. Их результаты также приведены в табл. 5.3.

Остановимся на работах, в которых исследовалось протекание через конечный массив случайных узлов. Скал и Шкловский [77] и Пайк и Сигер [66] решали задачу сфер, исследуя протекание между противоположными гранями куба, содержащего большое число случайных узлов  $n$  (до 3000 в [77] и до 8000 в [66]). В [66] исследовалась задача окружностей в квадратах, содержащих до 4000 узлов. Для проверки протекания вблизи противоположных граней куба (сторон квадрата) выделялось два пограничных контрольных слоя с толщиной, несколько меньшей среднего расстояния между узлами. Протекание регистрировалось тогда, когда обнаруживались два узла из разных контрольных слоев, принадле-

Таблица 5.3

Пороговые значения среднего числа связей на узел  $V_c$   
для задач окружностей и сфер

Метод	$V_c$ для окружностей	$V_c$ для сфер
Рядов	4,5 [67, 68]	2,7 [67, 68]
Монте-Карло	3,2 [69]	
1) Построение кластера, включающего центральный узел	3,82 [71]	2,4 [70]
	4,1 [74]	2,95 [72]
	4,4 [76]	2,4 [73]
		2,6 [75]
2) Протекание через конечный массив	4,5 [66]	2,7 [76]
		3,0 [77]
		2,7 [66]
		2,77 [35]

жащие одному и тому же кластеру, т. е. возникал путь с грани на грань. Для определения  $V_c$  здесь, как и в решеточных задачах, оказывается очень важным вопрос экстраполяции данных к бесконечному числу узлов. Пайк и Сигер [66] получили два значения усредненного по реализациям порога протекания  $\bar{V}_{cn}$ :  $\bar{V}_{c 1000} = 2,92 \pm 0,03$  и  $\bar{V}_{c 3000} = 2,80 \pm 0,02$ .

Экстраполируя по закону

$$\bar{V}_{cn} = V_c + \text{const} \cdot n^{-1/3\lambda}, \quad (18.10)$$

где  $\lambda = 1$ , они нашли, что  $V_c = 2,7$ . В работе [77] получены результаты  $V_{c 1500} = 3,0 \pm 0,1$  и  $V_{c 3000} = 2,9 \pm 0,1$ . Ввиду невысокой точности данных авторы отказались от экстраполяции и пользовались значением  $V_c = 3,0$ . Видно, что результаты [77] и [66] близки, хотя налицо систематическое расхождение на величину порядка 0,1.

Куркьярви [35] рассматривал протекание в конечном объеме с более сложными граничными условиями. После того, как в кубе генерировались координаты  $n$  случайных узлов, весь куб периодически продолжался во всех направлениях. Протекание регистрировалось тогда, когда хотя бы один узел куба оказывался связанным со своим двойником в одном фиксированном направлении. Кажется весьма вероятным, что такое определение протекания при  $n \rightarrow \infty$  дает правильное предельное значе-

ние порога  $B_c$ . Во всяком случае результаты оказываются близкими к результатам для протекания с грани на грань. Анализируя расчеты для  $n = 32, 64, 108, 256, 1000$  и  $2000$  случайных узлов, Куркьярви нашел, что в формуле (18.10)  $B_c = 2,77$  и  $\lambda = 0,6 \pm 0,25$ . Кроме того, Куркьярви исследовал зависимость дисперсии  $W_n^2$  величины  $B_{cn}$ . Он получил

$$W_n \approx \frac{\delta \xi_c}{\xi_c} = \frac{\delta r_c}{r_c} \approx n^{-1/3\nu}, \quad (18.11)$$

где  $\delta \xi_c$  и  $\delta r_c$  — среднеквадратичные флуктуации  $\xi_c$  и  $r_c$ , а  $\nu = 0,83 \pm 0,13$  (погрешность оценена нами по графику работы [35]). Легко показать, что так же, как и для решеточных задач, индекс  $\nu$ , найденный по зависимости (18.11), есть не что иное, как индекс радиуса корреляции. В результате мы видим, что значение индекса радиуса корреляции  $\nu = 0,83 \pm 0,13$  для задачи сфер в пределах погрешности определения совпадает со значением индекса  $\nu_3$  для решеточных задач (см. табл. 5.2). Пока это совпадение является единственным свидетельством в пользу того, что универсальность индексов охватывает и задачи на случайных узлах.

Перейдем теперь к задачам, в которых критерий связности выполняется не внутри сферы, а внутри более сложной замкнутой поверхности  $Q_\xi$ , линейные размеры которой подобным образом увеличиваются с увеличением некоторого параметра  $\xi$ . Эти задачи играют важную роль в теории прыжковой проводимости. Критерий связности для них можно записать в виде

$$\xi_{ij} = \varphi(r_{ij}) \leq \xi, \quad (18.12)$$

где  $\xi > 0$ ,  $\varphi$  — однородная, возрастающая и положительная функция  $x_{ij}, y_{ij}, z_{ij}$ , а уравнение поверхности  $Q_\xi$  имеет вид  $\xi_{ij} = \varphi(r_{ij}) = \xi$ . Задача состоит в нахождении  $\xi_c$  — нижней границы значений  $\xi$ , для которых существует бесконечный кластер связанных узлов. Геометрический образ такой задачи мы получим, если построим около каждого узла одинаковую поверхность  $Q_\xi$  и будем, увеличивая  $\xi$ , искать значение  $\xi = \xi_c$ , соответствующее образованию первой цепочки, в которой каждый узел лежит внутри поверхности, построенной на предыдущем (построение охватывающих поверхностей).

Переход к построению перекрывающихся поверхностей для анизотропной  $Q_{\xi}$  не всегда столь очевиден, как в случае сфер, а иногда (вогнутые поверхности) вообще невозможен.

Пусть  $V_{\xi}$  — объем, заключенный внутри поверхности  $Q_{\xi}$ . Тогда удобно ввести безразмерную величину  $B = NV_{\xi}$ , имеющую смысл среднего числа связей на узел. Критическому значению  $\xi = \xi_c$  соответствует критическое значение

$$B_c = NV_{\xi_c}. \quad (18.13)$$

В частном случае, когда  $Q_{\xi}$  — сфера,  $B_c$  определяется формулой (18.3). Представляет интерес выяснить, каким образом  $B_c$  зависит от формы поверхности  $Q_{\xi}$ . Здесь можно доказать следующую теорему, сформулированную Я. Г. Синаем. *Если поверхность  $Q'_{\xi}$  получается из поверхности  $Q_{\xi}$  с помощью линейного преобразования координат*

$$x'_k = \sum_l A_{kl} x_l, \quad (18.14)$$

*содержащего вращение и растяжение по осям ( $A_{kl}$  — константы), то значения  $B_c$  для них совпадают.* Из этой теоремы, в частности, следует, что для любого эллипсоида  $B_c$  такое же, как для сферы. Для доказательства необходимо заметить, что если вместе с поверхностями, построенными на каждом узле, подвергнуть преобразованию координаты узлов, то распределение узлов в пространстве по-прежнему останется абсолютно случайным (измениться может только их концентрация). Действительно, коррелятор плотностей  $\rho(r)$ , описывающий систему случайных точек, есть

$$\langle \rho(r_1) \rho(r_2) \rangle = N \delta(r_1 - r_2). \quad (18.15)$$

Замена (18.14) дает

$$\langle \rho(r'_1) \rho(r'_2) \rangle = \frac{N}{\text{Det}A} \delta(r'_1 - r'_2), \quad (18.16)$$

т. е. характер распределения не меняется.

При одновременном преобразовании координат узлов и поверхностей, очевидно, не меняются никакие топологические взаимоотношения. Если до преобразования не было протекания по охватывающим соседние узлы по-



верхностям, то оно и не возникнет после. Если протекание было, то оно не исчезнет. Пороговой ситуации в исходном пространстве соответствует пороговая ситуация в преобразованном. Кроме того, среднее число связей на узел  $B$  также не изменяется этим преобразованием.

Поэтому, если  $B_c$  — порог для поверхности  $Q_\xi$ , а  $B'_c$  — порог для  $Q'_\xi$ , то  $B_c = B'_c$ .

Для ответа на вопрос, сколь чувствительна величина  $B_c$  к изменениям формы поверхности  $Q_\xi$ , например к вогнутости, в работе [77] вычислялось значение  $B_c$  для нескольких поверхностей. Вычисления производились для 1500 узлов. Напомним, что для сферы  $B_{c\ 1500} = 3 \pm 0,1$ . Для куба было получено  $B_{c\ 1500} = 3 \pm 0,1$  и для тетраэдра  $B_{c\ 1500} = 2,9 \pm 0,1$ . На основании этих данных можно, по-видимому, утверждать, что с точностью до 0,1 величины  $B_c$  одинаковы для всех выпуклых поверхностей.

Исследовались также две вогнутые поверхности. Одна из них возникает при изучении прыжковой проводимости  $n$ -германия и представляет собой огибающую четырех эллипсоидов вращения одинаковой формы (см. рис. 6.5). Центр каждого эллипсоида находится в начале координат, а ось вращения направлена по одной из диагоналей куба. Полуось, совпадающая с осью вращения, у каждого эллипсоида в 4,45 раза меньше двух других его полуосей. Для такой «четырёхэллипсоидной» поверхности  $B_{c\ 1500} = 2,8 \pm 0,1$ . Таким образом, вогнутость немного понижает  $B_c$ . Второй расчет был произведен для «трехмерного креста», образованного тремя параллелепипедами, пересекающимися в начале координат и вытянутыми по осям. Уравнения граней параллелепипедов имели вид: 1)  $|x| = 1, |y| = 1, |z| = 5$ ; 2)  $|x| = 1, |y| = 5, |z| = 1$ ; 3)  $|x| = 5, |y| = 1, |z| = 1$ . Для такой «наиболее вогнутой» поверхности  $B_{c\ 1500} = 2,4 \pm 0,1$ , т.е. всего на 20% меньше, чем у сферы. Поэтому можно думать, что для всех вогнутых поверхностей уменьшение  $B_c$  заключено в этих пределах.

Пайк и Сигер [66] исследовали зависимость  $B_c$  от формы для плоских фигур. Для квадрата они получили  $B_c = 4,4$ , т.е. всего на 2% меньше, чем для окружности. Таким образом, в рамках одной размерности наблюдается

весьма слабая чувствительность  $B_c$  к форме поверхности, определяемой критерием связности (18.12).

Интересно сравнить эти приближенные инварианты со значениями  $B_c$  в решеточных задачах (см. (16.20)). Оказывается, что для задач на случайных узлах среднее число связей одного узла должно быть значительно больше, чем для решеточных задач. На наш взгляд, это можно объяснить тем, что в первом случае многие связи используются не эффективно. Речь идет о парах или тройках близких друг к другу узлов, которые участвуют в подсчете связей, но мало способствуют протеканию.

Можно попытаться найти характеристику задач на случайных узлах, близкую по смыслу к решеточному инварианту Заллена и Шера  $\phi_c$ . В построении перекрывающихся сфер для этого можно вычислить долю объема, попадающего внутрь сфер на пороге протекания. Возьмем произвольную точку  $A$  пространства и вычислим вероятность того, что она не попадает ни в одну из сфер с радиусом  $r_c/2$ , построенных на случайных узлах. Эта вероятность совпадает с вероятностью  $\exp\left\{-\frac{4\pi}{3} N \left(\frac{r_c}{2}\right)^3\right\}$  того, что в сфере с радиусом  $r_c/2$  и с центром в точке  $A$  нет ни одного случайного узла. Вероятность точке  $A$  попасть в одну из сфер дает искомую долю объема и равна

$$\phi_c = 1 - \exp\left\{-\frac{4\pi}{3} N \left(\frac{r_c}{2}\right)^3\right\} = 1 - \exp\left(-\frac{B_c}{8}\right) = 0,29 \quad (18.17)$$

(мы использовали значение  $B_c = 2,7$ ). Величина (18.17) сильно отличается от критического значения доли разрешенного объема  $\phi_c = 0,15$  для решеток. На наш взгляд, причина этого — в более экономном использовании «разрешенного объема» на решетках. В этом случае все связи устанавливаются касанием разрешенных сфер (см. рис. 5.8). В построении же перекрывающихся сфер участки с малой «толщиной» пути очень редки. Подтверждением такой точки зрения можно считать результаты опытов Фитцпатрика и др. [78], которые исследовали случайную плотную упаковку смеси твердых проводящих и непроводящих шаров одинакового размера. По резкому возрастанию электропроводности системы они

определили пороговое значение доли проводящих шаров, которое соответствует образованию бесконечного кластера соприкасающихся проводящих шаров. Плотность случайной плотной упаковки известна:  $f = 0,65$ . Оказалось, что в этом случае произведение  $fx_c$  близко к 0,15.

Таким образом, решающую роль в сильном увеличении доли разрешенного объема (18.17) для задач на случайных узлах играет, по-видимому, возможность частого сближения случайных узлов.

## § 19. Теория критических индексов

В § 16 мы ввели критические величины теории протекания и соответствующие критические индексы  $\beta$ ,  $\gamma$  и  $\nu$ . В табл. 5.2 приведены численные значения этих индексов, полученные с помощью различных экспериментов на ЭВМ. В этом параграфе излагается теория, цель которой состоит в том, чтобы указать связь между различными индексами. Хорошо известно, что в реализации аналогичной программы в теории термодинамических фазовых переходов второго рода достигнуты большие успехи. Поэтому естественно попытаться использовать аналогию между задачами протекания и задачей о фазовом переходе второго рода, например о ферромагнитном переходе.

Напомним, что  $P(x)$  эквивалентно параметру порядка, например намагниченности. Величина  $x$  отвечает температуре, причем порог  $x_c$  соответствует температуре Кюри  $T_c$ , область  $x < x_c$  соответствует парамагнитной области  $T > T_c$ , а область  $x > x_c$  — ферромагнитной области  $T < T_c$ .

Для углубления аналогии в теорию протекания следует ввести аналог магнитного поля и свободной энергии, зависящей от этого поля и параметра  $x$ . Это было сделано Кастелайном и Фортуином [79]. Покажем, как это делается, на примере задачи узлов.

Введем дополнительный черный узел, не принадлежащий рассматриваемой решетке (рис. 5.11). Такой узел имеет ряд необычных свойств, поэтому мы назвали его демоном Кастелайна и Фортуина. Пусть по определению демон связан с каждым из черных узлов решетки с вероятностью  $1 - \exp(-h)$ , где  $h$  — параметр, который

играет роль безразмерного магнитного поля  $\mu H/kT$  ( $\mu$  — спиновый магнитный момент,  $H$  — магнитное поле). Действительно, при  $h \neq 0$  и сколь угодно малом  $x$  величина  $P(x)$  оказывается отличной от нуля за счет связывания черных узлов в бесконечный кластер через посредство демона (рис. 5.11). Например, при  $x \ll 1$  и  $h \ll 1$ , когда черные узлы решетки изолированы друг от друга, легко получить, что  $P(x) = hx$ .

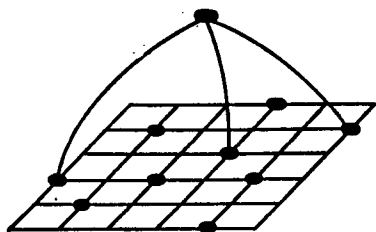


Рис. 5.11. Образование бесконечного кластера с помощью демона. Связи демона с черными узлами показаны выходящими из него дугами.

При наличии демона число конечных кластеров на один узел решетки может быть записано в виде

$$F(x, h) = \sum_s n_s \exp(-sh), \quad (19.1)$$

где  $n_s$  — число конечных кластеров из  $s$  черных узлов, приходящееся на узел решетки при  $h = 0$ . Множитель  $\exp(-sh) \equiv [\exp(-h)]^s$  имеет смысл доли кластеров из  $s$  узлов, у которых ни один из узлов не связан с демоном.

Величина  $F(x, h)$  аналогична свободной энергии ферромагнетика. Действительно, величины  $P(x)$  и  $S(x)$  при  $h = 0$  выражаются через первую и вторую производные от  $F(x, h)$  по  $h$ . Параметр порядка  $P(x)$  определяется производной  $\frac{\partial F}{\partial h}$  при  $h = 0$ :

$$P(x) = x - \sum_s s n_s = x + \left. \frac{\partial F}{\partial h} \right|_{h=0}. \quad (19.2)$$

Среднее число узлов конечного кластера определяется производной  $\frac{\partial^2 F}{\partial h^2}$  при  $h = 0$ :

$$S(x) = \frac{\sum_s s^2 n_s}{\sum_s s n_s} = \frac{1}{x - P(x)} \left. \frac{\partial^2 F}{\partial h^2} \right|_{h=0}. \quad (19.3)$$

Согласно (19.3) величина  $S(x)$  аналогична магнитной восприимчивости. В этом можно убедиться и непосредственно, подсчитывая приращение мощности бесконечного

кластера  $dP$  за счет малого изменения магнитного поля  $dh$ . Вероятность того, что некоторый новый кластер из  $s$  узлов присоединится к демону, есть  $s dh$ . При этом к бесконечному кластеру добавится  $s$  узлов. Суммируя по всем кластерам и деля на полное число узлов решетки, получим

$$\frac{dP}{dh} \Big|_{h=0} = \sum_s s^2 n_s = [x - P(x)] S(x). \quad (19.4)$$

Чтобы подчеркнуть аналогию с термодинамикой, для критических индексов величин  $P(x)$ ,  $S(x)$  и  $L(x)$  мы использовали те же обозначения  $\beta$ ,  $\gamma$  и  $\nu$ , которые принято использовать для их термодинамических аналогов. Дифференцируя функцию  $F(x, h)$  по обоим аргументам, можно ввести набор функций и критических индексов, аналогичных всем термодинамическим функциям и индексам. Например, можно ввести аналог теплоемкости

$$C(\tau, h=0) = \frac{\partial^2 F}{\partial \tau^2} \Big|_{h=0} \propto |\tau|^{-\alpha}, \quad (19.5)$$

где  $\tau = x - x_c$ . Записывая  $P(\tau, h)$  при  $\tau = 0$  и  $h \ll 1$  в виде

$$P(\tau=0, h) \propto h^{l^0}, \quad (19.6)$$

получаем определение еще одного критического индекса  $\delta$ . Определения других индексов мы пока выписывать не будем.

Естественный путь установления соотношений между индексами состоит в формулировке гипотезы подобия в области малых значений  $|\tau|$  и  $h$  ( $|\tau|, h \ll 1$ ). Гипотеза подобия состоит в том, что при каждом  $\tau$  и  $h$  сингулярная часть свободной энергии определяется кластерами с одним характерным размером и числом узлов. При увеличении  $\tau$  и  $h$  эти критические кластеры уменьшаются. Предполагается, что существуют такие индексы  $y$  и  $z$ , что при переходе от  $\tau$  и  $h$  к  $\tau' = \tau l^y$  и  $h' = h l^z$  характерные размеры критических кластеров уменьшаются в  $l$  раз, т. е. свободная энергия, которая по определению равна числу кластеров в единице объема, возрастает в  $l^d$  раз, где  $d$  — размерность пространства. Иначе говоря, свободная энергия удовлетворяет соотношению

$$F(\tau, h) = l^{-d} F(\tau l^y, h l^z). \quad (19.7)$$

Аналогичное соотношение предполагается и для корреляционной функции:

$$G(r, \tau, h) = l^{-d+2-\eta} G(rl^{-1}, \tau l^\nu, hl^z), \quad (19.8)$$

где  $\eta$  — новый критический индекс. Предположения (19.7), (19.8) необходимо дополнить простым соотношением, следующим из (16.5), (16.6), (16.9):

$$S(x)[x - P(x)] = \sum_r G(r, x). \quad (19.9)$$

Оно совершенно аналогично соотношению между магнитной восприимчивостью и корреляционной функцией спинов в теории ферромагнетизма. После этого формулировка гипотезы подобия становится тождественной формулировке, принятой в теории фазовых переходов (см. книгу Стенли [82]). Выражая с помощью (19.7) — (19.9) все критические индексы через  $\nu$  и  $z$ , можно получить обычные соотношения между ними:

$$\alpha = 2 - d\nu = 2 - 2\beta - \gamma, \quad (19.10)$$

$$\delta = \frac{\gamma}{\beta} + 1, \quad (19.11)$$

$$\eta = 2 - \frac{\gamma}{\nu}, \quad (19.12)$$

$$d\nu = 2\beta + \gamma. \quad (19.13)$$

Первая частичная формулировка гипотезы подобия для задач протекания, эквивалентная (19.7), но никак не затрагивающая корреляционную функцию, принадлежит Эссаму и Гвилиму [80]. В полном объеме гипотеза подобия была сформулирована в работах [81, 13].

Наибольшее значение в теории протекания имеет соотношение (19.13), связывающее между собой самые изученные критические индексы. Покажем, как можно получить его, делая предположения, эквивалентные (19.7) и (19.8), но не привлекая представления о магнитном поле или свободной энергии. Будем исходить из того, что в отсутствие магнитного поля среднее число узлов и размер критического кластера (радиус корреляции) определяются формулами (16.8) и (16.10), т. е.

$$s_c(\tau) \propto |\tau|^{-\Delta} \quad \text{и} \quad L(\tau) \propto |\tau|^{-\nu}. \quad (19.14)$$

Предположим также, что число критических кластеров на узел решетки, определенное, например, как  $\sum_{s=s_c}^{2s_c} n_s$ ,

имеет порядок величины  $[L(\tau)]^{-d}$ . Иначе говоря, мы предположим, что расстояние между двумя критическими кластерами порядка размера каждого из них и в этом смысле критические кластеры плотно заполняют пространство. Оценивая  $S(\tau)$ , получим из (19.3) и (19.14)

$$S(\tau) \approx L^{-d} s_c^2 \infty |\tau|^{d\nu-2\Delta}, \quad (19.15)$$

т. е.

$$\gamma = 2\Delta - d\nu. \quad (19.16)$$

Плотность критических кластеров по порядку величины равна  $s_c L^{-d} \approx |\tau|^{d\nu-\Delta}$ . Естественно считать, что плотность бесконечного кластера, «вобравшего» в себя практически все кластеры, бóльшие критических, имеет тот же порядок величины, т. е.

$$P(\tau, 0) = \tau^{d\nu-\Delta}. \quad (19.17)$$

Таким образом,

$$\beta = d\nu - \Delta. \quad (19.18)$$

Исключая индекс  $\Delta$  из (19.18) и (19.16), получаем (19.13).

Посмотрим теперь, удовлетворяют ли известные численные значения индексов соотношениям подобия. Начнем с соотношения (19.13), связывающего индексы  $\beta$ ,  $\gamma$  и  $\nu$ , значения которых приведены в табл. 5.2. Для двумерного случая, приняв  $\beta_2 = 0,14$  и  $\gamma_2 = 2,40$ , из (19.13) получим  $\nu_2 = 1,34$ , что находится в хорошем согласии со значениями  $\nu_2$  из табл. 5.2. В трехмерном случае, исходя из значений  $\beta_3 = 0,45$ ,  $\gamma_3 = 1,70$ , получаем  $\nu_3 = 0,87$ , что также хорошо согласуется с имеющимися численными данными.

Критический индекс  $\alpha$  вычислялся, насколько нам известно, только для двумерного случая. Согласно Киркпатрику [28],  $\alpha_2 = 0,65 \pm 0,01$ . Подставляя  $\nu_2 = 1,34$  в (19.10), находим, что  $\alpha = 0,62$ ; это находится в разумном согласии с расчетом. Индекс  $\delta_2$  был вычислен Гаунтом

и Сайксом [83] для задач узлов и связей на различных двумерных решетках; он оказался равным  $18,0 \pm \pm 0,75$ . Если в (19.11), согласно табл. 5.2, подставить  $\gamma_2 = 2,40$  и  $\beta_2 = 0,14$ , то получается очень близкая величина:  $\delta_2 = 18,2$ .

Важную проверку гипотезы подобия предприняли Дан и др. [27]. С помощью метода рядов они изучили критическое поведение моментов корреляционной функции  $G(r, x)$ :

$$\mu_n(x) = \sum_r r^n G(r, x). \quad (19.19)$$

Согласно (19.9) с точностью до несущественного множителя  $x - P(x)$  величина  $\mu_0(x)$  совпадает с  $S(x)$ , т. е. имеет критическое поведение вида

$$\mu_0(\tau) \propto |\tau|^{-\gamma}. \quad (19.20)$$

Моменты  $\mu_n(x)$  с  $n \neq 0$  также обращаются в бесконечность при  $|\tau| \rightarrow 0$ :

$$\mu_n(\tau) \propto |\tau|^{-\gamma^{(n)}}. \quad (19.21)$$

Согласно гипотезе подобия для корреляционной функции (19.8) при  $h = 0$  ее можно записать в виде

$$G(r, \tau) = r^{-d+2-\eta} g\left[\frac{r}{L(\tau)}\right]. \quad (19.22)$$

Поэтому при увеличении  $n$  на единицу  $\gamma^{(n)}$  должно увеличиваться на  $\nu$ :

$$\gamma^{(n)} = \gamma + n\nu \quad (n = 1, 2, \dots). \quad (19.23)$$

Дан и др. [27] вычислили индексы  $\gamma$ ,  $\gamma^{(2)}$  и  $\gamma^{(4)}$  для треугольной и гранецентрированной кубической решеток. Оказалось, что найденные ими значения удовлетворяют условию эквидистантности, подтверждая тем самым гипотезу подобия. Значения индексов  $\gamma$  и  $\nu$ , полученные в [27], приведены в табл. 5.2.

Подводя итоги проверки выполнения соотношений между индексами, вытекающих из гипотезы подобия, можно сказать, что известные в настоящее время численные значения индексов не противоречат этой гипотезе.

Перейдем теперь к теоретическим построениям, позволяющим не только связывать индексы друг с другом,



но и находить их численные значения. Прежде всего, интересно представить себе, какие значения критических индексов дает разложение свободной энергии  $F(\tau, h)$  по степеням параметра порядка  $P$ , которое используется в теории Ландау. В отличие от теории ферромагнетизма, в теории протекания симметрия не запрещает кубический член в этом разложении. Поэтому значения критических индексов в приближении самосогласованного поля оказываются следующими:

$$\beta = 1, \quad \gamma = 1, \quad \nu = 1/2. \quad (19.24)$$

(В теории Ландау  $\beta = 1/2$ .) Индексы (19.24) сильно отличаются от известных численных значений для двух и трех измерений (см. табл. 5.2). В теории ферромагнетизма индексы самосогласованного поля также не совпадают с индексами для двух и трех измерений. Однако существует критическая размерность  $d_c = 4$ , начиная с которой работает теория Ландау. Ее можно найти, как то значение  $d$ , при котором индексы теории Ландау удовлетворяют соотношению (19.13), следующему из гипотезы подобия.

Тулуз [84] предположил, что в теории протекания также существует критическая размерность, причем найти ее можно таким же способом. Подставляя индексы (19.24) в (19.13), он получил, что  $d_c = 6$ . Это предположение было подтверждено Киркпатриком [28], который вычислил индексы  $\beta$  и  $\gamma$  методом Монте-Карло для задачи узлов на простых «кубических» решетках размерности  $d = 3, 4, 5$  и  $6$ .

В последнее время появились работы, посвященные вычислению критических индексов методами ренормализационной группы (ренормгруппы) [32, 85—87]. Большинство авторов исходит из доказанной Кастелайном и Фортуином [79] гамильтоновой формулировки задач теории протекания. Это позволяет использовать методы, разработанные в теории фазовых переходов. Значения индексов, найденные методом ренормгруппы, как правило, разумно согласуются с результатами численных расчетов, приведенными в табл. 5.2. Особенно хорошее совпадение было достигнуто в работе [87] для двумерного случая.

## § 20. Электропроводность случайных сеток из проводящих и непроводящих элементов и топология бесконечного кластера

В § 16 мы сформулировали задачу о вычислении электропроводности  $\sigma(x)$  решетки из проводящих и непроводящих связей (см. рис. 5.7). После этого мы практически не касались вопроса о величине критического индекса  $t$  в формуле

$$\sigma(x) \propto (x - x_c)^t \quad (x \geq x_c) \quad (20.1)$$

и его связи с другими критическими индексами. В частности, для индекса  $t$  не нашлось аналога среди термодинамических индексов. Напомним, однако, что наша цель состоит в том, чтобы с помощью теории протекания научиться вычислять электропроводность сильно неоднородных систем. Задача о вычислении  $\sigma(x)$  — простейшая из задач, где соприкасаются электропроводность и протекание. Уже поэтому она заслуживает особого внимания. Электропроводность  $\sigma(x)$  удобно изучать непосредственным моделированием задачи, особенно в двумерном случае.

Первая работа такого рода принадлежит Ласту и Таулессу [88], которые на квадратном листе проводящей графитовой бумаги пробивали круглые отверстия с центрами, лежащими в случайных узлах квадратной решетки, и измеряли сопротивление между противоположными сторонами квадрата как функцию числа отверстий (отверстия, соответствующие ближайшим узлам, перекрывались).

Затем Ватсон и Лит [36] измерили электропроводность квадратной металлической сетки, в которой случайным образом вырезались узлы. Для задачи узлов они получили  $t_2 = 1,38 \pm 0,12$ . Левинштейн [37] исследовал электропроводность проводящей решетки, нарисованной на непроводящей бумаге аквадагом (графитовой суспензией). Для моделирования задачи связей на случайно выбранных связях решетки вырезались отверстия, прерывающие графитовую линию. Для моделирования задачи узлов делались отверстия с центром в узлах. Каждое из этих отверстий уничтожало все контакты графитовых линий в своем узле. Работая с квадратами

до:  $120 \times 120$ , Левинштейн получил, что  $t_2$  одинаково для задачи узлов и связей и равно  $1,15 \pm 0,15$ . Для трехмерной решетки задача узлов моделировалась с помощью куба, собранного из  $16 \times 16 \times 16$  стандартных сопротивлений [39]. В случайно выбранных узлах куба разрывались контакты всех сопротивлений. Исследование зависимости электропроводности от доли разорванных узлов привело к значению  $t_3 = 2$ .

Второй метод нахождения  $t$  состоит в вычислении сопротивления на ЭВМ с помощью решения системы уравнений Кирхгофа. Киркпатрик [89] для кубов  $25 \times 25 \times 25$  получил  $t_3 = 1,6 \pm 0,1$  для задачи связей и  $t_3 = 1,5 \pm 0,2$  для задачи узлов. Для плоских решеток, согласно Киркпатрику [28],  $t_2 = 1,1 \pm 0,1$ . Онизука [40] для задачи узлов в кубе  $50 \times 50 \times 50$  получил  $t_3 = 1,725 \pm 0,005$ . Страли [38] для квадратов  $100 \times 100$  нашел  $t_2 = 1,10 \pm 0,05$  для задачи связей и  $t_2 = 1,25 \pm 0,05$  для задачи узлов. В кубах  $25 \times 25 \times 25$  и  $30 \times 30 \times 30$  он получил  $t_3 = 1,70 \pm 0,05$  и  $t_3 = 1,75 \pm 0,05$  для задач связей и узлов соответственно.

Из перечисленных результатов можно сделать вывод, что индексы  $t$  для задач узлов и связей близки, т. е. в этом случае нет явного противоречия с гипотезой универсальности. Естественно поэтому считать, что и для континуального аналога рассматриваемой задачи (§ 17) индекс  $t$  имеет такие же значения.

Пусть в двухфазной системе одна фаза обладает локальной электропроводностью, равной единице, а другая совсем не проводит ток. Тогда можно записать зависимость эффективной электропроводности среды от доли пространства  $\phi$ , занятой проводящей фазой:

$$\sigma(\phi) \propto (\phi - \phi_c)^t \quad (\phi \geq \phi_c). \quad (20.2)$$

Существует значительное число физических систем, электропроводность которых интерпретируют либо с помощью представления о решетке из проводящих и непроводящих элементов (формула (20.1)), либо с помощью представления о континуальной двухфазной системе (формула (20.2)). К первым относятся вольфрамовые бронзы [90], ко вторым — ртуть при температурах выше критической [91], щелочно-аммиачные растворы [91, 41], гранулярные металлы — керметы [92],

островковые металлические пленки [93]. Во всех этих системах происходит переход металл — диэлектрик при изменении состава и плотности. В работах [90, 92, 93] индекс  $t$  определялся из экспериментальных данных. Он оказался равным  $1,8 \pm 0,2$  и  $1,9 \pm 0,2$  для трехмерных систем [90, 92] и  $1,15 \pm 0,20$  для двумерной системы [93], что находится в неплохом согласии с приведенными выше численными значениями.

Кроме непосредственных применений, связанных с задачами об электропроводности, имеется ряд других физических задач, которые сводятся к вычислению величины  $\sigma(x)$ . Во-первых, это задача о спектре длинноволновых спиновых волн в гайзенберговских разбавленных ферромагнетиках [89]. Во-вторых, это связанная с предыдущей и уже упоминавшаяся в § 16 задача о поведении критической температуры  $T_c(x)$  разбавленного гайзенберговского ферромагнетика при приближении к порогу протекания [10—12]. В-третьих, это задача об эластичности гелей (упругих полимерных растворов) [94, 95].

Мы ограничились лишь кратким перечислением физических приложений  $\sigma(x)$ . Несмотря на очевидную значимость этих применений, более подробный анализ не входит в наши намерения. В этом параграфе величина  $\sigma(x)$  будет интересовать нас главным образом с точки зрения той информации, которую она может дать о топологии бесконечного кластера вблизи порога протекания. Топология бесконечного кластера представляет интерес потому, что, как мы увидим ниже, вычисление электропроводности сильно неоднородной среды сводится к вычислению электропроводности некоторого бесконечного кластера.

О некоторых характеристиках бесконечного кластера уже говорилось выше. Так, мы уже видели, что его плотность равна  $P(x)$ . Размер его пор, или «период», равен радиусу корреляции  $L(x)$ . Однако для вычисления  $\sigma(x)$  этих сведений самих по себе недостаточно. Поэтому можно попытаться сделать некоторые предположения о структуре бесконечного кластера или задаться некоторой его моделью, а затем рассчитать  $\sigma(x)$ , выразив тем самым индекс  $t$  через другие индексы, характеризующие бесконечный кластер, прежде всего индекс  $\nu$ . Затем, сравнивая результат с известным численным значением  $t$ ,

можно было бы выбрать модель бесконечного кластера. О некоторых шагах в этом направлении мы расскажем ниже.

**а) Представление о мертвых концах.** Рассмотрим задачу об электропроводности большого куба со стороны, которая в единицах постоянной решетки равна  $l$ . Предположим, что бесконечный кластер напоминает простую кубическую решетку (или, скорее, сверхрешетку, построенную на исходной решетке). Тогда  $1/3$  связей, принадлежащих бесконечному кластеру, образует цепочки,

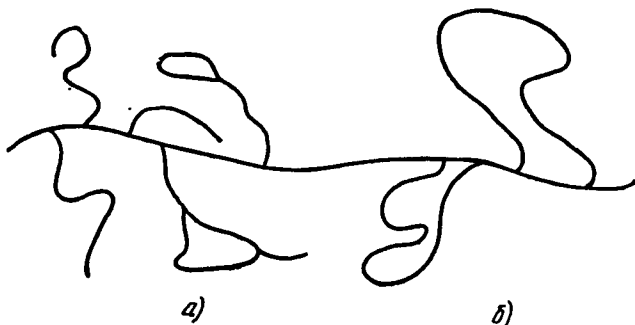


Рис. 5.12. Мертвые концы (а) и неэффективные с точки зрения электропроводности дублирования (б) на токопроводящей цепочке.

соединяющие две противоположные грани куба, т. е. число связей в этих цепочках порядка  $1/3 P(x) l^3$ . В каждой цепочке  $l$  связей, и, следовательно, число параллельных цепочек есть  $1/3 P(x) l^2$ . Сопротивление каждой цепочки пропорционально ее длине  $l$ . Поэтому сопротивление всего куба  $R$  пропорционально  $l [1/3 P(x) l^2]^{-1}$ . В результате для удельной электропроводности имеем

$$\sigma(x) \propto (Rl)^{-1} \propto P(x) \propto (x - x_c)^\beta, \quad (20.3)$$

т. е.  $t = \beta$ . В действительности  $t$  намного больше  $\beta$  (см. табл. 5.2), т. е. электропроводность  $\sigma(x)$  вблизи порога много меньше, чем следует из (20.3). Это означает, что подавляющая часть бесконечного кластера не проводит ток. Ласт и Таулесс [88] сделали вывод, что основная масса бесконечного кластера сосредоточена в цепочках, заканчивающихся тупиками. Они назвали такие цепочки мертвыми концами (рис. 5.12, а). Эти цепочки дают

вклад в  $P(x)$ , но не дают вклада в  $\sigma(x)$ . Вывод Ласта и Таулесса однако нельзя считать однозначным. Неэффективность бесконечного кластера с точки зрения электропроводности может быть связана и с большим числом длинных цепочек, дублирующих малые участки коротких цепочек (рис. 5.12, б).

б) Модель одножильной сетки. Дальнейшая конкретизация модели бесконечного кластера вблизи порога

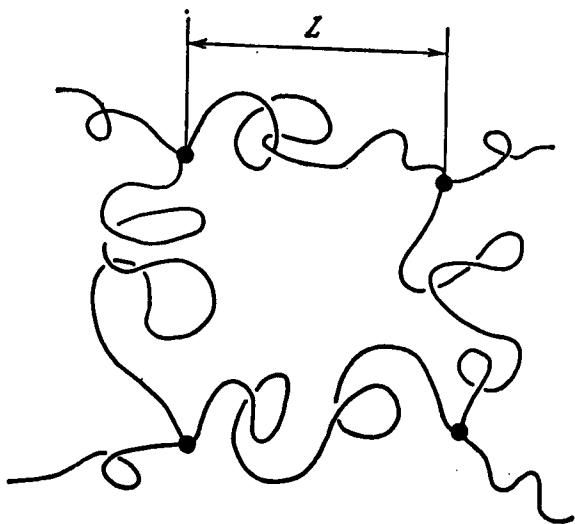


Рис. 5.13. Схематическое изображение модели одножильной сетки (без мертвых концов). Длина макросвязи вдоль кривой равна  $L$ .

протекания была предпринята Скал и Шкловским [53] и независимо Де Женом [94]. Эти авторы предположили, что в бесконечном кластере можно выделить мертвые концы и несущую их одномасштабную скелетную сетку с характерным размером (средним расстоянием между узлами) порядка радиуса корреляции  $L = (x - x_c)^{-\nu}$  (см. рис. 5.13). Участок скелетной сетки между узлами был назван макросвязью. Предполагалось, что скелетная сетка может быть выделена таким образом, что по крайней мере половина макросвязей не дублирована. В этом смысле макросвязь можно назвать одножильной. Допускалось, что макросвязь может быть сильно извилиста, т. е. представлять собой нечто вроде полимерного клубка (рис. 5.13). Иначе говоря, считалось, что длина ма-

кросвязи  $\mathcal{L}$  может быть много больше расстояния между ее концами  $L$ , т. е. при  $x - x_c \ll 1$  в единицах постоянной решетки

$$\mathcal{L} \approx (x - x_c)^{-\zeta}, \quad \text{где } \zeta \geq v. \quad (20.4)$$

Сформулированную выше модель мы будем называть моделью одножильной сетки. Попытки проверить эту модель непосредственным исследованием бесконечного кластера [53] не смогли ни опровергнуть, ни достаточно убедительно подтвердить ее. Подчеркнем поэтому, что модель одножильной сетки нельзя считать обоснованной. Бесконечный кластер содержит структуры всех масштабов, и априори не известно, можно ли из него выделить такой простой объект, как одножильную сетку.

Попытаемся все-таки без дальнейших предположений вычислить величину  $\sigma(x)$  одножильной сетки. Для этого сначала найдем индекс  $\zeta$  с помощью следующего строгого рассуждения [53]. Пусть имеется одножильная сетка, соответствующая некоторому значению  $x > x_c$ . Разорвем каждую связь сетки с вероятностью  $(x - x_c)x^{-1}$ . В результате доля целых связей окажется равно  $x[1 - (x - x_c)/x] = x_c$ , т. е. мы приходим к порогу протекания на исходной периодической решетке. Ввиду одножильности макросвязи для ее разрыва достаточно разрыва одной из составляющих ее связей решетки. Поэтому вероятность разрыва каждой макросвязи пропорциональна ее длине и равна  $\mathcal{L}(x - x_c)/x$ . Чтобы скелетная сетка пришла к порогу протекания, нужно разорвать определенную долю макросвязей  $y_c$ . Эта величина порядка нескольких десятых, как и все пороги протекания. Записав условие разрыва скелетной сетки

$$\frac{x - x_c}{x} \mathcal{L} = y_c, \quad (20.5)$$

получаем

$$\mathcal{L} \propto (x - x_c)^{-1}, \quad (20.6)$$

т. е.  $\zeta = 1$ . В двумерном случае  $v > 1$  и, следовательно,  $v > \zeta$ , т. е. модель одножильной сетки неприменима. В этом случае очень важную роль играет дублирование макросвязей. Если  $d \geq 3$ , то  $v < 1$  и никаких противоречий не возникает. В частности, при  $d = 3$   $v_3 \approx 0,9$  и из (20.6) получаем  $\mathcal{L} \approx L^{1,1}$ . Извилистость макросвяз

оказывается очень слабой по сравнению с извилистостью пути случайных блужданий ( $\mathcal{L} \approx L^2$ ) и даже по сравнению с извилистостью самоизбегающих путей ( $\mathcal{L} \approx L^{5/3}$ ).

Теперь мы можем вычислить  $\sigma(x)$ . Если сопротивление одной связи  $R_0$ , то сопротивление макросвязи  $\mathcal{L}R_0$ . Поэтому сопротивление кубика со стороной  $L$  по порядку величины равно

$$R = R_0 \mathcal{L}. \quad (20.7)$$

Поскольку в масштабах, больших  $L$ , бесконечный кластер с хорошей точностью однороден,  $\sigma(x)$  можно вычислить как удельную электропроводность кубика со стороной, равной радиусу корреляции. В этом параграфе мы измеряем радиус корреляции  $L$  и длину  $\mathcal{L}$  в единицах постоянной решетки  $l_0$  и оперируем с ними, как с безразмерными величинами. В размерных единицах радиус корреляции равен  $Ll_0$  и для  $\sigma(x)$  получаем

$$\sigma(x) = \frac{1}{R (Ll_0)^{d-2}}. \quad (20.8)$$

Подставляя (20.7) и (20.6) в (20.8), получим

$$\sigma(x) = R_0^{-1} \mathcal{L}^{-1} (Ll_0)^{2-d} = R_0^{-1} (x - x_c)^{1+\nu(d-2)} l_0^{2-d},$$

или

$$t = 1 + \nu(d - 2). \quad (20.9)$$

Таким образом, в трехмерном случае модель одножильной сетки дает [53]

$$t_3 = 1 + \nu_3. \quad (20.10)$$

Подстановка значения  $\nu_3 = 0,9$  дает значение  $t_3 = 1,9$ , что несколько больше расчетной величины  $t_3 = 1,7$ . Таким образом, модель одножильной сетки не точна. Однако в силу своей простоты и наглядности она имеет эвристическое значение. С ее помощью мы проиллюстрируем некоторые выводы в теории прыжковой проводимости. Модель одножильной сетки использовалась и в теории разбавленных ферромагнетиков [96].

в) Применение гипотезы подобия к вычислению электропроводности  $\sigma(x)$ . В работе Левинштейна и др. [97] был предложен подход, основанный на углублении гипотезы подобия. Авторы предположили, что крупномасштабная структура с приближением к порогу протекания



остается подобной самой себе в том смысле, что топология не изменяется, а все линейные размеры растут пропорционально радиусу корреляции  $L = |\tau|^{-\nu}$ . Грубо говоря, они сочли, что фотографии двух бесконечных кластеров при  $\tau = \tau_1$  и  $\tau = \tau_2$  совпадут, если масштаб одного из них изменить в  $(\tau_1/\tau_2)^\nu$  раз. Разумеется, это не относится к мелкомасштабной структуре, которая содержит минимальную длину, равную постоянной решетки. (Заметим, что если попытаться совместить предположение о существовании единственной длины  $L$  с моделью одножильной сетки, то оно сразу же отвергнет возможность извилистости макросвязей, приводящую к возникновению второй длины.)

Второе предположение, сделанное в [97], состоит в том, что сопротивление бесконечного кластера определяется крупномасштабной структурой. Оба предположения приводят к простой связи между индексами  $l$  и  $\nu$ . Рассмотрим нашу систему при значениях  $\tau = \tau_1$  и  $\tau = \tau_2$ , причем  $\tau_2 < \tau_1$ . Согласно сказанному выше мы заменяем ее системой произвольно перепутанных и связанных друг с другом проводов. Уменьшение  $\tau$  соответствует увеличению всех масштабов (но не толщины проводов!) в  $L(\tau_2)/L(\tau_1) = (\tau_1/\tau_2)^\nu$  раз. При изучении электропроводности мы должны рассматривать конечную систему, находящуюся между двумя плоскими металлическими контактами. Примем, что при изменении  $\tau$  расстояние между контактами также возрастает в  $(\tau_1/\tau_2)^\nu$  раз и во столько же раз возрастает разность потенциалов между контактами. Поскольку нас будет интересовать удельная электропроводность, изменение полного объема рассматриваемой системы не играет никакой роли.

При изменении масштабов системы никакого перераспределения токов не произойдет. Так как электрическое поле осталось неизменным, ток внутри каждого провода также не изменится. Однако макроскопическая плотность тока при уменьшении  $\tau$  уменьшится.

Рассмотрим элемент площади в плоскости, перпендикулярной среднему направлению тока. Число путей, пронизывающих этот элемент, уменьшится за счет увеличения всех масштабов в  $[L(\tau_1)/L(\tau_2)]^2 = (\tau_2/\tau_1)^{2\nu}$  раз. Следовательно, ток, проходящий через этот элемент, и удельная электропроводность также уменьшатся в

$(\tau_2/\tau_1)^{2\nu}$  раз. Поэтому в трехмерном случае

$$\sigma(x) \propto \tau^{2\nu}, \quad \text{т. е. } t = 2\nu. \quad (20.11)$$

Аналогичные рассуждения справедливы и в двумерном случае. Однако здесь нужно рассматривать ток, проходящий не через элемент площади, а через элемент длины. Поэтому при  $d = 2$

$$\sigma(x) \propto \tau^\nu, \quad \text{т. е. } t = \nu. \quad (20.12)$$

Подставляя значения  $\nu_2 = 1,33$  и  $\nu_3 = 0,90$  в (20.12) и (20.11), имеем  $t_2 = 1,33$  и  $t_3 = 1,8$ , что близко к экспериментальным значениям, но несколько больше их.

В заключение заметим, что совсем недавно появилось первое вычисление индекса  $t$  методом ренормгруппы. Ватсон и Стинчкомб [98] получили  $t_2 = 1,15$ , что находится в согласии с данными табл. 5.2.

## § 21. Теория протекания и электропроводность сильно неоднородных сред

В § 15 мы видели, что вычисление прыжковой проводимости сводится к задаче об электропроводности случайной сетки (см. рис. 4.6) с экспоненциально широким спектром сопротивлений.

Аналогичная задача может быть поставлена и в непрерывном пространстве. Пусть, например, требуется найти эффективную электропроводность среды с локальной удельной электропроводностью  $\sigma(\mathbf{r})$ , неоднородность которой экспоненциально велика, т. е.

$$\sigma(\mathbf{r}) = \sigma \exp[-\xi(\mathbf{r})], \quad (21.1)$$

причем

$$\langle [\xi(\mathbf{r}) - \langle \xi \rangle]^2 \rangle \gg 1. \quad (21.2)$$

Такой вид имеет, например, локальная электропроводность полупроводника с крупномасштабными флуктуациями потенциала, искривляющими дно зоны проводимости (см. рис. 3.5). В этом случае  $\xi(\mathbf{r}) = \epsilon(\mathbf{r})/kT$ , где  $\epsilon(\mathbf{r})$  — расстояние от уровня Ферми до дна зоны проводимости.

Задачу с экспоненциально широким спектром сопротивлений можно сформулировать и на решетке [99]. Представим себе простую кубическую решетку, между

ближайшими узлами которой включены случайные сопротивления. Запишем эти сопротивления в виде

$$R = R_0 \exp \xi', \quad (21.3)$$

где  $R_0$  — постоянная величина, а случайная переменная  $\xi'$  равномерно распределена в интервале  $-\xi_0 \leq \xi' \leq \xi_0$ , причем  $\xi_0 \gg 1$ . Требуется найти удельную электропроводность решетки и, в частности, определить ее зависимость от параметра  $\xi_0$ .

Все перечисленные выше системы мы назовем сильно неоднородными средами, подчеркивая тем самым главное их свойство — экспоненциально широкий спектр локальных значений электропроводности. Настоящий параграф посвящен разработке метода вычисления электропроводности сильно неоднородных сред. Как мы уже говорили в § 15, этот метод основан на теории протекания. Для первоначальной формулировки метода протекания и всестороннего его обсуждения мы выберем последнюю из поставленных выше задач — задачу на простой кубической решетке. Эта задача наиболее наглядна. Кроме того, нам кажется, что оперировать с решеточными задачами на первых порах легче, чем с задачами на случайных узлах и континуальными задачами, которые возникают в других случаях. Наконец, очень важным является также тот факт, что для избранной нами задачи выполнены детальные машинные расчеты. Это позволит подвергнуть прямой проверке предлагаемый метод вычисления электропроводности.

План параграфа состоит в следующем. Сначала мы приведем вывод, позволяющий определить только показатель экспоненты эффективной электропроводности. Затем мы сравним результат с численными расчетами. Далее мы приведем еще один вывод, позволяющий найти не только показатель экспоненты электропроводности, но и предэкспоненциальный множитель с точностью до численного коэффициента. Затем мы проиллюстрируем полученные результаты с помощью модели одножильной сетки.

Чтобы найти показатель экспоненты электропроводности простой кубической решетки с сопротивлениями (21.3), поступим следующим образом. Зададим некоторое число  $\xi$  в интервале  $-\xi_0 < \xi < \xi_0$  и заменим все

сопротивления с  $\xi' > \xi$  бесконечными сопротивлениями, т. е., попросту говоря, разорвем их. Пусть электропроводность решетки, соответствующая заданному значению  $\xi$ , есть  $\sigma(\xi)$ . Интересующая нас электропроводность всей сетки, очевидно, есть  $\sigma(\xi_0)$ . Задание  $\xi$  определяет вероятность того, что произвольно выбранное сопротивление не разорвано:

$$x(\xi) = \int_{-\xi_0}^{\xi} F(\xi') d\xi', \quad (21.4)$$

где  $F(\xi')$  — функция распределения  $\xi$ , по условию задачи имеющая вид

$$F(\xi') = \begin{cases} (2\xi_0)^{-1}, & |\xi| \leq \xi_0, \\ 0, & |\xi| > \xi_0. \end{cases} \quad (21.5)$$

Из (21.4) и (21.5) получаем

$$x(\xi) = \frac{\xi_0 + \xi}{2\xi_0}. \quad (21.6)$$

При  $\xi$ , близких к  $-\xi_0$ , величина  $x(\xi)$  мала и целые сопротивления образуют изолированные кластеры. При этом  $\sigma(\xi) = 0$ . Будем постепенно увеличивать  $\xi$ . Когда  $\xi$  достигнет порогового значения  $\xi_c$ , при котором  $x(\xi)$  оказывается равным порогу протекания  $x_c(b)$  для задачи связей, образуется бесконечный кластер из целых сопротивлений. Согласно (21.6) критическое значение  $\xi_c$  определяется условием

$$\frac{\xi_0 + \xi_c}{2\xi_0} = x_c(b). \quad (21.7)$$

При дальнейшем увеличении  $\xi$  от  $\xi_c$  до  $\xi_c + 1$  вследствие образования бесконечного кластера электропроводность  $\sigma(\xi)$  становится отличной от нуля и быстро возрастает (рис. 5.14). Дело в том, что в этой области значений  $\xi$  очень быстро уменьшается радиус корреляции

$$L(\xi) \approx l_0 [x(\xi) - x_c]^{-\nu} \sim l_0 (\xi - \xi_c)^{-\nu} \xi_0^{\nu}, \quad (21.8)$$

где  $l_0$  — период кубической решетки. Следовательно, быстро увеличивается число параллельных проводящих цепочек в сетке бесконечного кластера. В то же время,

пока  $\xi$  меняется на величину, меньшую единицы, изменением максимальной величины индивидуальных сопротивлений, входящих в кластер, можно пренебречь. Поэтому при  $\xi - \xi_c \ll 1$  рост величины  $\sigma(\xi)$  происходит степенным образом:

$$\sigma(\xi) \propto (\xi - \xi_c)^b, \quad \text{где } b > 0.$$

Назовем критической подсеткой бесконечный кластер, возникший при разности  $\xi - \xi_c$  порядка единицы. Сопротивление критической подсетки определяется наиболее высокоомными ее сопротивлениями (по определению эти сопротивления не могут быть шунтированы более низкоомными, иначе протекание возникло бы при  $\xi < \xi_c$ ). Поэтому для электропроводности критической подсетки имеем

$$\sigma(\xi_c + 1) \approx \sigma_0 \exp(-\xi_c). \quad (21.9)$$

Здесь показатель экспоненты мы написали с точностью до слагаемых порядка единицы. Дальнейшее увеличение  $\xi$  уже не может существенно изменить  $\sigma(\xi)$ , несмотря на продолжающееся увеличение густоты сетки бесконечного кластера. Дело в том, что теперь при изменении  $\xi - \xi_c$  в несколько раз будут подключаться сопротивления экспоненциально большие, чем  $\exp(\xi_c)$ . Новые цепочки, несмотря на их большое число, оказываются зашунтированными критической подсеткой и практически не меняют  $\sigma(\xi)$  (см. рис. 5.14). Таким образом, критическая подсетка по порядку величины определяет электропроводность случайной сетки, т. е.

$$\sigma(\xi_0) = \sigma_0 \exp(-\xi_c). \quad (21.10)$$

Мы пришли к очень важному и, как увидим ниже, общему для всех сильно неоднородных систем выводу: если последовательно в порядке возрастания сопротивлений включать элементы среды, то показатель экспоненты эффективной электропроводности определится

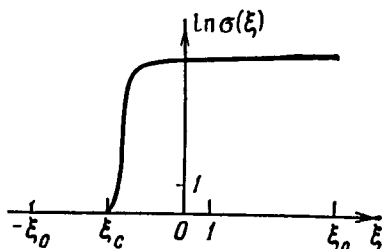


Рис. 5.14. Схематическая зависимость электропроводности  $\sigma(\xi)$  от максимальной величины показателей экспоненты включенных сопротивлений  $\xi$ .

теми элементами, которые впервые создадут протекание. Этот вывод, впервые сформулированный в работах Амбегаокара, Гальперина и Ланжера, Поллака и Шкловского и Эфроса (см. [16—18] к гл. 4), послужил основой теории прыжковой проводимости. Величина предэкспоненциального множителя  $\sigma_0$  оценена ниже (см. (21.12)).

Доведем до конца нашу модельную задачу. Согласно табл. 5.1 для простой кубической решетки  $x_c(b) = 0,25$ . Тогда из (21.7) получаем  $\xi_c = -\xi_0/2$ , и, следовательно, экспоненциальная зависимость электропроводности от  $\xi_0$  имеет вид \*)

$$\sigma(\xi_0) = \sigma_0 \exp(\xi_0/2). \quad (21.11)$$

Формула (21.11) по смыслу своего вывода справедлива лишь при  $\xi_0 \gg 1$ . Только в этом случае благодаря сильному разбросу сопротивлений из всей решетки можно выделить критическую подсетку и записать ее электропроводность в виде (21.9). Что представляет собой множитель  $\sigma_0$ ? Можно сказать сразу из соображений размерности, что

$$\sigma_0 = c(\xi_0) [R_0 l_0]^{-1}, \quad (21.12)$$

где  $l_0$  — период решетки и  $c(\xi_0)$  — безразмерный коэффициент, который может содержать некоторую степень  $\xi_0$ . Какая именно это степень, мы узнаем несколько ниже. Сейчас для нас важно лишь то, что при больших  $\xi_0$  зависимость  $\sigma_0$  от  $\xi_0$  степенная и никак не может конкурировать с главным экспоненциальным множителем в (21.11).

С целью проверки идей, лежащих в основе вывода формулы (21.11), Киркпатрик предпринял расчет эффективной электропроводности куба, имеющего  $15 \times 15 \times 15$  узлов, как функции  $\xi_0$ . Случайная переменная  $\xi'$  задавалась датчиком случайных чисел. Вычислялся ток при единичном напряжении на противоположных гранях куба. Для этого с помощью ЭВМ решалась система уравнений Кирхгофа, выражающая законы сохранения

\*) Отрицательное значение  $\xi_c$  в этой задаче связано с тем, что величина  $\xi'$  изменяется в интервале  $-\xi_0 \leq \xi' \leq \xi_0$ . В теории прыжковой проводимости величины  $\xi_i$  в (15.34) положительны и в результате  $\xi_c > 0$ ,

тока в каждом узле. Для проверки формулы (21.11) можно графически изобразить зависимость  $\ln[\sigma(\xi_0)/\sigma(0)]$  от  $\xi_0$  (здесь  $\sigma(0) \equiv (R_0 l_0)^{-1}$  — электропроводность куба при  $\xi_0 = 0$ ). Согласно (21.11) при  $\xi_0 \gg 1$  эта зависимость должна приближаться к прямой с наклоном  $1/2$ . В первой работе [99] Киркпатрику удалось продвинуться до  $\xi_0 \approx 7$ . При этом последние точки лучше ложились

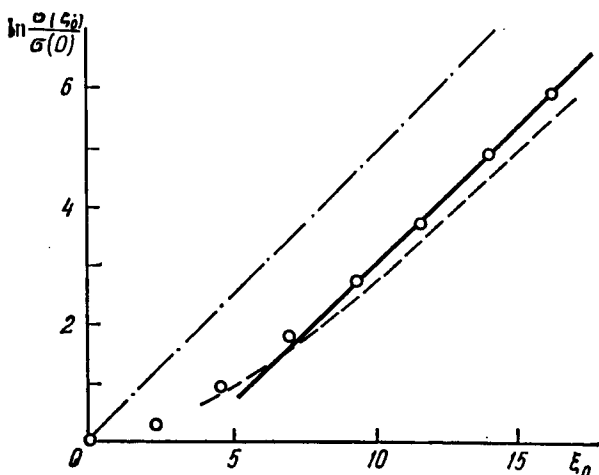


Рис. 5.15. Зависимость электропроводности  $\sigma$  модельной решеточной задачи от параметра  $\xi_0$ , характеризующего разброс показателей экспоненты сопротивлений. Точки — расчет Сигера и Пайка [101]. Прямая, проведенная через последние точки, почти параллельна биссектрисе координатного угла (штрих-пунктир), т. е. имеет наклон, близкий к теоретической величине  $1/2$ . Штриховая линия соответствует формуле (21.18) при  $A = 1$ .

на прямую с наклоном  $1/3$ , чем с наклоном  $1/2$ . Это послужило основанием для недоверия к методу протекания. Однако последующие расчеты, проведенные самим Киркпатриком [100] и Сигером и Пайком [101], показали, что при достаточно больших  $\xi_0$  ( $\xi_0 \geq 9$ ) согласие расчетов с результатом метода протекания становится вполне удовлетворительным. Это хорошо видно на рис. 5.15, построенном по данным работы [101].

Биссектриса координатного угла согласно указанным на осях масштабам имеет наклон  $1/2$ . Видно, что расчетные точки постепенно приближаются к прямой,

которая имеет очень близкий наклон. Если в области сравнительно малых  $\xi_0$  попытаться поместить некоторые точки на прямую, то полученные наклоны могут быть существенно меньше  $1/2$ .

Таким образом, мы видим, что метод протекания хорошо предсказывает показатель экспоненты электропроводности. Однако он количественно применим лишь при очень большом разбросе сопротивлений элементов.

Теперь для той же самой задачи мы приведем еще один вывод эффективной электропроводности  $\sigma(\xi_0)$ , который одновременно с показателем экспоненты позволяет определить степенную зависимость предэкспоненты  $\sigma_0$  от  $\xi_0$ . Для подавляющей части экспериментальных данных интерес представляет только показатель экспоненты. Поэтому, если читателю первый вывод кажется достаточно убедительным и его специально не интересует предэкспоненциальный множитель прыжковой проводимости, то он уже здесь может перейти к следующему параграфу.

Второй вывод  $\sigma(\xi_0)$  был предложен в работах [35, 102]. Представим себе, что в бесконечной решетке вырезан конечный куб со стороной  $l$ . Разорвем мысленно все сопротивления куба, а затем будем включать их в порядке возрастания (начиная с самых маленьких) до тех пор, пока не возникнет протекание с грани на грань. Пусть  $\xi_{cl}$  есть величина  $\xi'$  самого большого сопротивления, которое для этого нужно включить. Если размер куба достаточно мал, то включенные сопротивления будут сильно отличаться друг от друга. В частности, максимальное сопротивление с  $\xi' = \xi_{cl}$  окажется много больше следующего по величине включенного сопротивления и целиком определит сопротивление всего куба, которое окажется равным

$$R = R_0 \exp \xi_{cl}. \quad (21.13)$$

Вычисляя удельную электропроводность куба со стороной  $l$ , из (21.13) получим

$$\sigma_l = (R_0 l)^{-1} \exp(-\xi_{cl}). \quad (21.14)$$

Ввиду конечности куба величина  $\xi_{cl}$  флуктуирует от реализации к реализации (или, иначе говоря, в зависимости от места, где в бесконечной решетке вырезан куб).



Величина  $\xi_{cl}$  связана с помощью (21.6) с порогом протекания конечного куба  $x_{cl}$ . Среднеквадратичная флуктуация  $x_{cl}$  обсуждалась нами в § 16 и определяется формулой (16.12). Из (16.12) следует, что среднеквадратичная флуктуация  $\xi_{cl}$  равна

$$\delta\xi_{cl} = B'\xi_0 \left(\frac{l_0}{l}\right)^{1/\nu}. \quad (21.15)$$

Здесь  $\nu$  — индекс радиуса корреляции трехмерной задачи ( $\nu \approx 0,9$ ),  $B'$  — численный коэффициент порядка единицы и, в отличие от (16.12), где  $l$  выражалось в единицах постоянной решетки  $l_0$ , длина  $l$  выражена в размерных единицах.

Будем постепенно увеличивать размер куба  $l$ . Тогда наступит момент, когда максимальное сопротивление перестанет играть определяющую роль. Вслед за этим число существенных сопротивлений станет большим и будет расти пропорционально объему куба. При этом величина  $\sigma_l$  перестанет зависеть от  $l$  и будет равна искомой удельной электропроводности макроскопической системы. Другим критерием перехода от уравнения (21.14) к макроскопической электропроводности, происходящего при увеличении  $l$ , является исчезновение больших относительных флуктуаций  $\sigma_l$ . Это происходит, когда среднеквадратичная флуктуация  $\xi_{cl}$  сравнивается с единицей, т. е., согласно (21.15), при

$$l \approx l_0 \xi_0^\nu \equiv L_0. \quad (21.16)$$

Удельное сопротивление куба со стороной  $L_0$  можно получить из (21.14), если учесть, что в таком кубе с хорошей точностью можно пренебречь отличием  $\xi_{cl}$  от  $\xi_c$ :

$$\sigma_{L_0} \approx (R_0 L_0)^{-1} \exp(-\xi_c). \quad (21.17)$$

Но, согласно сказанному выше, эффективная электропроводность  $\sigma$  по порядку величины равна  $\sigma_{L_0}$ , т. е.

$$\sigma(\xi_0) \approx (R_0 L_0)^{-1} \exp(-\xi_c), \quad (21.18)$$

или

$$\sigma(\xi_0) = A (R_0 l_0)^{-1} \xi_0^{-\nu} \exp(-\xi_c),$$

где  $A$  — численный коэффициент порядка единицы. Сравнивая (21.18) с (21.10), мы видим, что результат для

показателя экспоненты подтвердился. Кроме того, найдена степень  $\xi_0$  в предэкспоненциальном множителе  $\sigma_0$ . Заметим, что, согласно (21.8), существенная в этом выводе длина  $L_0$  есть не что иное, как радиус корреляции бесконечного кластера, который возникает, если максимальный показатель экспонент включенных сопротивлений  $\xi$  превышает  $\xi_c$  на единицу. Этот кластер был назван нами критической подсеткой. Таким образом, мы вновь приходим к выводу, что именно критическая подсетка определяет эффективную электропроводность решетки.

Из-за неизвестного численного множителя  $A$  нельзя достаточно точно сравнить формулу (21.18) с расчетными данными. Тем не менее на рис. 5.15 мы построили зависимость  $\sigma(\xi_0)$  от  $\xi_0$ , которую дает формула (21.18) при  $A = 1$  (штриховая линия). Видно, что как абсолютная величина  $\sigma(\xi_0)$ , так и характер кривой хорошо передаются формулой (21.18). Максимальное отклонение расчетных значений  $\sigma(\xi_0)$  от формулы (21.18) при  $A=1$  в большом интервале изменения  $\xi_0$  ( $3 \leq \xi_0 \leq 16$ ) составляет 50%. Таким образом, согласие (21.18) с расчетами можно считать удовлетворительным.

В заключение этого параграфа проиллюстрируем результат (21.18) с помощью конкретной модели бесконечного кластера — модели одножильной сетки. Сначала мы в рамках этой модели опишем критическую подсетку, а затем вычислим ее сопротивление. Критическая подсетка — это часть исходной решетки сопротивлений, которая несет практически весь ток и определяет электропроводность решетки. Согласно сказанному выше, для построения критической подсетки нужно включить все сопротивления с  $\xi' \leq \xi = \xi_c + 1$ . Мы уже говорили, что радиус корреляции критической подсетки есть  $L_0 \approx l_0 \xi_0^{\nu}$ . Каждая макросвязь критической подсетки представляет собой согласно модели одножильной сетки цепочку последовательно включенных сопротивлений. Длина этой цепочки  $\mathcal{L}_0$  может быть вычислена, если в формулу

$$\mathcal{L} = l_0 (\xi - \xi_c)^{-1} \xi_0, \quad (21.19)$$

вытекающую из (20.6) и (21.6), подставить  $\xi - \xi_c = 1$ . В результате получаем

$$\mathcal{L}_0 \approx l_0 \xi_0. \quad (21.20)$$

Из (21.20) видно, что в цепочке макросвязи критической подсетки имеется порядка  $\xi_0$  последовательно включенных сопротивлений. Величины показателей экспоненты  $\xi'$  этих сопротивлений по условию распределены в интервале от  $-\xi_0$  до  $\xi_c + 1 = -1/2\xi_0 + 1$ .

Можно показать, что распределение значений  $\xi'$  в этом интервале является равномерным. Таким образом, если величины  $\xi'$  одной макросвязи пронумеровать в порядке убывания:

$$\xi'_1 \geq \xi'_2 \geq \dots \geq \xi'_n, \quad (21.21)$$

то средняя разность между последовательными значениями  $\xi'$  будет порядка единицы. Это значит, что первое сопротивление, грубо говоря, в  $e$  раз больше второго, второе — в  $e$  раз больше третьего и т. д. Сопротивление последовательной цепочки с таким распределением сопротивлений определяется первым (самым большим) сопротивлением:

$$R \approx R \exp(\xi_1) \approx R_0 \exp(\xi_c). \quad (21.22)$$

Таким образом, мы нашли сопротивление макросвязи. Это сопротивление одновременно является сопротивлением кубика со стороной  $L_0$ . Поскольку в масштабах, больших  $L_0$ , критическая подсетка однородна, макроскопическая проводимость по порядку величины совпадает с удельной электропроводностью упомянутого кубика. Следовательно,

$$\sigma \approx (RL_0)^{-1} \approx (R_0L_0)^{-1} \exp(-\xi_c),$$

т. е. мы снова приходим к формуле (21.18).

## ГЛАВА 6

### ЗАВИСИМОСТЬ ПРЫЖКОВОЙ ПРОВОДИМОСТИ ОТ КОНЦЕНТРАЦИИ ПРИМЕСЕЙ И ДЕФОРМАЦИИ КРИСТАЛЛА

#### § 22. Удельное сопротивление $\rho_3$ для полупроводников с изотропными волновыми функциями примесных состояний

Как мы уже неоднократно отмечали, вычисление прыжковой проводимости сводится к расчету эквивалентной сетки сопротивлений Миллера и Абрахамса, построенной на случайных узлах (см. рис. 4.6). Сопротивление, включенное между каждой парой узлов этой сетки, имеет вид

$$R_{ij} = R_{ij}^0 e^{\xi_{ij}}, \quad (22.1)$$

где в случае изотропных волновых функций примесей

$$\xi_{ij} = \frac{2r_{ij}}{a} + \frac{e_{ij}}{kT}, \quad (22.2)$$

$R_{ij}^0$  определяется формулой (15.36),  $e_{ij}$  — формулой (15.33). Сетка Миллера и Абрахамса — частный случай сильно неоднородных систем, общий подход к которым, названный нами методом протекания, подробно рассмотрен в § 21. Там, в частности, показано, что метод протекания хорошо работает при достаточно сильном разбросе сопротивлений. С другой стороны, мы еще в § 15 убедились, что спектр значений сопротивлений  $R_{ij}$  чрезвычайно широк в случаях, интересных с точки зрения эксперимента. Таким образом, прыжковая проводимость представляет идеальный объект для применения метода протекания,

Для вычисления показателя экспоненты прыжковой проводимости кратко повторим первый вывод предыдущего параграфа. Включим сначала сопротивления с  $\xi_{ij} \leq \xi$  и будем увеличивать  $\xi$  до тех пор, пока не возникнет протекание по включенным сопротивлениям. Это произойдет, очевидно, когда  $\xi = \xi_c$ , где  $\xi_c$  — порог протекания известной нам из § 18 задачи на случайных узлах с критерием связности

$$\xi_{ij} \leq \xi. \quad (22.3)$$

Увеличение  $\xi$  от  $\xi_c$  до  $\xi_c + 1$  приведет к формированию критической подсетки сопротивлений. Эта сетка будет шунтировать сопротивления с существенно большими, чем  $\xi_c + 1$ , значениями  $\xi_{ij}$ . С другой стороны, электропроводность критической подсетки определяется сопротивлениями с  $\xi_{ij}$ , близкими к  $\xi_c$ . В результате экспоненциальная зависимость прыжковой проводимости от всех параметров задачи заключена в величине порога протекания  $\xi_c$  и описывается формулой

$$\rho = \rho_0 e^{\xi_c}, \quad (22.4)$$

которая является исходной для всех теоретических исследований прыжковой проводимости. Чтобы увидеть, как из, казалось бы, довольно нефизичной записи (22.4) возникают физические результаты, рассмотрим задачу о вычислении экспоненциальной концентрационной зависимости величины  $\rho_3$ , определяемой формулой (14.5). Для нахождения этой величины следует пренебречь энергетическим слагаемым в выражении (22.2) для  $\xi_{ij}$ . Тогда получаем сетку с сопротивлениями

$$R_{ij} = R_{ij}^0 \exp\left(\frac{2r_{ij}}{a}\right). \quad (22.5)$$

Об этой задаче уже шла речь в § 15, где рассматривались попытки решить ее, не используя метод протекания. Критерий связности (22.3) теперь имеет очень простой вид:

$$\frac{2r_{ij}}{a} \leq \xi, \quad \text{или} \quad r_{ij} \leq \frac{a\xi}{2},$$

т. е. связность пары примесей зависит только от расстояния между ними. Это означает, что задача свелась к

задаче сфер (см. рис. 5.10). Поэтому, если  $r_c$  — перколяционный радиус, т. е. порог протекания в задаче с условием связности  $r_{ij} \leq r$ , то для  $\xi_c$  получаем

$$\xi_c = \frac{2r_c}{a}. \quad (22.6)$$

Таким образом, оказывается, что прыжки на расстояние перколяционного радиуса целиком определяют показатель экспоненты электропроводности. О вычислении перколяционного радиуса говорилось в § 18. Согласно приведенным в табл. 5.3 результатам расчетов,

$$B_c \equiv \frac{4\pi}{3} N r_c^3 = 2,7 \pm 0,1, \quad (22.7)$$

т. е.

$$r_c = (0,865 \pm 0,015) N^{-1/3}, \quad (22.8)$$

здесь  $N$  — концентрация узлов сетки Миллера и Абрахамса, равная концентрации основных примесей. Подставляя (22.8) в (22.6), из (22.4) получаем

$$\rho_3 = \rho_{03} \exp\left(\frac{\alpha}{N^{1/3} a}\right), \quad (22.9)$$

где

$$\alpha = 1,73 \pm 0,03. \quad (22.10)$$

Таким образом, мы пришли к важной экспоненциальной зависимости прыжковой проводимости от концентрации основных примесей (см. [16—18] к гл. 4). Сравнивая (22.9) с (14.6), мы видим, что найдена определенная в § 14 функция  $f(N)$ .

Сигер и Пайк (см. [101] к гл. 5) предприняли машинный расчет электропроводности по законам Кирхгофа с целью проверки того, насколько точной является формула (22.9). В отличие от действительной ситуации, в которой, согласно (15.36) и (15.18),  $R_{ij}^0$  степенным образом зависит от  $r_{ij}$ , они рассмотрели более простой случай, когда зависимость от  $r_{ij}$  целиком сосредоточена в показателе экспоненты, т. е.

$$R_{ij} = R_0 \exp\left(\frac{2r_{ij}}{a}\right), \quad (22.11)$$

где  $R_0$  — константа. Однако ясно, что для обсуждаемого сейчас вопроса о показателе экспоненты это различие не

имеет никакого значения. Ведь при выводе (22.9) мы вообще не конкретизировали зависимость  $R_{II}^0$  от  $r_{II}$ .

Расчет производился для куба, содержащего 1000 случайных узлов. Произвольным образом фиксировался борковский радиус, и концентрация менялась таким образом, что электропроводность убывала на 14 порядков (рис. 6.1). Оказалось, что зависимость  $\lg \sigma$  от концентрации очень хорошо описывается формулой (22.9), причем коэффициент  $\alpha$  в точности равен 1,73.

Перейдем теперь к сравнению предсказанной концентрационной зависимости (22.9) с экспериментальными данными. Напомним, однако, что формула (22.9) применима только к случаю волновых функций, спадающих на больших расстояниях по водородоподобному закону,  $F(r) \propto \exp(-r/a)$ . Радиус состояния  $a$ , фигурирующий в (22.9), в общем случае должен, согласно (4.12), вычисляться по формуле ( $E_0$  — энергия ионизации примеси)

$$a = \frac{\hbar}{\sqrt{2mE_0}}. \quad (22.12)$$

При этом он может быть несколько меньше борковского радиуса  $\frac{\hbar^2 \kappa}{me^2}$ .

Прежде всего рассмотрим арсенид галлия, у которого невырожденный минимум зоны проводимости расположен в центре зоны Бриллюэна и волновые функции доноров имеют водородоподобный вид. Авторы работы [1] нанесли на один график данные о зависимости  $\rho_3$  от

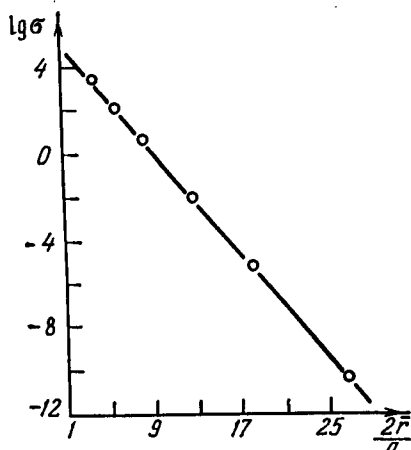


Рис. 6.1. Зависимость логарифма электропроводности от «среднего» расстояния между примесями  $\bar{r} = \left[ \frac{4\pi}{3} N \right]^{-1/3}$  для задачи, в которой сопротивление, включенное между двумя узлами, зависит только от расстояния между ними по закону (22.11).

$(N_D^{1/3} a)^{-1}$ , полученные в их работе и в работах других авторов [2—4] (рис. 6.2). Они исходили из значения  $a = 97 \text{ \AA}$ . Практически такое же значение получается из формулы (22.12), если, следуя табл. 1.1, подставить в нее значения  $m = 0,066m_0$  и  $E_0 = 6 \text{ мэв}$ .

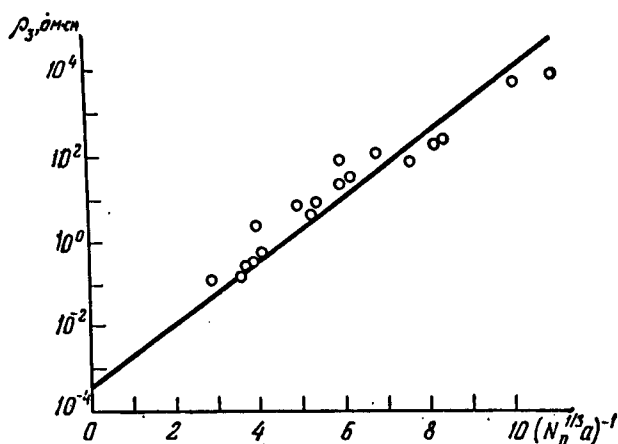


Рис. 6.2. Удельное сопротивление  $\rho_3$  слабо легированного  $n$ -GaAs как функция концентрации доноров [1—4].

На рис. 6.2 через экспериментальные точки нами проведена прямая, соответствующая формуле (22.9) с теоретическим значением коэффициента  $\alpha$  ( $\alpha_{\text{теор}} = 1,73$ ). Видно, что эта прямая неплохо описывает экспериментальные данные. Однако достаточно точно определить значение коэффициента  $\alpha$ , в силу разброса экспериментальных точек, не удастся. Скорее всего,  $\alpha = 1,7 \pm 0,2$ . Причина разброса точек на рис. 6.2 заключается, по-видимому, в не очень точном определении степени компенсации и концентрации доноров.

Кахлерт и др. [5] также исследовали зависимость  $\rho_3$  от  $N_D^{-1/3}$  в слабо легированных образцах арсенида галлия  $n$ -типа. Они получили значение  $\alpha = 1,88$ . Недавно появилась работа Лемуэна и др. [6], посвященная эпитаксиальным пленкам арсенида галлия. Авторы констатировали, что зависимость  $\rho_3$  от  $N_D^{-1/3}$  хорошо описывается формулой (22.9) при  $\alpha = 1,8$ . В фосфиде индия



спектр электронов вблизи дна зоны проводимости также изотропен. Подробное исследование прыжковой проводимости в этом полупроводнике было предпринято в работе Емельяненко и др. [7]. Авторы проанализировали зависимость  $\rho_3$  от  $(N_D^{1/3}a)^{-1}$ , исходя из значения радиуса состояния  $a = 77 \text{ \AA}$ . (Очень близкое значение,  $a = 76 \text{ \AA}$ , получается, если подставить в (22.12) данные для  $n\text{-InP}$  из табл. 1.1.) Значение  $\alpha$  оказалось равным 1,9, причем погрешность определения  $\alpha$  составляла, по нашей оценке, приблизительно 0,2. Таким образом, теоретическое значение 1,73 попадает в пределы погрешностей.

До сих пор мы говорили о полупроводниках, где примесные состояния изотропны. Однако имеются ситуации, когда, примесные состояния, строго говоря, анизотропны, но зависимость от углов настолько слабая, что при сравнении экспериментальных данных с формулой (22.9) анизотропией в первом приближении можно пренебречь. Здесь речь идет прежде всего об акцепторных состояниях в  $p\text{-Ge}$ . Асимптотика волновых функций этих состояний, необходимая для вычисления энергетического интеграла перекрытия (см. § 15), подробно обсуждалась в конце § 4. Там показано, что, вообще говоря, характерная длина спада волновой функции  $q^{-1}$  зависит от направления. Она максимальна для направлений [111] и минимальна в направлениях [100]. Для примеси галлия максимальное и минимальное значения  $q^{-1}$  равны 92 и 87  $\text{\AA}$ . Различие между этими значениями невелико, и поэтому можно пользоваться сферической моделью, которая пренебрегает гофрировкой поверхностей постоянной энергии. Эта модель дает для Ga значение боровского радиуса  $a = 90 \text{ \AA}$ .

Наиболее полные данные о прыжковой проводимости  $p\text{-Ge}$  были получены в работе Фрицше и Кейваса (см. [6] к гл. 4), где исследовалась серия образцов, в которой основной примесью был галлий. Температурные зависимости удельного сопротивления этой серии приведены на рис. 4.2.

Благодаря тому, что легирование образцов донорами и акцепторами производилось за счет ядерных реакций, вызванных облучением чистых кристаллов медленными

нейтронами, степень компенсации этих образцов определялась соотношением сечений ядерных реакций и была известна довольно точно ( $K=0,4$ ). Поэтому значения концентрации основных примесей могли быть точно определены по концентрации дырок в области зонной электропроводности. На

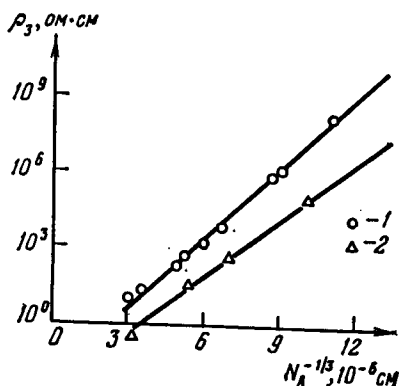


Рис. 6.3. Удельное сопротивление  $\rho_3$   $p$ -Ge с примесью Ga как функция концентрации Ga. Степень компенсации всех образцов  $K=0,4$ ; 1 — недеформированные кристаллы по данным работы Фрицше и Кейваса (см. [6] к гл. 4), 2 — образцы, подвергнутые сильному сжатию вдоль оси [100] [19] (см. § 23).

рис. 6.3 кружками показана зависимость  $\rho_3$  от  $N_A^{-1/3}$ . Видно, что все экспериментальные точки, в соответствии с законом (22.9), ложатся на одну прямую. Определение наклона прямой дает  $\alpha = 1,9 \pm 0,1$ .

В этом случае получается небольшое, но выходящее за рамки погрешностей отклонение  $\alpha$  от теоретического значения  $\alpha = 1,73 \pm 0,03$ . В остальных работах, посвященных  $p$ -германию, либо был велик разброс точек, либо исследовалось слишком мало образцов, что

не позволяло определить экспериментальное значение  $\alpha$  с желаемой точностью. Например, анализ данных Фрицше [8] для слабо компенсированного германия с примесью индия приводит к значению  $\alpha = 1,75 \pm 0,25$ . Из других данных по  $p$ -германию следует остановиться на работе [9], в которой исследовалась прыжковая проводимость по сравнительно глубоким примесным состояниям меди ( $E_0 = 40$  мэв). С помощью (22.12) можно легко убедиться, что радиус такого состояния, описывающий спад волновой функции на больших расстояниях, равен  $a = 47$  А. Он почти в два раза меньше, чем для мелких акцепторных примесей. Поэтому прыжковая проводимость имеет место при существенно больших концентрациях акцепторов. Однако и в этом случае сравнение

(22.9) с зависимостью  $\rho_3$  от  $N_A^{-1/3}$ , приведенной в статье [9], дает разумное значение  $\alpha = 2,0 \pm 0,3$ .

В кремнии поверхность постоянной энергии легкой дырки гофрирована, хотя и сильнее, чем в германии, но относительная амплитуда колебаний характерной длины спада  $q^{-1}$  как функции направления, согласно (4.25) и (4.26), все же невелика:

$$\frac{q_{[111]}^{-1} - q_{[100]}^{-1}}{q_{[111]}^{-1}} = \frac{(\gamma_1 + 2\gamma_3)^{1/2} - (\gamma_1 + 2\gamma_2)^{1/2}}{(\gamma_1 + 2\gamma_3)^{1/2}} \approx 0,20. \quad (22.13)$$

Поэтому можно надеяться, что сферическое приближение (4.27) довольно хорошо применимо и в этом случае.

Например, для акцепторной примеси бора, подставляя в (4.27) постоянные  $\gamma_1$ ,  $\gamma_2$ ,  $\gamma_3$  и энергию связи  $E_0$  из табл. 1.3 и 1.4 соответственно, получим  $a = 21,2 \text{ \AA}$ . Прыжковая проводимость кремния  $p$ -типа подробно изучена в работе Рэя и Фэна [10]. На рис. 6.4 приведена зависимость  $\rho_3$  от  $N_A^{-1/3}$  для кремния, легированного бором, построенная по данным этой работы. Видно, что экспоненциальный закон (22.9) хорошо выполняется. Определенный наклон прямой дает значение  $\alpha = 1,8 \pm 0,2$ .

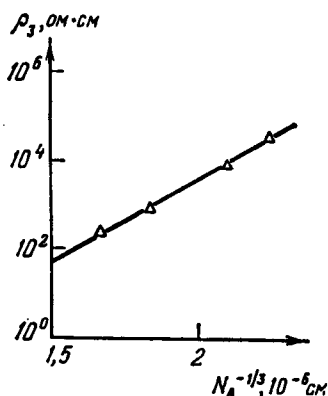


Рис. 6.4. Удельное сопротивление  $\rho_3$  Si с примесью В как функция концентрации В [10].

Резюмируем сравнение формулы (22.9) с экспериментальными данными. Прежде всего нужно констатировать, что во всех случаях предсказываемая этой формулой линейная зависимость  $\ln \rho_3$  от  $N^{-1/3}$  действительно наблюдается. В табл. 6.1 сведены все экспериментальные и теоретические значения  $\alpha$ . Видно, что в целом согласие теории и эксперимента удовлетворительное. Тем не менее часто имеется небольшое (обычно в пределах погрешностей) превышение экспериментальных значений над теоретическими.

Такую тенденцию можно пытаться объяснить, учитывая небольшое отклонение в распределении примесей

Т а б л и ц а 6.1

Экспериментальные и теоретические значения коэффициента  $\alpha$ 

$\alpha_{\text{эксп}}$		$\alpha_{\text{эксп}}$		$\alpha_{\text{теор}}$
<i>n</i> -GaAs [1]	1,7±0,2	<i>p</i> -Ge(Ga) ([6] гл. 4)	1,9±0,1	1,73±0,3
<i>n</i> -GaAs [5]	1,88	<i>p</i> -Ge(In) [8]	1,75±0,25	
<i>n</i> -GaAs [6]	1,9±0,2	<i>p</i> -Ge(Cu) [9]	2,0±0,3	
<i>n</i> -InP [7]	1,9±0,2	<i>p</i> -Si [10]	1,8±0,2	

от чисто случайного в образцах, изготовлявшихся из расплава. Дело в том, что в расплаве все доноры ионизованы и подвижны. Поэтому отталкивание доноров друг от друга может внести в их расположение определенную корреляцию. Коэффициент диффузии доноров в кристаллической решетке активационно зависит от температуры, и при некоторой температуре  $T_0$  среднее диффузионное смещение донора за характерное время охлаждения становится много меньше среднего расстояния между донорами. Можно сказать, что замораживается распределение доноров, равновесное при температуре  $T_0$ . Оценки, полученные с помощью эмпирических данных для коэффициента диффузии, дают для  $T_0$  величины порядка 1000°К. Если сравнивать энергию взаимодействия двух доноров, находящихся на среднем расстоянии, с величиной  $kT_0$ , то оказывается, что при концентрациях порядка  $10^{16} \div 10^{18} \text{ см}^{-3}$  и обычных значениях диэлектрической проницаемости,  $\kappa = 12 \div 16$ ,

$$\frac{e^2 N_D^{1/3}}{\kappa k T_0} = 0,01 \div 0,03. \quad (22.14)$$

Поэтому корреляция, вызванная отталкиванием, является очень слабой. Вопрос о том, на сколько процентов она может увеличить перколяционный радиус  $r_c$  и, следовательно, коэффициент  $\alpha$ , может быть разрешен с помощью расчетов методом Монте-Карло на ЭВМ.

Особое беспокойство у нас вызывает тот факт, что для серии образцов *p*-Ge со степенью компенсации  $K = 0,4$ , для которой процедура определения коэффи-

циента  $\alpha$ , по-видимому, оказывается наиболее точной, найденное значение  $\alpha$  достоверно отличается от теоретического. Дело в том, что в этих экспериментах примеси создавались в результате ядерных реакций в случайных узлах решетки при температурах, много меньших  $T_0$ , когда они уже не могли принять коррелированное расположение. Производившийся впоследствии отжиг дефектов при  $T = 400^\circ\text{C}$  также не мог привести к перемещению атомов галлия на расстояния, большие постоянной решетки. В этом можно убедиться, используя эмпирическую формулу для коэффициента диффузии галлия в германии. Таким образом, в рассматриваемом случае, по-видимому, отсутствует возможность отнести имеющееся небольшое расхождение с теорией за счет корреляции.

### § 23. Удельное сопротивление $\rho_3$ для полупроводников с анизотропными примесными состояниями

В предыдущем параграфе мы рассмотрели случай, когда экспоненциальную асимптотику волновой функции примеси можно считать изотропной. Для ряда объектов это приближение неприменимо. Сюда прежде всего относится  $n$ -германий, для которого имеется значительное количество экспериментальных данных. С анизотропными волновыми функциями приходится иметь дело и когда речь идет о  $n$ - и  $p$ -германии в условиях сильной одноосной деформации.

В настоящем параграфе мы применим метод протекания для вычисления экспоненциальной зависимости  $\rho_3$  от концентрации примесей в этих ситуациях. Идея метода протекания при этом остается неизменной: если имеется сетка сопротивлений (22.1), то необходимо решить задачу протекания с условием связности (22.3), и, согласно (22.4), порог протекания этой задачи  $\xi_c$  определит показатель экспоненты электропроводности. Однако в анизотропном случае первое слагаемое в  $\xi_{ij}$  не имеет простого вида  $2r_{ij}/a$  (см. (22.2)) и должно быть заново вычислено для каждой конкретной ситуации. Начнем с такого вычисления для волновых функций доноров в германии.

Как видно из § 15, первое слагаемое в  $\xi_{ij}$ , зависящее от вектора  $r_{ij}$ , соединяющего доноры  $i$  и  $j$ , целиком определяется показателем экспоненты энергетического интеграла перекрытия (15.4). Для  $n$ -германия волновые функции  $\Psi_i$  и  $\Psi_j$ , которые нужно подставлять в (15.4), согласно (2.27), (2.23) и (2.2) имеют вид

$$\Psi_i = \frac{1}{2} \sum_{g=1}^4 u_{n, k_g}(r) \exp(ik_g r) F_g(r), \quad (23.1)$$

$$\Psi_j = \frac{1}{2} \sum_{g'=1}^4 u_{n, k_{g'}}(r - r_{ij}) \exp(ik_{g'}(r - r_{ij})) F_{g'}(r - r_{ij}). \quad (23.2)$$

Здесь мы приняли  $r = r_i$  за начало координат;  $g$  — номер минимума зоны проводимости,  $u_{n, k_g}$  — периодическая функция в точке  $k_g$  зоны Бриллюэна,  $F_g(r)$  — огибающая функция примеси  $i$ , которая, согласно (4.16), (4.17), на больших расстояниях от примеси имеет вид

$$F_g(r) \propto \exp(-q_g(n)r) = \exp\left\{-\left(\frac{x^2 + y^2}{a^2} + \frac{z^2}{b^2}\right)^{1/2}\right\}, \quad (23.3)$$

где

$$a = \frac{\hbar}{\sqrt{2m_t E_0}}, \quad b = \frac{\hbar}{\sqrt{2m_l E_0}}. \quad (23.4)$$

Очень важно подчеркнуть, что декартовы системы координат  $(x, y, z)$  различны для разных минимумов. Для каждого данного минимума ось  $z$  совпадает с направлением вектора  $k_g$ , а оси  $x, y$  перпендикулярны ему.

Подставляя (23.1) и (23.2), например, в первый интеграл формулы (15.4) и воспользовавшись тем, что  $u_{nk}(r - r_{ij}) = u_{nk}(r)$ , получим

$$\begin{aligned} & \frac{1}{4} \sum_{g, g'} \exp(-ik_{g'} r_{ij}) \times \\ & \times \int u_{nk_g}(r) u_{nk_{g'}}(r) \exp(i(k_{g'} - k_g)r) \frac{e^{\alpha F_g(r)} F_{g'}(r - r_{ij})}{x |r - r_{ij}|} dr. \end{aligned} \quad (23.5)$$

Слагаемыми с  $g \neq g'$ , ввиду быстрых осцилляций подинтегрального выражения, можно пренебречь. Затем ши-

теграл можно разбить на интегралы по отдельным ячейкам и воспользоваться тем, что огибающие функции и кулоновский множитель медленно меняются на расстоянии порядка постоянной решетки. Тогда интеграл по ячейке от  $u_{n, k_g}^2$  можно заменить единицей и снова преобразовать сумму в интеграл:

$$\frac{1}{4} \sum_g \exp(-ik_g r_{ij}) \int \frac{e^2 F_g(r) F_g(r - r_{ij})}{x |r - r_{ij}|} dr. \quad (23.6)$$

К аналогичному виду можно привести и остальные интегралы (15.4). Вычисление слагаемых (23.6) удобно производить отдельно в системе координат  $(x, y, z)$  каждого эллипсоида. В результате получается, что

$$|I|^2 = \left(\frac{e^2}{6\pi\alpha^2}\right)^2 \sum_{g=1}^4 r_g^2 \exp\left(-\frac{2r_g}{a}\right), \quad (23.7)$$

где

$$\frac{r_g}{a} = \left(\frac{x_{ij}^2}{a^2} + \frac{y_{ij}^2}{a^2} + \frac{z_{ij}^2}{b^2}\right)^{1/2}, \quad (23.8)$$

$x_{ij}, y_{ij}$  и  $z_{ij}$  — проекции вектора  $r_{ij}$  на оси системы координат, связанной с минимумом  $g$ . В интересной для прыжковой проводимости области малых концентраций значения  $2r_g/a$  в (23.7) велики. Поэтому при заданном  $r_{ij}$  в сумме (23.7), как правило, господствует одно слагаемое, соответствующее минимальному из четырех значению  $2r_g/a$ . Следовательно, можно записать

$$|I|^2 \propto \exp\left\{-\frac{2 \ln r_g}{a}\right\}, \quad (23.9)$$

что для сопротивления перехода  $i \rightarrow j$  дает

$$R_{ij} = R_{ij}^0 \exp \xi_{ij}, \quad (23.10)$$

где

$$\xi_{ij} = 2 \frac{\ln r_g}{a}. \quad (23.11)$$

Имея в виду вычисление  $\rho_3$ , мы в (23.11) сразу же отбросили энергетическое слагаемое  $e_{ij}/kT$ . Для нахождения показателя экспоненты  $\rho_3$  теперь необходимо

решить задачу протекания с критерием связности

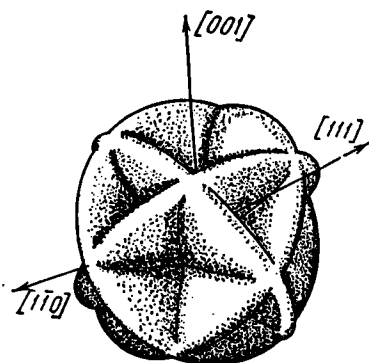
$$2 \frac{\min r_g}{a} \leq \xi. \quad (23.12)$$

Поверхность  $Q_\xi$ , соответствующая уравнению  $2 \frac{\min r_g}{a} = \xi$ , является огибающей четырех поверхностей  $Q_{\xi}^{(g)}$ , представляющих собой эллипсоиды одинаковой формы (рис. 6.5).

Каждый эллипсоид имеет ось вращения, направленную вдоль одной из диагоналей куба, и в своих главных осях описывается уравнением

$$\frac{2r_g}{a} = 2 \left( \frac{x_{ij}^2}{a^2} + \frac{y_{ij}^2}{a^2} + \frac{z_{ij}^2}{b^2} \right)^{1/2} = \xi. \quad (23.13)$$

Рис. 6.5. Четырехэллипсоидная поверхность  $Q_\xi$  для  $n$ -Ge.



Эллипсоиды сплющены вдоль оси вращения. Согласно (23.13) полуоси каждого эллипсоида равны

$$A = \frac{\xi a}{2}, \quad B = \frac{\xi b}{2}, \quad (23.14)$$

а их отношение имеет вид

$$\frac{A}{B} = \frac{a}{b} = \left( \frac{m_l}{m_t} \right)^{1/2} = 4,45. \quad (23.15)$$

Исходя из общей методики, разработанной в § 18, для вычисления порога протекания  $\xi_c$  можно вычислить объем  $V_\xi$ , заключенный внутри поверхности  $Q_\xi$ , а затем найти  $\xi_c$  с помощью условия

$$V_{\xi_c} N = B_c, \quad (23.16)$$

где  $B_c$  — пороговая величина среднего числа случайных узлов, связанных с данным узлом. Для интересующей нас четырехэллипсоидной поверхности в работе [11] было получено значение  $B_c$  несколько меньшее, чем для



сферы. С учетом экстраполяции к бесконечному массиву точек оно примерно равно 2,6.

Объем четырехэллипсоидного тела можно записать в виде  $V_{\xi} = G(a/b) V_{\text{эл}}$ , где

$$V_{\text{эл}} = \frac{4\pi}{3} A^2 B = \frac{4\pi}{3} a^2 b \left(\frac{\xi}{2}\right)^3 \quad (23.17)$$

— объем одного эллипсоида. Если бы отношение  $a/b$  было велико, то эллипсоиды напоминали бы очень тонкие диски и объемом областей, в которых они перекрываются, можно было бы пренебречь. Тогда  $G(a/b)$  было бы равно 4. С другой стороны, случаю  $a/b = 1$  соответствовало бы значение  $G(a/b) = 1$ . Функция  $G(x)$  найдена в работе [12] и имеет вид

$$G(x) = \frac{12}{\pi} \left\{ \arcsin \frac{x}{\sqrt{2+x^2}} - \arcsin \frac{\sqrt{3}x}{2\sqrt{2+x^2}} + \arcsin \frac{\sqrt{2}x}{\sqrt{1+2x^2}} - \arcsin \frac{\sqrt{3}x}{\sqrt{2}\sqrt{1+2x^2}} \right\}. \quad (23.18)$$

Для  $n$ -Ge отношение  $a/b = 4,45$  и  $G(a/b) = 2,6$ . Подставляя  $V_{\xi} = 2,6V_{\text{эл}}$  в (23.16) и используя цитированное выше значение  $B_c$ , можно найти  $\xi_c$ . Затем, подставляя  $\xi_c$  в (22.4), получим

$$\rho_3 = \rho_{03} \exp f(N_D), \quad (23.19)$$

где

$$f(N_D) = \left[ \frac{6B_c}{\pi G(a/b) N_D a^2 b} \right]^{1/3} = \frac{1,24}{N_D^{1/3} (a^2 b)^{1/3}}. \quad (23.20)$$

Результат (23.19), (23.20) можно сравнить с экспериментальными данными. На рис. 6.6 приведены зависимости  $\rho_3$  от  $N_D^{-1/3}$  для германия с примесью сурьмы (кривая 1) и для германия с примесью фосфора (кривая 2).

Проводя прямые через точки, соответствующие наиболее слабо легированным образцам, мы получаем для сурьмы

$$f(N_D) = (3,0 \pm 0,1) \cdot 10^6 N_D^{-1/3} \text{ см}^{-1} \quad (23.21)$$

и для фосфора

$$f(N_D) = (3,6 \pm 0,1) \cdot 10^6 N_D^{-1/3} \text{ см}^{-1}. \quad (23.22)$$

С другой стороны, подставляя значения энергии  $E_0 = 10,3$  и  $12,9$  мэв в (23.4), а затем полученные значения  $a$  и  $b$  в (23.20), находим для сурьмы и фосфора соответственно теоретические зависимости

$$f(N_D) = 3,0 \cdot 10^6 N_D^{-1/3} \text{ см}^{-1} \quad \text{и} \quad f(N_D) = 3,4 \cdot 10^6 N_D^{-1/3} \text{ см}^{-1}. \quad (23.23)$$

Таким образом, для слабо легированных образцов имеется хорошее согласие между экспериментальными дан-

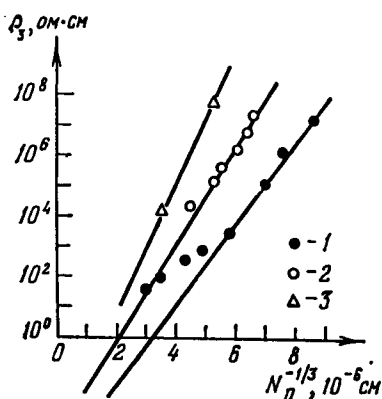


Рис. 6.6. Зависимость удельного сопротивления  $\rho_3$  от среднего расстояния между донорами  $N_D^{-1/3}$  для  $n$ -Ge: 1 — образцы Ge с Sb [13], 2 — Ge с P [14], 3 — Ge с Sb при предельном давлении [14].

ными и теоретической формулой (23.20). Для более легированных образцов наблюдается отклонение от закона (23.20). Интересно отметить, что такие отклонения являются общим свойством слабо компенсированных образцов. При промежуточной компенсации они практически не заметны (см. рис. 6.3).

До сих пор мы имели дело с весьма сложным случаем анизотропии — четырехэллипсоидным донорным состоянием. Анизотропный, но существенно более простой вид имеют волновые функции доноров в германии, подвергнутом сильному одноосному сжатию вдоль

оси [111]. Как известно [15], под действием такого сжатия энергия одного минимума зоны проводимости уменьшается, а остальных трех увеличивается. При давлениях порядка  $10^9$  дн/см<sup>2</sup> расщепление уже настолько сильно, что основное состояние электрона на доноре оказывается связанным не с четырьмя эллипсоидами, как без давления, а лишь с одним эллипсоидом. Давления, большие  $10^9$  дн/см<sup>2</sup>, не приводят к существенным новым изменениям волновой функции. Поэтому давление порядка  $10^9$  дн/см<sup>2</sup> можно назвать предельным или «максималь-

ным» давлением. При «максимальном» давлении в формулах (23.1), (23.2) и (23.7) остается одно слагаемое, соответствующее выделенному эллипсоиду. В результате мы получаем (23.10) со значением  $\xi_{ij}$ , равным

$$\xi_{ij} = 2 \left( \frac{x_{ij}^2}{a^2} + \frac{y_{ij}^2}{a^2} + \frac{z_{ij}^2}{b^2} \right)^{1/2}. \quad (23.24)$$

Здесь ось  $z$  совпадает с направлением деформации. Поверхность  $Q_\xi$ , соответствующая уравнению  $\xi_{ij} = \xi$ , в этом случае оказывается просто эллипсоидом, объем которого определяется формулой (23.17). В § 18 было доказано, что пороговая величина  $B_c$  для эллипсоида имеет то же самое значение, что и для сферы ( $B_c = 2,7$ ). Подставляя это значение  $B_c$  и формулу (23.17) в (23.16), находим  $\xi_c$  и  $f(N_D)$ :

$$f(N_D) = \xi_c = \frac{\alpha}{N_D^{1/3} (a^2 b)^{1/3}}, \quad (23.25)$$

где  $\alpha = 1,73 \pm 0,03$  — тот же самый численный коэффициент, что и в (22.9). Таким образом, переход от сферической поверхности «постоянного перекрытия» к эллипсоидальной эквивалентен замене радиуса состояния на среднее геометрическое от характерных длин спада волновой функции по трем взаимно перпендикулярным осям. Сравнивая формулы (23.20) и (23.25) для  $f(N_D)$ , мы видим, что значение числителя во втором случае заметно больше, чем в первом. Этот эффект связан, очевидно, с уменьшением объема состояния \*) за счет «отмирания» трех эллипсоидов и сам по себе должен приводить к экспоненциальному увеличению сопротивления под действием деформации.

Необходимо, однако, иметь в виду и другой эффект. Дело в том, что, благодаря уменьшению числа эллипсоидов, в 4 раза уменьшается значение волновой функции в ячейке, занятой донором. При этом в 4 раза уменьшается химический сдвиг уровня [15], т. е. энергия связи  $E_0$  должна убывать. В результате, согласно (23.4),

\*) Объемом состояния мы для краткости называем объем области пространства, в которой волновая функция не мала экспоненциально. Например, если волновая функция имеет вид (2.17), то объем состояния по порядку величины равен  $a^3$ .

возрастают характерные длины  $a$  и  $b$  и, следовательно, объем состояния. Влияние этого эффекта противоположно первому: сам по себе он должен приводить к уменьшению сопротивления.

В германии с примесью сурьмы химический сдвиг мал. Уменьшение его в 4 раза мало меняет энергию  $E_0$ . Поэтому сопротивление должно сильно возрасти под действием давления. На рис. 6.6 для двух образцов германия с примесью сурьмы,  $N_D = 6,5 \cdot 10^{15} \text{ см}^{-3}$  и  $N_D = 2,2 \cdot 10^{15} \text{ см}^{-3}$ , представлены результаты измерения удельного сопротивления  $\rho_z$  при «максимальном» одноосном сжатии [14]. Сопротивление первого образца под действием давления возросло примерно в 30 раз, а сопротивление второго — в  $10^4$  раз! Такое пьезосопротивление с полным правом можно назвать гигантским. Если через две точки провести прямую, то наклон ее оказывается в удовлетворительном согласии с формулой (23.25).

В германии с примесями фосфора или мышьяка уменьшение энергии связи должно играть более заметную роль, поскольку химический сдвиг для этих примесей существенно больше, чем для сурьмы. Однако легко убедиться в том, что даже в этом случае эффект уменьшения числа эллипсоидов должен быть сильнее, так что результирующее действие давления должно выражаться в увеличении сопротивления, хотя, конечно, и более слабо, чем для германия с сурьмой. Нам не известны работы, в которых изучалось бы пьезосопротивление германия с примесями мышьяка и фосфора в области столь же малых концентраций, как в обсуждавшихся выше образцах германия с сурьмой [14]. В области больших концентраций такие эксперименты проводились [16], причем оказалось, что сопротивление не растет, а, наоборот, падает. Это явление остается непонятным, даже если учесть, что концентрации примесей были сравнительно близкими к критической концентрации перехода металл — диэлектрик, так что существенную роль играла  $\epsilon_2$ -проводимость.

В отличие от недеформированного германия, где электропроводность является скалярной величиной, деформированный вдоль оси [111] германий теряет кубическую симметрию и становится одноосным кристаллом.

В этом случае в системе координат с осью  $z$ , совпадающей с направлением деформации, тензор удельного сопротивления должен иметь две независимые компоненты:  $\rho_{zz}$  и  $\rho_{xx} = \rho_{yy}$ . Напрашивается естественный вопрос: к какой из этих компонент относится результат (23.25)? По своему смыслу метод протекания дает одинаковый показатель экспоненты для обеих компонент тензора удельного сопротивления. Действительно, мы постоянно исходим из того, что при изменении  $\xi$  существует единственный порог протекания  $\xi_c$ , ниже которого нет никакого бесконечного кластера, а выше — имеется один бесконечный кластер, который в достаточно больших масштабах равномерно заполняет пространство. Поэтому критическая подсетка определяет ток во всех направлениях и показатель экспоненты удельного сопротивления независимо от направления внешнего поля равен  $\xi_c$ .

Если не оперировать понятиями метода протекания, то на первый взгляд может показаться, что наш вывод об изотропии экспоненциального члена парадоксален. Действительно, волновая функция донора при «максимальном» давлении имеет вид сильно сплющенного вдоль оси [111] эллипсоида вращения (отношение полуосей  $a/b = 4,45$  (см. рис. 6.7)). Естественно, казалось бы, считать, что за проводимость в некотором направлении ответственны перескоки электронов именно в этом направлении. Вероятности перескоков на одно и то же расстояние вдоль оси [111] и в перпендикулярном направлении отличаются экспоненциально. Поэтому, казалось бы, проводимость должна обладать анизотропией, причем такой сильной, что отношение величин  $\rho_{zz}/\rho_{xx}$  экспоненциально велико.

Однако, внимательно посмотрев на рис. 6.7, на котором изображены сильно вытянутые в направлении  $x$  (ось  $z$  соответствует оси сжатия) волновые функции случайно расположенных доноров, мы видим, что у фиксированного донора ближайший сосед в достаточно узком интервале углов  $(-\theta, \theta)$  с плоскостью  $(x, y)$  находится очень далеко. Следовательно, ток по оси  $x$  обязательно связан с прыжками под конечными, хотя, безусловно, малыми углами с плоскостью  $(x, y)$ . Вероятность таких перескоков экспоненциально превышает

вероятность перескоков под углами  $\theta$ , сравнимыми с  $\pi/2$ . Поэтому и для передвижения вдоль оси  $z$  электрону легче всего использовать те же наиболее вероятные, скользящие вдоль плоскости  $(x, y)$  перескоки, что и для передвижения вдоль оси  $x$  (см. рис. 6.7). При этом главный экспоненциальный член удельного сопротивления  $\exp \xi_c$  для направлений  $x$  и  $z$  окажется одинаковым. Сопротивление  $\rho_{zz}$  вдоль оси  $z$  несколько превышает  $\rho_{xx}$  за счет

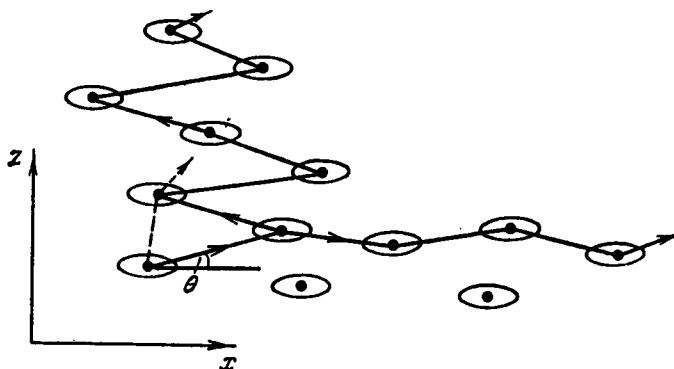


Рис. 6.7. Эллипсы представляют собой поверхности постоянной волновой функции, построенные вокруг случайно расположенных доноров. Ось  $z$  — направление сжатия  $[111]$ ; в этом направлении волновая функция убывает быстрее, чем в других. Сплошными линиями показаны пути перескоков электрона, определяющие сопротивление по осям  $z$  и  $x$ ; пунктир — путь движения «в лоб» по оси  $z$ , сопротивление которого во много раз больше.

того, что проекция типичного перескока на ось  $z$  меньше, чем на ось  $x$ . Это, однако, должно отразиться лишь на предэкспоненциальном факторе  $\rho_{0z}$ .

Сейчас мы покажем, что для случая, когда  $\xi_{ij}$  имеет вид (23.24), отношение  $\rho_{zz}/\rho_{xx}$  можно вычислить с помощью простого математического приема [17]. Предположим для начала, что величина  $R_{ij}^0$  в формуле  $R_{ij} = R_{ij}^0 \exp \xi_{ij}$  не зависит от вектора  $r_{ij}$  (в дальнейшем мы увидим, что это предположение несущественно). Не изменяя сопротивлений  $R_{ij}$ , соединяющих пары узлов сетки Миллера и Абрахамса, совершим переход из лабораторной системы координат  $(x, y, z)$  в новую систему координат  $(x', y', z')$ , которую мы назовем изотропизирующей или И-системой. И-система координат, по опре-

делению, отличается от лабораторной только изменением масштабов:

$$x = \left(\frac{a}{b}\right)^{1/3} x', \quad y = \left(\frac{a}{b}\right)^{1/3} y', \quad z = \left(\frac{b}{a}\right)^{2/3} z'. \quad (23.26)$$

Другими словами, в И-системе мы измеряем координаты узлов той же самой сетки сопротивлений с помощью новых линеек, причем линейки, используемые для измерения координат по осям  $z$  и  $x, y$ , различны.

Благодаря тому, что (23.26) сохраняет неизменным элементарный объем ( $dx dy dz = dx' dy' dz'$ ), концентрация узлов сетки в И-системе не изменяется ( $N_D = N'_D$ ). Кроме того, если в лабораторной системе распределение примесей было случайным, то преобразование (23.26) не создает никакой корреляции (см. § 18). Однако самое важное свойство И-системы состоит в том, что если сопротивления  $R_{ij}$  выразить через координаты  $x', y', z'$ , т. е. подставить (23.26) в (23.24), то оказывается, что в И-системе  $R_{ij}$  зависят только от расстояния между примесями:

$$R_{ij} = R_{ij}^0 \exp\left(\frac{2r'_{ij}}{a^*}\right), \quad a^* \equiv (a^2 b)^{1/3}. \quad (23.27)$$

Следовательно, в И-системе координат удельное сопротивление изотропно и определяется формулой, аналогичной (22.9):

$$\rho' = \rho'_0 \exp\left(\frac{\alpha}{N_D^{1/3} a^*}\right). \quad (23.28)$$

Теперь следует возвратиться в лабораторную систему координат. Для этого в лабораторной системе в сетке Миллера и Абрахамса выделим макроскопический параллелепипед со сторонами  $x, y, z$ . В И-системе его стороны соответственно равны  $x', y', z'$ . Поскольку при переходе из лабораторной системы в И-систему сопротивления  $R_{ij}$  не менялись, сопротивление выделенного параллелепипеда в обеих системах координат одинаково. В частности, для сопротивления в направлении  $z$  имеем

$$\frac{\rho_{zz} \cdot z}{xy} = \frac{\rho'_{zz} \cdot z'}{x'y'}. \quad (23.29)$$

Подставляя (23.26) в (23.29), получаем

$$\rho_{zz} = \left(\frac{a}{b}\right)^{4/3} \rho'_{zz} = \left(\frac{a}{b}\right)^{4/3} \rho'_0 \exp\left(\frac{a}{N^{1/3} D a^*}\right). \quad (23.30)$$

Аналогичным образом, инвариантность сопротивления вдоль оси  $x$  дает

$$\rho_{xx} = \left(\frac{b}{a}\right)^{2/3} \rho'_{xx} = \left(\frac{b}{a}\right)^{2/3} \rho'_0 \exp\left(\frac{a}{N^{1/3} D a^*}\right). \quad (23.31)$$

Из (23.30) и (23.31) видно, что показатели экспоненты для  $\rho_{zz}$  и  $\rho_{xx}$  одинаковы и совпадают с (23.25). Для отношения  $\rho_{zz}/\rho_{xx}$  получаем

$$\frac{\rho_{zz}}{\rho_{xx}} = \left(\frac{a}{b}\right)^2 = \frac{m_l}{m_t}. \quad (23.32)$$

Выведем теперь формулу (23.32) еще раз с помощью представления о критической подсетке сопротивлений.

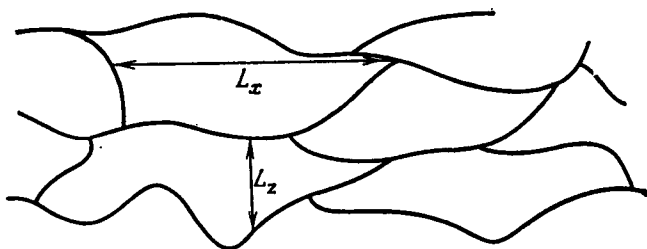


Рис. 6.8. Схематическое изображение критической подсетки сопротивлений в случае, когда в направлении  $z$  волновые функции спадают быстрее, чем в направлениях  $x$ ,  $y$ . Линии изображают цепочки сопротивлений. «Мертвые» концы не показаны.

Поскольку в  $I$ -системе критическая подсетка изотропна, в лабораторной системе координат она должна быть анизотропной (рис. 6.8). Ячейка сетки должна характеризоваться не одним радиусом корреляции, как в изотропном случае, а двумя различными длинами  $L_x$  и  $L_z$ , причем  $L_x/L_z = a/b$ .

Рассмотрим параллелепипед, у которого длины ребер, параллельных  $x$ ,  $y$  и  $z$ , равны соответственно  $L_x$ ,  $L_y$  и  $L_z$ . В любом направлении сопротивление этого параллелепипеда определяется сопротивлением одной макро-



связи, которое в свою очередь приблизительно равно величине критического сопротивления  $R_{ij}^0 \exp \xi_c$ . Таким образом, сопротивление параллелепипеда во всех направлениях в среднем одинаково. Однако при пересчете к удельному сопротивлению мы легко приходим к результату

$$\frac{\rho_{zz}}{\rho_{xx}} = \left( \frac{L_x}{L_z} \right)^2 = \left( \frac{a}{b} \right)^2.$$

Нетрудно убедиться, что учет анизотропии предэкспоненциального множителя  $R_{ij}^0$  в (22.1) не изменит наших выводов. Дело в том, что зависимость  $R_{ij}^0$  от координат практически не меняет формы поверхности  $R_{ij}(r_{ij}) = \text{const}$  и, следовательно, мало влияет на отношение длин  $L_z$  и  $L_x$ .

К сожалению, предсказание о том, что в сильно сжатом по оси [111]  $n$ -германии  $\rho_{zz}/\rho_{xx} = (m_l/m_t)^2 \approx 20$ , еще не подвергалось экспериментальной проверке.

Выше мы обсуждали влияние одноосного сжатия на  $n$ -германий. Посмотрим, что дает метод протекания для  $p$ -германия. Как известно, одноосное сжатие по оси [001] приводит к снятию вырождения валентной зоны в точке  $k = 0$ . При больших давлениях спектр двух возникающих ветвей имеет вид [18]

$$\varepsilon_1(k) = -\frac{\hbar^2}{2} \left[ \frac{k_z^2}{m_{l1}} + \frac{k_x^2 + k_y^2}{m_{t1}} \right], \quad (23.33)$$

$$\varepsilon_2(k) = -C\eta - \frac{\hbar^2}{2} \left[ \frac{k_z^2}{m_{l2}} + \frac{k_x^2 + k_y^2}{m_{t2}} \right], \quad (23.34)$$

где  $C\eta$  — пропорциональное деформации  $\eta$  расщепление в точке  $k = 0$ , а

$$\frac{m_0}{m_{l1,2}} = \gamma_1 \pm 2\gamma_2, \quad \frac{m_0}{m_{t1,2}} = \gamma_1 \mp \gamma_2. \quad (23.35)$$

Когда расщепление  $C\eta$  становится много больше энергии связи акцептора, состояние акцептора оказывается целиком связанным с верхней зоной (23.33). Спектр (23.33) имеет простой эллипсоидальный вид, и поведение волновой функции акцептора на больших расстояниях от центра описывается формулой (23.3). Если в (23.4)

подставить экспериментальное значение  $E_0$  для примеси Ga при предельно больших деформациях и вычисленные с помощью табл. 1.3 значения  $m_{11}$  и  $m_{12}$ , то получается [19], что  $a = 84 \text{ \AA}$  и  $b = 132 \text{ \AA}$ .

Выше мы вывели формулу (23.25) для показателя экспоненты  $\rho_3$  в случае примесных состояний, связанных с одним эллипсоидом. Подставляя в нее указанные значения  $a$  и  $b$ , можно найти экспоненциальную зависимость электропроводности от концентрации для сильно деформированного  $p$ -Ge с примесью Ga. Заранее можно предсказать, что электропроводность таких кристаллов должна возрасти. Действительно, результат (23.25) отличается от формулы (22.9) для случая изотропных волновых функций только тем, что роль радиуса состояния играет величина  $(a^2b)^{1/2}$ . Для рассматриваемого случая  $(a^2b)^{1/2} = 98 \text{ \AA}$ , т. е. больше, чем равный  $90 \text{ \AA}$  радиус акцепторного состояния галлия в недеформированном кристалле. Поэтому, в отличие от обсуждавшейся выше ситуации для  $n$ -германия, одноосное сжатие по оси [100] в  $p$ -германии должно привести не к увеличению, а к уменьшению сопротивления.

Прыжковая проводимость  $p$ -германия под действием деформации вдоль оси [100] исследована в работе [19]. Экстраполированные к бесконечным давлениям значения  $\rho_3$  показаны точками 2 на рис. 6.3. Видно, что, в согласии с предсказанием теории, деформация заметно уменьшила сопротивление. Наклон прямой, проведенной через эти точки, позволяет определить коэффициент  $\alpha$  в (23.25). Он оказывается равным 1,7—1,8, что хорошо согласуется с теорией и найденными выше экспериментальными значениями.

При сильной одноосной деформации анизотропия электропроводности  $p$ -германия должна определяться формулой (23.32). К сожалению, соответствующих экспериментальных данных пока нет.

## ГЛАВА 7

### ПРЫЖКОВАЯ ПРОВОДИМОСТЬ В МАГНИТНОМ ПОЛЕ

#### § 24. Сопротивление элемента $R_{ij}$ в магнитном поле

В физике полупроводников исследование магнитосопротивления является одним из наиболее широко распространенных методов. В тех случаях, когда за проводимость ответственны носители, находящиеся в разрешенных зонах, сопротивление степенным образом зависит от магнитного поля. Теория этого явления хорошо разработана и основывается на учете влияния магнитного поля с помощью кинетического уравнения или уравнения для матрицы плотности.

Характерной особенностью прыжкового механизма проводимости является гигантское положительное магнитосопротивление, экспоненциальным образом зависящее от напряженности магнитного поля. Такое магнитосопротивление наблюдалось в  $n$ -InSb [1, 2],  $p$ -InSb [3],  $n$ -Ge [4, 5],  $p$ -Ge [6—8],  $n$ -InP [9, 10] и  $n$ -GaAs [11—14]. Сладек [1] в  $n$ -InSb и Гальбо и Сладек [11] в  $n$ -GaAs в магнитных полях напряженностью 28 и 140 кэ соответственно достигли увеличения сопротивления в  $10^5$  раз. На рис. 7.1 в качестве примера приведена экспериментальная зависимость удельного сопротивления  $n$ -GaAs от напряженности магнитного поля для двух значений температуры.

Природа гигантского магнитосопротивления была выяснена уже в первой работе Сладека [1]. Как известно, сильное магнитное поле сжимает волновые функции примесных электронов в поперечном направлении,

превращая их в простейшем случае из сферически симметричных в сигарообразные. При этом перекрытие «хвостов» волновых функций соседних примесей в среднем резко уменьшается, что приводит к экспоненциальному возрастанию сопротивления. Обнаружение экспоненциального положительного магнитосопротивления практически однозначно свидетельствует о роли «хвостов» волновых функций и тем самым может быть одним из доказательств того, что в исследуемом веществе проводимость имеет прыжковый характер.

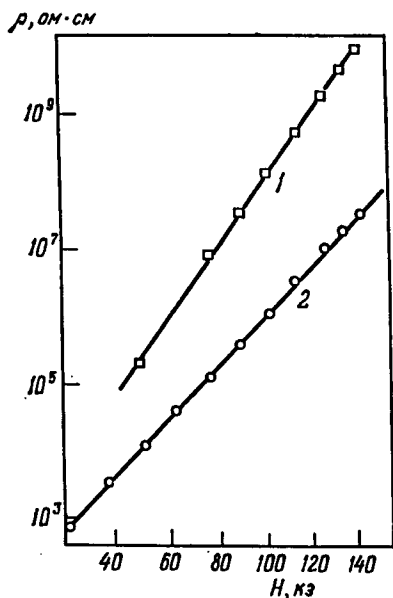


Рис. 7.1. Зависимость удельного сопротивления  $\rho$  для  $n$ -GaAs с концентрацией электронов  $n = 5 \cdot 10^{16} \text{ см}^{-3}$  от напряженности магнитного поля  $H$  [11].  $T$ , °К: 1 — 1,9; 2 — 4,2.

Важный вклад в теорию магнитосопротивления внес Микошиба [15]. Он впервые показал, что экспоненциальная магнитополевая зависимость сопротивления различна в двух областях значений магнитного поля, которые мы ниже будем называть областями слабого и сильного поля. Микошиба вычислил магнитосопротивление и в той и в другой областях.

Теория магнитосопротивления, изложенная в этой главе, разработана Шкловским [16, 17] и основана на методе протекания, который появился почти через 10 лет после работы Микошибы. Полученные ниже результаты в случае слабого поля отличаются от результатов Микошибы численным множителем в показателе экспоненты сопротивления, а в случае сильного поля приводят к иным зависимостям от магнитного поля и концентрации примесей.

Как и в предыдущей главе, мы будем интересоваться только показателем экспоненты  $\rho_3$  <sup>промежутое сопротивление</sup>. Вычисление  $\rho_3$  естественным образом разбивается на два этапа. На первом

этапе следует найти асимптотику волновых функций примеси в магнитном поле и вычислить частоту переходов между парой доноров и сопротивление элемента  $R_{ij}(r_{ij}, H)$ . Решению этой части задачи посвящен настоящий параграф. Затем следует методом протекания найти сопротивление сетки. В результате мы должны получить удельное сопротивление  $\rho_3$  как функцию магнитного поля и концентрации примесей. Эта часть задачи решается в § 25. Там же производится сравнение с экспериментальными данными.

Начнем с изучения изолированного водородоподобного атома в магнитном поле. Пусть донор помещен в центр цилиндрической системы координат  $(\rho, \varphi, z)$ , ось  $z$  которой направлена по магнитному полю  $H$ . Выберем векторный потенциал следующим образом:

$$A = \frac{1}{2} [H, r],$$

т. е.

$$A_\varphi = \frac{1}{2} H\rho, \quad A_\rho = 0, \quad A_z = 0. \quad (24.1)$$

Тогда уравнение Шредингера для огибающей волновой функции электрона  $F(r)$  в кулоновском поле донора и в магнитном поле примет вид (см. [8, стр. 525] к гл. 1)

$$-\frac{\hbar^2}{2m} \left[ \frac{1}{\rho} \frac{\partial}{\partial \rho} \left( \rho \frac{\partial F}{\partial \rho} \right) + \frac{\partial^2 F}{\partial z^2} + \frac{1}{\rho^2} \frac{\partial^2 F}{\partial \varphi^2} \right] - \\ - \frac{ie\hbar H}{2mc} \frac{\partial F}{\partial \varphi} + \frac{e^2 H^2}{8mc^2} \rho^2 F - \frac{e^2 F}{r\kappa} = EF. \quad (24.2)$$

В дальнейшем нас будет интересовать волновая функция основного состояния. Поскольку оно обладает цилиндрической симметрией, все члены, содержащие производные по  $\varphi$ , обращаются в нуль, и уравнение принимает вид

$$-\frac{\hbar^2}{2m} \left[ \frac{1}{\rho} \frac{\partial}{\partial \rho} \left( \rho \frac{\partial F}{\partial \rho} \right) + \frac{\partial^2 F}{\partial z^2} \right] + \frac{\hbar^2 \rho^2}{8m\lambda^4} F - \frac{e^2}{\kappa r} F = EF. \quad (24.3)$$

Здесь мы ввели магнитную длину

$$\lambda = \left( \frac{c\hbar}{eH} \right)^{1/2}. \quad (24.4)$$

В (24.3) влияние магнитного поля свелось к появлению дополнительной «магнитной» потенциальной энергии  $\hbar^2 \rho^2 / 8m\lambda^4$ , препятствующей уходу электрона от оси  $z$  (рис. 7.2). Она создает осцилляторную потенциальную

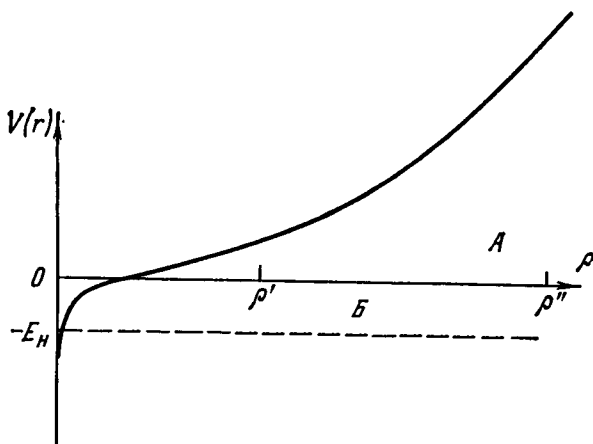


Рис. 7.2. Потенциальная энергия электрона как функция расстояния от оси  $z$  при  $z = 0$ . Пунктиром показан уровень энергии электрона. Изображен случай  $\lambda \gg a$ . Видно, что асимптотика определяется туннелированием под магнитный барьер  $A$  и барьер  $B$ , связанный с отрицательным значением энергии состояния. Вид асимптотики определяется тем, какой из двух барьеров мощнее. Если мощность барьера  $A$  больше, чем барьера  $B$  (точка  $\rho''$ ), то работает асимптотика «сильного поля» (24.13). Если, наоборот, барьер  $B$  мощнее (точка  $\rho'$ ), то работает асимптотика «слабого поля» (24.11).

яму и сама по себе, в отсутствие кулоновского потенциала, приводит к возникновению эквидистантного спектра энергий поперечного движения — уровней Ландау:

$$E_n = \left(n + \frac{1}{2}\right) \hbar \Omega,$$

где

$$\Omega = \frac{eH}{mc} = \frac{\hbar}{m\lambda^2} \quad (24.5)$$

— ларморова частота. Согласно (24.5) магнитная длина имеет смысл амплитуды нулевых колебаний свободного электрона на нулевом уровне Ландау, т. е. наименьшего расстояния, на котором магнитное поле может локализовать электрон,

Аналитическое решение уравнения (24.3) невозможно. Ниже мы займемся отысканием асимптотического поведения  $F(\rho, z)$  при больших  $\rho$  и  $z$ . Однако предварительно мы остановимся на необходимых для дальнейшего сведениях об энергии основного состояния. При этом нам будет удобно говорить о положительной величине  $E_H$  — энергии ионизации основного состояния, равной разности между энергией дна нулевого уровня Ландау и наименьшим собственным значением уравнения (24.3).

Рассмотрим сначала случай достаточно слабых полей, соответствующих неравенству  $\lambda \gg a_B$ , где  $a_B$  — радиус состояния в нулевом магнитном поле. В этом случае, как легко видеть из (24.3), на расстоянии порядка  $a_B$  от донора «магнитный потенциал» мал по сравнению с кулоновским потенциалом и энергией связи и, следовательно, не влияет на волновую функцию электрона в той области, где она велика. Поэтому магнитное поле не влияет на энергию ионизации основного состояния  $E_H$ :

$$E_H = E_0 = \frac{\hbar^2}{2ma_B^2}. \quad (24.6)$$

В противоположном предельном случае сильных магнитных полей, соответствующем неравенству  $\lambda \ll a$ , магнитное поле локализует электрон в поперечном направлении значительно сильнее, чем кулоновский потенциал. В результате, как видно из сравнения (24.5) и (24.6), кинетическая энергия поперечного движения  $1/2\hbar\Omega$  оказывается значительно больше энергии связи, которая обеспечивается кулоновским потенциалом. Это дало возможность Эллиоту и Лоудону [18] и Хасегаве и Ховарду [19] разработать метод решения уравнения (24.3) в адиабатическом по поперечному движению приближении. Они предложили искать огибающую волновой функции  $F(\rho, z)$  в виде

$$F(\rho, z) = \psi_0(\rho) F(z), \quad (24.7)$$

где

$$\psi_0(\rho) = \frac{1}{\lambda \sqrt{2\pi}} \exp\left(-\frac{\rho^2}{4\lambda^2}\right) \quad (24.7a)$$

— радиальная волновая функция электрона на нулевом уровне Ландау. Для нахождения  $F(z)$  следует подставить (24.7) в уравнение (24.3), домножить его на  $\psi_0(\rho)$

и проинтегрировать по  $\rho$ . В результате получается усредненное по поперечному движению одномерное уравнение, описывающее движение по оси  $z$ . Его решение дает

$$E_H = E_0 \ln^2 \left( \frac{a}{\lambda} \right)^2 = E_0 \ln^2 \left( \frac{H}{H_0} \right), \quad (24.8)$$

где

$$H_0 = \frac{cm^2 e^3}{\kappa^2 \hbar^3} \quad (24.9)$$

— поле, в котором  $\lambda = a$ . В промежуточной области магнитных полей, где  $\lambda \approx a$ , энергию находят с помощью прямого вариационного расчета, причем волновую функцию ищут в виде [20]

$$F(\rho, z) = [(2\pi)^{3/2} a_{\perp}^2 a_{\parallel}]^{-1/2} \exp \left[ -\frac{x^2 + y^2}{4a_{\perp}^2} - \frac{z^2}{4a_{\parallel}^2} \right], \quad (24.10)$$

где  $a_{\perp}$  и  $a_{\parallel}$  — вариационные параметры. В результате получают зависимость  $E_H$  от  $H$ , которая при  $\lambda \gg a$  ( $H \ll \ll H_0$ ) переходит в (24.6), а при  $\lambda \ll a$  ( $H \gg H_0$ ) приближается к асимптотическому закону (24.8). Установлено, что в широкой области полей  $H_0 < H < 100 H_0$  вычисленная зависимость  $E_H$  от  $H$  хорошо аппроксимируется законом  $E_H \propto H^{1/2}$ . Зависимость  $E_H$  от  $H$  многократно изучалась экспериментально. Величина  $E_H$  измерялась по энергии активации константы Холла в области температур, где проводимость связана с забросом электронов с донорных уровней в разрешенную зону. Оказалось, что экспериментальные величины находятся в хорошем согласии с теоретическими расчетами.

Перейдем теперь к обсуждению асимптотического поведения волновой функции основного состояния водородоподобного атома в магнитном поле. Поскольку ниже мы будем вычислять только показатель экспоненты электропроводности, нам достаточно найти лишь экспоненциальный множитель волновой функции. Для читателя, который не хочет следовать всем деталям расчета, мы отложим их до конца параграфа и сразу же приведем результаты и их качественную интерпретацию.

Рассмотрим сначала случай сравнительно слабых магнитных полей, для которых выполняется неравенство  $\lambda \gg a$ . Тогда на не очень больших расстояниях от оси  $z$  ( $a \ll \rho \ll \lambda^2/a$ ) волновая функция  $F(r)$  спадает по



закону

$$F(r) \propto \exp\left(-\frac{r}{a} - \frac{\rho^2 r a}{24\lambda^4}\right). \quad (24.11)$$

Первое слагаемое в скобках (24.11) имеет такой же вид, как для водородоподобного атома в отсутствие магнитного поля (см. (2.17)). Все влияние поля полностью сосредоточено во втором слагаемом, которое в пределах применимости формулы (24.11), т. е. при  $\rho \ll \lambda^2/a$ , много меньше первого.

Функция  $F(r)$  описывает амплитуду вероятности туннелирования электрона от донора в удаленную точку  $r$ , причем первое слагаемое (24.11) связано с туннелированием под почти прямоугольный барьер  $B$  (см. рис. 7.2), обусловленный отрицательным значением энергии. Второе слагаемое связано с наличием дополнительного магнитного барьера  $A$  для движения в направлении, поперечном магнитному полю. При этом расстояние  $\lambda^2/a$  имеет ясный физический смысл: когда  $\rho \approx \lambda^2/a$ , магнитная потенциальная энергия достигает величины энергии связи  $E_0$  и, следовательно, мощность магнитного барьера становится сравнимой с мощностью барьера  $B$ . В результате при  $\rho \ll \lambda^2/a$  магнитный барьер создает лишь небольшую поправку в показателе экспоненты волновой функции. Наоборот, при  $\rho \gg \lambda^2/a$  следует ожидать, что спад волновой функции в поперечном направлении будет в основном определяться магнитным полем, т. е. иметь такой же вид, как для свободного электрона в магнитном поле. Как мы покажем ниже, количественный расчет подтверждает эту точку зрения. При  $\rho \gg \lambda^2/a$  и  $\lambda \gg a$  волновая функция убывает по закону

$$F(r) \propto \exp\left(-\frac{\rho^2}{4\lambda^2} - \frac{|z|}{a}\right). \quad (24.12)$$

В поперечном направлении  $F(r)$  убывает таким же образом, как  $\psi_0(\rho)$  (см. (24.7a)). В продольном направлении асимптотика (24.12), так же как и (24.11), совпадает с не возмущенной магнитным полем. Формулу (24.11) мы будем в дальнейшем называть асимптотикой слабого поля, а формулу (24.12) — асимптотикой сильного поля. Подчеркнем, что слово «слабое» означает лишь то, что второе слагаемое в скобках (24.11) много меньше первого. Тем не менее предполагается, что второе слагае-

мое много больше единицы, так что  $F(r)$  нельзя разложить в ряд по полю. Таким образом, слабое в нашей терминологии поле приводит к экспоненциальному ослаблению «хвостов» волновых функций.

До сих пор мы говорили о ситуации, когда  $\lambda \gg a$ . По мере приближения  $\lambda$  к  $a$  с асимптотикой волновой функции происходят следующие изменения. «Окно»  $a \ll \rho \ll \ll \lambda^2/a$ , в котором работает асимптотика слабого поля (24.11), постепенно сужается и «закрывается» при  $\lambda \approx a$ . Одновременно энергия связи  $E_H$  начинает зависеть от  $H$ , в результате чего асимптотика сильного поля (24.12) несколько изменяется. Она принимает вид

$$F(r) \propto \exp\left(-\frac{\rho^2}{4\lambda^2} - \frac{|z|}{a_H}\right), \quad (24.13)$$

где

$$a_H = \frac{\hbar}{\sqrt{2mE_H}} \quad (24.14)$$

— характерная длина спада волновой функции в направлении  $z$ . При  $\lambda \gg a$ , когда  $E_H = E_0$  и, согласно (24.6),  $a_H = a = \hbar/\sqrt{2mE_0}$ , формула (24.13) переходит в (24.12). Таким образом, асимптотику сильного поля при любых  $\lambda$  можно описывать формулой (24.13).

Формула (24.13), конечно, никоим образом не противоречит тому, что хорошее совпадение с экспериментальными данными для  $E_H$  было достигнуто на основе вариационной волновой функции (24.10). Ведь энергия практически «не реагирует» на форму экспоненциального «хвоста» волновой функции, и поэтому успех использования вариационной функции (24.10) ничего не говорит о действительной асимптотике.

Следующий этап состоит в вычислении энергетического интеграла перекрытия (15.5) волновых функций двух доноров  $i$  и  $j$ . Здесь, однако, возникает следующее осложнение. Поместим донор  $i$  в начало цилиндрической системы координат и выберем вектор-потенциал в виде (24.1). Тогда волновая функция  $F_i(r)$  донора  $i$  будет определяться формулами (24.11) и (24.13). Что касается волновой функции  $F_j(r)$  донора  $j$ , то для нее эти выражения непосредственно неприменимы, поскольку вектор-потенциал  $A$  в соответствующем уравнении Шредингера

отсчитывается по-прежнему от точки  $r_i$ . Иначе говоря, в уравнении Шредингера для  $F_j$  вектор-потенциал имеет дополнительное по сравнению с уравнением для  $F_i$  постоянное слагаемое  $1/2 [Hr_{ij}]$ . Наличие такого слагаемого, как легко видеть непосредственно из уравнения Шредингера, приводит к появлению у  $F_j(r)$  дополнительного фазового множителя [21]:

$$F_j(r) = \exp \left\{ -\frac{ie}{2\hbar c} [Hr_{ij}] r \right\} F(r - r_j), \quad (24.15)$$

где функция  $F(r)$  определяется (24.11) или (24.13).

Показатели экспонент интегралов, входящих в энергетический интеграл перекрытия (15.5), легко вычисляются (см. Приложение 1). Повторяя затем ход рассуждений § 15, можно определить экспоненциальную зависимость сопротивления, включенного между узлами  $i$  и  $j$ , от вектора  $r_{ij}$  и магнитного поля.

При  $\lambda \gg a$  и  $(x_{ij}^2 + y_{ij}^2)^{1/2} \ll \lambda^2/a$  эта зависимость имеет вид  $R_{ij} = R_{ij}^0 \exp \xi_{ij}$ , где

$$\xi_{ij} = \frac{2r_{ij}}{a} + \frac{(x_{ij}^2 + y_{ij}^2) ar_{ij}}{12\lambda^4}. \quad (24.16)$$

С другой стороны, при  $\lambda \gg a$  и  $(x_{ij}^2 + y_{ij}^2)^{1/2} \gg \lambda^2/a$  или при  $\lambda \ll a$  и любых  $r_{ij}$ , для которых  $R_{ij}$  экспоненциально велико, можно получить

$$\xi_{ij} = \frac{x_{ij}^2 + y_{ij}^2}{2\lambda^2} + \frac{2|z_{ij}|}{a_H}. \quad (24.17)$$

Теперь у нас есть все, что необходимо для того, чтобы применить метод протекания и вычислить показатель экспоненты прыжковой проводимости. Мы сделаем это в следующем параграфе. Конец настоящего параграфа мы посвятим математическому выводу формул (24.11) и (24.13).

Для нахождения показателя экспоненты волновой функции на больших расстояниях можно пренебречь кулоновским потенциалом  $e^2/\kappa r$ , практически не влияющим на мощность барьера, под который туннелирует электрон. Экспоненциальная асимптотика  $F(r)$  определяется лишь энергией электрона и величиной магнитного

поля и, следовательно, имеет такой же вид, как для произвольного короткодействующего потенциала, создающего уровень с той же самой энергией  $E$ . Выбрав радиус короткодействующего потенциала достаточно малым, мы приходим к выводу, что экспоненциальный множитель волновой функции на больших расстояниях совпадает с экспоненциальным множителем функции Грина электрона  $G_E(r, r')$  при  $r' = 0$  и  $E < 0$ , представляющей собой решение уравнения (24.3), у которого кулоновское слагаемое заменено на  $\delta$ -функцию от  $r$ . Здесь полезно заметить, что при  $H = 0$  функция Грина  $G_E(r, 0) \propto \propto r^{-1} \exp(-r/a)$ , а согласно (2.17), волновая функция основного состояния  $F(r) \propto \exp(-r/a)$ . Таким образом, показатели экспонент обеих величин, в согласии со сказанным выше, совпадают.

Функция Грина  $G_E(r, 0)$  может быть записана в виде

$$G_E(r, 0) = \sum_{n=0}^{\infty} \int_{-\infty}^{\infty} \frac{dp_z}{2\pi} \frac{\psi_{np_z}(\rho, z) \psi_{np_z}(0, 0)}{\hbar\Omega \left(n + \frac{1}{2}\right) + \frac{p_z^2}{2m} - E}, \quad (24.18)$$

где  $\psi_{np_z}(\rho, z)$  — собственные функции электрона в магнитном поле, соответствующие состояниям с нулевым моментом количества движения относительно оси  $z$ :

$$\psi_{np_z} = \frac{1}{\lambda \sqrt{2\pi}} L_n \left( \frac{\rho^2}{2\lambda^2} \right) \exp \left( -\frac{\rho^2}{4\lambda^2} \right) \exp \left( -\frac{i}{\hbar} p_z z \right), \quad (24.19)$$

$L_n(x) = e^x \frac{d^n}{dx^n} (e^{-x} x^n) \frac{1}{n!}$  sic!  $\Rightarrow$  (24.21)  $\mathcal{L}_n \frac{1}{n!}$  — полиномы Лагерра. В (24.18) отсутствуют функции, соответствующие ненулевым значениям момента количества движения, поскольку эти функции равны нулю в начале координат. Используя (24.19) и вводя  $E_H = = 1/2 \hbar \Omega - E$ , перепишем (24.18) в виде

$$G_E(r, 0) \propto \exp \left( -\frac{\rho^2}{4\lambda^2} \right) \int_{-\infty}^{\infty} dp_z \int_0^{\infty} dt \exp \left( -\frac{i}{\hbar} p_z z - \frac{p_z^2}{2m} t - E_H t \right) \times \times \sum_{n=0}^{\infty} \exp(-n\hbar\Omega t) L_n \left( \frac{\rho^2}{2\lambda^2} \right). \quad (24.20)$$

Для вычисления суммы можно использовать производящую функцию полиномов Лагерра

$$\frac{1}{1-x} \exp\left(\frac{x\xi}{x-1}\right) = \sum_{n=0}^{\infty} L_n(\xi) x^n \quad \text{при } |x| < 1. \quad (24.21)$$

В нашем случае  $x = \exp(-\hbar\Omega t)$  и  $\xi = \rho^2/2\lambda$ , так что

$$\begin{aligned} \sum_{n=0}^{\infty} \exp(-n\hbar\Omega t) L_n\left(\frac{\rho^2}{2\lambda^2}\right) &= \\ &= \exp\left(\frac{\rho^2/2\lambda^2}{1 - e^{-\hbar\Omega t}}\right) [1 - \exp(-\hbar\Omega t)]^{-1}. \end{aligned} \quad (24.22)$$

Подставляя (24.22) в (24.20) и вычисляя интеграл по  $\rho_z$ , получим

$$\begin{aligned} G_E(r, 0) &\propto \exp\left(-\frac{\rho^2}{4\lambda^2}\right) \int_0^{\infty} \frac{dt}{\sqrt{t}} \exp\left\{-\frac{mz^2}{2t\hbar^2} - E\hbar t\right\} \times \\ &\times \exp\left\{\frac{\rho^2}{2\lambda^2} [1 - \exp(\hbar\Omega t)]^{-1}\right\} [1 - \exp(-\hbar\Omega t)]^{-1}. \end{aligned} \quad (24.23)$$

Дальнейшее вычисление интеграла по  $t$  должно вестись методом перевала. Введем переменную  $\eta \equiv \hbar\Omega t$ . Тогда необходимо в зависимости от существенных в интеграле (24.23) значений  $t \approx \bar{t}$  рассмотреть два случая: 1)  $\bar{\eta} \equiv \hbar\Omega \bar{t} \ll 1$  и 2)  $\bar{\eta} \gg 1$ . Первый случай, как мы увидим ниже, возможен лишь при  $\lambda \gg a$ . При  $\eta \ll 1$  разложим (24.23) по  $\eta$  и сохраним члены порядка  $\eta$ ,  $\eta^2$  и  $\eta^3$  в первой квадратной скобке (24.23). Тогда, с учетом (24.5) и (24.6), получаем

$$G_E(r, 0) \propto \int_0^{\infty} \eta^{-3/2} \exp\left\{-\frac{r^2}{2\lambda^2\eta} - \frac{\lambda^2\eta}{2a^2} - \frac{\rho^2\eta}{24\lambda^2}\right\} d\eta. \quad (24.24)$$

Показатель экспоненты (24.24) имеет максимум при  $\eta = \bar{\eta} = ra/\lambda^2$ . Подставляя  $\bar{\eta}$  в фигурную скобку, получаем, что экспоненциальный множитель функции  $G(r, 0)$  и, следовательно, функции  $F(r)$  имеет вид (24.11). Из условия  $\bar{\eta} \ll 1$  и условия экспоненциальной малости (24.11) следует, что эта формула применима лишь при  $\lambda \gg a$  и на не очень больших расстояниях  $a \ll r \ll \lambda^2/a$ .

Вычислим теперь интеграл (24.23) в другом предельном случае:  $\bar{\eta} = \hbar\Omega\bar{t} \gg 1$ . В этом случае имеем

$$G_E(r, 0) \propto \exp\left(-\frac{\rho^2}{4\lambda^2}\right) \int_0^\infty \eta^{-1/2} \times \\ \times \exp\left(-\frac{z^2}{2\lambda^2\eta} - \frac{\lambda^2\eta}{2a_H^2} - \exp(-\eta) \frac{\rho^2}{2\lambda^2}\right) d\eta, \quad (24.25)$$

где  $a_H$  определяется (24.14). Показатель экспоненты (24.25) имеет максимум при  $\eta = \bar{\eta} = |z| a_H / \lambda^2$ . Подставляя  $\bar{\eta}$  в подынтегральное выражение (24.25), мы получаем, что экспоненциальный множитель функции  $G_E(r, 0)$  и, следовательно, функция  $F(r)$  имеют вид (24.13).

### § 25. Вычисление магнитосопротивления и обсуждение экспериментальных данных

Согласно общему рецепту (22.4), для вычисления показателя экспоненты удельного сопротивления  $\rho_z$  нужно найти порог протекания  $\xi_c$  в задаче на случайных узлах с условием связности  $\xi_{ij} \leq \xi$ . Сделаем это сначала для случая «слабого поля». В этом случае уравнение  $\xi_{ij} = \xi$  поверхности  $Q_\xi$ , согласно (24.16), имеет вид

$$\frac{2r}{a} + \frac{(x^2 + y^2)ra}{12\lambda^4} = \xi. \quad (25.1)$$

Как мы уже подчеркивали, в условиях «слабого поля» второе слагаемое в левой части (25.1), описывающее влияние магнитного поля и согласно (24.4) пропорциональное  $H^2$ , мало по сравнению с первым по параметру  $\delta = (ra/\lambda^2)^2 \ll 1$ . Поэтому поверхность  $Q_\xi$  мало отличается от сферы с радиусом  $a\xi/2$ . Покажем, что в первом порядке по параметру  $\delta$  уравнение (25.1) совпадает с уравнением эллипсоида. Для этого возведем обе стороны (25.1) в квадрат и, отбрасывая квадратичный по  $\delta$  член, получим

$$\frac{4r^2}{a^2} + \frac{4r^2(x^2 + y^2)}{12\lambda^4} = \xi^2. \quad (25.2)$$

Поскольку второе слагаемое в левой части (25.2) пропорционально  $\delta$ , мы можем заменить в нем  $r^2$  на  $\left(\frac{a\xi}{2}\right)^2$ .

Это вносит в (25.2) несущественную погрешность порядка  $\delta^2$ . В результате мы получаем уравнение эллипсоида

$$\frac{z^2}{A^2} + \frac{x^2 + y^2}{C^2} = 1, \quad (25.3)$$

где

$$A = \frac{a\xi}{2} \quad \text{и} \quad C = \frac{a\xi}{2} \left( 1 - \frac{a^4\xi^2}{96\lambda^4} \right). \quad (25.4)$$

Объем этого эллипсоида равен

$$V_\xi = \frac{4\pi}{3} AC^2 = \frac{4\pi}{3} \left( \frac{a\xi}{2} \right)^3 \left[ 1 - \frac{a^4\xi^2}{96\lambda^4} \right]^2. \quad (25.5)$$

Для нахождения  $\xi_c$  теперь надо приравнять среднее число связей на один узел  $B = V_\xi N$  критическому значению этого параметра  $B_c$ :

$$\frac{4\pi}{3} N \left( \frac{a\xi_c}{2} \right)^3 \left[ 1 - \frac{a^4\xi_c^2}{96\lambda^4} \right]^2 = B_c. \quad (25.6)$$

Решая (25.6) с помощью разложения по малому параметру, стоящему рядом с единицей в квадратных скобках, получим

$$\xi_c = \frac{\alpha}{N^{1/3}a} + \frac{B_c a}{24\pi N \lambda^4}, \quad (25.7)$$

где  $\alpha = 2(3B_c/4\pi)^{1/3}$ . Напомним, что, как показано в § 18, для эллипсоидов критическая величина  $B_c$  имеет то же самое значение, что и для сфер ( $B_c = 2,7$ ). Поэтому в (25.7)  $\alpha = 1,73$  и первое слагаемое (25.7) совпадает со значением  $\xi_c$  при  $H = 0$ . Второе слагаемое описывает относительно малую поправку, связанную с магнитным полем. Это слагаемое тем не менее может быть больше единицы и поэтому приводит к экспоненциальной зависимости электропроводности от магнитного поля. Подставляя (24.4) и значение  $B_c = 2,7$  в уравнение (25.7), получим

$$\frac{\rho_3(H)}{\rho_3(0)} = \exp \left( t \frac{ae^2}{Nc^2\hbar^2} H^2 \right), \quad (25.8)$$

где

$$t \equiv B_c/24\pi = 0,036. \quad (25.8a)$$

Значение (25.8a) для  $t$  было найдено в работе [17]. Ранее в работе [22] был получен иной результат:  $t = B_c/16\pi = 0,06$ . Для его вывода использовались

предположения о том, что слабое магнитное поле не изменяет путей протекания и вклад в магнитосопротивление дают только пары примесей с векторами  $r_{ij}$ , перпендикулярными  $H$ . Эти предположения, на наш взгляд, не соответствуют действительности.

Поскольку в слабом поле прыжки происходят так же, как и в отсутствие поля, на расстояние порядка  $N^{-1/3}$ , условия применимости (25.8) имеют вид

$$\lambda \gg a \quad \text{и} \quad \frac{\lambda^2}{a} \gg N^{-1/3} \gg a. \quad (25.9)$$

Эти условия могут выполняться только в слабо легированном полупроводнике, в котором  $Na^3 \ll 1$ , и в не очень сильном магнитном поле. Учитывая, что  $Na^3 \ll 1$ , получим, что в (25.9) второе из ограничений на  $\lambda$  ( $\lambda > a/(Na^3)^{1/6}$ ) является более жестким. Оно по порядку величины определяет характерное поле  $H_c$  перехода от случая слабого поля к сильному:

$$H_c = \frac{N^{1/3} c \hbar}{ae}. \quad (25.10)$$

Вычислим теперь  $\rho_z$  для случая сильного поля. В этом случае уравнение  $\xi_{ij} = \xi$  поверхности  $Q_\xi$ , согласно (24.17), имеет вид

$$\frac{2|z|}{a_H} + \frac{x^2 + y^2}{2\lambda^2} = \xi. \quad (25.11)$$

Оно определяет поверхность, состоящую из двух сложенных основаниями усеченных параболоидов вращения с осью, направленной вдоль  $H$ . Мы назовем эту поверхность двойным параболоидом (рис. 7.3). Высота  $h$  и радиус основания  $b$  каждого параболоида соответственно равны:

$$h = \frac{a_H}{2} \xi, \quad b = \lambda \sqrt{2\xi}. \quad (25.12)$$

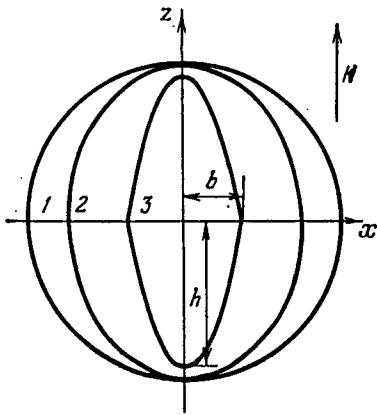


Рис. 7.3. Эволюция поверхности  $Q_\xi$  постоянного показателя экспоненты сопротивления  $R_{ij}$  при увеличении магнитного поля: 1 —  $H = 0$ ; 2 — слабое; 3 — сильное магнитное поле.



Нетрудно вычислить и объем всего двойного параболоида:

$$V_{\xi} = \pi \lambda^2 a_H \xi^2. \quad (25.13)$$

Условие  $V_{\xi} N = B_c$  для определения  $\xi_c$  имеет вид

$$\pi \lambda^2 a_H N \xi_c^2 = B_c. \quad (25.14)$$

Решая (25.14) относительно  $\xi_c$ , получаем

$$\xi_c = \frac{q}{(N a_H \lambda^2)^{1/2}}, \quad (25.15)$$

где

$$q = \left( \frac{B_c}{\pi} \right)^{1/2}. \quad (25.16)$$

Используя (24.4) и (24.14), получим окончательный результат для сильного поля:

$$\rho_3 = \rho_0 \exp \left\{ \frac{q}{(N a_H \lambda^2)^{1/2}} \right\} = \rho_0 \exp \left\{ q \frac{2^{1/4} m^{1/4} e^{1/2}}{c^{1/2} \hbar} \frac{H^{1/2} E_H^{1/4}}{N^{1/2}} \right\}. \quad (25.17)$$

В § 18 говорилось, что с помощью численных методов не удастся заметить отличие значения  $B_c$  для двойного параболоида от значения для сферы. Подставляя  $B_c = 2,7$  в (25.16), мы получаем  $q = 0,92$ .

Формула сильного поля (25.17) в слабо легированном полупроводнике работает при  $H \gg H_c$ . Поле  $H_c$ , определенное согласно (25.10), меньше поля  $H_0$ , при котором  $\lambda = a$  (см. (24.9)). Поэтому в область применимости (25.17) включаются оба случая,  $\lambda \gg a$  и  $\lambda \ll a$ , в которых  $E_H$  по-разному зависит от поля. Как мы говорили в § 24, при  $\lambda \gg a$  величина  $E_H$  не зависит от  $H$  и равна своему невозмущенному значению  $E_0$ . Таким образом, формула (25.17) при  $\lambda \gg a$  ( $H \ll H_0$ ) дает

$$\rho_3(H) = \rho_0 \exp \{ \text{const } H^{1/2} \}. \quad (25.18)$$

При  $\lambda \ll a$  ( $H \gg H_0$ ) для  $E_H$  можно использовать формулу (24.8). В результате получаем, что при  $H \rightarrow \infty$

$$\rho_3 = \rho_0 \exp \left\{ \text{const } H^{1/2} \ln^{1/2} \left( \frac{H}{H_0} \right) \right\}. \quad (25.19)$$

В довольно широком промежуточном интервале полей  $H_0 < H < 100H_0$  в (25.17) следует подставлять величины

$E_H$ , найденные численным расчетом [20], или же значения, извлеченные из экспериментов по эффекту Холла в области температур, где проводимость определяется электронами, активированными в зону. В § 24 мы уже говорили, что в этом интервале и рассчитанные и экспериментальные значения  $E_H$  пропорциональны  $H^{1/2}$ . Подстановка такой зависимости в (25.17) дает  $\rho_3 = \rho_0 \exp \{ \text{const } H^{7/12} \}$ . Таким образом, можно считать, что при  $H \gg H_c$  зависимость от магнитного поля всегда имеет вид

$$\rho_3 = \rho_0 \exp \{ \text{const } H^m \}, \quad (25.20)$$

где показатель  $m$  немного превосходит 0,5.

Рассмотрим несколько подробнее переход от слабого магнитного поля к сильному, т. е. от формулы (25.8) к формуле (25.17). Для этого разделим (25.7) и (25.15) на  $\xi_c^0 = \alpha/N_D^{1/3}a$  — значение  $\xi_c$ , соответствующее  $H = 0$  (см. (22.9)). В результате получим

$$\frac{\xi_c}{\xi_c^0} = \begin{cases} 1 + \frac{t}{\alpha} \left( \frac{H}{H_c} \right)^2 & (H \ll H_c), \\ \frac{q}{\alpha} \left( \frac{H}{H_c} \right)^{1/2} & (H \gg H_c), \end{cases} \quad (25.21)$$

где  $H_c$  определено с помощью (25.10) и мы пренебрегли отличием  $a_H$  от  $a$ . Асимптоты (25.21) изображены на рис. 7.4 штриховыми линиями. Верхняя соответствует слабому полю, нижняя — сильному. К сожалению, универсальная зависимость  $\frac{\xi_c}{\xi_c^0} \left( \frac{H}{H_c} \right)$ , описывающая всю область изменения  $H/H_c$ , пока не найдена. Сплошная линия представляет интерполяцию «на глаз» между двумя предельными случаями.

Видно, что область применимости формулы (25.8) весьма ограничена. Практически формула (25.8) может описывать увеличение  $\xi_c$  не более чем на 10%. Поскольку  $\xi_c^0$  обычно порядка 20, формула (25.8) должна работать до значений  $\rho_3(H)/\rho_3(0)$  порядка 10. Ясно также, что между областями применимости (25.8) и (25.17) имеется протяженная переходная область, в которой, в частности, может быть значительный участок с зависимостью  $\xi_c$  от  $H$ , близкой к линейной. Наконец,

с помощью рис. 7.4 можно приближенно оценить значение магнитного поля  $H_{1/2}$ , при котором величина  $\xi_c - \xi_c^0 = \ln \frac{\rho_3(H)}{\rho_3(0)}$  оказывается в два раза меньше, чем следует из (25.8) (соответствующая точка указана стрелкой на оси абсцисс):

$$H_{1/2} \approx 4H_c = 4 \frac{N^{1/3} c \hbar}{ae}. \quad (25.22)$$

Как видно из рис. 7.4, использование поля  $H_{1/2}$  может дать более удобный критерий для разделения областей слабого и сильного поля, чем использование поля  $H_c$ .

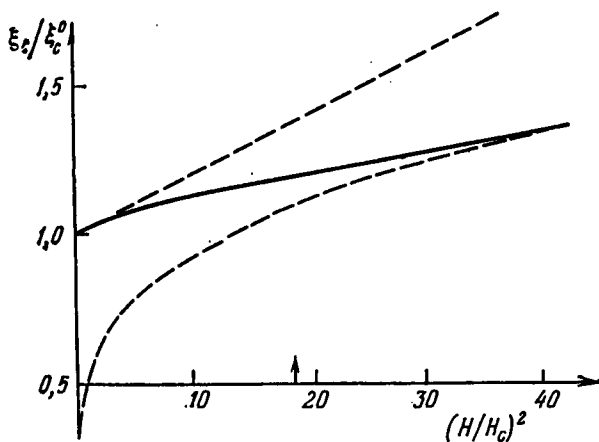


Рис. 7.4. Зависимость  $\xi_c$  от безразмерного магнитного поля  $H/H_c$

При выводе формул (25.8) и (25.17) нигде не говорилось, о каком направлении тока и о каком магнитосопротивлении — продольном или поперечном — идет речь. В § 23 в связи с экспериментами по одноосному сжатию германия мы подробно объяснили, что даже при сильной анизотропии показателя экспоненты волновой функции главный экспоненциальный член удельного сопротивления  $\rho_3$  не зависит от направления тока. Таким образом, формулы (25.8), (25.17) равным образом применимы для продольного и поперечного магнитосопротивлений.

Что касается предэкспоненциального множителя, то его различие для  $\rho_{\perp}$  и  $\rho_{\parallel}$  связано с геометрической ани-

зотропией критической подсетки, т. е. с наличием у нее двух характерных длин (см. § 23). Анизотропия критической подсетки в свою очередь обусловлена тем, что продольные и поперечные магнитному полю проекции векторов  $r_{ij}$  принадлежащих ей сопротивлений в среднем не равны друг другу. В случае слабого поля векторы  $r_{ij}$  сопротивлений критической подсетки заканчиваются внутри поверхности  $Q_{\xi_c}$ , которая представляет собой эллипсоид, мало отличающийся от сферы с радиусом  $r_c$  (см. (25.3), (25.4)). Поэтому в слабом поле относительная анизотропия  $(\rho_{\perp} - \rho_{\parallel})/\rho_{\parallel}$  должна быть малой по сравнению с единицей. В случае сильного поля векторы  $r_{ij}$  сопротивлений критической подсетки заканчиваются внутри двойного параболоида.

Подставляя (25.15) в (25.12), легко показать, что при выполнении условий сильного поля  $\lambda^2/a \ll N^{-1/2}$  высота  $h_c$  и радиус основания  $b_c$  этого параболоида удовлетворяют неравенству

$$h_c \gg N^{-1/3} \gg b_c. \quad (25.23)$$

Таким образом, в сильном поле основная часть прыжков происходит под малыми углами к магнитному полю на расстояния, много большие среднего расстояния между примесями. Локальный характер оптимальных путей и структуру критической подсетки в целом в сильном поле можно представить себе с помощью рис. 6.7 и 6.8 соответственно, если вообразить, что магнитное поле направлено по оси  $x$ . Ясно, что сильная геометрическая анизотропия критической подсетки должна приводить к существенной анизотропии предэкспоненциального множителя сопротивления. Мы ограничимся здесь лишь приведенными выше качественными соображениями.

Количественные оценки анизотропии сопротивления в магнитном поле, использующие некоторые дополнительные приближения, можно найти в работе одного из авторов ([17] к гл. 6).

Сравним формулы (25.8) и (25.17) с результатами Микошибы [15]. В случае слабого поля Микошиба получил формулу (25.8) с коэффициентом  $t = 1/12 = 0,083$  вместо 0,036. В случае сильного поля Микошиба предположил, что магнитосопротивление определяется пере-

скоками на расстояние  $N^{-1/2}$  в направлении, поперечном полю. Это дает

$$\rho_3(H) = \rho_0 \exp \left\{ \frac{(N^{-1/2})^2}{4\lambda^2} \right\} = \rho_0 \exp \left\{ \frac{eH}{4N^{2/3}c\hbar} \right\}. \quad (25.24)$$

При  $H \gg H_c$  формула (25.24) приводит к сопротивлению существенно большему, чем (25.17). Это не удивительно, так как мы уже говорили, что в сильных полях прыжки под малыми углами к магнитному полю на расстояния, большие  $N^{-1/2}$ , являются более выгодными, чем прыжки под большими углами на среднее расстояние.

Приступим теперь к сравнению выводов теории с экспериментальными данными. Формула (25.8) справедлива при малых концентрациях доноров и в не очень сильных полях. Именно такие условия имеют место в наименее легированных образцах из исследованных в работах [3, 5, 7, 8, 10, 12—14] образцов Ge, InSb, InP и GaAs. Строго говоря, для сравнения теории и эксперимента нужно с помощью экстраполяции экспериментальных зависимостей удельного сопротивления

$$\rho(H, T) = \rho_3(H) \exp \left\{ \frac{e_3(H)}{kT} \right\} \quad (25.25)$$

к  $T^{-1} = 0$  найти отношение  $\frac{\rho_3(H)}{\rho_3(0)}$ . В некоторых работах данные представлены именно в таком виде. В других работах приводятся графики  $\frac{\rho(H)}{\rho(0)}$ . Однако, как мы увидим в § 26, в достаточно слабо легированных образцах энергия активации  $e_3$  практически не зависит от магнитного поля. Поэтому  $\frac{\rho_3(H)}{\rho_3(0)} = \frac{\rho(H)}{\rho(0)}$  и, следовательно, сравнение с теорией можно производить и прямо по графикам  $\frac{\rho(H)}{\rho(0)}$ .

Во всех цитированных работах было обнаружено, что относительное магнитосопротивление  $\frac{\rho_3(H)}{\rho_3(0)}$  или  $\frac{\rho(H)}{\rho(0)}$ , в соответствии с (25.8), растет с магнитным полем по закону

$$\frac{\rho_3(H)}{\rho_3(0)} = \exp(AH^2), \quad (25.26)$$

где  $A$  — постоянная величина.

На рис. 7.5 в качестве примера приведена зависимость  $\frac{\rho(H)}{\rho(0)}$  от  $H^2$  для образца  $n$ -GaAs с концентрацией доноров  $N_D = 2,1 \cdot 10^{15} \text{ см}^{-3}$  [12]. Обычно, как и в этом примере, закон (25.26) справедлив до значений  $\frac{\rho(H)}{\rho(0)}$

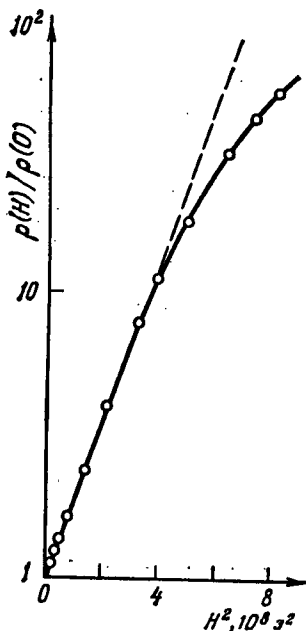


Рис. 7.5. Относительное магнитосопротивление слабо легированного  $n$ -GaAs при  $T = 2,6 \text{ }^\circ\text{K}$  [12].

порядка нескольких десятков, что разумно согласуется со сделанной выше оценкой области применимости (25.8). Предсказываемая формулой (25.8) зависимость от концентрации примесей  $A \propto N^{-1}$ , по-видимому, также всегда выполняется. Поэтому для тех случаев, когда волновые функции точно или приближенно изотропны, используя значения радиусов состояний, приведенные в гл. 6, можно найти экспериментальные значения коэффициента  $t$  в формуле (25.8). Эти значения собраны в табл. 7.1.

Рассмотрим сначала данные для кристаллов  $A_3B_5$   $n$ -типа. Видно, что для GaAs данные различных авторов заметно отличаются друг от друга. Однако в большинстве случаев экспериментальные значения  $t$  не слишком сильно отличаются от теоретического  $t_{\text{теор}} = 0,036$ .

С другой стороны, из табл. 7.1 видно, что для  $p$ -Ge и  $p$ -InSb \*) имеется сильное (в три раза) расхождение значений с теорией. В связи с этим подчеркнем еще раз, что при нахождении  $t$  для образцов  $p$ -типа мы, согласно формуле (4.20), использовали значения радиуса состояния  $a = \hbar / (2m_t E_0)^{1/2}$ , определяющиеся легкой дыркой. В то время как концентрационная зависимость прыжковой проводимости находится в разумном согла-

\*) Для анализа данных по  $p$ -InSb использовался рис. 3 из работы [3]. На рис. 4 этой работы искажен масштаб.

Таблица 7.1

Экспериментальные значения коэффициента  $t$  в формуле (25.8)

Вещество	$t_{\text{эксп}}$	Вещество	$t_{\text{эксп}}$
<i>n</i> -GaAs [12]	0,036—0,046	<i>p</i> -Ge [7]	0,013 <sup>*</sup> )
<i>n</i> -GaAs [13]	0,06	<i>p</i> -Ge [8]	0,01 <sup>*</sup> )
<i>n</i> -GaAs [14]	0,026—0,032	<i>p</i> -InSb [3]	0,01 <sup>*</sup> )
<i>n</i> -InP [10]	0,04		

<sup>\*</sup>) Значения, найденные нами для наименее легированных образцов непосредственно по графическим данным, приведенным в оригинальных статьях.

сии с таким значением  $a$  (см. § 22), магнитосопротивление оказывается аномально слабым. Логично считать, что эта аномалия как-то связана со структурой акцепторного состояния. Эта аномалия не нашла еще объяснения, несмотря на постановку специального исследования [23].

В работе [12] на слабо легированном *n*-GaAs исследовался также вопрос об отличии величин продольного и поперечного магнитосопротивления в слабом магнитном поле. Оказалось, что сопротивление вдоль поля  $\rho_{\parallel}$  лишь на 8% меньше сопротивления поперек поля, в то время как каждое из них возросло в 4 раза. Этот результат находится в согласии с теоретическим предсказанием метода протекания о незначительном различии  $\rho_{\parallel}$  и  $\rho_{\perp}$ .

Область сильных полей  $H \gg H_{1/2}$ , где, согласно теории, должен работать закон (25.17), менее исследована, чем область слабых полей. Кахлерт и др. [13] измеряли магнитосопротивление *n*-GaAs в полях до 60 кэ. Даже для самого слабо легированного образца ( $N_D = 1,35 \cdot 10^{15} \text{ см}^{-3}$ ) формула (25.22) дает  $H_{1/2} \approx 27 \text{ кэ}$ . Таким образом, в этих опытах не приходится рассчитывать на выполнение сильного неравенства  $H \gg H_{1/2}$  и на наблюдение закона (25.17) в чистом виде. Тем не менее в области полей  $30 < H < 60 \text{ кэ}$  авторы добились хорошего согласия наблюдаемой зависимости магнитосопротивления от поля с формулой (25.17) с помощью единственного подгоночного параметра  $q$  (значения  $E_H$  они

брали из расчета Ларсена [24]). Наилучшее описание экспериментальных кривых было достигнуто при  $q = 0,7$ , что несколько меньше теоретического значения  $q = 0,92$ .

В более широком интервале полей, до 140 кэ, арсенид галлия исследовался в работе Гальбо и Сладека [11]. На рис. 7.1 приведены заимствованные из [11] экспериментальные зависимости удельного сопротивления от магнитного поля. Анализ показывает [22], что эти зависимости с хорошей точностью можно описать формулой

$$\rho \propto \exp(\text{const } H^m), \quad \text{где } m \approx 0,6. \quad (25.27)$$

Однако из формулы (25.27), строго говоря, нельзя сделать окончательный вывод о зависимости  $\rho_3(H)$ . Дело в том, что исследовавшиеся в [11] образцы  $n\text{-GaAs}$  с  $N_D \approx 10^{16} \text{ см}^{-3}$  уже настолько сильно легированы, что перекрытие волновых функций существенно уменьшает величину  $\epsilon_3$ . Об этом явлении уже говорилось в § 14. Более подробно о нем пойдет речь в гл. 8. Магнитное поле сжимает волновые функции, ослабляет перекрытие и увеличивает  $\epsilon_3$ . Это приводит к дополнительному очень существенному вкладу в магнитосопротивление, количественная теория которого пока отсутствует. Важная роль зависимости  $\epsilon_3(H)$  видна непосредственно на рис. 7.1. С ней связано дополнительное увеличение  $\rho$  при  $T = 1,9^\circ\text{К}$ . Таким образом, для проверки закона (25.17) большой интерес представляло бы исследование менее легированных, чем в [11], образцов в столь же сильных магнитных полях.

Интересно отметить, что экспоненциальное магнитосопротивление должно иметь место не только в слабо легированном ( $Na^3 \ll 1$ ), но и в сравнительно сильно легированном полупроводнике ( $Na^3 \gtrsim 1$ ). Действительно, магнитное поле, сжимая волновые функции, должно привести к тому, что полупроводник эффективно станет слабо легированным и металлическая проводимость сменится прыжковой проводимостью. Это явление называют магнитным вымораживанием электронов. Вымораживание происходит тогда, когда объем состояния электрона на изолированном доноре становится меньше, чем объем  $N^{-1}$ , приходящийся на один донор. Магнитное поле существенно изменяет объем состояния только при  $\lambda \ll a$ , когда состояние имеет сигарообразную форму с попереч-



ным размером  $\lambda$  и продольным  $a_H$  (см. формулу сильного поля (24.13)). Таким образом, условие образования локализованных состояний в магнитном поле имеет вид  $Na_H\lambda^2 \approx 1$ . При дальнейшем увеличении поля сопротивление будет возрастать за счет уменьшения перекрытия «хвостов» волновых функций. Поскольку в этой области полей заведомо выполняется неравенство  $\lambda \ll a$ , с самого начала возрастание сопротивления должно происходить по закону сильного поля (25.17). Заметим, что при выполнении условия  $Na_H\lambda^2 \approx 1$  показатель экспоненты в (25.17) становится порядка единицы, что свидетельствует о переходе к металлической проводимости.

Срыв металлической проводимости и экспоненциальное уменьшение прыжковой проводимости с ростом магнитного поля неоднократно наблюдались в  $n$ -InSb [1, 25—27]. Работа Сладека [1] дает возможность проверить вывод теории об изотропии показателя экспоненты сопротивления в сильном магнитном поле. При наибольшем значении  $H = 28$  кэ удельные сопротивления  $\rho_{\perp}$  и  $\rho_{\parallel}$  образца В-1 возросли почти в  $10^5$  раз, в то время как их отношение увеличилось лишь до 2,5.

До сих пор мы рассматривали полупроводники, в которых в отсутствие магнитного поля волновые функции примесей изотропны (или практически изотропны). Если волновая функция примеси существенно анизотропна, то магнитосопротивление может зависеть от ориентации магнитного поля по отношению к кристаллическим осям. Природа этого явления заключается в том, что в зависимости от направления магнитного поля в поперечных ему направлениях электрон обладает разной эффективной массой, т. е., образно говоря, волновая функция имеет разную упругость по отношению к магнитному сжатию. Следует специально подчеркнуть, что это явление принципиально отличается от обсуждавшейся выше анизотропии сопротивления в магнитном поле. Там речь шла о зависимости предэкспоненциального множителя от направления тока по отношению к магнитному полю. Здесь мы говорим о зависимости показателя экспоненты (одинакового для любого направления тока) от направления магнитного поля по отношению к осям кристалла.

В качестве простейшего примера этого явления рассмотрим сильно деформированный по [111]  $n$ -германий,

в котором примесное состояние связано только с одним эллипсоидом электронного спектра, так что гамильтониан свободного электрона в магнитном поле можно записать в виде

$$\mathcal{H} = \frac{\left(p_x - \frac{e}{c} A_x\right)^2}{2m_t} + \frac{\left(p_y - \frac{e}{c} A_y\right)^2}{2m_t} + \frac{\left(p_z - \frac{e}{c} A_z\right)^2}{2m_t}, \quad (25.28)$$

где  $A = \frac{1}{2} [\mathbf{H}, \mathbf{r}]$  — векторный потенциал. Перейдем в И-систему координат  $(x', y', z')$  с помощью соотношений (23.26), в которых  $a$  и  $b$  определяются (23.4). Легко проверить, что в И-системе гамильтониан примет изотропный вид [28]:

$$\mathcal{H} = \frac{\left(p' - \frac{e}{c} A'\right)^2}{2m^*}, \quad (25.29)$$

где  $m^* = (m_t^2 m_l)^{1/3}$ ,  $A' = \frac{1}{2} [\mathbf{H}', \mathbf{r}']$ , а эффективное магнитное поле  $\mathbf{H}'$ , действующее в И-системе, связано с полем  $\mathbf{H}$  соотношениями

$$H'_x = \left(\frac{b}{a}\right)^{1/3} H_x, \quad H'_y = \left(\frac{b}{a}\right)^{1/3} H_y, \quad H'_z = \left(\frac{a}{b}\right)^{2/3} H_z. \quad (25.30)$$

Как говорилось в § 24, показатель экспоненты волновой функции на большом расстоянии от центра определяется функцией Грина гамильтониана  $\mathcal{H}$ , соответствующей отрицательной энергии  $-E_0$ . Так как гамильтониан  $\mathcal{H}$  в И-системе имеет такой же вид, как для полупроводника с изотропным спектром, а концентрация примесей и их распределение не изменяются при переходе в И-систему, выражения для показателя экспоненты  $R_{ij}$  и удельного сопротивления в И-системе получаются из формул (25.8) и (25.17), если величины  $H$ ,  $m$  и  $a$  заменить на  $H'$ ,  $m^*$  и  $a^* = (a^2 b)^{1/3} = \frac{\hbar}{\sqrt{2m^* E_0}}$  соответственно.

Возвратимся теперь в лабораторную систему координат. В § 23 мы видели, что переход из И-системы в лабораторную отражается лишь на предэкспоненциальном множителе удельного сопротивления и не меняет пока-

зателей экспоненты. Поэтому указанная выше замена в (25.8) и (25.17) дает окончательные результаты для показателя экспоненты магнитосопротивления. Например, в слабом магнитном поле  $H' < 4 \frac{N^{1/3} c \hbar}{a^* e}$  получаем

$$\begin{aligned} \ln \frac{\rho_3(H)}{\rho_3(0)} &= t \frac{a^* e^2}{N \hbar^2 c^2} (H')^2 = \\ &= t \frac{a^{2/3} b^{1/3} e^2}{N \hbar^2 c^2} \left[ \left( \frac{b}{a} \right)^{2/3} (H_x^2 + H_y^2) + \left( \frac{a}{b} \right)^{4/3} H_z^2 \right], \quad (25.31) \end{aligned}$$

где  $H_x$ ,  $H_y$ ,  $H_z$  — проекции поля  $\mathbf{H}$  на оси эллипсоида. Из (25.31), в частности, видно, что показатель экспоненты магнитосопротивления в случае, когда магнитное поле направлено по оси вращения эллипсоида, в  $(a/b)^2 = m_i/m_t \approx 20$  раз больше, чем в случае, когда магнитное поле лежит в плоскости  $(x, y)$ . К сожалению, нам не известны экспериментальные данные, с которыми можно было бы сравнить этот вывод.

Покажем теперь, что в важном частном случае, когда речь идет о действии слабого магнитного поля на кристалл кубической симметрии, например на недеформированный германий, показатель экспоненты магнитосопротивления не может зависеть от направления магнитного поля. Действительно, квадратичную по магнитному полю добавку к показателю экспоненты сопротивления можно записать в виде

$$\Delta \xi_c = \xi_c - \xi_c^0 = A_{ik} H_i H_k, \quad (25.32)$$

где  $\xi_c$  и  $\xi_c^0$  — пороги протекания в магнитном поле и без него соответственно,  $A_{ik}$  — тензор второго ранга. В кубическом кристалле  $A_{ik} = A \delta_{ik}$  и, следовательно,  $\Delta \xi_c = A H^2$ .

Таким образом, в полях, где  $\lg \frac{\rho(H)}{\rho(0)} \propto H^2$ , величина

$\lg \frac{\rho(H)}{\rho(0)}$  не может зависеть от направления поля. Этот вывод, разумеется, не относится к области сильных полей. Более того, магнитное поле, в котором наступает отклонение от закона (25.26), может быть различным для разных направлений.

Экспериментальные данные для  $n$ -Ge качественно согласуются со сделанными выше выводами. На рис. 7.6

приведены зависимости  $\frac{\rho(H)}{\rho(0)}$  от  $H^2 N_D^{-1}$  для двух направлений магнитного поля, полученные Хробочекком и Сладеком [4]. Видно, что в достаточно слабых магнитных

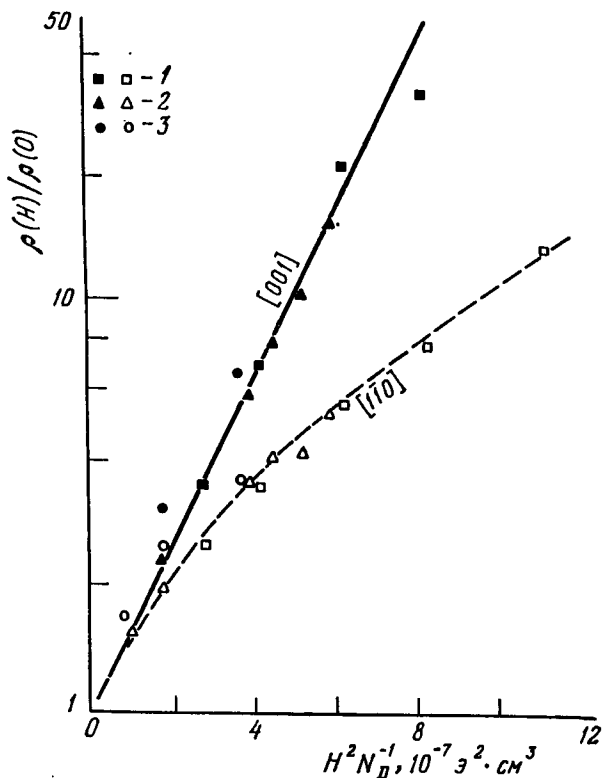


Рис. 7.6. Магнитосопротивление трех образцов  $n$ -Ge как функция  $H^2 N_D^{-1}$  для направлений поля по [001] и  $[\bar{1}\bar{1}0]$  по данным [4].  $N_D, \text{см}^{-3}$ :  
1 —  $5,4 \cdot 10^{15}$ ; 2 —  $9,0 \cdot 10^{15}$ ; 3 —  $2,0 \cdot 10^{16}$ .

полях магнитосопротивление для обоих направлений практически одинаково. При увеличении магнитного поля магнитосопротивление, соответствующее направлению поля по [001], продолжает расти по закону (25.26), в то время как для направления  $[\bar{1}\bar{1}0]$  быстро наступают отклонения от него. В результате в достаточно сильных полях магнитосопротивление оказывается сильно анизотропным. Различие в характере кривых можно каче-

ственно объяснить, если вспомнить, что поверхность  $Q_{\xi}$  представляет собой огибающую четырех сплюснутых эллипсоидов  $Q_{\xi}^{(g)}$ , каждый из которых напоминает блин, нанизанный на одну из диагоналей куба (см. рис. 6.5). Если магнитное поле направлено по оси  $[001]$ , то все четыре эллипсоида находятся в эквивалентных по отношению к нему позициях, причем компонента  $H_z$ , оказывающая, согласно (25.30), наиболее сильное сжимающее действие, равна  $H/\sqrt{3}$ . С другой стороны, в случае ориентации по  $[\bar{1}\bar{1}0]$  поле перпендикулярно к осям вращения двух эллипсоидов и слабо на них действует. Зато с осями вращения двух других эллипсоидов поле составляет сравнительно малый угол и для них  $H_z$  слабо отличается от  $H$ .

Как показано выше, благодаря кубической симметрии такое перераспределение ролей между эллипсоидами в слабом магнитном поле не приводит к зависимости  $\Delta\xi_c$  от направления поля  $\mathbf{H}$ . Однако при ориентации  $\mathbf{H}$  по  $[\bar{1}\bar{1}0]$  большее значение эффективного поля  $H'$  приводит к сравнительно более раннему нарушению неравенства

$$H' < 4 \frac{N^{1/3} c \hbar}{a^* e} \text{ и ранним отклонениям от закона (25.26).}$$

Для вычисления магнитосопротивления  $n$ -Ge в области действия закона (25.26) рассмотрим случай, когда магнитное поле ориентировано по оси  $[001]$ . Изучим сначала, как действие поля изменит поверхность  $Q_{\xi}^{(1)}$ , которая при  $H = 0$  была эллипсоидом с осью вращения, направленной по  $[111]$ . Свяжем с этим эллипсоидом систему координат  $(x, y, z)$ , направив ось  $z$  по  $[111]$ , ось  $x$  — по  $[\bar{1}\bar{1}0]$  и ось  $y$  — по  $[11\bar{2}]$ . Проекции магнитного поля на эти оси равны  $H_x = 0$ ,  $H_y = -\sqrt{\frac{2}{3}} H$ ,

$H_z = \frac{1}{\sqrt{3}} H$ . Согласно (25.31) сжимающее действие поля  $H_z$  сильнее, чем действие поля  $H_y$ , примерно в  $(H_z/H_y)^2 m_t/m_l \approx 10$  раз. Поэтому в дальнейшем мы будем считать, что поле  $H_y = 0$ . Впервые такое приближение было предложено Микошибой [15].

Перейдем в  $I$ -систему координат рассматриваемого эллипсоида с помощью преобразований (23.26). Как мы

видели выше, в  $I$ -системе координат величина  $\xi_{II}$  получается из результатов для случая изотропных волновых функций с помощью замены  $H$ ,  $m$  и  $a$  на  $H'$ ,  $m^*$  и  $a^*$ . Согласно (25.30) и принятому выше приближению  $H_y = 0$  в нашем случае  $H'_x = H'_y = 0$ ,  $H'_z = \left(\frac{a}{b}\right)^{2/3} \frac{H}{\sqrt{3}}$ . Сделав указанную замену в (24.16), получим, что в  $I$ -системе уравнение поверхности  $Q_{\xi}^{(1)}$  имеет вид

$$\frac{2r'}{a^*} + \frac{r' a^* (x'^2 + y'^2)}{36\lambda^4} \left(\frac{a}{b}\right)^{4/3} = \xi \quad (25.33)$$

или в системе координат  $(x, y, z)$

$$2 \sqrt{\frac{x^2 + y^2}{a^2} + \frac{z^2}{b^2}} \left[ 1 + \frac{(x^2 + y^2) a^{*2}}{72\lambda^4} \left(\frac{a}{b}\right)^{2/3} \right] = \xi. \quad (25.34)$$

Уравнение (25.34) отличается от уравнения (23.13) наличием второго слагаемого в квадратных скобках. В условиях слабого поля это слагаемое мало по сравнению с единицей. Поэтому, пренебрегая членами порядка  $\lambda^{-8}$  и действуя так же, как при переходе от (25.1) к (25.3), уравнение (25.34) можно преобразовать к уравнению эллипсоида:

$$2 \left( \frac{x^2 + y^2}{a^2(H)} + \frac{z^2}{b^2} \right)^{1/2} = \xi, \quad (25.35)$$

где  $a(H) \equiv a - \Delta a(\xi, H)$  и

$$\Delta a(\xi, H) = \frac{\xi^2 a^{*2} a^3}{288\lambda^4} \left(\frac{a}{b}\right)^{2/3}. \quad (25.36)$$

Из (25.35) и (25.36) видно, что магнитное поле, сжимая эллипсоид в направлении  $x, y$ , делает его менее сплюснутым, но не изменяет его главных осей. Поскольку магнитное поле одинаково ориентировано по отношению ко всем эллипсоидам, изменение поверхности каждого из них в своих осях описывается уравнениями (25.35), (25.36). Таким образом, в магнитном поле поверхность  $Q_{\xi}$  остается огибающей четырех одинаковых эллипсоидов, ориентированных так же, как при  $H = 0$ . Единственное отличие состоит в замене  $a$  на  $a(H)$ . Поэтому объем четырехэллипсоидного тела в магнитном поле  $V_{\xi}(H)$  можно записать в виде

$$V_{\xi}(H) = G \left( \frac{a(H)}{b} \right) V_{эл}(H),$$

где функция  $G(x)$  определяется формулой (23.18), а  $V_{эл}(H) = 4/3\pi a^2(H)b(\xi/2)^3$  — объем эллипсоида. Линеаризуя  $V_{\xi}(H)$  по  $\Delta a$  и подставляя результат в (23.16), получим условие для определения  $\xi_c$ :

$$\frac{4\pi}{3}Na^2b\left(\frac{\xi_c}{2}\right)^3\left\{G\left(\frac{a}{b}\right) + \frac{\Delta a(\xi_c, H)}{a}\left[\frac{a}{b}G'\left(\frac{a}{b}\right) + 2G\left(\frac{a}{b}\right)\right]\right\} = B_c, \quad (25.37)$$

где  $G'(x)$  — производная  $G(x)$  по аргументу. Пороговая величина  $B_c$  благодаря изменению формы поверхности  $Q_{\xi}$ , в принципе, может отличаться от значения  $B_c$  при  $H=0$ :  $B_c(H) = B_c + g\frac{\Delta a}{a}$ . Здесь, однако, необходимо вспомнить, что зависимость  $B_c$  от формы поверхности очень слаба (см. § 18). Действительно, для четырехэллипсоидной поверхности  $n$ -Ge значение  $B_c = 2,6$ , а для сферы  $B_c = 2,7$ . Поэтому при  $\Delta a/a \approx 1$ , когда поверхность  $Q_{\xi}$  приближается к сфере, поправка  $g\Delta a/a$  не может превышать 0,1. Таким образом, коэффициент  $g$  должен быть очень малым и отличием  $B_c(H)$  от  $B_c$  можно пренебречь. Тогда, находя  $\xi_c$  из (25.37) с точностью до членов первого порядка по  $\Delta a/a$ , получим, что  $\xi_c = \xi_c^0 + \Delta\xi_c$ , где  $\xi_c^0$  определяется выражением (23.20), а

$$\Delta\xi_c = \frac{1}{3} \frac{(\xi_c^0)^2 a^{*2} a^2}{288\lambda^4} \left(\frac{a}{b}\right)^{2/3} \left[\frac{a}{b} \frac{G'(a/b)}{G(a/b)} + 2\right]. \quad (25.38)$$

Вычисляя с помощью (23.18) выражение в квадратных скобках и используя (23.20), получаем

$$\ln \frac{\rho_3(H)}{\rho_3(0)} = \Delta\xi_c = \tau \frac{a^* e^2 H^2}{Nc^2 \hbar^2}, \quad (25.39)$$

где  $a^* = (a^2 b)^{1/3}$  и  $\tau = 0,039$ .

До сих пор мы пренебрегали проекцией  $H$  на ось  $y$  эллипсоида. Можно показать (см. [12] к гл. 6), что поле  $H_y$  увеличивает  $\tau$  до значения  $\tau = 0,042$ .

Сравним этот результат с данными для образцов германия, легированного фосфором [4]. Данные о зависимости  $\epsilon_3(H)$  для этих образцов отсутствуют. Если предположить, что этой зависимостью можно пренебречь, то

$\frac{\rho_z(H)}{\rho_z(0)} = \frac{\rho(H)}{\rho(0)}$  и для определения  $\tau$  можно использовать данные, приведенные на рис. 7.6. Измеряя наклон прямой, соответствующей направлению  $[001]$ , и учитывая, что для фосфора  $a^* = 38 \text{ \AA}$ , получаем  $\tau = 0,048$ . Небольшое превышение экспериментального значения  $\tau$  над теоретическим связано, возможно, с наличием слабой зависимости  $\varepsilon_z(H)$ . Таким образом, для  $n$ -Ge, так же как и для других полупроводников  $n$ -типа, согласно теории магнитосопротивления с экспериментом оказывается более или менее удовлетворительным.



## ГЛАВА 8

### ЭНЕРГИЯ АКТИВАЦИИ ПРЫЖКОВОЙ ПРОВОДИМОСТИ

#### § 26. Энергия активации $e_3$ при слабой компенсации

В гл. 4 уже говорилось, что обычно температурная зависимость удельного сопротивления слабо легированного полупроводника в области прыжковой проводимости имеет вид

$$\rho = \rho_3 \exp\left(\frac{e_3}{kT}\right). \quad (26.1)$$

До сих пор мы интересовались величиной  $\rho_3$  и поэтому в формуле (15.35) для  $\xi_{ij}$  отбрасывали энергетическое слагаемое  $e_{ij}/kT$ . В этой главе, наоборот, сосредоточим внимание на экспоненциальной температурной зависимости прыжковой проводимости, целиком связанной с этим слагаемым. Запишем  $\xi_{ij}$  в виде  $\xi_{ij} = \xi_{ij}^0 + \Delta\xi_{ij}$ , где  $\Delta\xi_{ij} \equiv e_{ij}/kT$ , а  $\xi_{ij}^0$  — не зависящая от температуры часть  $\xi_{ij}$  (в изотропном случае  $\xi_{ij}^0 = 2r_{ij}/a$ ). Согласно идее метода протекания, чтобы найти экспоненциальную зависимость сопротивления от температуры, следует найти порог протекания  $\xi_c$  в задаче с критерием связности

$$\xi_{ij} \leq \xi. \quad (26.2)$$

Запишем  $\xi_c$  в виде

$$\xi_c = \xi_c^0 + \Delta\xi_c, \quad (26.3)$$

где  $\xi_c^0$  — порог протекания задачи с критерием связности  $\xi_{ij}^0 \leq \xi$ , который нам уже известен (см. гл. 6), а  $\Delta\xi_c$  — добавка за счет слагаемых  $\Delta\xi_{ij}$ . Тогда задача состоит в определении  $\Delta\xi_c$ .

Начнем с изложения теории для случая слабой компенсации  $K \ll 1$  (см. [1] к гл. 3). На первых порах отвлечемся от существования крупномасштабного потенциала. Тогда, согласно сказанному в § 11, энергии большинства доноров оказываются очень близкими к невозмущенному уровню, принятому нами за нуль отсчета. Ширина пика плотности состояний много меньше расстояния от него до уровня Ферми  $\mu$ . Поэтому в выражении (15.33) для  $\epsilon_{ij}$  у подавляющего большинства пар соседних доноров можно пренебречь энергиями  $\epsilon_i$  и  $\epsilon_j$  по сравнению с  $\mu$ . В результате это выражение приобретает очень простой вид:  $\epsilon_{ij} = \mu$ . Таким образом, оказывается, что для каждого сопротивления показатель экспоненты  $\xi_{ij}$  получает одинаковое энергетическое слагаемое  $\Delta \xi_{ij} = \mu/kT$ . В этом случае определение добавки к порогу протекания не вызывает затруднений.

Действительно, критерий связности (26.2) можно записать в виде  $\xi_{ij}^0 \leq \xi'$ , где

$$\xi' = \xi - \frac{\mu}{kT}. \quad (26.4)$$

Тогда ясно, что критическое значение  $\xi'_c$  величины  $\xi'$ , при котором возникает бесконечный кластер, равно порогу протекания  $\xi_c^0$ . Переходя с помощью (26.4) от  $\xi'_c$  к  $\xi_c$ , получаем

$$\xi_c = \xi_c^0 + \frac{\mu}{kT}. \quad (26.5)$$

Подставляя  $\xi_c$  в формулу (22.4), находим

$$\rho = \rho_{03} \exp(\xi_c^0) \exp\left(\frac{\mu}{kT}\right) = \rho_3 \exp\left(\frac{\mu}{kT}\right). \quad (26.6)$$

Положение уровня Ферми при  $T=0$  было найдено в § 11. Согласно (11.11)  $\mu(T=0) = 0,61\epsilon_D$ , и, следовательно,

$$\epsilon_3 = 0,61\epsilon_D = 0,99 \frac{e^2 N^{1/3}}{\kappa}. \quad (26.7)$$

Результат (26.6) прозрачен с физической точки зрения. При  $K \ll 1$  состояния в окрестности уровня Ферми очень редки (см. рис. 3.3), перекрытие между ними очень слабое и их вкладом в электропроводность можно пренебречь. Можно считать, что вся электропроводность связана с движением дырок по донорным состояниям с

энергиями, близкими к нулю, т. е. к энергии изолированного дупора. Равновесная концентрация дырок на этих состояниях пропорциональна  $\exp(-\mu/kT)$ . Отсюда в (26.6) и возникает множитель  $\exp(\mu/kT)$ . Происхождение энергии активации  $\varepsilon_3$  в слабо компенсированном полупроводнике  $n$ -типа напоминает происхождение энергии активации  $\varepsilon_1$  в полупроводнике  $p$ -типа. В обоих случаях дырки связаны с акцепторами и нужно совершить определенную работу, чтобы активировать дырку в состоянии, в которых она дает заметный вклад в электропроводность. В первом случае эта работа равна  $\varepsilon_3$ , во втором —  $\varepsilon_1$ .

Формула (26.7) обладает важной особенностью. В нее не входят параметры, содержащие постоянную Планка, например радиус состояний  $a$ . Она содержит только концентрацию примесей и единственную характеристику кристалла — диэлектрическую проницаемость. Поэтому формулу (26.7) можно называть классической.

Амплитуда крупномасштабного потенциала при  $K \ll 1$  мала по сравнению с  $\mu$ . Поэтому он создает малую поправку к (26.7) (см. [1] к гл. 3):

$$\varepsilon_3 = 0,61\varepsilon_D (1 - 0,29K^{1/4}). \quad (26.8)$$

Строго говоря, поправка на компенсацию  $0,29K^{1/4}$  получена для  $K \leq 5 \cdot 10^{-4}$ , когда применимо описание крупномасштабного потенциала, данное в § 12. Однако можно надеяться, что с ее помощью удастся оценить влияние компенсации и при несколько больших значениях  $K$ .

Другая формула для  $\varepsilon_3$  была получена Миллером и Абрахамсом. Они исходили из формул (11.14), (11.15); которые мы критиковали в § 11. Результат Миллера и Абрахамса при  $K \ll 1$  имеет вид

$$\varepsilon_3 = \varepsilon_D (1 - 1,35K^{1/3}). \quad (26.9)$$

Формула (26.9) довольно сильно отличается от (26.8). При самых малых  $K$  отличие превышает 60%.

На рис. 8.1 построен график зависимости наблюдаемой энергии активации  $\varepsilon_3$  от концентрации основных примесей для слабо компенсированных образцов  $n$ - и  $p$ -германия [1—4].

Рассмотрим сначала наименее легированные образцы. Соответствующие им точки ложатся близко к сплош-

ной прямой. Таким образом, наблюдаемые значения  $\epsilon_3$  при малых  $K$  удовлетворительно согласуются с (26.7). Эти значения в принципе можно согласовать и с (26.9) ввиду сильной зависимости этого выражения от компенсации. Однако для этого необходимо предположить, что

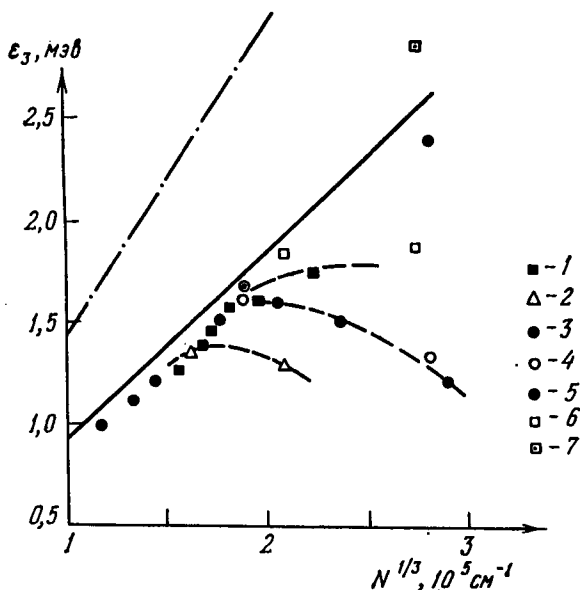


Рис. 8.1. Зависимость энергии активации прыжковой проводимости  $\epsilon_3$  от концентрации основных примесей в Ge при малой компенсации. 1 — серия образцов с P [1]; 2 — образцы с Ga [2]; 3 — образцы с Sb [3]; 4 — два образца с Sb [1]; 5 — те же два образца при «максимальном» одноосном сжатии; 6 — образцы с P [4]; 7 — один из этих образцов в очень сильном магнитном поле. Сплошная прямая — теоретическая зависимость (26.7). Штрих-пунктирная линия — результат Миллера и Абрахамса (26.9) при  $K = 0$ .

в образцах, исследованных в работах [1—4], случайная, неконтролируемая степень компенсации менялась в очень узких пределах (3—3,5%). Такое предположение представляется искусственным. Шэр и Лэкс [5], анализируя данные работы [6] для слабо легированного и слабо компенсированного кремния, также пришли к выводу, что формула (26.7) значительно лучше описывает эксперимент, чем формула (26.9).

При увеличении  $N_D$  наблюдается отклонение  $\epsilon_3$  от предсказанного формулой (26.7) классического значе-

ния. При некоторой концентрации  $N_D \equiv N_m$  функция  $\varepsilon_3(N_D)$  имеет максимум, а затем постепенно уменьшается до нуля. Это явление естественно связать с возрастанием перекрытия волновых функций доноров. Тогда поправки к классическому закону для  $\varepsilon_3$  должны содержать радиус состояний  $a$  и, следовательно, постоянную Планка. Зависимость от  $a$  действительно становится очевидной при сравнении данных для различных примесей в германии. Между эффективными боровскими радиусами галлия, сурьмы и фосфора существует соотношение  $a_{Ga} > a_{Sb} > a_P$ . Поэтому влияние перекрытия при той же концентрации примесей должно быть сильнее всего для галлия, слабее для сурьмы и еще слабее для фосфора. На рис. 8.1 видно, что этот вывод согласуется с экспериментом. По мере роста концентрации эффекты перекрытия становятся существенными в той же последовательности (Ga, Sb, P).

Еще более наглядно справедливость квантовой трактовки отклонений от предельного классического закона (26.7) видна из опытов по одноосному сжатию. Как мы уже говорили выше, в  $n$ -германии под давлением «объем» состояния уменьшается и перекрытие соседних донорных состояний ослабевает. Поэтому, если мы имеем образец с существенным отклонением величины  $\varepsilon_3$  от закона (26.7), то давление должно это отклонение сильно уменьшить. На рис. 8.1 видно, что именно так и произошло в эксперименте [1] с образцом с концентрацией сурьмы  $N_D = 2,2 \cdot 10^{16} \text{ см}^{-3}$  ( $N_D^{1/3} = 2,8 \cdot 10^5 \text{ см}^{-1}$ ). Под давлением величина  $\varepsilon_3$  сильно возросла и оказалась очень близкой к классическому закону (26.7). Сильное возрастание энергии активации наблюдалось также и для более легированных образцов (см. [15] к гл. 6).

С другой стороны, для менее легированных образцов, приближенно удовлетворяющих классическому закону (26.7), давление не должно изменять  $\varepsilon_3$ . Ведь и без давления перекрытие в них не играет роли и дальнейшее его уменьшение не может сказаться на  $\varepsilon_3$ . В таких условиях находится образец с концентрацией  $N_D = 6,5 \cdot 10^{15} \text{ см}^{-3}$  ( $N_D^{1/3} = 1,87 \cdot 10^5 \text{ см}^{-1}$ ). И действительно, как видно из рис. 8.1, давление ничтожно мало изменило величину  $\varepsilon_3$ .

Сильное магнитное поле также уменьшает перекрытие волновых функций. Поэтому его действие на величину  $\epsilon_3$  должно быть аналогичным действию одноосного сжатия на *n*-германий: при малых концентрациях примесей оно не должно менять  $\epsilon_3$ , а при более значительном легировании должно увеличивать  $\epsilon_3$ , ликвидируя отклонение от (26.7). В цитированных в гл. 7 экспериментальных работах по исследованию магнитосопротивления эти выводы качественно подтверждаются. При слабом легировании  $\epsilon_3$  не зависит от *H*, а при сильном — существенно возрастает. (Второй случай иллюстрируется данными Гальбо и Сладека для арсенида галлия (см. рис. 7.1), согласно которым при изменении *H* от 40 до 140 кэ  $\epsilon_3$  возрастает в два раза.)

Надо, однако, иметь в виду, что в большинстве цитированных в гл. 7 работ речь идет о полупроводниках с промежуточной степенью компенсации, для которых классический предел не должен совпадать с (26.7). Количественное сопоставление зависимости  $\epsilon_3(H)$  с формулой (26.7) можно сделать только по данным работы Хробочека и др. [4], где исследовались два слабо компенсированных образца германия с фосфором в сильных магнитных полях до 140 кэ (см. рис. 8.1). В отсутствие магнитного поля величина  $\epsilon_3$  слабо легированного образца хорошо согласовывалась с (26.7), а  $\epsilon_3$  более сильно легированного образца, как следовало ожидать, была меньше, чем предсказывает формула (26.7). Однако в сильном магнитном поле  $\epsilon_3$  второго образца выросла, но не до значений, соответствующих (26.7), как в экспериментах по одноосному сжатию [1], а до большей величины (см. рис. 8.1). Авторы [4] объяснили сильное возрастание  $\epsilon_3$  увеличением роли прыжков с переворотом спина. Это объяснение не представляется нам убедительным, а сам экспериментальный факт кажется непонятым.

Выше мы обсудили экспериментальные данные, свидетельствующие о том, что уменьшение  $\epsilon_3$  с ростом  $N_D$  обусловлено увеличением перекрытия волновых функций доноров. Поллак и Кнотек [7] заметили, что отклонения от закона  $\ln \rho_3 = \alpha/N^{1/3}a$ , которые хорошо видны на рис. 6.6, возникают при  $N \approx N_m$ , т. е. тогда же, когда начинается уменьшение  $\epsilon_3$ . Естественно поэтому считать,

что эти явления тесно связаны друг с другом и имеют общую причину. Обзор работ, посвященных этому очень важному вопросу, можно найти в статье Кнотека [8].

В некоторых слабо компенсированных полупроводниках [2, 3] в высокотемпературной части области прыжковой проводимости наблюдается отклонение зависимости  $\rho(T)$  от закона (26.1). Кривая зависимости  $\ln \rho$  от  $T^{-1}$  с ростом температуры постепенно становится более пологой (см. рис. 8.2). По-видимому, только переход к зонной проводимости прерывает эту тенденцию к образованию плато. Причина отклонения от закона (26.1) может быть понята на основе следующих качественных соображений.

При низких температурах большая часть дырок связана вблизи акцепторов и не участвует в прыжковой проводимости. Однако с повышением температуры концентрация оторванных от акцепторов дырок  $N_D \exp(-\epsilon_3/kT)$  растет и, наконец, при

$$T \approx T_s \equiv k^{-1} \epsilon_3 \ln^{-1} (N_D/N_A) \quad (26.10)$$

становится сравнимой с полной концентрацией дырок, равной  $N_A$ . В результате дальнейший рост числа оторванных дырок замедляется и постепенно прекращается. Это явление может быть названо эффектом насыщения прыжковой проводимости. Для количественного описания эффекта насыщения [9] следует найти зависимость химического потенциала  $\mu$  от температуры. После этого следует подставить  $\mu(T)$  вместо  $\mu(T=0)$  в формулу (26.6). Для нахождения  $\mu(T)$  следует обобщить метод, разработанный в гл. 3, на случай конечных температур. Основой этого метода служило то, что акцептор образует лишь 0-, 1- и 2-комплексы, концентрации которых можно вычислить, а  $\mu$  определяется условием электронейтральности  $N_0(\mu) = N_2(\mu)$ . Для нахождения зависимости  $\mu$  от  $T$  в условии электронейтральности необходимо учитывать дырки, оторвавшиеся от акцепторов. Очень существенным является то обстоятельство, что температура  $T_s$ , при которой возникает насыщение, мала по сравнению с характерными энергиями связи электронов в комплексах, имеющими порядок величины  $\epsilon_D$ , и по сравнению с химическим потенциалом  $\mu(T)$ . (Величина  $\mu(T)$  при

$T = 0$  также имеет порядок  $\varepsilon_D$ , а при возрастании  $T$ , как будет видно ниже, растет.) Это обстоятельство позволяет считать, что концентрация оторванных дырок равна  $N_D \exp[-\mu(T)/kT]$  и что, как и при  $T = 0$ , при  $T = T_s$

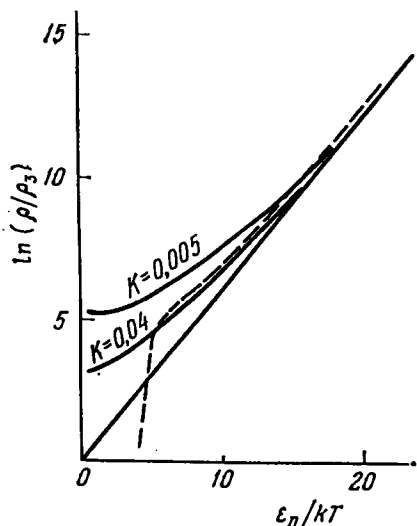


Рис. 8.2. Температурная зависимость удельного сопротивления в области насыщения прыжковой проводимости. Типичная экспериментальная кривая — штриховая линия; сплошные кривые — результат расчета величины  $\mu/kT = \ln[\rho(T)/\rho_3]$  с помощью уравнения (26.12) при разных степенях компенсации  $K$ ; прямая — асимптота расчетных кривых.

ных  $K$  табулированы в работе [9] и частично показаны на рис. 8.2. Согласно (26.6)  $\ln(\rho/\rho_3) = \mu/kT$ , поэтому графики на рис. 8.2 одновременно дают зависимость  $\ln\rho$  от  $T^{-1}$ . С учетом перехода к зонной проводимости они напоминают типичные экспериментальные кривые для слабо компенсированных образцов германия с сурьмой и галлием [2, 3]. Однако детальное сравнение с экспериментом до сих пор не производилось.

Из расчетов следует также важный вывод, касающийся метода обработки экспериментальных данных,

имеются только 0-, 1- и 2-комплексы, причем концентрации 0- и 2-комплексов, как функции  $\mu(T)$ , по-прежнему определяются формулами (11.8) и (11.10). Таким образом, при  $kT \ll \varepsilon_D$  уравнение для определения  $\mu(T)$  имеет вид

$$N_0[\mu(T)] = N_2[\mu(T)] + N_D \exp[-\mu(T)/kT], \quad (26.11)$$

или в развернутом виде

$$\begin{aligned} N_A \exp(-z^3) &= \\ &= 7,14 \cdot 10^{-4} z^6 N_A + \\ &+ N_D \exp[-\varepsilon_D/kTz], \end{aligned} \quad (26.12)$$

где  $z(T) \equiv \varepsilon_D/\mu(T)$ . Уравнение (26.12) решалось на ЭВМ при различных значениях  $\varepsilon_D/kT$  и  $K = N_A/N_D$ . Зависимости  $\mu(T)/kT$  от  $\varepsilon_D/kT$  при раз-



На рис. 8.2 видно, что выход расчетной кривой на прямую, соответствующую постоянной энергии активации  $\epsilon_3$ , происходит при  $T \leq \epsilon_D/15k$ . Для образцов германия исследованных в [1—3],  $k^{-1}\epsilon_D \approx 20 \div 30^\circ\text{K}$ . Поэтому энергию активации в них нужно определять при температурах ниже  $2^\circ\text{K}$ .

### § 27. Энергии активации $\epsilon_1$ и $\epsilon_3$ при сильной компенсации

Обратимся теперь к рассмотрению прыжковой проводимости при больших степенях компенсации, т. е. при  $1 - K \ll 1$ . Пренебрежем для начала крупномасштабным потенциалом и воспользуемся тем, что плотность состояний для сильно компенсированного полупроводника обладает теми же качественными особенностями, что и для слабо компенсированного (см. рис. 3.3 и 3.4). Вблизи уровня Ферми состояния редки, и при не очень низких температурах прыжками между ними можно пренебречь. Большая часть уровней сосредоточена в пике плотности состояний вблизи  $\epsilon = 0$ . Для типичной пары соседей энергии  $\epsilon_i$  и  $\epsilon_j$  имеют порядок величины  $\epsilon_D$ , т. е. они много меньше, чем расстояние до уровня Ферми (см. формулу (13.3)). В результате в выражении (15.33) можно пренебречь величинами  $\epsilon_i$  и  $\epsilon_j$  по сравнению с  $\mu$ . Тогда

$$\epsilon_{ij} = |\mu|. \quad (27.1)$$

(Напомним, что при сильной компенсации энергия Ферми  $\mu$ , отсчитываемая от уровня изолированной примеси, отрицательна.)

Как и в случае слабой компенсации, мы снова столкнулись с очень простой ситуацией, когда энергетическое слагаемое в  $\xi_{ij}$  одинаково для всех пар  $(i, j)$ . Повторяя рассуждения, приведенные в начале предыдущего параграфа, и используя (13.3), получаем

$$\epsilon_3 = |\mu| = \epsilon_D 2^{-1/3} (1 - K)^{-1/3}. \quad (27.2)$$

Смысл результата (27.2) очень прост. Главный вклад в прыжковую электропроводность дают электроны, активированные с уровня Ферми в пик плотности состояний. Равновесная концентрация активированных электронов

пропорциональна  $\exp[-|\mu|/kT]$ . В результате сопротивление обладает энергией активации (27.2).

В некоторых случаях, когда крупномасштабный потенциал ввиду корреляции в расположении примесей сильно подавлен, результат (27.2) является окончательным. Если же корреляция отсутствует, то крупномасштабный потенциал, как показано в § 13, имеет амплитуду порядка  $\mu$  и нельзя, как в случае слабой компенсации, считать, что он создает малую поправку к  $\epsilon_3$ .

Крупномасштабный потенциал плавно модулирует в пространстве энергетическое распределение уровней (см. рис. 3.5). В результате вблизи точки  $r$  энергии уровней большинства доноров определяются в основном потенциальной энергией электрона в крупномасштабном потенциале:

$$V(r) = -e\phi(r).$$

В силу плавности потенциала энергии соседних доноров оказываются практически одинаковыми. Подставляя  $\epsilon_i = \epsilon_j = V(r)$  в (15.33) и учитывая, что  $V(r) \geq \mu$  (или в безразмерных единицах, используемых в § 13,  $\chi \geq -\delta$ ), получаем

$$\epsilon_{ij} = |\mu - V(r)| = V(r) - \mu. \quad (27.3)$$

Таким образом, теперь энергетическое слагаемое в  $\xi_{ij}$  не является одинаковым для всех пар и плавно меняется в пространстве. Нахождение температурной поправки к  $\xi_c^0$  в общем случае представляет собой трудную задачу, которая будет решена в § 28. Однако при очень сильной компенсации для ее решения можно воспользоваться тем фактом, что характерная длина изменения потенциала оказывается очень большой. В этом случае можно применить подход, который мы назовем макроскопическим. Он основан на разбиении всего пространства на объемы, линейные размеры которых отвечают двум требованиям. С одной стороны, каждый объем должен быть настолько велик, чтобы его можно было считать макроскопическим в том смысле, что в нем малы флуктуации порога протекания и можно пользоваться методом протекания для вычисления его сопротивления. С другой стороны, объем должен быть настолько малым, чтобы можно было считать в нем величину  $V(r)$  и, следовательно, величину  $\epsilon_{ij}$  постоянными. Постоянство  $\epsilon_{ij}$  внутри объема позволяет

действовать при вычислении его удельного сопротивления  $\rho(r)$  так же, как и при выводе (26.6). В результате получаем

$$\rho(r) = \rho_3 \exp \left[ \frac{V(r) - \mu}{kT} \right]. \quad (27.4)$$

Задача теперь сводится к вычислению эффективного удельного сопротивления среды, в которой локальное значение  $\rho(r)$  флуктуирует в соответствии с заданным законом распределения потенциальной энергии  $V(r)$ . Для ее решения следует вторично применить метод протекания. Зафиксируем некоторое значение параметра  $\xi$  и, в согласии с терминологией § 17, мысленно покрасим области, в которых  $V(r)/kT < \xi$ , в черный цвет и остальные области — в белый. Представим себе, что включены только черные области, т. е. вместо белых областей вставлен диэлектрик. Будем теперь постепенно увеличивать  $\xi$ , включая все более высокоомные области. При  $\xi = \xi_c = V_c/kT$ , где  $V_c$  — уровень протекания по черному для потенциальной энергии  $V(r)$  (см. § 17 и рис. 3.5), впервые возникнет бесконечный кластер включенных областей. При  $\xi = \xi_c + 1$  формируется критическая подсетка из этих областей. Дальнейшее увеличение  $\xi$ , т. е. включение более высокоомных областей, не приводит к заметному изменению сопротивления. Поэтому эффективное удельное сопротивление определяется критической подсеткой и равно

$$\rho = \rho_3 \exp \left( \frac{-\mu + V_c}{kT} \right). \quad (27.5)$$

Таким образом,

$$\epsilon_3 = V_c - \mu \quad (27.6)$$

(см. рис. 3.5). В принципе свойства потенциала  $\phi(r)$  задаются уравнениями (13.15) и (13.17), которые, однако, никем не решались. Поэтому значение  $V_c$  точно не известно. Тем не менее можно сказать, что  $\epsilon_3$  по порядку величины равно амплитуде флуктуаций потенциальной энергии, т. е.

$$\epsilon_3 = C_1 \gamma(r_s) = C_1 \frac{e_D}{(1 - K)^{1/3}}, \quad (27.7)$$

где  $C_1$  — положительное число порядка единицы. Действительно, уровень протекания определяется из условия,

что доля пространства  $\vartheta_c$ , в котором  $V(r) < V_c$ , есть величина порядка единицы (например, для гауссовского потенциала  $\vartheta_c = 0,17$ ).

Функция распределения  $F(V)$  потенциальной энергии  $V(r)$  равна нулю при  $V < \mu$ . Поэтому условие для определения  $\vartheta_c$  имеет вид

$$\int_{\mu}^{V_c} F(V) dV = \vartheta_c, \quad (27.8)$$

и, чтобы его выполнить,  $V_c$  должно быть больше  $\mu$  на величину порядка амплитуды колебаний  $V(r)$  или, что то же самое, порядка характерного энергетического масштаба изменения функции  $F(V)$ .

Основной вывод, который мы можем сделать из формул (27.2) и (27.7), состоит в том, что при  $K \rightarrow 1$  углубление уровня Ферми приводит к возрастанию  $\varepsilon_3$  пропорционально  $(1 - K)^{-1/3}$ .

Прежде чем перейти к сравнению теории с экспериментальными данными, имеет смысл рассмотреть еще одно важное следствие углубления уровня Ферми, а именно возрастание энергии активации  $\varepsilon_1$  при  $K \rightarrow 1$ . Покажем сейчас, что при  $1 - K \ll 1$  энергия  $\varepsilon_1$  должна расти по закону

$$\varepsilon_1 = E_0 + \frac{C_1 \varepsilon_D}{(1 - K)^{1/3}}, \quad (27.9)$$

где  $C_1$  — та же самая постоянная, что и в (27.7). Если компенсация достаточно сильна, так что для вычисления  $\varepsilon_3$  применимо макроскопическое описание (27.4), приводящее к (27.7), то, согласно (27.9) и (27.7), между  $\varepsilon_1$  и  $\varepsilon_3$  существует простое соотношение:

$$\varepsilon_1 = E_0 + \varepsilon_3. \quad (27.10)$$

Начнем вычисление энергии активации  $\varepsilon_1$  с простейшего случая, когда крупномасштабный потенциал почему-либо отсутствует (см. рис. 3.4). Тогда концентрация свободных электронов в зоне проводимости пропорциональна  $\exp[-(E_0 - \mu)kT]$  и, следовательно, энергия активации  $\varepsilon_1$  равна просто расстоянию от уровня Ферми

до дна зоны проводимости:

$$\varepsilon_1 = E_0 - \mu = E_0 + |\mu|. \quad (27.11)$$

(Напомним, что мы отсчитываем  $\mu$  вверх от уровня изолированного донора.) В то же время, согласно (27.2), при отсутствии крупномасштабного потенциала  $\varepsilon_3 = |\mu|$  и соотношение (27.10), очевидно, выполняется. Если крупномасштабный потенциал имеется, то он одинаковым образом модулирует и дно зоны проводимости, и уровень изолированного донора, так что в каждой точке эти два энергетических рельефа отличаются точно на  $E_0$  (см. рис. 3.5). Чему же равна энергия  $\varepsilon_1$ , когда зона проводимости искривлена? Для ответа на этот вопрос оценим прежде всего туннельную прозрачность крупномасштабного рельефа. Вероятность туннелирования между соседними ямами на дне зоны проводимости под типичный барьер высотой  $\gamma(r_s)$  и шириной  $r_s$  по порядку величины равна

$$\exp\left[-\frac{r_s \sqrt{m\gamma(r_s)}}{\hbar}\right] = \exp\left[-(1-K)^{-5/6}(Na^3)^{-1/6}\right]. \quad (27.12)$$

(Здесь использованы формулы (13.9) и (13.8) для  $\gamma(r_s)$  и  $r_s$ .)

В слабо легированном ( $Na^3 \ll 1$ ) и сильно компенсированном ( $1 - K \ll 1$ ) полупроводнике эта вероятность очень мала, и поэтому туннельным эффектом можно пренебречь. Тогда электрон может двигаться лишь по классически доступным областям. Но для того, чтобы области, доступные для электрона с данным значением энергии, насквозь пронизывали кристалл и соединяли его контакты, необходимо, чтобы эта энергия была больше уровня протекания для энергетического рельефа дна зоны (см. рис. 3.5). Таким образом, подвижность как функция энергии обладает порогом, совпадающим с порогом протекания. С другой стороны, равновесная концентрация электронов экспоненциально убывает с ростом энергии. Поэтому основной вклад в электропроводность дают электроны в полосе энергий порядка  $kT$  вблизи порога протекания. Энергия активации  $\varepsilon_1$  равна расстоянию от зонного порога протекания до уровня Ферми:

$$\varepsilon_1 = V_c + E_0 - \mu. \quad (27.13)$$

Подставляя (27.6) в (27.13), получаем (27.10). Выполнение равенства (27.10) также наглядно видно на рис. 3.5.

Перейдем теперь к сравнению полученных выше законов возрастания  $\epsilon_1$  и  $\epsilon_3$  при  $K \rightarrow 1$  с экспериментальными данными. Начнем с данных для  $\epsilon_1$ . Специальное исследование германия с целью проверки (27.9) производилось в работе Шлимака и Емцева [10]. Исходные некомпенсированные образцы  $n$ -Ge с концентрацией мышьяка  $N_D = 10^{16} \text{ см}^{-3}$  облучались медленными нейтронами, которые, как известно, производят трансмутации ядер Ge в ядра Ga (акцептор) и ядра As и Se (доноры). Концентрация созданных облучением акцепторов в 2,5 раза превышает концентрацию возникших доноров. Поэтому степень компенсации возрастала при увеличении дозы облучения. Образец с наибольшей дозой облучения ( $K \approx 0,90$ ) для получения еще более точной компенсации дополнительно облучался  $\gamma$ -квантами, производящими дефекты с акцепторными свойствами. Результаты эксперимента [10] изображены на рис. 8.3.

Видно, что по мере компенсации энергия активации  $\epsilon_1$  возрастает от 12 до 28 мэв. Авторам [10] удалось достоверно определить степень компенсации лишь для образцов 1—6. Они пришли к выводу, что зависимость  $\epsilon_1(K)$  удовлетворительно согласуется с (27.9).

Посмотрим теперь, что дает эксперимент для зависимости  $\epsilon_3(K)$  при  $K \rightarrow 1$ . На рис. 4.3 мы воспроизвели зависимость  $\epsilon_3(K)$  для  $p$ -Ge, приведенную в обзоре Мотта и Туза ([7] к гл. 4). Видно, что при  $K \rightarrow 1$  энергия  $\epsilon_3$  действительно быстро растет.

Выше говорилось, что одновременно с  $\epsilon_3$  должна расти  $\epsilon_1$ . Из рис. 7 обзора Мотта и Туза видно, что упомянутое возрастание  $\epsilon_3$  действительно сопровождается увеличением  $\epsilon_1$  примерно на 1—2 мэв. К сожалению, в этих опытах компенсация недостаточно велика, так что  $\epsilon_3$  и приращение  $\epsilon_1$  по порядку величины не превосходят среднего разброса доноров по энергии  $\epsilon_D$ , в условиях этого эксперимента равного 1,85 мэв. Поэтому количественное сравнение этих данных с формулами (27.7), (27.9) и (27.10), строго говоря, невозможно. Увеличение  $\epsilon_1$  и  $\epsilon_3$  на величины порядка 1 мэв при сильной компенсации наблюдалось и в  $n$ -германии ([9] к гл. 4, [11]). Един-

ственной известной нам экспериментальной работой, в которой наблюдалось возрастание обеих энергий активации  $\epsilon_1$  и  $\epsilon_3$ , достаточное для сравнения с формулами

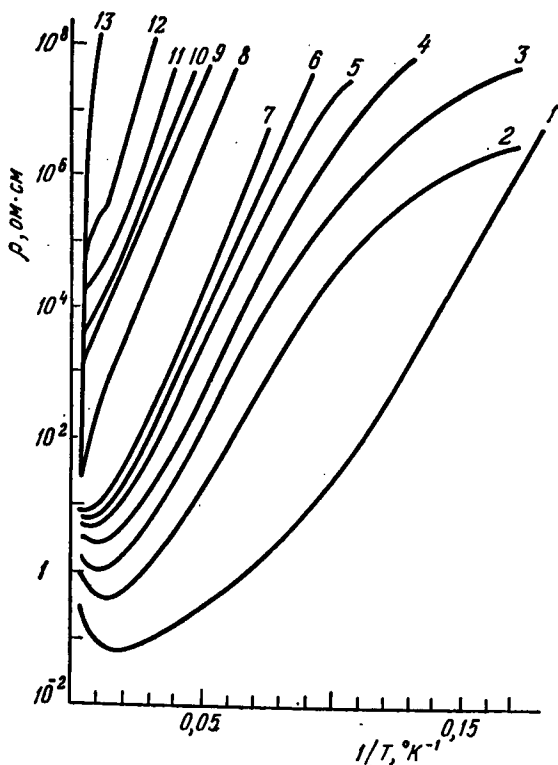


Рис. 8.3. Удельное сопротивление  $\rho$  для  $n$ -Ge с концентрацией доноров  $N_D = 10^{16} \text{ см}^{-3}$  как функция обратной температуры при различных степенях компенсации  $K = N_A/N_D$  [10]: 1—исходный некомпенсированный образец; 2—4 — компенсация за счет облучения тепловыми нейтронами; 5—13 — компенсация  $\gamma$ -лучами. Значения  $(1 - K)$ : 2 —  $3,3 \cdot 10^{-1}$ ; 3 —  $1,8 \cdot 10^{-1}$ ; 4 —  $1,03 \cdot 10^{-1}$ ; 5 —  $5,6 \cdot 10^{-2}$ ; 6 —  $4,4 \cdot 10^{-2}$ ; 7 — 13 — степень компенсации точно не известна.

(27.7), (27.9) и (27.10); является исследование проводимости  $p$ -InSb с примесями марганца, германия, железа и цинка [12]. Для некоторых образцов в работе [12] наблюдалось увеличение  $\epsilon_3$  в 10 раз и  $\epsilon_1$  — почти в 2 раза. При этом абсолютные величины прироста обеих энергий активации были одного порядка, а иногда совпадали и

количественно. Обработка зависимости  $\epsilon_1(K)$  показала, что она удовлетворительно согласуется с законом (27.9), а коэффициент  $C_1$  равен 0,7.

### § 28. Метод малых возмущений в теории протекания и общая теория энергии активации $\epsilon_3$

В предыдущих параграфах мы вычислили энергию активации  $\epsilon_3$  при очень слабой и очень сильной компенсации. В первом случае это удалось сделать благодаря тому, что слагаемое  $\epsilon_{ij}/kT$  было одинаково для всех пар. Во втором случае использовался макроскопический подход, основанный на постоянстве величин  $\epsilon_{ij}/kT$  в больших областях пространства. Ясно, однако, что, когда потенциал становится менее плавным, макроскопический подход делается неприменимым. Мы увидим ниже, что в случае сильной компенсации это происходит при  $K = K_1$ , где  $K_1$  удовлетворяет неравенству  $1 - K_1 \ll 1$ . Таким образом, вопрос о том, как вычислять энергию активации в промежуточной области компенсаций, остается открытым. Казалось бы, если слагаемое  $\epsilon_{ij}/kT$  в условии связности

$$\frac{2r_{ij}}{a} + \frac{\epsilon_{ij}}{kT} \leq \xi \quad (28.1)$$

нельзя считать постоянным, то при понижении температуры большую роль будут играть сопротивления  $R_{ij}$  со все меньшими значениями  $\epsilon_{ij}$  и энергия активации  $\epsilon_3$  не может быть постоянной, а должна убывать с понижением температуры. Почему же тем не менее существует постоянная энергия активации  $\epsilon_3$ ? Ниже мы покажем, что достаточным условием для существования постоянной энергии активации является выполнение для типичных соседних узлов неравенства

$$\frac{\epsilon_{ij}}{kT} \ll \frac{2r_{ij}}{a}. \quad (28.2)$$

Кратко пояснить роль неравенства (28.2) можно следующим образом. Выполнение (28.2) означает, что отбор сопротивлений с помощью условия связности (28.1) происходит в основном по величине расстояний  $r_{ij}$ . По



отношению к значениям  $\varepsilon_{ij}/kT$  этот отбор оказывается случайным. В результате в области (28.2) понижение температуры не приводит к повышенной роли сопротивлений с малым  $\varepsilon_{ij}$ , т. е. высказанные выше сомнения оказываются неосновательными. Неравенство (28.2) обычно хорошо выполняется в слабо легированных полупроводниках при не очень низких температурах. Действительно, оценим  $\varepsilon_{ij}$  как  $e^2 N^{1/3}/\kappa$  и  $r_{ij}$  — как  $N^{-1/3}$ , где  $N$  — концентрация основных примесей. Тогда для  $p$ -германия с концентрацией акцепторов  $N_A = 10^{15} \text{ см}^{-3}$  при температуре  $T = 3^\circ\text{K}$  получаем, что левая сторона (28.2) равна 3,6, а правая — 21. При достаточно низких температурах и в образцах с концентрациями примесей, сравнительно близкими к переходу диэлектрик — металл, неравенство (28.2) нарушается. Это приводит к тому, что энергия активации не остается постоянной, а степенным образом убывает при понижении температуры. Прыжковой проводимости с убывающей энергией активации посвящена следующая глава. В этом параграфе речь будет идти только о тех случаях, в которых неравенство (28.2) выполняется.

Ниже мы не только докажем существование энергии активации при промежуточных степенях компенсации, но и покажем, как ее вычислять по заданному распределению величин  $\varepsilon_{ij}$ . Для этого будет изложен метод малых возмущений в теории протекания ([14] к гл. 5). Он состоит в разложении порога протекания  $\xi_c$  в ряд по малым возмущениям  $\varepsilon_{ij}/kT\xi_c^0$  вблизи невозмущенного значения  $\xi_c^0$ . Если ограничиться первым приближением теории возмущений, то получается сравнительно простой рецепт вычисления энергии активации  $\varepsilon_3$ . Ниже показано, что область применимости этого рецепта является «дополнительной» к области применимости макроскопического подхода, т. е. везде, где применим второй рецепт, неприменим первый, и наоборот. В результате возникает полная картина температурной зависимости прыжковой проводимости слабо легированного полупроводника при температурах, для которых выполняется неравенство (28.2).

В этой главе для простоты мы рассмотрим только случай изотропных волновых функций, когда  $\xi_c^0 = 2r_c/a$ .

Случай анизотропных волновых функций рассмотрен в Приложении 2.

Прежде чем излагать метод малых возмущений, остановимся на статистических свойствах величин  $\Delta\xi_{ij} = \varepsilon_{ij}/kT$ . Простейшей мыслимой является ситуация, в которой энергии различных узлов независимы и распределены одинаковым образом. Иначе говоря, плотность вероятности того, что узел  $i$  имеет энергию  $\varepsilon_i$ , не зависит от энергий других узлов и от номера узла  $i$ . С такой ситуацией мы встречаемся в твердых растворах полупроводников [13], где разброс уровней может быть в основном связан с флуктуациями состава в объеме примесного состояния, а не с кулоновским взаимодействием соседних примесей. В этом случае, хотя распределение величин  $\varepsilon_{ij}$  для всех пар одинаково, величины  $\varepsilon_{ij}$  и  $\varepsilon_{jk}$  для двух пар, имеющих общий узел, оказываются, в соответствии с (15.33), коррелированными.

Более типичной является ситуация, в которой коррелированы энергии  $\varepsilon_i$  и  $\varepsilon_j$ , причем степень корреляции зависит от расстояния  $r_{ij}$  (см. гл. 3). Характерное расстояние  $r_{ij}$ , вплоть до которого имеется корреляция энергий  $\varepsilon_i$  и  $\varepsilon_j$ , назовем корреляционной длиной энергии и обозначим через  $r_k$ . При  $1 - K \ll 1$  эта величина совпадает с радиусом нелинейного экранирования  $r_s$  (см. (13.8)), а при  $K \ll 1$  она равна радиусу линейного экранирования  $r_0$  (см. (12.10)). В промежуточной области  $K$  величина  $r_k$  порядка среднего расстояния между примесями  $N^{-1/3}$ . В этом можно убедиться, подставляя  $K \approx 0,5$  в формулы (12.10) и (13.8). Благодаря корреляции энергий плотность вероятности  $f(\varepsilon_{ij})$  того, что паре  $(i, j)$  соответствует заданное значение  $\varepsilon_{ij}$ , или, короче, функция распределения  $\varepsilon_{ij}$ , оказывается различной для пар с разными  $r_{ij}$ . Кроме того, величины  $\varepsilon_{ij}$  для недостаточно удаленных друг от друга пар оказываются коррелированными между собой, даже если эти пары не имеют общего узла.

Имея в виду вернуться к реальной ситуации, рассмотрим сначала гипотетический случай, когда для всех пар величины  $\varepsilon_{ij}$  не зависят друг от друга и обладают одной и той же функцией распределения  $f(\varepsilon_{ij})$ . Тогда функция  $f(\varepsilon_{ij})$  целиком определяет влияние слагаемых  $\Delta\xi_{ij}$  на порог протекания  $\xi_c$ . Дальнейшее рассмотрение основано

на предположении о том, что  $\xi_c$  можно разложить в ряд по моментам функции распределения

$$\frac{\xi_c}{\xi_c^0} = 1 + A \left\langle \frac{\Delta \xi_{ij}}{\xi_c^0} \right\rangle + B \left\langle \left( \frac{\Delta \xi_{ij}}{\xi_c^0} \right)^2 \right\rangle + C \left\langle \frac{\Delta \xi_{ij}}{\xi_c^0} \right\rangle^2 \dots, \quad (28.3)$$

где знак  $\langle \dots \rangle$  означает усреднение с помощью функции распределения  $f(\epsilon_{ij})$ . Например,

$$\left\langle \frac{\Delta \xi_{ij}}{\xi_c^0} \right\rangle = \frac{1}{\xi_c^0} \int_0^{\infty} \frac{\epsilon_{ij}}{kT} f(\epsilon_{ij}) d\epsilon_{ij}. \quad (28.4)$$

Величина (28.4) является фактически тем малым параметром, по которому и ведется разложение в (28.3). Наибольший интерес представляет коэффициент  $A$  в разложении (28.3). Чтобы определить его, предположим, что все  $\epsilon_{ij}$  одинаковы и равны  $\epsilon_0$ , т. е.  $f(\epsilon_{ij}) = \delta(\epsilon_{ij} - \epsilon_0)$ . Тогда, пренебрегая квадратичными по  $\Delta \xi_{ij}/\xi_c^0$  членами (28.3), получим

$$\xi_c = \xi_c^0 + A \frac{\epsilon_0}{kT}. \quad (28.5)$$

С другой стороны, в § 26 мы уже показали, что добавление постоянного слагаемого к  $\xi_{ij}$  приводит к тому, что точно такое же слагаемое добавляется к  $\xi_c^0$ , т. е.

$$\xi_c = \xi_c^0 + \frac{\epsilon_0}{kT}. \quad (28.6)$$

Сравнивая (28.5) и (28.6), видим, что  $A = 1$ . Следовательно, для произвольной функции распределения  $f(\epsilon_{ij})$  в первом порядке по  $\Delta \xi_{ij}/\xi_c^0$  формула (28.3) имеет вид

$$\xi_c = \xi_c^0 + \langle \Delta \xi_{ij} \rangle = \xi_c^0 + \frac{\langle \epsilon_{ij} \rangle}{kT}. \quad (28.7)$$

Согласно (22.4) это означает, что температурная зависимость сопротивления имеет вид (26.1), причем

$$\epsilon_3 = \langle \epsilon_{ij} \rangle. \quad (28.8)$$

Покажем теперь, что метод теории возмущений применим и при учете следующих осложняющих обстоятельств: 1) функция распределения  $f(\epsilon_{ij})$  зависит от расстояния между донорами  $r_{ij}$ , и 2) величины  $\epsilon_{ij}$  для раз-

личных пар коррелированы друг с другом. Прежде всего заметим, что увеличение порога протекания от значения  $\xi_c^0$  до  $\xi_c$  при переходе от условия связности  $\xi_{ij}^0 \equiv \frac{2r_{ij}}{a} \leq \xi$  к условию  $\xi_{ij}^0 + \Delta\xi_{ij} \leq \xi$  происходит только за счет пар, для которых

$$\xi_{ij}^0 < \xi_c^0, \text{ но } \xi_{ij}^0 + \Delta\xi_{ij} > \xi_c^0. \quad (28.9)$$

В нашем случае  $\Delta\xi_{ij} \ll \xi_c^0$ . Следовательно, играют роль те пары, у которых  $\xi_{ij}^0$  очень близко к  $\xi_c^0$ , т. е.  $r_{ij}$  близко к перколяционному радиусу  $r_c$ . Число существенных пар мало, и, следовательно, расстояния между ними велики по сравнению со средним расстоянием между примесями (ниже мы дадим количественную оценку этого расстояния). Поэтому при не слишком большой величине корреляционной длины энергии  $r_k$  величины  $\varepsilon_{ij}$  существенных пар оказываются некоррелированными. Тогда мы приходим к очень важному выводу: сдвиг порога протекания целиком определяется функцией распределения величин  $\varepsilon_{ij}$  для существенных пар с  $\xi_{ij} = \xi_c^0$ , т. е. с  $r_{ij} = r_c$ .

Теперь можно произвести аналогичное (28.3) разложение порога протекания по моментам этой функции. В результате получаем, что

$$\Delta\xi_c = \xi_c - \xi_c^0 = \langle \Delta\xi_{ij} \rangle |_{\xi_{ij}^0 = \xi_c^0} \quad (28.10)$$

или для энергии активации

$$\varepsilon_3 = \langle \varepsilon_{ij} \rangle |_{r_{ij} = r_c}. \quad (28.11)$$

Обозначения, примененные в (28.10) и (28.11), имеют целью подчеркнуть, что усреднение производится по всем парам, для которых  $\xi_{ij} = \xi_c^0$ , т. е.  $r_{ij} = r_c$  \*).

\*) Строго говоря, определение существенной пары не исчерпывается условием  $r_{ij} = r_c$ . Необходимо, чтобы эта пара была элементом пути протекания, что накладывает определенные условия на конфигурацию доноров в ее окружении. В частности, эта пара не может находиться внутри полости большого радиуса, не содержащей других доноров. Если  $\varepsilon_{ij}$  определяются конфигурацией соседних доноров, то функция распределения величины  $\varepsilon_{ij}$  для существенных пар отличается от функции распределения для произвольных пар, имеющих  $r_{ij} = r_c$ . Тогда формула (28.11) является приближенной.

Скал и др. ([14] к гл. 5) предприняли численную проверку рецепта теории возмущений (28.11) для конкретного распределения величин  $\varepsilon_{ij}$ . Вероятность  $w(\varepsilon_i)$  того, что данный узел имеет энергию  $\varepsilon_i$ , предполагалась не зависящей от энергий других узлов и от номера узла и равной

$$w(\varepsilon_i) = \begin{cases} \frac{1}{2\Delta\varepsilon}, & |\varepsilon_i - \mu| \leq \Delta\varepsilon, \\ 0, & |\varepsilon_i - \mu| > \Delta\varepsilon. \end{cases} \quad (28.12)$$

В таком случае функция распределения  $f(\varepsilon_{ij})$  величин  $\varepsilon_{ij}$  не зависит от расстояния  $r_{ij}$ , формула (28.11) является точной и дает

$$\varepsilon_3 = \langle \varepsilon_{ij} \rangle = \int_{-\infty}^{\infty} d\varepsilon_i \int_{-\infty}^{\infty} d\varepsilon_j w(\varepsilon_i) w(\varepsilon_j) \varepsilon_{ij}. \quad (28.13)$$

Подставляя (28.12) в (28.13) и используя определение (15.33), получаем

$$\varepsilon_3 = \frac{5}{6} \Delta\varepsilon \quad (28.14)$$

или

$$\frac{\xi_c - \xi_c^0}{\xi_c^0} = \frac{5}{6} \frac{\Delta\varepsilon}{kT\xi_c^0}. \quad (28.15)$$

Цель расчета состояла в проверке равенства (28.15). Поскольку все выводы теории, заключенные в формулах (28.11), (28.14) и (28.15), в равной мере применимы к случаям двух и трех измерений, для простоты расчеты производились в двумерном случае. Зависимость порога протекания  $\xi_c$  от параметра  $\Delta\varepsilon/kT$  была найдена с помощью решения задачи с критерием связности (28.1) методом Монте-Карло на ЭВМ. Результаты расчета представлены кружками на рис. 8.4.

Прямая линия соответствует теоретическому закону (28.15). Видно, что при  $\Delta\varepsilon/kT\xi_c^0 \ll 1$ , т. е. при выполнении неравенства (28.2), расчетные точки близки к теоретической кривой. Согласно (26.1) и (22.9)

$$\frac{\xi_c - \xi_c^0}{\xi_c^0} = \frac{\ln[\rho(T)/\rho_3]}{\ln[\rho_3/\rho_{03}]}. \quad (28.16)$$

Поэтому рис. 8.4 описывает зависимость  $\ln \rho$  от  $T^{-1}$ . При малых  $\Delta \epsilon / kT \xi_c^0$  имеется постоянная энергия активации. С понижением температуры неравенство (28.2) перестает выполняться и происходит постепенный переход к прыжковой проводимости с убывающей энергией активации (см. гл. 9).

Посмотрим теперь, к каким новым результатам для энергии активации  $\epsilon_3$  слабо легированного полупроводника приводит формула (28.11) при сильной компенсации ( $1 - K \ll 1$ ). В случае сильной компенсации величины  $\epsilon_{ij}$ , согласно (27.3), определяются крупномасштабным потенциалом и не зависят от расстояния между примесями. Поэтому рецепт (28.11) дает

$$\epsilon_3 = \langle V \rangle - \mu =$$

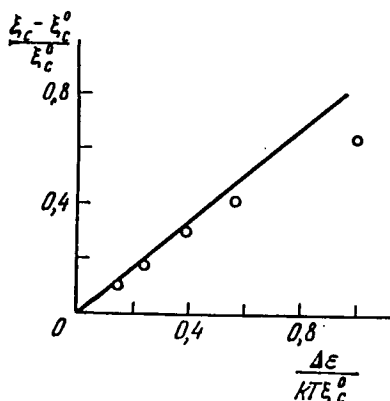
$$= \int_{\mu}^{\infty} F(V) \{V - \mu\} dV =$$

$$= \frac{C_2 \epsilon_D}{(1 - K)^{1/3}}, \quad (28.17)$$

Рис. 8.4. Сравнение уравнения (28.15) (прямая линия) и численного расчета порога протекания методом Монте-Карло (кружки).

где  $F(V)$  — функция распределения потенциальной энергии  $V(r)$ , а  $C_2$  — неизвестная численная константа. Сопоставляя (28.17) с (27.6) и (27.7), мы видим, что по порядку величины эти формулы дают одинаковый результат. Поскольку уровень протекания  $V_c$  ниже среднего значения потенциала  $\langle V \rangle$ , то имеет место неравенство  $C_2 > C_1$ . Таким образом, энергия активации  $\epsilon_3$ , вычисленная по теории возмущений, больше значения  $\epsilon_3$ , полученного с помощью макроскопического подхода, на величину порядка ее самой. Естественно, что в области применимости (28.17) соотношение (27.10) не должно выполняться.

Перейдем теперь к вопросу об областях применимости макроскопического подхода и теории возмущений. Как говорилось в § 27, условием применимости макро-



скопического подхода является возможность разбить пространство на объемы, в которых, с одной стороны, значение  $\varepsilon_{ij}$  можно считать постоянным, а с другой стороны, флуктуации порога протекания, связанные с конечностью этих объемов, малы по сравнению с величиной  $\Delta\xi_c$ . Согласно (18.11) среднеквадратичная флуктуация порога протекания объема с линейным размером  $l$  равна

$$\xi_c^0 (lN^{1/3})^{-1/\nu}. \quad (28.18)$$

Следовательно, наши объемы должны иметь линейные размеры, большие чем

$$L_1 = N^{-1/3} \left( \frac{\xi_c^0}{\Delta\xi_c} \right)^\nu. \quad (28.19)$$

С другой стороны, величины  $\varepsilon_{ij}$  меняются на длине порядка  $r_k$ . Таким образом, условие применимости макроскопического подхода имеет вид

$$L_1 \ll r_k. \quad (28.20)$$

Наша точка зрения состоит в том, что условием применимости теории возмущений является неравенство, обратное (28.20). Вместо того, чтобы доказывать это утверждение, мы проиллюстрируем его, представляя бесконечный кластер в виде одножильной сетки (см. § 20). Выясним сначала смысл длины  $L_1$  в формуле (28.19). Рассмотрим «невозмущенную» задачу на случайных узлах с условием связности  $\xi_{ij}^0 \leq \xi$ , имеющую пороговое значение  $\xi_c^0$ . Пусть  $\xi = \xi_c$ . Так как  $\xi_c > \xi_c^0$ , мы имеем разветвленный бесконечный кластер из связанных элементов, причем, согласно (18.9), длина  $L_1$  является радиусом корреляции этого кластера. На рис. 8.5 изображен «скелет» такого кластера в рамках модели одножильной сетки.

Введение в условие связности возмущающего слагаемого  $\Delta\xi_{ij}$ , т.е. возврат к задаче с условием связности (28.1), приводит к тому, что при  $\xi = \xi_c$  часть связей переходит в разряд разорванных, причем, по определению  $\xi_c$ , бесконечный кластер переходит в пороговое состояние. При этом на каждую макросвязь приходится в среднем примерно один разрыв. Как видно из рисунка, только

те пары узлов, которые находятся в макросвязях и между которыми прерывается связь, играют роль в увеличении порога от  $\xi_c^0$  до  $\xi_c$ . Среднее расстояние между

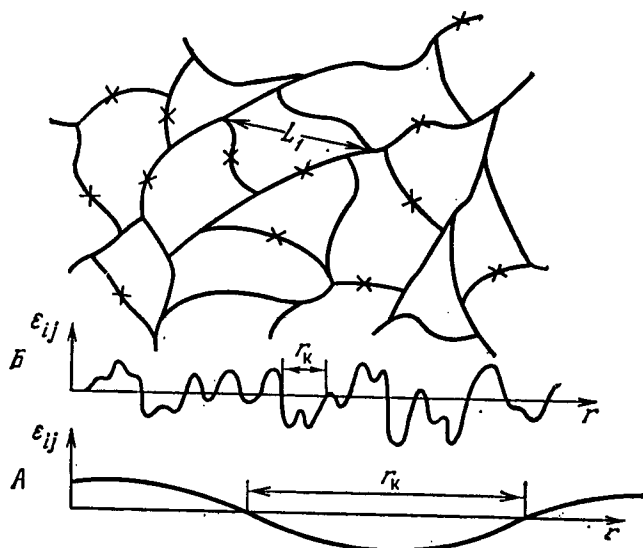


Рис. 8.5. Функции  $e_{ij}(r)$  (внизу) с одинаковыми амплитудами колебания  $\gamma$ , но с разными длинами корреляции  $r_k$ . Бесконечный кластер (вверху), состоящий из сопротивлений с  $\xi_{ij}^0 \leq \xi_c = \xi_c^0 + \Delta\xi_c$  (где  $\Delta\xi_c$  порядка  $\gamma/kT$ );  $L_1$  — его радиус корреляции. Крестики — существенные разрывы, переводящие кластер в пороговое состояние. Видно, что в случае А величины  $e_{ij}$  существенных рвущихся сопротивлений коррелированы, а в случае В — не коррелированы.

ними порядка  $L_1$ . С другой стороны, условием применимости теории возмущений является, как мы видели выше, отсутствие корреляции между значениями  $\Delta\xi_{ij}$ , существенными для протекания. Следовательно, теория возмущений применима, если  $L_1 \gg r_k$ . Граница же между областями применимости теории возмущений и макроскопического подхода задается равенством

$$L_1 = r_k. \quad (28.21)$$

При сильной компенсации, подставляя в (28.19)  $\xi_c^0 = \alpha/N^{1/3}a$  и  $\Delta\xi_c \approx \gamma(r_s)/kT$ , где  $\gamma(r_s)$  определяется (13.9),



получим

$$L_1 \approx N^{-1/3} \left[ \frac{kT(1-K)^{1/3}}{e_D(N^{1/3}a)} \right]^v. \quad (28.22)$$

Подставляя затем (28.22) и значение  $r_k = r_s$  из (13.8) в уравнение (28.21), находим, что с понижением температуры при

$$T = \frac{e_D}{k} (N^{1/3}a) (1-K)^{-\frac{1}{3} - \frac{2}{3v}} \quad (28.23)$$

происходит переход от формулы (28.17) к (27.7). Зафиксировав температуру, можно из (28.23) найти значение степени компенсации  $K_1(T)$ , при котором происходит переход от (28.17) к (27.7). Легко увидеть, что при  $T > k^{-1}e_D(N^{1/3}a)$ , т. е. в условиях, когда выполняется неравенство (28.2), имеет место неравенство  $1 - K_1 \ll 1$ , так что и при  $K < K_1$  и при  $K > K_1$  компенсация является сильной. Это оправдывает использование нами крупномасштабного потенциала при получении (28.17).

Итак, теория предсказывает, что кривая зависимости  $\ln \rho$  от  $1/T$  для сильно компенсированных полупроводников может иметь излом, причем низкотемпературная энергия активации должна быть меньше высокотемпературной. Нам не известны какие-либо экспериментальные данные, убедительно свидетельствующие о таком изломе. В связи с этим следует заметить, что, во-первых, нам не известны коэффициенты  $C_2$  и  $C_1$  и, следовательно, величина излома. Во-вторых, две разные энергии активации возникают благодаря крупномасштабному потенциалу. Если же этот потенциал подавлен, то должна существовать только одна энергия активации (27.2).

До сих пор в этой главе говорилось только о случаях сильной и слабой компенсации. Теперь мы обсудим промежуточные между этими случаями значения компенсации. При промежуточной компенсации крупномасштабный потенциал отсутствует, а энергии соседних примесей отличаются на величины порядка  $e_D$ . (Это можно увидеть из формул (12.10), (12.13), (13.8) и (13.9), подставив в них значения  $K = 0,5$  и  $1 - K = 0,5$ .) В результате во всей области температур  $T > e_D k^{-1} N^{1/3} a$ , где выполняется неравенство (28.2), энергия активации  $e_3$

определяется рецептом теории возмущений (28.11). Однако вычислению  $\epsilon_3$  препятствует то, что задача о распределении электронов по донорам, соответствующем минимуму энергии, при  $K \approx 0,5$  не содержит малых параметров. Поэтому функция распределения  $f(\epsilon_{ij})$  пар с расстояниями  $r_{ij} = r_c$  может быть определена лишь с помощью машинных расчетов, которые до сих пор не производились.

В такой ситуации теория дает только два качественных вывода. Во-первых, энергия активации  $\epsilon_3$  должна быть порядка  $\epsilon_D$ , и, во-вторых, согласно формулам (27.7), (28.17) и (26.8) на кривой  $\epsilon_3(K)$  может быть минимум. Как мы видели в гл. 4, зависимость  $\epsilon_3(K)$  действительно имеет минимум при  $K \approx 0,4$  (см. рис. 4.3). Однако убывание  $\epsilon_3(K)$  с ростом  $K$  при  $K \leq 0,3$  оказывается значительно более быстрым, чем это следует из (26.8). Кроме того, значение  $\epsilon_3$  при  $K = 0,4$  оказывается аномально малым ( $\epsilon_3 = 0,3 \epsilon_D$ ). Путь, на котором можно попытаться объяснить эти факты, мы обсудим в гл. 10.

Нам известен только один случай, когда при промежуточной компенсации распределение энергий  $\epsilon_i$  оказывается настолько простым, что вычисление классического предела  $\epsilon_3$  с помощью рецепта теории возмущений (28.11) удастся довести до конца. Речь идет о прыжковой проводимости в твердых растворах различных полупроводников, где флуктуации состава раствора в объеме примесного состояния могут приводить к более сильному разбросу примесных уровней, чем кулоновское взаимодействие [13]. Главное упрощающее обстоятельство состоит в том, что энергии уровней различных примесей в этом случае оказываются независимыми, а их функция распределения  $w(\epsilon)$  легко поддается вычислению. В работах [13, 14] сообщаются экспериментальные данные для кристаллов  $\text{Ge}_{1-x}\text{Si}_x$ , с помощью теории возмущений вычисляется энергия активации  $\epsilon_3$  и обсуждается сравнение теории с экспериментальными данными.

## ГЛАВА 9

# ПРЫЖКОВАЯ ПРОВОДИМОСТЬ С ПЕРЕМЕННОЙ ДЛИНОЙ ПРЫЖКА

### § 29. Закон Мотта

В 1968 г. Мотт [1] обратил внимание на то, что при достаточно низких температурах прыжковая проводимость определяется состояниями с энергиями в узкой полоске вблизи уровня Ферми. Он показал, что если плотность состояний на уровне Ферми отлична от нуля, то температурная зависимость удельного сопротивления имеет универсальный вид:

$$\rho(T) = \rho_0 \exp \left\{ \left( \frac{T_0}{T} \right)^{1/4} \right\}, \quad (29.1)$$

где

$$T_0 = \frac{\beta}{kg(\mu) a^3}, \quad (29.2)$$

$g(\mu)$  — плотность состояний на уровне Ферми,  $a$  — радиус близких к уровню Ферми локализованных состояний,  $\beta$  — численный коэффициент. Зависимость (29.1) теперь называют законом Мотта.

Мотт ограничился качественным выводом (29.1), не позволяющим определить точное значение коэффициента  $\beta$ . Впоследствии Амбегаокар и др. [2] вывели закон (29.1) с помощью метода протекания и сформулировали безразмерную задачу теории протекания, из решения которой можно найти значение  $\beta$ . Мы не будем сразу же прибегать к методу протекания во всей его полноте, а изложим сначала качественный вывод (29.1), в идейном смысле напоминающий оригинальный вывод Мотта и

поясняющий, как нам кажется, физическую сущность явления.

Представим себе систему с локализованными состояниями вблизи уровня Ферми. Естественно предположить, что благодаря множителю  $\exp(\epsilon_{ij}/kT)$  в выражении (15.34) для сопротивления  $R_{ij}$  при очень низких температурах основной вклад в электропроводность будут вносить сопротивления с очень малыми значениями  $\epsilon_{ij}$ . Согласно определению  $\epsilon_{ij}$  (15.33), это означает, что

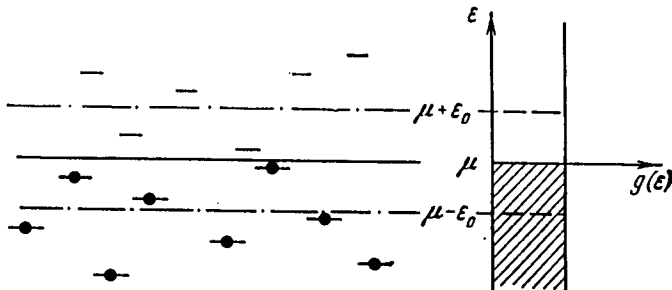


Рис. 9.1. Построение полосы, захватывающей состояния, уровни которых удалены от уровня Ферми не больше чем на  $\epsilon_0$ . Справа изображена плотность состояний, область занятых состояний заштрихована.

энергии  $\epsilon_i$  и  $\epsilon_j$  существенных состояний лежат в узкой полоске вблизи уровня Ферми, причем ширина этой полоски убывает при  $T \rightarrow 0$ . Если плотность состояний на уровне Ферми  $g(\mu)$  отлична от нуля, то во всей интересной при достаточно низких температурах области энергий плотность состояний  $g(\epsilon)$  можно считать постоянной и равной  $g(\mu)$  (рис. 9.1). По той же причине можно считать, что радиус локализованных состояний не зависит от энергии. Наконец, третье важное упрощение состоит в том, что, ввиду малой ширины полоски, состояния, попадающие в нее, находятся очень далеко друг от друга и их распределение в пространстве можно считать некоррелированным.

Для вывода (29.1) посмотрим, какую электропроводность могут обеспечить уровни, попадающие внутрь симметричной относительно уровня Ферми полоски, заданной неравенством

$$|\epsilon_i - \mu| \leq \epsilon_0 \quad (29.3)$$

(см. рис. 9.1). Концентрация состояний в полоске

$$N(\epsilon_0) = 2g(\mu)\epsilon_0. \quad (29.4)$$

Оценим с помощью (15.34) сопротивление, связанное с перескоком между двумя типичными узлами полоски. При этом вместо  $r_{ij}$  следует подставить типичное расстояние между узлами  $[N(\epsilon_0)]^{-1/2}$ , а вместо  $\epsilon_{ij}$  — величину  $\epsilon_0$ . Отсюда для удельного сопротивления, которое могут обеспечить прыжки по узлам, принадлежащим рассматриваемой полоске, получим

$$\rho = \rho_0 \exp \left\{ \frac{1}{N^{1/3}(\epsilon_0) a} + \frac{\epsilon_0}{kT} \right\} = \rho_0 \exp \left\{ \frac{1}{[g(\mu)\epsilon_0]^{1/3} a} + \frac{\epsilon_0}{kT} \right\}. \quad (29.5)$$

В этом выводе мы пренебрегаем численными множителями в показателе экспоненты удельного сопротивления. Проанализируем зависимость  $\rho(\epsilon_0)$ . Видно, что при больших  $\epsilon_0$  главную роль в показателе экспоненты (29.5) играет второе слагаемое. Оно приводит к уменьшению  $\rho(\epsilon_0)$  при убывании  $\epsilon_0$ . Однако при достаточно малых  $\epsilon_0$  первое и второе слагаемые сравниваются и дальнейшее уменьшение  $\epsilon_0$  приводит уже не к уменьшению  $\rho(\epsilon_0)$ , а к его росту. Это связано с тем, что в этом случае состояния в полоске оказываются очень редкими и определяющую роль играет уменьшение перекрытия между ними. Таким образом, благодаря конкуренции между перекрытием и активацией удельное сопротивление  $\rho(\epsilon_0)$  имеет острый минимум при

$$\epsilon_0 = \epsilon_0(T) \equiv \frac{(kT)^{3/4}}{[g(\mu)a^3]^{1/4}}. \quad (29.6)$$

В этом смысле полоску с шириной  $\epsilon_0(T)$  можно назвать оптимальной.

Естественно предположить, что электропроводность всей системы по порядку величины определяется оптимальной полоской. Тогда, подставляя (29.6) в (29.5), получим закон Мотта (29.1) и выражение (29.2) для  $T_0$ . Ясно, что численный коэффициент  $\beta$  таким способом определить нельзя.

Формула (29.6) оправдывает исходное предположение о том, что при достаточно малых температурах

ширина полоски состояний, ответственных за электропроводность, очень мала. Если производную  $\frac{d(\ln \rho)}{d(kT)^{-1}}$  назвать энергией активации, соответствующей данной температуре, то из закона Мотта следует, что энергия активации по порядку величины равна  $\epsilon_0(T)$  и монотонно убывает с понижением температуры пропорционально  $T^{3/4}$ . Поэтому электропроводность, следующую закону Мотта, иногда называют электропроводностью с монотонно убывающей энергией активации. Это название, конечно, охватывает не только закон (29.1), но и иные температурные зависимости, например все зависимости вида

$$\rho(T) = \rho_0 \exp \left\{ \left( \frac{T_0}{T} \right)^p \right\}, \quad \text{где } 0 < p < 1. \quad (29.7)$$

Посмотрим, как при изменении температуры изменяется средняя длина прыжка  $\bar{r}$ , т.е. среднее расстояние  $r_{ij}$  между состояниями в оптимальной полоске. Из (29.4) и (29.6), опуская численные множители, имеем

$$\bar{r} \approx [g(\mu) \epsilon_0(T)]^{-1/3} \approx a \left( \frac{T_0}{T} \right)^{1/4}. \quad (29.8)$$

Таким образом, при понижении температуры средняя длина прыжка растет пропорционально  $T^{-1/4}$ . Напомним, что в области применимости температурной зависимости (26.1) с постоянной энергией активации  $\epsilon_3$  средняя длина прыжка порядка перколяционного радиуса, или среднего расстояния между примесями, и при изменении температуры не меняется. Для того чтобы подчеркнуть характерное отличие электропроводности, следующей закону Мотта, ее называют электропроводностью с переменной длиной прыжка (*variable range hopping*).

Перейдем теперь к строгому выводу закона Мотта, основанному на применении метода протекания. Как всегда, следует начать с условия связности

$$\frac{2r_{ij}}{a} + \frac{\epsilon_{ij}}{kT} \leq \xi \quad (29.9)$$

и найти порог протекания  $\xi_c$ . Введем максимальные значения величины  $\epsilon_{ij}$  и расстояния  $r_{ij}$ , допускаемые

условием связности:

$$\varepsilon_{\max} = kT\xi \quad \text{и} \quad r_{\max} = \frac{a\xi}{2}. \quad (29.10)$$

Согласно (15.33)  $\varepsilon_{ij} = |\varepsilon_i - \varepsilon_j|$ , если энергии  $\varepsilon_i$  и  $\varepsilon_j$  лежат по разные стороны от уровня Ферми, и  $\varepsilon_{ij} = \max\{|\varepsilon_i - \mu|, |\varepsilon_j - \mu|\}$ , если энергии находятся по одну сторону. Поэтому из условия  $\varepsilon_{ij} \leq \varepsilon_{\max}$  вытекает, что  $|\varepsilon_i - \mu|, |\varepsilon_j - \mu| \leq \varepsilon_{\max}$ .

Введем теперь безразмерные переменные

$$s_i = \frac{r_i}{r_{\max}}, \quad \Delta_i = \frac{\varepsilon_i - \mu}{\varepsilon_{\max}}. \quad (29.11)$$

Тогда условие связности (29.9) примет вид

$$s_{ij} + \Delta_{ij} \leq 1, \quad (29.12)$$

где

$$s_{ij} = |s_i - s_j| = \frac{r_{ij}}{r_{\max}}, \quad (29.13)$$

$$\Delta_{ij} = \frac{1}{2} \{ |\Delta_i| + |\Delta_j| + |\Delta_i - \Delta_j| \}. \quad (29.14)$$

Концентрация узлов в безразмерном  $s$ -пространстве в полоске энергий  $|\Delta_i| < 1$ , согласно (29.4) и (29.10), равна

$$n(\xi) = 2g(\mu) \varepsilon_{\max} r_{\max}^3 = \frac{1}{4} g(\mu) kT a^3 \xi^4. \quad (29.15)$$

Нахождение  $\xi_c$  теперь свелось к решению следующей безразмерной задачи теории протекания. В пространстве заданы случайно расположенные узлы с концентрацией  $n(\xi)$ . Каждый узел в интервале  $|\Delta_i| < 1$  обладает случайной энергией  $\Delta_i$ , причем все значения  $\Delta_i$  равновероятны. Требуется найти пороговую концентрацию  $n(\xi) = n_c$ , при которой впервые возникает протекание, при условии, что критерий связности задан неравенством (29.12). Если  $n_c$  найдено, то, согласно (29.15),

$$\xi_c = \left[ \frac{4n_c}{g(\mu) kT a^3} \right]^{1/4}. \quad (29.16)$$

Подставляя (29.16) в (22.4), получим формулы (29.1) и (29.2) с

$$\beta = 4n_c. \quad (29.17)$$

В литературе имеется ряд оценок для  $n_c$ , полученных из различных соображений [1—5]. Эти оценки приводят к довольно сильно отличающимся значениям  $n_c$  в интервале от 2 до 7. Поставленная выше безразмерная задача теории протекания решалась методом Монте-Карло на ЭВМ в работе [6]. Среднее значение по 10 различным реализациям случайного массива из 1500 узлов оказалось равным  $5,7 \pm 0,3$ . Экстраполяция к бесконечному массиву в работе [6] не производилась. Если произвести экстраполяцию с помощью закона, аналогичного (16.13) и (18.10), то получаем  $n_c = 5,3 \pm 0,3$ , т. е.

$$\beta = 21,2 \pm 1,2. \quad (29.18)$$

Закон Мотта (29.1) легко допускает обобщение на случай пространства произвольной размерности  $d$ . В этом случае под  $g(\mu)$  в (29.4) следует понимать  $d$ -мерную плотность состояний, а величину  $[N(\epsilon_0)]^{1/3}a$  в формуле (29.5) заменить на  $[N(\epsilon_0)]^{1/d}a$ . Тогда можно повторить рассуждения, использованные в качественном выводе закона Мотта. В результате получим закон (29.7) с  $p = (d + 1)^{-1}$ . В частности, при  $d = 2$  [7]

$$\rho = \rho_0 \exp \left[ \left( \frac{\tilde{T}_0}{T} \right)^{1/3} \right], \quad (29.19)$$

где

$$\tilde{T}_0 = \frac{\bar{\beta}}{k\bar{g}(\mu) a^2}, \quad (29.20)$$

$\bar{g}(\mu)$  — двумерная плотность состояний на уровне Ферми.

Коэффициент  $\bar{\beta}$  был вычислен с помощью решения безразмерной задачи теории протекания, аналогичной задаче, рассматривавшейся в трехмерном случае. Согласно [6],

$$\bar{\beta} = 13,8 \pm 0,8. \quad (29.21)$$

В одномерном случае качественный вывод дает  $p = 1/2$ . Однако этот результат нельзя понимать буквально в связи с теми особенностями, которые имеет задача о протекании в одномерном случае. Вопрос о возможности применения формулы (29.7) с  $p = 1/2$  в одномерном случае разобран в работах [7—9].

До сих пор в этой главе рассматривалась модель, в которой плотность локализованных состояний не зави-



сит от энергии (см. рис. 9.1). В такой модели область применимости закона Мотта не ограничена ни при высоких, ни при низких температурах. Посмотрим теперь, во что переходит закон Мотта, если отказаться от предположения о постоянстве  $g(\epsilon)$ . Начнем со случая, когда плотность состояний скачком обращается в нуль на некотором расстоянии от уровня Ферми, т. е.

$$g(\epsilon) = \begin{cases} \frac{N}{2\Delta\epsilon}, & |\epsilon - \mu| \leq \Delta\epsilon, \\ 0, & |\epsilon - \mu| > \Delta\epsilon; \end{cases} \quad (29.22)$$

$N$  — число состояний в единице объема. Для повторения приведенного выше перколяционного вывода закона Мотта в этом случае достаточно, чтобы плотность состояний не зависела от энергии в области  $|\epsilon - \mu| \leq \leq \epsilon_{\max}$ , где  $\epsilon_{\max}$  определяется (29.10). Это позволяет сделать строгое утверждение, касающееся модели с плотностью состояний (29.22). Закон Мотта (29.1) справедлив в этом случае, если  $\epsilon_{\max} \leq \Delta\epsilon$ , т. е. при температурах  $T \leq T_c$ , где  $T_c$  определяется равенством

$$kT_c \xi_c(T_c) = \Delta\epsilon. \quad (29.23)$$

Используя (29.16) — (29.18), из (29.23) получаем

$$T_c = 0,29\Delta\epsilon (N^{1/3}a)^{-1}. \quad (29.24)$$

В § 28 показано, что если для типичной пары состояний выполняется неравенство (28.2), то прыжковая проводимость имеет постоянную энергию активации  $\epsilon_a$ . Легко убедиться, что для системы с плотностью состояний (29.22) неравенство (28.2) эквивалентно неравенству  $T \gg T_c$ . Таким образом, при изменении температуры в несколько раз в области  $T \geq T_c$  происходит переход от закона Мотта к температурной зависимости вида (26.1). Следует подчеркнуть, что такой переход не является уникальным свойством плотности состояний (29.22). Он возникает всегда, когда плотность состояний сосредоточена в конечной области энергий вблизи уровня Ферми.

Посмотрим теперь, как в переходной области изменяется локальная энергия активации

$$\epsilon_a = \frac{d \ln \rho}{d(kT)^{-1}} = \frac{d\xi_c}{d(kT)^{-1}}. \quad (29.25)$$

При  $T = T_c$  из (29.25) и (29.23) получаем, что  $\epsilon_a(T_c) = 1/4\Delta\epsilon$ . С другой стороны, плотности состояний (29.22) соответствует  $\epsilon_3 = 5/6\Delta\epsilon$  (см. (28.14)). Таким образом, в переходной области понижение температуры приводит к уменьшению  $\epsilon_a$  от  $\epsilon_3$  до  $0,3\epsilon_3$ . Если есть основания думать, что в рассматриваемой системе плотность состояний напоминает (29.22) и из эксперимента известна величина  $\epsilon_3$ , то с помощью условия  $\epsilon_a(T_c) = 0,3\epsilon_3$  на экспериментальной зависимости  $\ln \rho$  от  $T^{-1}$  можно указать температуру  $T_c$ , ниже которой должен быть справедлив закон Мотта.

Выше были проанализированы следствия обращения в нуль плотности состояний вне конечного интервала вокруг уровня Ферми. Исследуем теперь, какой вид имеет температурная зависимость прыжковой проводимости, если  $g(\epsilon)$  обращается в нуль на уровне Ферми. Поллак [3] и Гамильтон [10] рассмотрели случай, когда плотность состояний убывает по закону

$$g(\epsilon) = g' \left( \frac{|\epsilon - \mu|}{e'} \right)^n, \quad \text{где } n > 0. \quad (29.26)$$

В этом случае, заменяя формулу (29.4) соотношением

$$N(\epsilon_0) = g' \int_{\mu - \epsilon_0}^{\mu + \epsilon_0} \left( \frac{|\epsilon'' - \mu|}{e'} \right)^n d\epsilon'' = \frac{2}{n+1} \frac{g' \epsilon_0^{n+1}}{(e')^n} \quad (29.27)$$

и повторяя затем любой из вышесказанных выводов легко прийти к закону (29.7), в котором величина  $p$  связана с  $n$  соотношением

$$p = \frac{n+1}{n+4}. \quad (29.28)$$

При  $n = 0$ , т. е. при  $g(\epsilon) = \text{const}$ , из (29.28) мы снова получаем  $p = 0,25$ .

Перейдем теперь от рассмотрения различных моделей плотности состояний к обсуждению реальных физических систем — аморфных и кристаллических полупроводников. Обратимся сначала к экспериментальным данным о прыжковой проводимости с переменной длиной прыжка. Этому вопросу посвящено несколько десятков экспериментальных работ, весьма полную библиографию можно найти во взаимодополняющих статьях Хилла [11, 12] и Забродского [13].

Большая часть работ посвящена аморфным полупроводникам: германию, кремнию, углероду, арсениду галлия и т. д. Обычно в этих веществах при сравнительно высоких температурах электропроводность имеет энергию активации порядка половины ширины квазизапрещенной зоны, найденной по порогу оптического поглощения. Поэтому считают, что высокотемпературная электропроводность связана с возбуждением носителей в состояния, близкие к порогу подвижности (см. рис. 2.4). При уменьшении температуры энергия активации постепенно убывает. Разумно предположить, что после того, как энергия активации становится много меньше ширины квазизапрещенной зоны, наблюдается прыжковая проводимость с переменной длиной прыжка. С целью проверки закона Мотта строят зависимость  $\ln \rho$  от  $T^{-1/4}$  и смотрят, насколько она близка к прямой. Большинство экспериментальных зависимостей с некоторыми погрешностями удается «спрямить» в масштабе  $T^{-1/4}$ . Однако область изменения переменной  $T^{-1/4}$  обычно очень мала и трудно судить о том, насколько точно выполняется закон Мотта.

Хилл [12] и Забродский [13] проанализировали практически все опубликованные экспериментальные кривые другим методом. Они находили локальную энергию активации  $e_a = d \ln \rho / d(kT)^{-1}$  и анализировали зависимости  $\lg e_a$  от  $\lg T$ . Анализ показал, что у аморфных полупроводников встречаются различные значения показателя  $p$  — от 0,18 до 0,7. Однако большинство экспериментальных значений все же группируется около величины  $p = 0,25$ .

Перейдем теперь к наблюдениям прыжковой проводимости с переменной длиной прыжка в легированных кристаллических полупроводниках. На рис. 9.2 в качестве примера приведены температурные зависимости удельного сопротивления  $p$ -германия с  $K = 0,4$  при температурах до  $0,1^\circ\text{K}$ , полученные Шлимаком и Никулным [14].

Близкие по виду кривые были получены для слабо компенсированного  $n$ -Ge [15], для  $n$ -Ge с промежуточной компенсацией [16], для более сильно легированного и сильно компенсированного германия [17] и для арсенида галлия [18—20]. Анализ этих данных, выполненный

Забродским, показывает, что имеются любые значения  $\rho$  от 0,25 до 0,7, причем большинство значений группируется скорее около величины 0,5, нежели около 0,25, как у аморфных полупроводников.

В чем же состоит причина различия показателя  $\rho$  для аморфных и кристаллических полупроводников? Выше

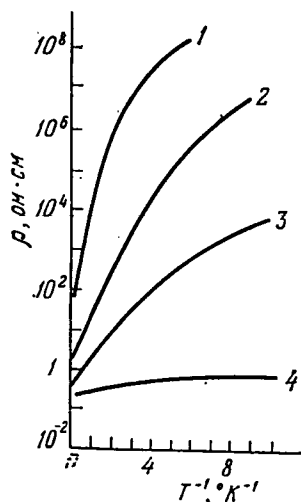


Рис. 9.2 Зависимость удельного сопротивления  $\rho$ -Ge с  $K = 0,4$  от обратной температуры [14] при различных концентрациях акцепторов (Ga).  $N_A$  (см $^{-3}$ ): 1— $3,5 \cdot 10^{16}$ ; 2— $8,2 \cdot 10^{16}$ ; 3— $1,3 \cdot 10^{17}$ ; 4— $2,0 \cdot 10^{17}$ .

(29.4)—(29.6) вместо  $g(\mu)$  фигурирует величина  $g(\mu + \epsilon_0(T))$ . Как показывают приведенные в гл. 10 оценки, в аморфных полупроводниках при понижении температуры сопротивление становится неизмеримо большим еще тогда, когда ширина оптимальной полосы  $\epsilon_0(T)$  во много раз больше  $\Delta$ . Таким образом, в аморфных полупроводниках кулоновская щель не сказывается на электропроводности.

Вне кулоновской щели плотность состояний аморфного полупроводника, строго говоря, тоже непостоянна.

говорилось, что значение  $\rho$  зависит от поведения плотности состояний вблизи уровня Ферми. В гл. 10 показано, что вследствие кулоновского взаимодействия электронов плотность локализованных состояний в непосредственной окрестности уровня Ферми убывает по закону (29.26) с  $n = 2$ . Этот минимум называют кулоновской щелью. Благодаря существованию кулоновской щели при достаточно низких температурах прыжковая проводимость с переменной длиной прыжка должна подчиняться закону (29.7) с  $\rho = 1/2$ . Асимптотическое поведение электропроводности достигается при таких температурах, когда ширина оптимальной полосы  $\epsilon_0(T)$  становится порядка ширины щели  $\Delta$ . При более высоких температурах  $\epsilon_0(T) \gg \Delta$ , так что кулоновская щель не влияет на электропроводность, а в формулах (29.2),

Она, по-видимому, монотонно убывает в глубь запрещенной зоны. Характерная энергия спада плотности состояний  $E$  определяется флуктуациями ближнего порядка и в сотни раз превышает ширину щели. Поэтому вполне может существовать большая температурная область, где ширина полосы  $\epsilon_0(T)$  меньше  $E$  и, следовательно, плотность состояний можно считать постоянной. В таких условиях должен выполняться закон Мотта. Из анализа экспериментальных данных можно сделать вывод, что во многих аморфных полупроводниках реализуется именно такая ситуация.

Плотность состояний легированного кристаллического полупроводника с учетом образования кулоновской щели изображена для разных степеней компенсации на рис. 10.1. Видно, что при промежуточной компенсации,  $K \approx 0,5$ , вообще не существует области энергий вокруг уровня Ферми, в которой плотность состояний можно было бы считать постоянной. Это означает, что при понижении температуры после области  $\epsilon_3$ -проводимости и беспараметрической переходной области электропроводность должна следовать закону (29.7) с  $\rho = 0,5$ . При предельно слабой и предельно сильной компенсации ширина кулоновской щели мала (см. рис. 10.1). Однако и вне щели плотность состояний сильно зависит от энергии.

Рассмотрим, например, случай сильной компенсации. В этом случае плотность состояний быстро убывает практически во всей области энергий от пика плотности состояний вблизи  $\epsilon = 0$  до кулоновской щели. В результате во всей области температур, соответствующей прыжковой проводимости с переменной длиной прыжка, показатель  $\rho$ , как и при промежуточной компенсации, должен быть больше, чем 0,25. Как мы видели выше, этот вывод подтверждается имеющимися экспериментальными данными.

### § 30. Магнитосопротивление в области действия закона Мотта

В § 24 было показано, что действие магнитного поля на волновую функцию усиливается с удалением от примесного центра. Поэтому в области прыжковой проводи-

мости с переменной длиной прыжка при достаточно низких температурах естественно ожидать большого положительного магнитосопротивления. Ниже мы изложим теорию магнитосопротивления в слабых и сильных (по терминологии § 24) полях в области действия закона Мотта ([22] к гл. 7). Обобщение на другие случаи прыжковой проводимости с переменной длиной прыжка не представляет труда.

Рассмотрим сначала область слабых магнитных полей. Согласно формуле (24.16) слабое поле создает относительно малую добавку в показателе экспоненты сопротивления  $\xi_{ij}$ :

$$\Delta\xi_{ij} = \frac{r_{ij}^3 a \sin^2 \phi}{12\lambda^4}, \quad (30.1)$$

где  $\phi$  — угол между вектором  $r_{ij}$  и магнитным полем,  $\lambda$  — магнитная длина. Слабое магнитное поле практически не меняет характерную длину прыжков. Поэтому, подставляя (29.8) в (30.1), получим, что добавка к  $\xi_c$  по порядку величины равна  $\frac{a^4}{\lambda^4} \left(\frac{T_0}{T}\right)^{3/4}$ , т. е. магнитосопротивление в слабом поле имеет вид

$$\ln \frac{\rho(H)}{\rho(0)} = t_1 \frac{e^2 a^4 H^2}{c^2 \hbar^2} \left(\frac{T_0}{T}\right)^{3/4}. \quad (30.2)$$

Здесь  $t_1$  — численный коэффициент. Согласно оценке, приведенной в Приложении 3,  $t_1 = 5/2016$ .

Перейдем теперь к сильному магнитному полю. Согласно (24.17) в этом случае критерий связности можно записать в виде

$$\frac{2|z_{ij}|}{a_H} + \frac{x_{ij}^2 + y_{ij}^2}{2\lambda^2} + \frac{e_{ij}}{kT} \leq \xi. \quad (30.3)$$

Введем безразмерные переменные

$$\tau = \frac{x}{x_{\max}}, \quad \eta = \frac{y}{y_{\max}}, \quad \zeta = \frac{z}{z_{\max}}, \quad \Delta_i = \frac{e_i - \mu}{e_{\max}}, \quad (30.4)$$

где

$$x_{\max} = y_{\max} = \sqrt{2} \lambda \xi^{1/2}, \quad z_{\max} = \frac{\xi a_H}{2} \quad (30.5)$$

и

$$e_{\max} = kT\xi$$

— максимальные значения соответствующих величин, допускаемые неравенством (30.3). В переменных (30.4) неравенство (30.3) принимает вид

$$\tau_{ij}^2 + \eta_{ij}^2 + |\zeta_{ij}| + \Delta_{ij} \leq 1, \quad (30.6)$$

где  $\tau_{ij} = \tau_i - \tau_j$ ,  $\eta_{ij} = \eta_i - \eta_j$ ,  $\zeta_{ij} = \zeta_i - \zeta_j$  и  $\Delta_{ij}$  определяется (29.14). Запишем число состояний с безразмерной энергией  $|\Delta_i| \leq 1$ , находящихся в единице объема безразмерного пространства  $(\tau, \eta, \zeta)$ , в виде

$$n(\xi) = 2g(\mu) \varepsilon_{\max} x_{\max} y_{\max} z_{\max} = 2g(\mu) \lambda^2 a_H k T \xi^3. \quad (30.7)$$

Нахождение  $\xi_c$  свелось теперь к решению новой безразмерной задачи теории протекания, которая состоит в следующем. В пространстве заданы случайно расположенные узлы с концентрацией  $n$ . Каждый узел обладает случайной энергией  $\Delta_i$ , равномерно распределенной в интервале  $|\Delta_i| < 1$ . Нужно найти пороговую безразмерную концентрацию  $n = n_{c1}$ , при которой впервые возникает протекание, определенное критерием связности (30.6). Согласно (30.7)  $\xi_c$  связано с  $n_{c1}$  следующим образом:

$$\xi_c = 2^{-1/3} n_{c1}^{1/3} [g(\mu) \lambda^2 a_H k T]^{-1/3}. \quad (30.8)$$

Согласно оценке, приведенной в Приложении 3,  $n_{c1} \approx 4,2$ . Подставляя это значение в (30.8), получаем

$$\rho(H) = \rho_0 \exp \left\{ \left[ \frac{T_0(H)}{T} \right]^{1/3} \right\}, \quad (30.9)$$

где

$$T_0(H) = \frac{2,1eH}{g(\mu) c h a_H k}.$$

Поскольку зависимость длины  $a_H$  от магнитного поля является довольно слабой (см. гл. 7), можно приближенно считать, что  $T_0(H) \propto H$ . Для того чтобы найти магнитное поле  $H_c$ , в котором происходит переход от (30.2) к (30.9), достаточно приравнять показатели экспонент этих формул. В результате находим

$$H_c \approx \frac{\hbar c}{e} \frac{[g(\mu) k T]^{1/4}}{a^{5/4}}. \quad (30.10)$$

Задача о магнитосопротивлении в области действия закона Мотта рассматривалась также Сапрапто и Бат-

чером [21]. Их результаты существенно отличаются от (30.2) и (30.9). Главная причина этого расхождения состоит в том, что они использовали вариационную формулу (24.10) для волновых функций на больших расстояниях от центра, где, как говорилось в гл. 7, эта формула неприменна.

Нам известна единственная экспериментальная работа, в которой наблюдалось экспоненциальное увеличение сопротивления с ростом магнитного поля в области прыжковой проводимости с переменной длиной прыжка. Гершензон и др. [22] исследовали магнитосопротивление слабо легированного и сильно компенсированного  $n$ -InSb. В слабых магнитных полях авторы констатировали хорошее согласие с формулой (30.2). В то же время в сильных полях температурная зависимость магнитосопротивления была более сильной, чем предсказывает формула (30.9).

### § 31. Зависимость прыжковой проводимости аморфных пленок от толщины

Интересное применение метода протекания связано с размерными эффектами в прыжковой электропроводности аморфных полупроводников. Аморфные полупроводники, в которых наблюдается прыжковая проводимость с переменной длиной прыжка, получают и исследуют в виде тонких пленок. Экспериментально установлено, что прыжковая электропроводность пленок сильно зависит от их толщины. Продольное удельное сопротивление пленки, как правило, растет с уменьшением толщины пленки, а поперечное, наоборот, убывает. Ниже мы будем рассматривать только продольное удельное сопротивление, поскольку для этого случая имеется больше экспериментальных данных и лучше разработана теория. В вопросе о поперечном сопротивлении мы ограничимся ссылками на оригинальные работы [23—25].

На рис. 9.3 в качестве примера приведены температурные зависимости продольного удельного сопротивления пленок углерода различной толщины, полученные Хаусером [26, 27]. Бросается в глаза сильная зависимость сопротивления от толщины. В принципе размерные эффекты могут быть связаны с зависимостью струк-



туры пленок от их толщины. В результате локальные, микроскопические электронные свойства пленок могут

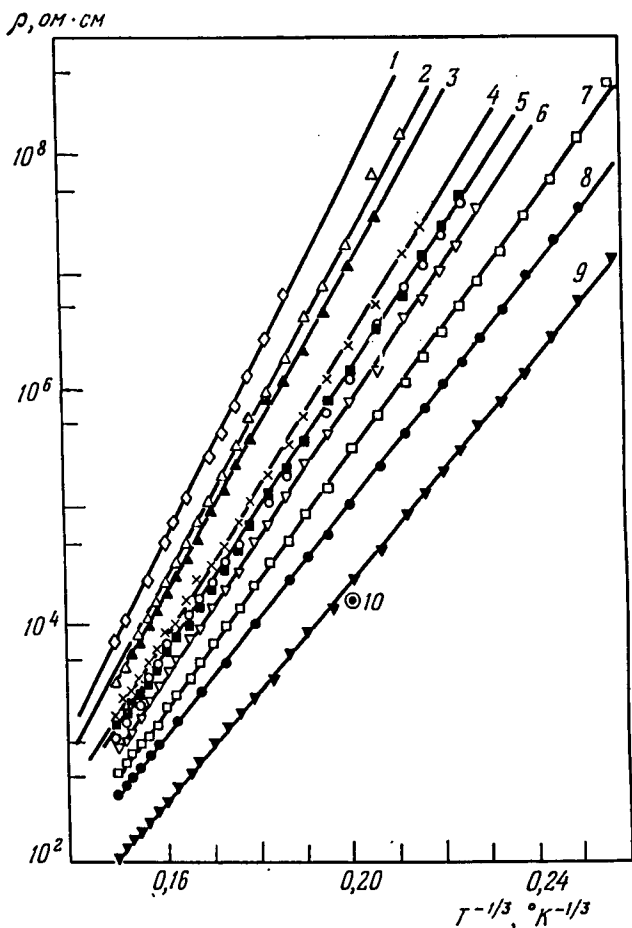


Рис. 9.3. Температурные зависимости удельного сопротивления аморфных пленок углерода различной толщины  $d$ , Å: 1 — 150; 2 — 200; 3 — 170; 4 — 225; 5 — 250, 300; 6 — 275; 7 — 400; 8 — 575; 9 — 1200. Значение 10 соответствует  $d = 6890$  Å [27].

меняться с толщиной. Например, существуют указания, что более толстые пленки имеют более упорядоченную структуру, нежели тонкие. Поэтому в них плотность состояний на уровне Ферми  $g(\mu)$  может быть меньше, что

может сильно увеличить прыжковое сопротивление. Такого рода структурные размерные эффекты мы не будем рассматривать. Наша цель состоит в изложении теории размерного эффекта, возникающего при неизменных локальных характеристиках вещества пленки благодаря ограниченности пространства, доступного для тока.

Постараемся сразу же объяснить основную идею теории. Прежде всего, ясно, что при нахождении продольного удельного сопротивления можно непосредственно использовать метод протекания. Другими словами, требуется найти порог протекания  $\xi_{cd}$ , т. е. такое значение  $\xi$ , при котором в задаче с критерием связности (29.9) возникает бесконечный кластер из сопротивлений, находящихся внутри пленки с толщиной  $d$ . Тогда

$$\rho(d) = \rho_0 \exp(\xi_{cd}). \quad (31.1)$$

Бесконечный кластер, образованный внутри пленки, мы будем называть пленочным бесконечным кластером. Представим себе мысленно, что материал пленки заполняет все пространство. Пусть  $\xi_c \equiv \xi_{c\infty}$  — порог возникновения бесконечного кластера в неограниченном пространстве (объемного бесконечного кластера). Как соотносятся друг с другом  $\xi_{cd}$  и  $\xi_c$ ?

На рис. 9.4 сплошными линиями изображен объемный бесконечный кластер. Ясно, что если пленочный бесконечный кластер существует, то он всегда является частью объемного. Однако не всегда часть объемного бесконечного кластера, попадающая внутрь пленки, является связной. Если объемный бесконечный кластер достаточно редкий, то внутри пленки он разбивается на не связанные между собой части, и пленочный бесконечный кластер отсутствует (именно такой случай изображен на рисунке). Таким образом, мы видим, что наличие границ пленки, запрещая часть связей между ее сопротивлениями, затрудняет образование пленочного бесконечного кластера по сравнению с объемным. Поэтому для создания протекания в пленке нужно включить больше сопротивлений, т. е. имеет место неравенство  $\xi_{cd} > \xi_c$ . Чем меньше толщина пленки, тем сильнее ограничения и тем больше  $\xi_{cd}$ . Это явление и должно приводить к экспоненциальному росту удельного сопротивления при уменьшении  $d$ .

Характер зависимости  $\xi_{ca}$  от  $d$  существенно определяется соотношением толщины пленки и средней длины прыжков для объемной электропроводности. Для определенности мы будем рассматривать ситуацию, в которой объемная электропроводность определяется законом Мотта. Тогда средняя длина прыжков, соответствующих сопротивлениям с  $\xi_{ij} \leq \xi_c$ , равна

$$\bar{r} = qr_{\max} = q \frac{a\xi_c}{2} = q \frac{a}{2} \left( \frac{T_0}{T} \right)^{1/4}, \quad (31.2)$$

где  $q$  — численный коэффициент, который, как можно показать, равен 0,5. В работах [10, 28—30] была найдена величина  $\xi_{ca}$  для случая очень тонких пленок, для

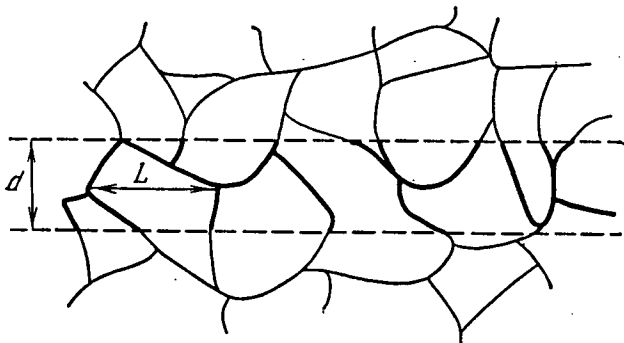


Рис. 9.4. К вопросу о пороге возникновения объемного и пленочного бесконечных кластеров. Границы пленки изображены штриховыми линиями, части объемного бесконечного кластера, попадающие внутрь пленки, — жирными линиями,  $L$  — радиус корреляции,  $d$  — толщина пленки.

которых  $d \ll \bar{r}$ . В этом случае все прыжки происходят практически параллельно плоскости пленки и проводимость имеет чисто двумерный характер. Поэтому температурная зависимость сопротивления тонких пленок должна описываться законом (29.19). Если задана трехмерная плотность состояний на уровне Ферми  $g(\mu)$ , то величина  $\bar{g}(\mu)$ , имеющая смысл числа состояний в единичном интервале энергии на единицу площади, связана с ней очевидным соотношением  $\bar{g}(\mu) = g(\mu)d$ . Подстановка этого соотношения в (29.20) дает для температуры

$\tilde{T}_0$  в законе (29.19)

$$\tilde{T}_0 = \frac{\tilde{\beta}}{kg(\mu) a^2 d}. \quad (31.3)$$

Согласно работам [26—28, 31], в тонких пленках германия, кремния и углерода действительно наблюдается переход от закона Мотта к закону (29.19). При этом для тонких пленок  $\tilde{T}_0$  растет с уменьшением  $d$  пропорционально  $d^{-1}$ , как предсказывает формула (31.3). Кнотек и др. [28] обратили внимание на то, что если из экспериментов с толстыми и очень тонкими пленками с помощью формул (29.1) и (29.19) найдены температуры  $T_0$  и  $\tilde{T}_0$ , то по этим температурам можно определить плотность состояний  $g(\mu)$  и радиус состояний  $a$ . Действительно, согласно (29.2) и (31.3) получаем

$$a = d \frac{\tilde{T}_0}{T_0} \frac{\beta}{\tilde{\beta}}. \quad (31.4)$$

Вычислив  $a$  с помощью (31.4), можно определить  $g(\mu)$ , подставляя  $a$  в любую из формул (29.2) и (31.3). В работе [28] считалось, что отношение  $\beta/\tilde{\beta} = 2$ , при этом для аморфного германия авторы получили  $a = 9,4 \text{ \AA}$ ,  $g(\mu) = 1,5 \cdot 10^{18} \text{ эв}^{-1} \cdot \text{см}^{-3}$ . Такой же анализ для аморфного кремния дает  $a = 3 \text{ \AA}$ ,  $g(\mu) = 3 \cdot 10^{19} \text{ эв}^{-1} \cdot \text{см}^{-3}$  и для аморфного углерода [27]  $a = 12 \text{ \AA}$ ,  $g(\mu) = 10^{18} \text{ эв}^{-1} \cdot \text{см}^{-3}$ . Согласно расчетам, выполненным в [6],  $\beta$  и  $\tilde{\beta}$  определяются формулами (29.18) и (29.21), а отношение  $\beta/\tilde{\beta} = 1,53$ . Использование этого значения  $\beta/\tilde{\beta}$  в (31.4) приводит к величинам  $a$ , составляющим 76% от цитированных выше, и в 2,3 раза увеличивает значения плотности состояний  $g(\mu)$ .

До сих пор говорилось об очень тонких пленках, для которых выполняется неравенство  $d \ll \bar{r}$ . Обратимся теперь к более толстым пленкам, для которых  $d \gg \bar{r}$  и электропроводность имеет трехмерный характер. В работе [32] были высказаны эвристические соображения, позволяющие определить зависимость  $\xi_{ca}$  от  $d$  в этом случае. Представим себе, как мы делали это выше, что материал пленки заполняет все трехмерное пространство. Зададим значение  $\xi$ , большее  $\xi_c$ , и включим все сопротивления с  $\xi_{ij} \leq \xi$ . В результате возникнет объем-

ный бесконечный кластер (см. рис. 9.4). Радиус корреляции этого бесконечного кластера запишем в виде

$$L(\xi) \approx \bar{r} \left( \frac{\xi_c}{\xi - \xi_c} \right)^{v_3}. \quad (31.5)$$

Величина  $\bar{r}$  в формуле (31.5) играет роль естественного масштаба длины, так же как и величина  $N^{-1/3}$  в формуле (18.9). И в том и в другом случаях этим масштабом является среднее расстояние между узлами. Только в случае прыжков между ближайшими соседями входит среднее расстояние между любыми узлами, а в случае закона Мотта — среднее расстояние между узлами, энергия которых попадает в разрешенную полосу  $\varepsilon_{ij} \leq \varepsilon_{\max}$ .

Если  $\xi$  настолько близко к  $\xi_c$ , что радиус корреляции  $L$  много больше толщины пленки, то объемный бесконечный кластер разбивается внутри пленки на изолированные, не связанные друг с другом части. Постепенно увеличивая  $\xi$ , можно сделать радиус корреляции  $L$  меньше толщины пленки. Тогда внутри пленки окажется много слоев сетки объемного бесконечного кластера. Несмотря на наличие границ, они образуют пленочный бесконечный кластер. Из этого рассуждения можно сделать вывод, что порог протекания  $\xi_{cd}$  по порядку величины определяется условием

$$L(\xi_{cd}) = d. \quad (31.6)$$

Подставляя (31.5) в (31.6), получим

$$\xi_{cd} = \xi_c \left[ 1 + D \left( \frac{\bar{r}}{d} \right)^{1/v_3} \right], \quad (31.7)$$

где  $D$  — неизвестная численная постоянная, а  $v_3 \approx 0,9$  — трехмерный индекс радиуса корреляции. Условие (31.6) можно поставить под сомнение, если исходить из картины бесконечного кластера в виде строго периодической, например простой кубической, решетки с периодом  $L$ . Тогда оказалось бы, что если одна из плоскостей решетки типа (100) параллельна поверхности пленки и попадает внутрь ее, то пленочный бесконечный кластер, образованный этой плоскостью, существует и при  $L \gg d$ . Однако здесь следует вспомнить, что бесконечный кластер является случайной сеткой, в которой такие выде-

ленные плоскости отсутствуют. Поэтому при  $L \gg d$  плечный бесконечный кластер не возникает.

Конечно, приведенные выше эвристические соображения требуют численной проверки. Косвенная проверка на примере задачи узлов была предпринята Шейнманом [33]. Представим себе простую кубическую решетку размера  $l \times l \times n$ , которая при  $l \gg n$  будет служить моделью пленки с толщиной  $n$ . Пусть

$$x_c(n) = \lim_{l \rightarrow \infty} x_{cl}(n), \quad (31.8)$$

где  $x_{cl}(n)$  — среднее по большому числу реализаций значение порога протекания между двумя противоположными малыми гранями решетки  $l \times n$ . Приведенные выше эвристические соображения нетрудно обобщить для нахождения зависимости  $x_c(n)$  от толщины пленки  $n$ . Радиус корреляции объемного бесконечного кластера в единицах постоянной решетки при  $x > x_c$  равен  $L(x) \approx (x - x_c)^{-\nu_3}$ , где  $x_c$  — порог протекания трехмерной задачи узлов. Величину  $x_c(n)$  следует находить из условия  $L[x_c(n)] \approx n$ , которое приводит к результату, очень похожему на (31.7):

$$x_c(n) = x_c [1 + D_1 n^{-1/\nu_3}], \quad (31.9)$$

где  $D_1$  — численный коэффициент. Шейнман вычислил  $x_c(n)$  для  $n = 1, 2, 3, 4, 5, 7$  и  $9$  методом Монте-Карло. Он нашел, что приближение  $x_c(n)$  к  $x_c$  с ростом  $n$  хорошо описывается формулой (31.9) со значениями  $\nu_3 \approx 1,00 \pm 0,08$ , близкими к известному из независимых расчетов индексу радиуса корреляции  $\nu_3 \approx 0,9$ . Результаты Шейнмана подтверждают соображения, на которых основан вывод формул (31.9) и (31.7).

Рассмотрим теперь следствия, вытекающие из формулы (31.7). Из этой формулы видно, что, несмотря на малость отношения  $\bar{r}/d$ , поправка к  $\xi_c$  за счет размерных эффектов оказывается больше единицы в довольно широкой области изменения толщин  $\bar{r}\xi_c^{\nu_3} > d > \bar{r}$  (типичное значение  $\xi_c$  в эксперименте равно 30). Во всей этой области толщин удельное сопротивление экспоненциально зависит от толщины. Используя (31.1), (31.2) и (31.7), получим, что эта экспоненциальная зависимость имеет

ВИД

$$\begin{aligned} \rho(d) &= \rho(\infty) \exp \left\{ D \left( \frac{T_0}{T} \right)^{1/4} \left( \frac{\bar{r}}{d} \right)^{1/\nu_3} \right\} = \\ &= \rho(\infty) \exp \left\{ C \left( \frac{T_0}{T} \right)^{\frac{\nu_3+1}{4\nu_3}} \left( \frac{a}{d} \right)^{1/\nu_3} \right\}, \quad (31.10) \end{aligned}$$

где  $\rho(\infty)$  — удельное сопротивление трехмерной среды, а  $C = D(q/2)^{1/\nu_3}$ . При  $d = \bar{r}$  размерная поправка к  $\xi_c$  оказывается порядка  $\xi_c$ . К такому же результату при  $d \approx \bar{r}$  приводят и формулы (29.19), (31.3), справедливые для очень тонких пленок  $d \ll \bar{r}$ . Таким образом, результаты обеих теорий «сшиваются» при  $d \approx \bar{r}$ . Толщина пленки  $\bar{r}\xi_c^{\nu_3}$ , начиная с которой возникает экспоненциальная зависимость, имеет ясный физический смысл. В § 21 мы видели, что прыжковая электропроводность определяется всегда критической подсеткой, которая возникает при включении всех сопротивлений с  $\xi_{ij} \leq \xi_c + 1$ . Радиус корреляции этой подсетки, согласно (31.5), равен  $\bar{r}\xi_c^{\nu_3}$ . Поэтому до тех пор, пока  $d$  много больше, чем  $\bar{r}\xi_c^{\nu_3}$ , пригодно макроскопическое описание и уменьшение толщины создает лишь малые поправки к удельному сопротивлению. Однако, начиная с  $d \approx \bar{r}\xi_c^{\nu_3}$ , уменьшение толщины начинает качественно перестраивать пути протекания тока и возникает экспоненциальная зависимость сопротивления от  $d$ .

Сравним теперь формулу (31.10) с экспериментальными данными Хаусера [26, 27], представленными на рис. 9.3. Из (31.10) видно, что для проверки зависимости  $\rho$  от  $d$  и определения экспериментального значения показателя  $\nu_3$  удобно строить график зависимости  $\lg \lg \frac{\rho(d)}{\rho(\infty)}$  от  $\lg d$  (рис. 9.5).

Основная трудность в этой процедуре заключается в выборе величины объемного удельного сопротивления  $\rho(\infty)$ . В качестве  $\rho(\infty)$  мы приняли значение удельного сопротивления для пленки с толщиной 6890 Å, приведенное в [27]. На рисунке приведены результаты обработки данных при  $T = 123^\circ\text{K}$ . Видно, что практически все точки находятся вблизи прямой линии. Определение ее наклона дает значение  $\nu_3 = 0,85$ , хорошо согласую-

щеся с известными значениями трехмерного индекса корреляции (см. табл. 5.2). Точка, соответствующая  $d = 1200 \text{ \AA}$ , не попадает на прямую. Нужно, однако, учесть, что эта точка наиболее чувствительна к выбору  $\rho(\infty)$ . Небольшим уменьшением  $\rho(\infty)$  (примерно в 2 раза) эту точку можно уложить на одну прямую с осталь-

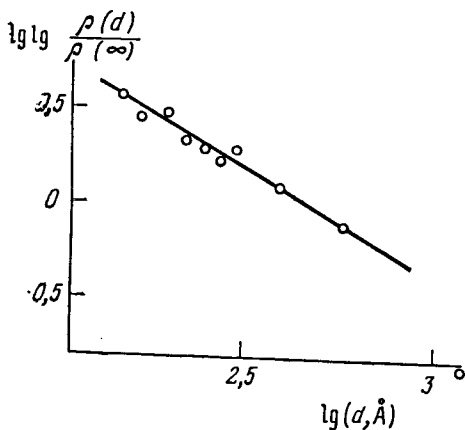


Рис. 9.5. Обработка экспериментальных зависимостей удельного сопротивления от толщины пленки при  $T = 123 \text{ }^\circ\text{K}$ .

ными. Более подробное обсуждение экспериментальных данных для пленок углерода, германия и кремния можно найти в работе [34].

### § 32. Предэкспоненциальный множитель прыжковой проводимости

До сих пор мы все время интересовались только экспоненциальными зависимостями прыжковой проводимости. Как мы видели в § 21, метод протекания позволяет также с точностью до численного коэффициента вычислить предэкспоненциальный множитель электропроводности сильно неоднородной среды. Согласно (21.18) для решеточной задачи с сопротивлениями (21.3) удельное сопротивление можно записать в виде

$$\rho = \rho_0 \exp \xi_c, \quad \rho_0 \approx R_0 L_0, \quad (32.1)$$



где  $R_0$  — предэкспоненциальный множитель (21.3), а  $L_0$  — радиус корреляции критической подсетки, т. е. бесконечного кластера, соответствующего  $\xi = \xi_c + 1$ . Вывод формул (21.18) и (32.1) легко повторить для задачи на случайных узлах с сопротивлениями  $R_{ij} = R_0 \exp \xi_{ij}$ , где  $R_0$  — постоянная величина.

В теории прыжковой проводимости предэкспоненциальный множитель  $R_{ij}^0$  сопротивления элемента сетки Миллера и Абрахамса, согласно (15.36) и (15.18), зависит от  $r_{ij}$  и энергий  $\epsilon_i$  и  $\epsilon_j$ . Поэтому возникает вопрос, какие значения  $r_{ij}$ ,  $\epsilon_i$  и  $\epsilon_j$  окажутся существенными при выводе выражения, аналогичного (32.1). По смыслу вывода формул (21.18) и (32.1) сопротивления, определяющие электропроводность кубика со стороной  $l = L_0$  и, следовательно, макроскопическую электропроводность, имеют значения  $\xi_{ij}$ , близкие к  $\xi_c$ . Поэтому при вычислении предэкспоненты прыжкового удельного сопротивления  $\rho_0$  вместо  $R_0$  в (32.1) нужно подставлять величину  $R_{ij}^0$  для типичных сопротивлений с  $\xi_{ij} = \xi_c$ . Мы обозначим ее  $(R_{ij}^0)_{\xi_{ij} = \xi_c}$ . Тогда

$$\rho_0 \approx L_0 \cdot (R_{ij}^0)_{\xi_{ij} = \xi_c}. \quad (32.2)$$

Значения  $\xi_c$  и  $L_0$  различны для случаев электропроводности с постоянной и переменной длиной прыжка. Поэтому выражения для  $\rho_0$  в этих случаях сильно отличаются друг от друга.

Рассмотрим сначала случай  $\epsilon_3$ -проводимости. В этом случае в выражения (15.36) и (15.18), определяющие  $R_{ij}^0$ , надо подставлять  $|\epsilon_i - \epsilon_j| \approx \epsilon_3$  и  $r_{ij} \approx r_c \approx N^{-1/3}$ . Величина  $L_0$  находится подстановкой  $\xi = \xi_c = 1$  в (18.9) и равна  $N^{-1/3} \xi_c^v$ . Используя формулу (22.6) для  $\xi_c$ , из (32.2), (15.36) и (15.18) получаем

$$\rho_{03} = \eta_1 \rho_0' Q(N) (N^{1/3} a)^{1-v} \frac{kT}{\epsilon_3}, \quad (32.3)$$

где

$$\rho_0' = \frac{9\pi}{4} \frac{a^3 ds^5 \hbar^4 \chi^2}{E_1^2 e^6} \quad \text{и} \quad Q(N) = \left[ 1 + \left( \frac{\eta_2 \epsilon_3 a}{2\hbar s} \right)^2 \right]^4, \quad (32.4)$$

а  $\eta_1$  и  $\eta_2$  — неизвестные численные коэффициенты порядка единицы.

В (32.3) выделены множители, зависящие от температуры и концентрации примесей ( $\epsilon_3 \approx e^2 N^{1/2} / \kappa$ ), и постоянная величина  $\rho'_0$ . Особое внимание нам хочется обратить на множитель  $Q(N)$ , который обычно опускают. Напомним, что, по смыслу вывода (15.18), этот множитель описывает ослабление взаимодействия фононов с примесными электронами при  $qa \gtrsim 1$ , связанное с осцилляциями смещения, создаваемого акустической волной в объеме, где волновая функция электрона не мала (см. (15.15)). Хотя параметр  $\epsilon_3 a / 2\hbar s$  обычно порядка единицы, множитель  $Q(N)$  может быть порядка 100. С другой стороны, множитель  $(N^{1/2} a)^{1-\nu}$  можно всегда считать величиной порядка единицы, так как, согласно табл. 5.2, трехмерный индекс радиуса корреляции  $\nu$  близок к 0,9.

До сих пор обработка экспериментальных данных с учетом степенных зависимостей  $\rho_{03}$  от  $T$  и  $N$ , вытекающих из (32.3), не производилась. Емельяненко и др. (см. [7] к гл. 6) обработали экспериментальные данные для пяти полупроводников с помощью формул (26.1) и (22.9), в которых предэкспонента  $\rho_{03}$  считалась величиной, не зависящей от  $T$  и  $N$ . Для этих же полупроводников они по формуле (32.4) вычислили величины  $\rho'_0$ . Оказалось, что во всех случаях экспериментальное значение  $\rho_{03}$  в 10—20 раз больше, чем расчетное значение постоянной  $\rho'_0$ , что в принципе можно отнести за счет больших значений  $Q(N)$ . С другой стороны, Кахлерт и др. (см. [13] к гл. 7) констатировали, что для GaAs экспериментальное значение  $\rho_{03}$  близко к  $\rho'_0$ .

Перейдем теперь к электропроводности, подчиняющейся закону Мотта (29.1). В этом случае в выражение для  $(R_{ij}^0)_{\xi_{ij} - \xi_c}$ , согласно (29.9), (29.10), следует подставлять  $|\epsilon_i - \epsilon_j| \approx \epsilon_{\max} = kT\xi_c$ ,  $r_{ij} \approx r_{\max} = a\xi_c/2$ , где  $\xi_c = (T_0/T)^{1/4}$ . При этом, учитывая, что типичное значение величины  $|\epsilon_i - \epsilon_j| \propto T^{3/4}$  и убывает при  $T \rightarrow 0$ , можно пренебречь вторым слагаемым в квадратной скобке (15.18). Радиус корреляции критической подсетки  $L_0$  определяется подстановкой  $\xi - \xi_c = 1$  в (31.5) и равен  $\bar{r}\xi_c^\nu \approx a(T_0/T)^{(1+\nu)/4}$ . Подставляя все эти оценки в (32.2),

получаем для величины  $\rho_0$  в формуле (29.1)

$$\rho_0 \approx \rho'_0 \xi_c^{\nu-2} = \rho'_0 \left( \frac{T_0}{T} \right)^{(\nu-2)/4}. \quad (32.5)$$

Формула (15.18), которую мы использовали при получении результата (32.5), справедлива, если огибающая волновая функция локализованных состояний  $F_l(r)$  имеет водородоподобный вид (2.17).

В аморфных полупроводниках локализованные состояния вблизи уровня Ферми, по-видимому, связаны с мелкомасштабными структурными дефектами. По этой причине асимптотика их волновых функций скорее всего имеет вид, характерный для короткодействующих потенциалов:  $F(r) \sim r^{-1} \exp(-r/a)$ . Если с такой волновой функцией заново произвести расчеты, ранее приводившие к формуле (15.18), то получится выражение, отличающееся от (15.18) дополнительным множителем порядка  $(a/r_l)^4$ . Наличие этого множителя изменяет температурную зависимость предэкспоненты. По аналогии с выводом (32.5) легко получить, что для аморфного полупроводника

$$\rho_0 \approx \rho'_0 \left( \frac{T_0}{T} \right)^{(\nu+2)/4}. \quad (32.6)$$

Выше, так же как при обсуждении предэкспоненциального множителя в § 21, мы следовали работам Куркьярви ([35] к гл. 5) и авторов ([102] к гл. 5 и [35]). Ранее предэкспоненциальный множитель другими методами вычисляли Поллак ([17] к гл. 4), Киркпатрик ([100] к гл. 5 и [36]) и Скал и Шкловский ([53] к гл. 5). Недавно этот вопрос был снова рассмотрен Батчером [37]. Результаты разных авторов несколько отличаются друг от друга. К сожалению, мы не имеем возможности дать здесь сравнительный анализ цитированных работ.

## ГЛАВА 10

### ВЛИЯНИЕ КОРРЕЛЯЦИОННЫХ ЭФФЕКТОВ НА ПЛОТНОСТЬ СОСТОЯНИЙ И ПРЫЖКОВУЮ ПРОВОДИМОСТЬ

В предыдущих главах теория плотности состояний и прыжковой проводимости строилась в приближении самосогласованного поля. Суть приближения состояла в том, что потенциал, действующий на некоторый донор, считался не зависящим от энергии этого донора и определялся с помощью некоторой самосогласованной процедуры. Кроме того, в теории прыжковой проводимости игнорировались временные корреляции чисел заполнения (см. гл. 4). Ниже будет показано, что эти методы хорошо работают в предельных случаях сильной и слабой компенсации, однако при промежуточной компенсации они неприменимы или нуждаются в дополнительном обосновании.

К моменту написания книги исследование корреляционных эффектов далеко от завершения. Некоторые вопросы неясны, в точках зрения разных авторов имеются значительные противоречия. Поэтому в книге мы уделили этой фундаментальной проблеме сравнительно скромное место. Тем не менее мы убеждены в том, что без детальной разработки теории корреляционных эффектов полное понимание экспериментальных данных по низкотемпературной прыжковой проводимости кристаллических полупроводников невозможно (по крайней мере при промежуточных степенях компенсации). В этой главе изложены идеи и результаты, представляющиеся наиболее четкими, хотя весь ее материал вряд ли можно считать устоявшимся.

### § 33. Кулоновская щель в плотности состояний

а) Теория кулоновской щели. Задача о плотности состояний слабо легированного и компенсированного полупроводника в области примесной зоны была сформулирована и частично рассмотрена в гл. 3. При малой концентрации примесей квантовые эффекты, связанные с перекрытием соседних состояний, можно считать малыми, а сами состояния — строго локализованными (см. гл. 2). Тогда рассматриваемая система состоит из узлов трех разных видов, хаотически расположенных в пространстве. Эти узлы суть доноры, отдавшие свои электроны акцепторам и заряженные положительно, нейтральные доноры и акцепторы, которые при рассматриваемых температурах всегда имеют электроны и заряжены отрицательно (для определенности мы рассматриваем полупроводник  $n$ -типа). Все примеси считаются неподвижными, однако электроны могут переходить с одного донора на другой, вследствие чего электронная система приходит в термодинамическое равновесие. При нулевой температуре распределение электронов по донорам определяется минимумом полной электростатической энергии

$$H = \frac{e^2}{x} \left[ \frac{1}{2} \sum_k^{\text{дон}} \sum_{k' \neq k}^{\text{дон}} \frac{(1-n_k)(1-n_{k'})}{r_{kk'}} - \sum_k^{\text{дон}} \sum_l^{\text{акц}} \frac{(1-n_k)}{r_{kl}} + \frac{1}{2} \sum_l^{\text{акц}} \sum_{l' \neq l}^{\text{акц}} \frac{1}{r_{ll'}} \right]. \quad (33.1)$$

Здесь  $n_k$  — числа заполнения доноров;  $n_k = 0$ , если донор  $k$  ионизован, и  $n_k = 1$ , если донор  $k$  нейтрален. Суммирование ведется по всем координатам доноров и акцепторов в соответствии с указаниями над знаками сумм. Набор чисел заполнения  $\{n_k\}$  должен быть найден из условия минимума (33.1) при фиксированном значении полного числа электронов  $\sum_k n_k$ . Вместо этого можно искать безусловный минимум функционала

$$\tilde{H} = H - \mu \sum_k^{\text{дон}} n_k, \quad (33.2)$$

где  $\mu$  — химический потенциал (уровень Ферми), который следует затем определить из условия нейтральности. В первую очередь нас будут интересовать не сами числа заполнения, а одноэлектронные энергии  $\varepsilon_i$ . Величина  $\varepsilon_i$  есть потенциальная энергия электрона на узле  $i$ , создаваемая всеми прочими заряженными узлами системы. Это определение справедливо и для заполненных, и для пустых узлов. То же самое определение использовалось и в гл. 3. Ниже, однако, нам будет удобно отсчитывать энергии не от уровня изолированной примеси, как в гл. 3, а от уровня Ферми  $\mu$ , т. е. наряду с  $\varepsilon_i$  ввести величины

$$\bar{\varepsilon}_i \equiv \varepsilon_i - \mu = - \sum_k^{\text{дон}} \frac{(1 - n_k)}{r_{ik}} + \sum_l^{\text{акц}} \frac{1}{r_{il}} - \mu. \quad (33.3)$$

Вычисление наборов  $\{n_i\}$  и  $\{\bar{\varepsilon}_i\}$ , соответствующих основному состоянию (минимуму энергии (33.1)), представляет крайне сложную многоэлектронную задачу. Действительно, заполнение узла  $i$  определяется энергией  $\bar{\varepsilon}_i$ , которая в свою очередь зависит от всех прочих чисел заполнения. В гл. 3 эта задача была решена в предельных случаях слабой ( $K \ll 1$ ) и сильной ( $1 - K \ll 1$ ) компенсации. Оказалось, что при слабой компенсации можно выделить 0- и 2-комплексы, определяющие положение уровня Ферми и слабо взаимодействующие друг с другом. При сильной компенсации решающим обстоятельством было наличие большой корреляционной длины, что позволило ввести самосогласованное поле, усредненное по флуктуациям малого масштаба.

В этом параграфе речь пойдет главным образом о случае промежуточной компенсации, когда в сформулированной выше классической задаче заведомо нет никаких малых параметров. Единственной величиной с размерностью энергии является  $\varepsilon_D = e^2/\kappa r_D$ , т. е. энергия взаимодействия на среднем расстоянии между примесями. Эта энергия и определяет ширину примесной зоны. Естественно думать, что набор энергий  $\varepsilon_i$  описывается колоколообразной функцией распределения, а уровень Ферми при промежуточной компенсации находится где-то вблизи максимума. Однако эта точка зрения неправоверна.

Рассматривая взаимодействие локализованного электрона с ближайшими к нему зарядами, Поллак и Кнотек [1, 2] показали, что плотность состояний должна иметь минимум вблизи уровня Ферми, т. е. заполненные состояния отделены от пустых щелью. Аналогичный результат получил Сринивасан [3]. Куросава и Сагимото [4] обнаружили такого рода минимум, моделируя систему доноров и акцепторов методом Монте-Карло. Мы показали [5], что благодаря дальнедействующему характеру кулоновского потенциала одночастичная плотность состояний обращается в нуль на уровне Ферми. Эти результаты были подтверждены численными экспериментами [6, 7].

Вычисленная плотность состояний примесной зоны слабо легированного полупроводника изображена на рис. 10.1. Минимум плотности состояний в окрестности уровня Ферми получил название «кулоновской щели».

Заметим, что плотность состояний имеет минимум и в модели Лифшица, описанной в § 9. Эта модель не учитывает взаимодействия электронов, а принимает во внимание разброс за счет квантового перекрытия состояний. Поэтому указанное сходство носит скорее случайный характер: за ним не стоят никакие физические причины. Более того, сходство является довольно

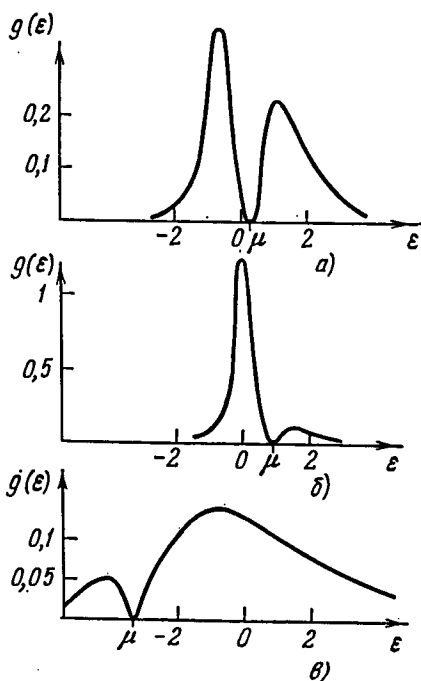


Рис. 10.1. Плотность состояний слабо легированного компенсированного полупроводника при различных степенях компенсации с учетом образования кулоновской щели. Энергия  $\epsilon$  дана в единицах  $e^2 N_D^{1/3} \kappa^{-1}$ , плотность состояний — в единицах  $N_D (e^2 N_D^{1/3} \kappa^{-1})^{-1}$ . а)  $K = 0,5$ ; б)  $K = 0,1$ ; в)  $K = 0,9$ .

поверхностным. Так, в модели Лифшица плотность состояний в минимуме не обращается в нуль, а сам минимум привязан к энергии изолированного уровня, в то время как учет электрон-электронного взаимодействия приводит к тому, что плотность состояний обращается в нуль при энергии, равной энергии Ферми.

В отсутствие малого параметра единственный возможный подход к минимизации выражения (33.2) состоит в том, чтобы произвести минимизацию сначала по отношению к изменению одного (любого) числа заполнения, потом по отношению к одновременному изменению двух чисел заполнения и т. д. Численные эксперименты, выполненные на ЭВМ [7], показывают, что эта процедура сходится очень хорошо.

Легко убедиться в том, что положительность приращения  $\bar{H}$  при изменении числа заполнения  $n_i$  на доноре  $i$  достигается лишь при условии, что

$$n_i = \begin{cases} 0, & \text{если } \bar{\epsilon}_i > 0, \\ 1, & \text{если } \bar{\epsilon}_i < 0. \end{cases} \quad (33.4)$$

Условие (33.4) означает просто, что состояния с энергиями ниже уровня Ферми должны быть заполненными, а выше уровня Ферми — пустыми. Численный эксперимент, проделанный в [7], показывает, что если удовле-

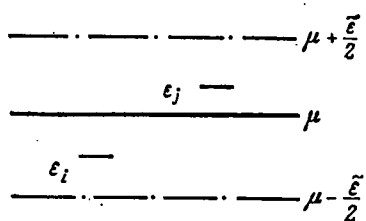


Рис. 10.2. Энергетические уровни в полосе около уровня Ферми.

творены только условия (33.4) для всех доноров, то энергия  $\bar{H}$  отличается от энергии основного состояния всего на несколько процентов.

Перейдем к следующему приближению и рассмотрим процесс, состоящий в переносе электрона с заполненного в основном состоянии донора  $i$  на пустой донор  $j$  (рис. 10.2). С помощью (33.1) можно получить приращение  $\Delta_i^j$  энергии  $\bar{H}$ , связанное с этим процессом, и написать условие его положительности:

$$\Delta_i^j = \bar{\epsilon}_j - \bar{\epsilon}_i - e^2/\kappa r_{ij} > 0. \quad (33.5)$$



Выражение (33.5) для  $\Delta_i^j$  можно получить также из следующих соображений. Разобьем процесс переноса электрона на два этапа. На первом этапе перенесем его с заполненного донора  $i$  на бесконечность (считая для определенности, что потенциальная энергия на бесконечности равна нулю), а на втором этапе перенесем его из бесконечности на пустой донор  $j$ . Работа, которую нужно произвести на первом этапе, по определению равна  $-\bar{\epsilon}_i$ . Работа, необходимая на втором этапе, равняется  $\bar{\epsilon}_j$  при условии, что заполнение всех доноров соответствует основному состоянию. Однако, на втором этапе донор  $i$  пуст, тогда как в основном состоянии он должен быть заполнен. Следовательно, при вычислении работы нужно считать, что на узле  $i$  находится положительно заряженная дырка, притяжение к которой уменьшает искомую работу на величину, равную  $e^2/\kappa r_{ij}$ .

Это уменьшение можно объяснить также следующим образом. Величина  $\Delta_i^j$  равна работе, затраченной на перенос электрона, или равна разности потенциальных энергий переносимого электрона на донорах  $j$  и  $i$ . Однако по определению в энергию  $\bar{\epsilon}_j$  включен потенциал, создаваемый самим переносимым электроном, находящимся на доноре  $i$ . Эту часть потенциальной энергии следует исключить, чтобы получить искомую работу. В некотором смысле эту процедуру можно назвать исключением самодействия. Наконец, еще один способ интерпретации (33.5) состоит в том, что это есть энергия, нужная для рождения электронно-дырочной пары с учетом взаимодействия электрона и дырки, находящихся в локализованных состояниях.

Согласно (33.4) неравенство (33.5) должно выполняться для любой пары доноров с  $\bar{\epsilon}_j > 0$  и  $\bar{\epsilon}_i < 0$ . Оно и приводит к кулоновской щели в плотности состояний  $g(\bar{\epsilon})$ . Из (33.5) следует, что состояния, близкие к уровню Ферми, должны находиться на значительном расстоянии друг от друга. Рассмотрим доноры, чьи энергии попадают в узкую полосу от  $-\bar{\epsilon}/2$  до  $+\bar{\epsilon}/2$  в окрестности уровня Ферми (см. рис. 10.2). Согласно (33.5) расстояние  $r_{ij}$  между двумя любыми донорами этой полосы, лежащими по разные стороны от уровня Ферми ( $\bar{\epsilon}_i \cdot \bar{\epsilon}_j < 0$ ), не может быть меньше  $e^2/\kappa \bar{\epsilon}$ . Следовательно, концентрация доноров  $n(\bar{\epsilon})$  в полосе шириной  $\bar{\epsilon}$  не может превы-

шать  $\bar{\epsilon}^3 \kappa^3 / e^6$ , а плотность состояний  $g(\bar{\epsilon}) = dn(\bar{\epsilon})/d\bar{\epsilon}$  обращается в нуль при  $\bar{\epsilon} \rightarrow 0$  не медленнее, чем по закону  $\bar{\epsilon}^2$ .

Легко понято, что более резкого закона в приближении, основанном на неравенстве (33.5), быть не может. Действительно, уменьшение плотности состояний связано с сильным взаимодействием между состояниями, имеющими малую энергию. Закон  $g(\bar{\epsilon}) \sim \bar{\epsilon}^2$  означает, что среднее расстояние между уровнями с энергией, меньшей  $\bar{\epsilon}$ , порядка  $e^2/\kappa\bar{\epsilon}$ , а энергия взаимодействия на этом расстоянии порядка собственной энергии уровней, т. е. порядка  $\bar{\epsilon}$ . Допустим, что  $g(\bar{\epsilon})$  спадает резче, чем  $\bar{\epsilon}^2$ . Тогда среднее расстояние оказалось бы больше  $e^2/\kappa\bar{\epsilon}$ , а энергия взаимодействия меньше  $\bar{\epsilon}$ . Ясно, что такое слабое взаимодействие не может быть причиной для уменьшения плотности состояний. На основании этих соображений нами была предложена формула

$$g(\bar{\epsilon}) = \frac{\alpha_3 \bar{\epsilon}^2 \kappa^3}{e^6}, \quad (33.6)$$

где  $\alpha_3$  — численный коэффициент.

В двумерном случае совершенно аналогичные рассуждения приводят к закону

$$g(\bar{\epsilon}) = \frac{\alpha_2 |\bar{\epsilon}| \kappa^2}{e^4}. \quad (33.7)$$

Здесь  $g$  относится к единице площади и потому имеет другую размерность, чем в трехмерном случае;  $\alpha_2$  — численный коэффициент. Следует отметить, что выражения (33.6), (33.7) не зависят явно от концентрации примесей или степени компенсации. Каждое из них представляет единственно возможную по соображениям размерности комбинацию из энергии и эффективного заряда электрона  $e/\sqrt{\kappa}$ . Это обстоятельство побудило авторов [5] выдвинуть предположение о том, что в приближении, учитывающем лишь неравенства (33.5), плотность состояний в окрестности уровня Ферми имеет универсальный вид, зависящий лишь от числа измерений пространства, т. е. константы  $\alpha_2$  и  $\alpha_3$  одинаковы для любой неупорядоченной системы с локализованными электронами и с кулоновским взаимодействием.

Как видно из предыдущих рассуждений, плотность состояний с энергией порядка  $\bar{\epsilon}$  определяется кулоновским взаимодействием на расстоянии порядка  $e^2/\kappa\bar{\epsilon}$ , т. е. особенность плотности состояний вблизи уровня Ферми связана с дальнодействием. Эта особенность в свою очередь приводит к ослаблению электронного экранирования. (Напомним, что радиус экранирования линейной теории обращается в бесконечность при стремлении к нулю плотности состояний на уровне Ферми.) Таким образом, в целом картина оказывается согласованной.

Выражения (33.6), (33.7) справедливы, если энергия  $\bar{\epsilon}$  столь мала, что расстояние  $e^2/\kappa\bar{\epsilon}$  значительно превосходит среднее расстояние между донорами. Отсюда получаем неравенство  $\bar{\epsilon} \ll \epsilon_D$ . При промежуточной степени компенсации, естественно, оказалось, что ширина кулоновской щели порядка всей ширины примесной зоны.

Эфрос [8] предложил самосогласованный метод для описания плотности состояний в окрестности уровня Ферми, учитывающий только неравенства (33.5). Он получил интегральные уравнения, из которых в пределе малых энергий следуют выражения (33.6), (33.7).

Численные коэффициенты, найденные с помощью этих уравнений в работе [7], оказались следующими:  $\alpha_2 = 2/\pi$ ,  $\alpha_3 = 3/\pi$ .

В работе Барановского и др. [7] плотность состояний определялась с помощью численного эксперимента на ЭВМ. Рассматривалась модель, согласно которой доноры находятся в узлах простой решетки, акцепторы отсутствуют вовсе, электронов вдвое меньше, чем доноров, а для соблюдения нейтральности заряд донора принят равным половине заряда электрона. Каждому донору присваивается случайное значение энергии, лежащее в интервале от  $-A$  до  $A$ . Если  $A = 0$ , то все доноры эквивалентны и электроны заполняют их так, что в основном состоянии заполненные и пустые доноры образуют решетку типа NaCl. При отличных от нуля значениях  $A$  распределение электронов по донорам в основном состоянии оказывается непериодическим, причем при малых энергиях образуется кулоновская щель. Согласно гипотезе универсальности вид плотности состояний в окрестности уровня Ферми не должен зависеть от модели. В работе [7] было показано также, что плотность

состояний при малых энергиях удовлетворяет законам (33.6), (33.7), причем величины  $\alpha_2$  и  $\alpha_3$  в пределах точности расчета не отличались от полученных с помощью самосогласованного уравнения. В двумерном случае было также показано, что коэффициент  $\alpha_2$  не зависит от  $A$ .

Следует иметь в виду, что предыдущие рассуждения относились лишь к системам с локализованными электронными состояниями. В хороших металлах кулоновская щель не образуется: электрон-электронное взаимодействие слабо влияет на структуру элементарных возбуждений в окрестности уровня Ферми.

Обсудим теперь следующее приближение, состоящее в минимизации выражения (33.2) по отношению к одновременному переносу двух электронов. Этот вопрос подробно рассматривался в работе [8]. Результат состоит в следующем: в двумерном случае ограничений на плотность состояний более существенных, чем неравенства (33.5), не возникает. Поэтому есть все основания считать, что результат (33.7) является окончательным. В трехмерном случае ситуация гораздо сложнее. Условия минимума полной энергии по отношению к одновременному переносу двух электронов, по-видимому, приводят к тому, что при достаточно малых энергиях плотность состояний убывает не по закону (33.6), а значительно более резко:

$$g(\bar{\epsilon}) \sim \exp\left(-\beta \sqrt{\frac{\Delta}{|\bar{\epsilon}|}}\right). \quad (33.8)$$

Здесь  $\Delta$  — ширина кулоновской щели (при промежуточной компенсации  $\Delta \approx \epsilon_D$ ),  $\beta$  — неизвестный численный коэффициент. Ниже мы дадим качественное объяснение этого результата.

Возбуждение с энергией  $\bar{\epsilon}_i$  возникает при переносе электрона с донора  $i$  на бесконечность (или наоборот, если в основном состоянии донор  $i$  пуст). Однако в системе существует еще одна ветвь элементарных возбуждений с малой энергией, соответствующая переносу электрона на близлежащий донор. Как уже говорилось, энергия, нужная для этого процесса, равна  $\Delta_i$  и выражается формулой (33.5).

Пусть  $F(\omega)d\omega$  — число пар доноров в единице объема, у которых энергия  $\Delta_i^j$  лежит в интервале от  $\omega$  до  $\omega + d\omega$ . Возникает вопрос: как ведет себя функция  $F(\omega)$  при малых значениях  $\omega$ ? Строго обоснованного ответа на этот вопрос не существует, однако в работе [8] приводятся аргументы в пользу того, что при  $\omega \rightarrow 0$  функция  $F(\omega)$  стремится к отличному от нуля пределу порядка  $N_D/\epsilon_D$ . Благодаря кулоновской щели в функции  $g(\bar{\epsilon})$ , основной вклад в  $F(\omega)$  при малых  $\omega$  дают не состояния с малыми  $\bar{\epsilon}$ , а состояния с  $\bar{\epsilon}$  порядка  $\epsilon_D$ . При этом, согласно (33.5), расстояние между донорами, осуществляющими возбуждение с малой энергией, не может быть большим, так как член  $e^2/\kappa r_{ij}$  должен точно компенсировать большую положительную величину  $\bar{\epsilon}_j - \bar{\epsilon}_i$  (напомним, что  $\bar{\epsilon}_j > 0$ ,  $\bar{\epsilon}_i < 0$ ).

Назовем пары доноров с малой энергией возбуждения резонансными парами. Резонансные пары сильно взаимодействуют с одночастичными возбуждениями, имеющими малую энергию  $\bar{\epsilon}$ , что и приводит к существенному изменению плотности состояний  $g(\bar{\epsilon})$ . Допустим, что дополнительный электрон помещен на узел  $j$ . Потенциал, создаваемый этим электроном, вызовет перемещение электронов внутри большого числа резонансных пар, окружающих узел  $j$  (рис. 10.3). Вследствие этого в окрестности рассматриваемого донора возникает дополнительная поляризация, создаваемая дополнительными моментами возбужденных пар. Поляризация создает потенциал, понижающий энергию электрона на доноре  $j$ .

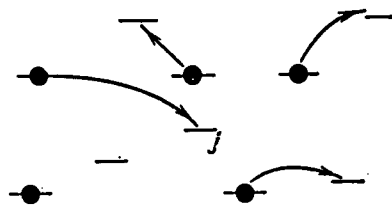


Рис. 10.3. Образование «поляронной атмосферы» при помещении электрона на узел  $j$ .

Описанное явление совершенно аналогично полярному эффекту, существующему в ионных кристаллах. В нашем случае поляризация возникает, однако, не за счет смещения ионов, а за счет электронных переходов внутри резонансных пар. (Представление об «электронных поляронах» было введено независимо в работах [8,

9].) Поляронный эффект создает «жесткую» щель в одночастичной плотности состояний. Действительно, допустим для простоты, что конфигурации резонансных пар и поляронные сдвиги для всех узлов с малой энергией  $\bar{\epsilon}_j$  одинаковы. Тогда сразу же получим, что наименьшая энергия пустого донора  $\bar{\epsilon}_i$  должна быть выше наименьшей энергии заполненного донора  $\bar{\epsilon}_i$  на величину, равную поляронному сдвигу. В противном случае было бы энергетически выгодно перенести электрон с донора  $i$  на донор  $j$  и произвести переполаризацию «поляронных атмосфер» в окрестности каждого донора. Таким образом, в рассматриваемом идеализированном случае должна существовать абсолютная щель, разделяющая заполненные и пустые состояния (плотность состояний в некоторой области точно равна нулю). Если учесть, что конфигурации резонансных пар в окрестностях разных доноров разные, то вместо абсолютной щели возникнет закон (33.8).

Итак, в трехмерном случае, наряду с одноэлектронными возбуждениями, имеет смысл вводить многоэлектронные возбуждения (полярны). Энергию полярона  $\psi_i$ , находящегося на узле  $i$ , следует определить как взятую со знаком минус работу, которую нужно совершить, чтобы снять электрон с узла  $i$ , перенести его на бесконечность (где его энергия равна  $\mu$ ) и перестроить поляронную атмосферу в окрестности образовавшейся дырки. Аналогичным образом, энергия полярона  $\psi_j$ , который может появиться на пустом узле  $j$ , равна работе, нужной, чтобы перенести электрон из бесконечности на узел  $j$  и перестроить атмосферу узла  $j$ . В основном состоянии для всех заполненных узлов  $\psi_i < 0$ , а для пустых  $\psi_j > 0$ .

Найдем теперь работу, требуемую для переноса полярона с заполненного узла  $i$  на пустой узел  $j$ , находящийся далеко от узла  $i$ . Так как каждая пара атмосферы создает лишь дипольный потенциал, то взаимодействие образовавшихся электрона и дырки можно считать чисто кулоновским. Поэтому, аналогично (33.5), получим

$$\Delta_i^j = \psi_j - \psi_i - \frac{e^2}{\kappa r_{ij}}, \quad (33.9)$$

причем в основном состоянии  $\Delta'_i > 0$ . Из этого неравенства следует, что при малых значениях  $\psi$  плотность поляронных состояний  $g(\psi) \propto \psi^2$ , точно так же как из (33.5) следует (33.6). Таким образом, плотность поляронных состояний должна иметь «мягкую» щель.

Как следует из (33.8), поляронные эффекты должны проявляться в одночастичной плотности состояний при  $\bar{\epsilon} < \Delta$ . Тем не менее в численных экспериментах [7] в трехмерном случае наблюдался закон (33.6), а не (33.8) вплоть до  $\bar{\epsilon} \approx 0,15\Delta$ . Видимо, это означает, что коэффициент  $\beta$  мал и поляронный эффект имеет место лишь при очень малых энергиях.

Для определенности в предыдущих рассуждениях фигурировал полупроводник с промежуточной степенью компенсации. Однако, по существу, все сказанное относится к любой неупорядоченной системе с локализованными состояниями. Специфической характеристикой системы является лишь ширина щели  $\Delta$ .

Перейдем теперь к случаю, когда энергия кулоновского взаимодействия между двумя локализованными носителями тока, расположенными на среднем расстоянии друг от друга, мала по сравнению с характерной энергией разброса уровней. По разным причинам такая ситуация реализуется в примесной зоне кристаллических полупроводников при сильной и слабой компенсации, а также в аморфных полупроводниках. Рассмотрим модель, в которой разброс уровней имеет некулоновскую природу и приводит к тому, что плотность состояний на уровне Ферми без учета кулоновского взаимодействия равна  $g_0$ . Как было показано выше, неравенства (33.5) понижают плотность состояний вблизи уровня Ферми вплоть до значений, определяемых (33.6), (33.7). Однако такие рассуждения имеют смысл, пока эти значения малы по сравнению с  $g_0$ .

Определим ширину кулоновской щели  $\Delta$ , приравняв правые части (33.6), (33.7) к  $g_0$ . При  $\bar{\epsilon} \ll \Delta$  справедливы все соображения, высказанные выше для полупроводника с промежуточной компенсацией. Однако можно показать, что при  $\bar{\epsilon} \gg \Delta$  кулоновские эффекты не приводят ни к каким ограничениям на плотность состояний и она равна  $g_0$ . В результате плотность состояний в рассматриваемой модели выглядит, как показано на рис. 10.4, причем в

трехмерном случае

$$\Delta = \frac{e^3 g_0^{1/2}}{\kappa^{3/2}}. \quad (33.10)$$

Именно так обстоит дело в аморфном полупроводнике при условии, что характерная энергия спада плотности состояний в запрещенной зоне значительно больше  $\Delta$ . Для аморфного германия, используя значение  $g_0 = 1,5 \cdot 10^{18} \text{ эВ}^{-1} \cdot \text{см}^{-3}$  ([28] к гл. 9), получаем  $\Delta \approx 1 \text{ мэВ}$ , так что указанное условие хорошо выполняется.

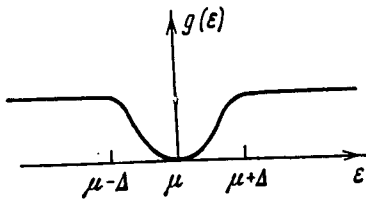


Рис. 10.4. Плотность состояний аморфного полупроводника вблизи уровня Ферми.

Рассмотрим теперь случай слабо компенсированного полупроводника. Строго говоря, весь разброс донорных уровней связан здесь с кулоновским взаимодействием, а существенные расстояния велики по сравне-

нию с межатомными. Однако и в этом случае можно выделить эффекты, приводящие к кулоновской щели, от эффектов, создающих разброс уровней, которые рассмотрены в гл. 3.

Допустим, что кулоновский потенциал обрезан на расстояниях, больших по сравнению со средним расстоянием между донорами. Энергетическая схема донорных уровней, описанная в § 11, изменится при этом очень мало. Согласно (11.16) плотность состояний на уровне Ферми порядка  $g_0 \approx KN_D/\epsilon_D$ . Вернем теперь кулоновскому потенциалу дальнююдействующую часть. Тогда возникнет кулоновская щель, ширину которой можно оценить по формуле (33.10). Легко видеть, что ширина  $\Delta$  мала по сравнению с энергией  $\epsilon_D$  и с амплитудой крупномасштабного потенциала. Поэтому наличие щели мало сказывается на результатах §§ 11, 12. В случае сильной компенсации ситуация аналогична. В непосредственной близости от уровня Ферми возникает кулоновская щель, однако ее ширина мала по сравнению с амплитудой крупномасштабного потенциала  $\gamma(r_s)$ . Поэтому существование кулоновской щели не влияет на теорию нелинейного экранирования, развитую в § 13. Плотность со-



стояний с учетом кулоновской щели изображена на рис. 10.1 при разных степенях компенсации.

б) Возможные проявления кулоновской щели. Обсудим теперь вопрос о том, в каких физических явлениях должна проявляться кулоновская щель плотности состояний. Одним из наиболее прямых способов ее обнаружения является туннельная спектроскопия. Эксперименты с барьерами Шоттки действительно указывают на то, что плотность состояний в полупроводнике с локализованными электронными состояниями имеет резкий минимум, находящийся точно на уровне Ферми [10]. Существенно, что если уровень Ферми смещается (например, под действием компенсации), то минимум следует за ним. Возможно, что это есть проявление кулоновской щели (впервые такая идея была высказана в работе Куросавы и Сагimoto [4]). Как уже говорилось, благодаря кулоновской щели радиус экранирования линейной теории обращается в бесконечность и экранирование слабых потенциалов носит нелинейный характер [8]. В принципе это обстоятельство также может быть проверено экспериментом.

Кулоновская щель должна ярко проявляться в прыжковой проводимости с переменной длиной прыжка. Однако прежде, чем перейти к рассмотрению прыжковой проводимости при промежуточной компенсации, следует устранить важный дефект в формулировке задачи о сетке сопротивлений Миллера и Абрахамса, данной в гл. 4. В рамках использованного в этой главе самосогласованного приближения каждому донору соответствует нефлуктуирующая во времени энергия  $\epsilon_i$ , связанная со средними числами заполнения уравнением (15.21). Через энергии  $\epsilon_i$  выражается число переходов в единицу времени, а затем и сопротивление  $R_{ij}$ . В частности, при образовании (15.20) в (15.22) энергия фонона  $\Delta_i^j$ , поглощаемого при переходе  $i \rightarrow j$ , заменяется на  $\epsilon_j - \epsilon_i$ . Такая замена справедлива, когда состояния  $i$  и  $j$  находятся по одну сторону от уровня Ферми. Однако если в основном состоянии донор  $i$  занят, а донор  $j$  пуст, то, заменив  $\Delta_i^j$  на  $\epsilon_j - \epsilon_i$ , мы учитываем нефизическое самодействие перескакивающего электрона. Как следует из обсуждения, связанного с неравенством (33.5), действительное

значение энергии фонона, в котором самодействие исключено, равно

$$\Delta_i^j = \varepsilon_j - \varepsilon_i - \frac{e^2}{\kappa r_{ij}} \quad (33.11)$$

при условии, что  $\varepsilon_j > \mu$ , а  $\varepsilon_i < \mu$ .

Покажем теперь, каким образом можно сформулировать модель сетки сопротивлений, учитывая исключение самодействия. Введем величины  $F(n_i, n_j)$ , определив их как вероятности того, что числа заполнения доноров  $i$  и  $j$  равны  $n_i$  и  $n_j$ , в то время как на остальных донорах числа заполнения  $\tilde{n}_k$  соответствуют основному состоянию. Если электрическое поле  $E$  отсутствует, эти вероятности равны

$$\begin{aligned} F^0(n_i, n_j) &= \\ &= C^0 \exp \left\{ -\frac{1}{kT} \left[ \varphi_i n_i + \varphi_j n_j - \mu(n_i + n_j) + \frac{e^2}{\kappa r_{ij}} n_i n_j \right] \right\}, \end{aligned} \quad (33.12)$$

где

$$\varphi_i = \frac{e^2}{\kappa} \left[ \sum_l^{\text{акц}} \frac{1}{|r_i - r_l|} - \sum_{k \neq i, j}^{\text{дон}} \frac{1 - \tilde{n}_k}{|r_i - r_k|} - \frac{1}{|r_i - r_j|} \right] \quad (33.13)$$

— потенциальная энергия электрона на доноре  $i$  при условии, что донор  $j$  пуст,  $\varphi_j$  — аналогичная величина для донора  $j$ , а  $C$  — нормировочная постоянная. Индекс 0 означает, что соответствующая величина относится к случаю  $E = 0$ .

Электрическое поле изменяет распределение электронов. При  $E \neq 0$  наиболее общее выражение для вероятностей  $F(n_i, n_j)$  имеет вид

$$\begin{aligned} F(n_i, n_j) &= C \exp \left\{ -\frac{1}{kT} \left[ \varphi_i n_i + \varphi_j n_j - \mu(n_i + n_j) + \right. \right. \\ &\quad \left. \left. + \frac{e^2}{\kappa r_{ij}} n_i n_j + \delta\mu_i n_i + \delta\mu_j n_j + \delta\mu_{ij} n_i n_j \right] \right\}, \end{aligned} \quad (33.14)$$

причем в слабом поле величины  $\delta\mu_i$ ,  $\delta\mu_j$  и  $\delta\mu_{ij}$  пропорциональны  $E$ . Кроме того, электрическое поле изменяет энергию фонона:

$$\Delta_i^j = \varepsilon_j - \varepsilon_i - \frac{e^2}{\kappa r_{ij}} + eEr_{ij} = \varphi_j - \varphi_i + eEr_{ij}. \quad (33.15)$$

Ток между донорами  $i$  и  $j$  равен

$$\begin{aligned} \mathcal{J}_{ij} &= -e [\langle \gamma_{ij} n_i (1 - n_j) \rangle - \langle \gamma_{ji} n_j (1 - n_i) \rangle] = \\ &= -e [\gamma_{ij} F(1, 0) - \gamma_{ji} F(0, 1)]. \end{aligned} \quad (33.16)$$

Подставляя (33.15) в выражение (15.17) для  $\gamma_{ij}$  и аналогичное выражение для  $\gamma_{ji}$ , используя (33.14) и линеаризуя (33.16) по электрическому полю, можно получить сетку сопротивлений, которая отличается от сетки, определенной формулами (15.33)–(15.35), лишь значениями  $\varepsilon_{ij}$ :

$$\varepsilon_{ij} = \begin{cases} |\varepsilon_j - \varepsilon_i| - \frac{e^2}{\chi r_{ij}}, & (\varepsilon_i - \mu)(\varepsilon_j - \mu) < 0, \\ \max[|\varepsilon_i - \mu|, |\varepsilon_j - \mu|], & (\varepsilon_i - \mu)(\varepsilon_j - \mu) > 0. \end{cases} \quad (33.17)$$

Сравнивая (33.17) с (15.33), видим, что исключение самодействия, как и следовало ожидать, приводит к уменьшению  $\varepsilon_{ij}$  на величину  $e^2/\chi r_{ij}$  в том случае, когда энергии  $\varepsilon_i$  и  $\varepsilon_j$  находятся по разные стороны уровня Ферми.

Посмотрим теперь, как это сказывается на вычислении электропроводности. Начнем с области температур, где прыжковая проводимость имеет постоянную энергию активации  $\varepsilon_3$ . В этой области при сильной и слабой компенсации электроны, как правило, совершают прыжки по одну сторону уровня Ферми. Поэтому исключение самодействия не изменяет результатов, полученных для этих случаев в гл. 8. С другой стороны, при промежуточной компенсации, когда электрон часто пересекает уровень Ферми, исключение самодействия должно играть важную роль и может заметно уменьшить значение  $\varepsilon_3$ .

Перейдем теперь к области прыжковой проводимости с переменной длиной прыжка. В этом случае за проводимость ответственна полоска состояний с энергиями, лежащими по обе стороны уровня Ферми, и прыжки с  $(\varepsilon_i - \mu)(\varepsilon_j - \mu) < 0$  происходят часто. Однако легко убедиться, что если ширина оптимальной полоски  $\varepsilon_0(T)$  больше ширины кулоновской щели  $\Delta$ , то для типичной пары доноров  $i, j$  кулоновское слагаемое  $-e^2/\chi r_{ij}$  в (33.17) пренебрежимо мало по сравнению с  $|\varepsilon_j - \varepsilon_i|$ . Таким образом, в тех случаях, когда справедлив закон Мотта, исключение самодействия не играет существенной роли.

Рассмотрим теперь случай, когда ширина оптимальной полоски  $\epsilon_0(T)$  меньше ширины кулоновской щели  $\Delta$ . Как говорилось в гл. 9, правильный, с нашей точки зрения, результат для электропроводности получится, если, имея в виду формулу (33.6), подставить в (29.28)  $n = 2$ . Тогда

$$\ln \sigma \propto \exp\left[-\left(\frac{T_1}{T}\right)^{1/2}\right]. \quad (33.18)$$

Повторяя выкладки, приведенные в § 29, с учетом явного выражения (33.6) для плотности состояний найдем

$$T_1 = \frac{\beta_1 e^2}{\kappa a}, \quad (33.19)$$

где  $\beta_1$  — численный коэффициент.

Нам представляется, что формулы (33.18), (33.19) описывают прыжковую проводимость в пределе низких температур. Ниже мы попробуем аргументировать эту точку зрения. Нетривиальный момент вывода состоял в том, что мы использовали плотность состояний (33.6) так, как если бы она была результатом решения некоей одноэлектронной задачи с заданным внешним потенциалом. Многоэлектронная природа задачи проявляется в том, что, согласно (33.11), энергия электронного перехода между парой доноров не всегда определяется разностью энергий  $\epsilon_j - \epsilon_i$ . Более того, как говорилось выше, существуют резонансные пары, для которых энергия перехода мала, несмотря на большое значение разности  $\epsilon_j - \epsilon_i$ . Именно такими парами определяется поглощение низкочастотного излучения. Однако они не могут давать существенный вклад в статическую проводимость, поскольку доноры, образующие пару, близки друг к другу и изолированы от других пар.

Рассмотрим оптимальную полоску, которая возникает при выводе формулы (33.18) с помощью (33.6) по рецепту, данному в § 29. Легко показать, что среднее расстояние между донорами, попавшими в эту полоску,  $\bar{r}$ , связано с ее энергетической шириной  $\epsilon_0(T)$  соотношением  $\bar{r} \approx e^2/\kappa\epsilon_0(T)$ . Отсюда следует, что кулоновский член  $e^2/\kappa r_{ij}$  (33.11) порядка типичных значений разности  $\epsilon_j - \epsilon_i$ . Однако он не может быть близок к  $\epsilon_j - \epsilon_i$  для всех пар доноров в полоске. Поэтому мы считаем, что учет этого

члена сказывается лишь на численном коэффициенте в формуле (33.19), но не на зависимостях, которые следуют из (33.18).

Предыдущие рассуждения основывались на формуле (33.6). Однако, как говорилось выше, в трехмерном случае необходимо учитывать поляронный эффект, приводящий к «экспоненциальной щели» (33.8). В работе [8] приводятся аргументы в пользу того, что перенос заряда должен осуществляться поляронами, но благодаря тому, что плотность состояний поляронов имеет щель вида (33.6), формулы (33.18) и (33.19) справедливы и в этом приближении.

Продолжим теперь начатое в гл. 9 обсуждение экспериментальных данных о прыжковой проводимости с переменной длиной прыжка. Прежде всего заметим, что в аморфных полупроводниках переход от закона Мотта (29.1) к закону (33.18) должен происходить при очень низких температурах. Для того чтобы оценить температуру перехода, следует приравнять ширину оптимальной полоски (29.6) и ширину щели (33.10). Для аморфного германия температура, найденная из этого уравнения при условии, что  $g(\mu) = 1,5 \cdot 10^{18} \text{ эв}^{-1} \cdot \text{см}^{-3}$ ,  $a = 9,4 \text{ \AA}$  ([28] к гл. 9), оказывается порядка  $1^\circ\text{К}$ . При таких температурах сопротивление аморфных полупроводников, как правило, столь велико, что не поддается измерению.

Перейдем к обсуждению экспериментальных данных по низкотемпературной проводимости кристаллических полупроводников. В некоторых случаях (см. [12, 13] к гл. 9) найденные экспериментально значения  $p$  в (29.7) близки к 0,5 во всей области температур и тенденции к уменьшению  $p$  при  $T \rightarrow 0$  не наблюдается. Поэтому кажется разумным объяснить эти данные наличием кулоновской щели на уровне Ферми. Здесь, однако, надо иметь в виду следующее обстоятельство. Температура  $T_1$ , определенная из эксперимента, сильно зависит от концентрации примесей и степени компенсации. Наличие концентрационной зависимости  $T_1$  при фиксированной степени компенсации  $K = 0,4$  хорошо видно на рис. 9.2. Если эти данные анализировать с помощью формулы

$$T_1(N) \propto \left( \frac{N_M - N}{N_M} \right)^{\eta}, \quad (33.20)$$

где  $N_M$  — концентрация, соответствующая переходу Мотта, то получается  $\eta \approx 1,5$ . Как видно из формулы (33.19), убывание  $T_1$  при  $N \rightarrow N_M$  в принципе может быть связано с двумя причинами: возрастанием диэлектрической проницаемости  $\kappa$  или возрастанием радиуса состояний  $a$ . Действительно, имеются соображения как микроскопического ([13] к гл. 4), так и феноменологического [11, 12] характера, указывающие на то, что в точке перехода диэлектрическая проницаемость должна обращаться в бесконечность. Такое поведение диэлектрической проницаемости уже было обнаружено в кремнии [13]. С другой стороны, по мере приближения к переходу Мотта за счет сильного перекрытия волновых функций происходит постепенная делокализация волновых функций и возрастание радиуса состояний  $a$ . Это явление аналогично обсуждавшемуся в гл. 2 возрастанию радиуса состояний при приближении уровня Ферми к порогу подвижности в одноэлектронной картине перехода Андерсона (см. (7.4)). Для количественного объяснения зависимости (33.20) необходимо, по-видимому, учитывать увеличение  $\kappa$  и  $a$ .

### § 34. Роль многочастичных корреляций в прыжковой проводимости

В предыдущих главах, излагая теорию прыжковой проводимости, мы исходили из модели эквивалентной сетки сопротивлений Миллера и Абрахамса. Как мы видели в § 15, для вывода этой модели пришлось перейти к самосогласованному приближению типа Хартри. Математически этот переход состоит прежде всего в том, что выражение (15.20) для среднего по времени потока электронов с узла  $i$  на узел  $j$  заменяется на произведение средних:

$$\langle \gamma_{ij} n_i (1 - n_j) \rangle \rightarrow \langle \gamma_{ij} \rangle \langle n_i \rangle (1 - \langle n_j \rangle) = \langle \gamma_{ij} \rangle f_i (1 - f_j). \quad (34.1)$$

Здесь  $n_i = 0$  или  $1$  — число заполнения донора  $i$ , а  $f_i = \langle n_i \rangle$ . Напомним, что, благодаря связанной с кулоновским взаимодействием зависимости энергий  $\epsilon_i$  и  $\epsilon_j$  от чисел заполнения остальных электронов, вероятность перехода  $\gamma_{ij}$  также зависит от чисел заполнения  $n_k$ . Если  $\{n\}$  — набор чисел заполнения, то  $\gamma_{ij}$  является его функ-

ционалом:  $\gamma_{ij} = \gamma_{ij}\{n\}$ . Кроме расщепления, произведенного в (34.1), самосогласованное приближение состоит еще в том, что средняя вероятность перехода  $\langle \gamma_{ij} \rangle$  вычисляется в предположении, что энергии узлов  $i$  и  $j$  не флуктуируют во времени и равны своим средним значениям, выражающимся через функции распределения других узлов согласно (15.21). Другими словами, усреднение функционала  $\gamma_{ij}\{n\}$  сводится к замене  $\langle \gamma_{ij}\{n\} \rangle = \gamma_{ij}\{f\}$ .

В предыдущем параграфе схема Миллера—Абрахамса была усложнена введением двухузельной функции распределения (33.14), описывающей заполнение пар доноров, между которыми происходит переход электрона. Однако и эта модифицированная схема является приближенной, поскольку предполагается, что до и после электронного перехода заполнение всех прочих доноров соответствует основному состоянию.

Ясно, что лежащее в основе самосогласованного приближения пренебрежение корреляцией флуктуаций чисел заполнения не может давать точные результаты. Поэтому принципиальное значение имеет вопрос о том, как изменятся результаты, полученные в гл. 6—9, при выходе за рамки самосогласованного приближения, или, говоря другими словами, при учете корреляционных эффектов.

Нетрудно сразу же сделать важный общий вывод. Он состоит в том, что результаты вычисления показателя экспоненты  $\rho_3$ , полученные в гл. 6 и 7, вообще нечувствительны к процедуре временного усреднения. Действительно, все эти результаты основаны на существовании связанного с перекрытием волновых функций и не зависящего от времени множителя типа  $\exp(-2r_{ij}/a)$  в вероятностях перехода  $\gamma_{ij}$  и  $\gamma_{ji}$ . При вычислении  $\rho_3$  можно считать, что значение  $kT$  больше разброса уровней примесей. Тогда вероятность заполнения узла будет слабо зависеть от его энергии и корреляционные эффекты будут малы. Поэтому при вычислении показателя экспоненты удельного сопротивления можно использовать эквивалентную сетку Миллера и Абрахамса с сопротивлениями  $R_{ij} = R_{ij}^0 \exp(2r_{ij}/a)$ .

Однако, когда мы обращаемся к экспоненциальной температурной зависимости прыжковой проводимости,

сразу становится ясным, что здесь роль корреляционных эффектов может быть значительной. Цель этого параграфа состоит в том, чтобы, во-первых, дать классификацию корреляционных эффектов в электропроводности и, во-вторых, выделить те из полученных в гл. 8 и 9 результаты, которые нечувствительны к корреляционным эффектам. Наша классификация состоит в разделении возможных корреляций на два вида, один из которых мы условно называем хаббардовским, а другой — кулоновским.

**а) Хаббардовские корреляции.** Пренебрежем сначала всеми эффектами, связанными с кулоновским взаимодействием электронов, находящихся на разных примесях. Физическая ситуация, в которой при вычислении  $\varepsilon_3$  допустимо такое приближение, имеет место в твердых растворах различных полупроводников (см. § 28), если разброс уровней, связанный с флуктуациями состава, превышает энергию кулоновского взаимодействия электронов на соседних центрах. Для проводимости с переменной длиной прыжка такое приближение допустимо в аморфных полупроводниках. В отсутствие кулоновского взаимодействия энергии  $\varepsilon_i$ ,  $\varepsilon_j$  и вероятность перехода  $\gamma_{ij}$  не зависят от чисел заполнения. Поэтому  $\gamma_{ij}$  можно вынести за знак усреднения:

$$\langle \gamma_{ij} n_i (1 - n_j) \rangle = \gamma_{ij} \langle n_i (1 - n_j) \rangle. \quad (34.2)$$

Тогда для тока между узлами  $i$  и  $j$  получаем

$$\mathcal{J}_{ij} = \mathcal{J}_{ij}^S + \mathcal{J}_{ij}^C, \quad (34.3)$$

где согласно (15.24)

$$\mathcal{J}_{ij}^S = -e \gamma_{ij} f_i (1 - f_j) - \gamma_{ji} f_j (1 - f_i) \quad (34.4)$$

— ток в самосогласованном приближении,

$$\mathcal{J}_{ij}^C = -e (\gamma_{ji} - \gamma_{ij}) (\langle n_i n_j \rangle - f_i f_j) \quad (34.5)$$

— корреляционная поправка к току. В равновесии  $n_i$  и  $n_j$  флуктуируют независимо ( $\langle n_i n_j \rangle = f_i f_j$ ) и корреляционная поправка к току равна нулю. Однако если имеется внешнее электрическое поле и через пару  $(i, j)$  течет ток, то коррелятор  $\langle n_i n_j \rangle$ , вообще говоря, не должен расцепляться. Разность  $(\langle n_i n_j \rangle - f_i f_j)$  может ока-



заться пропорциональной внешнему полю, и тогда слабое  $\mathcal{J}_{ij}^C$  может быть не менее важным, чем  $\mathcal{J}_{ij}^S$ . Наличие тока  $\mathcal{J}_{ij}^C$  связано, очевидно, с тем, что движение каждого электрона коррелировано с движением других. Первопричина этой корреляции состоит в сильном отталкивании электронов на одном центре, благодаря которому наличие двух электронов на центре в (34.2) считается невозможным. Поэтому такую корреляцию мы условно назовем хаббардовской. Кикучи [14] и Ричардс [15] исследовали влияние хаббардовских корреляций на энергию активации ионной электропроводности в периодических решетках, состоящих из двух подрешеток с различными энергиями узлов. Исключительно ясная формулировка проблемы хаббардовских корреляций дана Ричардсом [15].

Посмотрим, как влияет хаббардовская корреляция на энергию активации  $\epsilon_3$  при различных степенях компенсации. Начнем со случая сильной компенсации. В этом случае при низких температурах электроны, как правило, прыгают по донорам, которые практически все время пусты (см. рис. 3.4). Поэтому с подавляющей вероятностью  $n_i = 0$  и в формуле (34.2) можно пренебречь отличием выражения  $(1 - n_i)$  от единицы. В результате для тока  $\mathcal{J}_{ij}$  получаем выражение

$$\mathcal{J}_{ij} = -e[\gamma_{ij}f_i - \gamma_{ji}f_j], \quad (34.6)$$

которое при учете неравенств  $f_i, f_j \ll 1$  совпадает с (34.4). Таким образом, хаббардовская корреляция не влияет на  $\epsilon_3$  при сильной компенсации. Аналогичная ситуация имеет место и при слабой компенсации,  $K \ll 1$ . В этом случае с подавляющей вероятностью все доноры, дающие вклад в проводимость, заполнены. Поэтому в (34.2) можно заменить  $n_i$  на единицу. Тогда число переходов  $i \rightarrow j$  в единицу времени окажется пропорциональным  $(1 - f_j)$  — вероятности возникновения дырки на узле  $j$ . Легко убедиться, что при учете неравенств  $(1 - f_i), (1 - f_j) \ll 1$  получающийся результат совпадает с (34.4).

В отличие от случаев сильной и слабой компенсации, при промежуточной компенсации электроны часто пересекают уровень Ферми, т. е. переходят из состояний над уровнем Ферми в состояния под уровнем Ферми и

наоборот. При этом важную роль играют оба множителя  $n_i$  и  $(1-n_j)$  в (34.2). В результате априори нет оснований думать, что ток  $\mathcal{J}_{ij}^C$  мал. То же самое относится к области действия закона Мотта в аморфных полупроводниках.

Чтобы найти корреляторы  $\langle n_i n_j \rangle$  и в конечном счете вычислить ток  $\mathcal{J}_{ij}^C$ , следует ввести многоузельную функцию распределения  $F\{n_i\}$ , которая представляет собой вероятность осуществления заданного набора чисел заполнения  $\{n_i\}$ . Затем нужно решить систему уравнений для  $F\{n_i\}$ , вытекающих из ее стационарности. Вообще говоря, это очень сложная задача и нам не удалось найти в ней существенных упрощений. Поэтому в качестве простейшего примера мы рассмотрели цепочку из четырех примесей 1, 2, 3, 4, в которой переходы возможны только между соседними примесями. Мы считали, что к концам цепочки, т. е. к примесям 1 и 4, приложена разность потенциалов  $U$ , причем числа заполнения этих примесей не флуктуируют и равны своим средним значениям:

$$\begin{aligned} f_1 &= \left[ \frac{1}{2} \exp\left(\frac{e_1 - eU - \mu}{kT}\right) + 1 \right]^{-1}, \\ f_4 &= \left[ \frac{1}{2} \exp\left(\frac{e_4 - \mu}{kT}\right) + 1 \right]^{-1}. \end{aligned} \quad (34.7)$$

Полное описание такой цепочки в стационарных условиях предполагает знание вероятностей  $F(n_2, n_3)$  для четырех различных комбинаций чисел заполнения примесей 2 и 3:  $F(0, 0)$ ,  $F(0, 1)$ ,  $F(1, 0)$  и  $F(1, 1)$ . Эти четыре вероятности удовлетворяют, конечно, условию нормировки

$$F(0, 0) + F(0, 1) + F(1, 0) + F(1, 1) = 1. \quad (34.8)$$

Таким образом, имеется три независимых величины, в качестве которых можно выбрать одноузельные функции распределения

$$f_2 = F(1, 0) + F(1, 1), \quad f_3 = F(0, 1) + F(1, 1) \quad (34.9)$$

и коррелятор  $\langle n_2 n_3 \rangle \equiv F(1, 1)$ . Условия стационарности этих величин дают три уравнения для их определения. За недостатком места мы не будем здесь ни выписывать уравнения, ни исследовать их решение. Сформулируем

только результат. Вообще говоря, коррелятор  $\langle n_2 n_3 \rangle$  оказывается немалым. В результате ток через цепочку не описывается самосогласованным приближением. Например, когда все множители  $\exp(-2r_{ij}/a)$  в  $\gamma_{ij}$  равны друг другу, учет коррелятора  $\langle n_2 n_3 \rangle$  изменяет энергию активации электропроводности цепочки. Однако имеется важный случай, когда корреляционные поправки оказываются малыми. Это происходит, когда для всех пар  $\frac{2r_{ij}}{a} \gg \frac{e_{ij}}{kT}$  и  $r_{23} > r_{12}, r_{34}$ , т. е. переход  $2 \rightarrow 3$  обладает наибольшим сопротивлением. Физическая причина малости  $\langle n_2 n_3 \rangle$  состоит в том, что при  $\gamma_{12}, \gamma_{34} \gg \gamma_{23}$  коррелятор  $\langle n_2 n_3 \rangle$  быстро релаксирует. Из рассмотрения этого случая может возникнуть надежда, что при выполнении неравенства (28.2) для максимальных в критической подсетке сопротивлений с  $r_{ij} = r_c$ , определяющих электропроводность системы, хаббардовские корреляции не существенны. Это означало бы, что хаббардовские корреляции не влияют на энергию активации  $\epsilon_3$  даже при промежуточной степени компенсации. Разумеется, рассмотренный нами пример может служить лишь наводящим аргументом и доказательной силой не обладает.

**б) Кулоновские корреляции.** Все нехаббардовские корреляции, т. е. корреляции, связанные с взаимодействием электронов на разных узлах, мы назовем кулоновскими. Впервые на роль кулоновских корреляций в прыжковой проводимости обратили внимание Поллак и Кнотек [2].

Простейший кулоновский корреляционный эффект — это исключение самодействия при одноэлектронном переходе между состояниями, находящимися по разные стороны от уровня Ферми. Исключение самодействия и следствия, к которым оно приводит, были подробно рассмотрены в § 33. Однако кроме этого эффекта существуют другие корреляционные эффекты, связанные с кулоновским взаимодействием двух и более электронов.

Следуя Кнотеку и Поллаку [2], рассмотрим двухэлектронные кулоновские корреляции на примере системы из четырех доноров, расположенных в пространстве так, как это показано на рис. 10.5.

Пусть переход  $a \rightleftharpoons b$  принадлежит критической подсетке, а переход  $c \rightleftharpoons d$  лежит вне ее и ток через него

практически равен нулю. Пусть в основном состоянии на донорах  $a$  и  $c$  находятся электроны, а доноры  $b$  и  $d$  пусты. В рамках самосогласованного приближения мы пренебрегаем флуктуациями чисел заполнения доноров, окружающих доноры  $a$  и  $b$ , т. е. считаем остальные электроны замороженными. Тогда для вероятности перехода между донорами  $a$  и  $b$  с поглощением фонона, согласно (15.17), имеем

$$\overline{d} \quad \overline{c} \quad \gamma_{ab} \sim \exp \left[ -\frac{\Delta_a^b}{kT} \right]. \quad (34.10)$$

Если выйти за рамки самосогласованного приближения, то можно представить, что в те редкие моменты, когда за счет термических флуктуаций электрон переходит с донора  $c$  на донор  $d$ , энергия фонона, необходимая для перехода  $a \rightarrow b$ , сильно уменьшается. Именно так обстоит



Рис. 10.5. Четыре примесных центра с двумя электронами (черные кружки).

дело для конфигурации доноров, изображенной на рисунке. Существуют такие конфигурации, когда, несмотря на экспоненциальную малость времени, которое электрон с донора  $c$  проводит на доноре  $d$ , выигрыш в энергии фонона настолько велик, что средняя по времени вероятность  $\langle \gamma_{ab} \rangle$  имеет энергию активации меньшую, чем в формуле (34.10). Следовательно, в принципе двухэлектронные корреляции могут существенно изменить энергию активации электропроводности. Однако сейчас мы покажем, что при сильной и слабой компенсации двухэлектронными корреляциями при вычислениях  $\epsilon_3$  можно пренебречь.

Дело в том, что при  $K \ll 1$  и  $1 - K \ll 1$  для подавляющего числа доноров, имеющих энергии вблизи пика плотности состояний, сдвиг уровня определяется крупномасштабным потенциалом. За создание этого потенциала на фиксированном доноре ответственно большое число заряженных примесей, попадающих в радиус корреляции потенциала  $r_0$  или  $r_s$  (см. (12.10) и (13.8)). Несмотря на то, что вклад каждой заряженной примеси в потенциал фиксированного донора флуктуирует во времени, сумма большого числа флуктуирующих слагаемых испытывает лишь малые относительные флуктуации, т. е.

изменением во времени уровней большинства доноров можно пренебречь. Таким образом, при  $K \ll 1$  и  $1 - K \ll 1$  кулоновские корреляции, так же как и хаббардовские, не изменяют результатов, полученных в §§ 26, 27.

Иная ситуация имеет место при промежуточной степени компенсации, когда энергия донора определяется кулоновским потенциалом его ближайших соседей. В этом случае, как показывает рассмотренный выше пример конфигурации из четырех примесей (см. рис. 10.5), многоэлектронные корреляционные эффекты могут уменьшить энергию активации  $\varepsilon_3$ . Возможно, в совместном действии этих эффектов и эффекта исключения самодействия и заключается причина упомянутой в § 28 аномальной малости величины  $\varepsilon_3$  при  $K = 0,5$  (см., однако, [6]). Количественной теории, которая позволяла бы учитывать многоэлектронные кулоновские корреляции и заменила модель сетки сопротивлений Миллера и Абрахамса, до сих пор не существует.

Кнотек и Поллак [2] рассмотрели также иной механизм влияния электрон-электронного взаимодействия на прыжковую проводимость. Они ввели представление об одновременных многоэлектронных переходах. Примером такого перехода может быть процесс одновременного перехода электронов с узлов  $a$  и  $c$  на узлы  $b$  и  $d$  (см. рис. 10.5). В отличие от рассмотренного выше случая последовательных переходов, вероятность одновременного перехода содержит произведение множителей  $\exp(-2r_{ab}/a)$  и  $\exp(-2r_{cd}/a)$ . Поэтому при достаточно слабом легировании такими переходами можно пренебречь. Однако при промежуточном легировании, когда слагаемые  $\varepsilon_{ij}/kT$  становятся сравнимыми с  $2r_{ij}/a$ , выигрыш за счет того, что для некоторых конфигураций энергия активации двухэлектронного перехода меньше, чем у одноэлектронного, может оказаться более важным. Исходя из этой идеи, Поллак и Кнотек дали качественное объяснение некоторых экспериментальных фактов, относящихся к случаю промежуточного легирования. В частности, они считают, что уменьшение  $\varepsilon_3$  при увеличении легирования объясняется одновременными двухэлектронными переходами (см. [7] к гл. 8).

## ЧАСТЬ II

# СИЛЬНО ЛЕГИРОВАННЫЕ ПОЛУПРОВОДНИКИ

### Введение

Сильно легированным полупроводником (СЛП) принято называть полупроводник с концентрацией примесей  $N$ , удовлетворяющей неравенству  $Na^3 \gg 1$ , где  $a$  — борковский радиус примесного состояния\*). Выполнением этого неравенства мы будем оправдывать различные приближения, которые делаются в теории. В полупроводниках с малой эффективной массой ( $n$ -InSb,  $n$ -InAs и др.) условие сильного легирования выполняется в довольно широкой области концентраций. Например, в  $n$ -InSb значение  $Na^3 = 1$  достигается при  $N \approx 5 \cdot 10^{15} \text{ см}^{-3}$ . С другой стороны, в  $n$ -Ge  $Na^3 = 1$  при  $N \approx 3 \cdot 10^{19} \text{ см}^{-3}$ . Это означает, что область сильного легирования в германии весьма мала.

Как уже говорилось в гл. 3, физическим критерием, разделяющим области слабого и сильного легирования, является локализация электронов на уровне Ферми. В слабо легированном полупроводнике электронные состояния на уровне Ферми локализованы и низкотемпературная проводимость носит активационный характер, а в СЛП они делокализованы и проводимость носит металлический характер. Экспериментальный факт состоит в том, что переход от металлической проводимости к активационной (переход Мотта) происходит обычно при концентрациях, удовлетворяющих условию  $Na^3 \approx 0,02$ . Поэтому возникает широкая промежуточная область концентраций,  $0,02 < Na^3 < 1$ , которую по причине делокации электронов никак нельзя отнести к области слабого легирования и в которой не выполняется условие сильного легирования. С точки зрения физических

\*) Первая монография, обобщающая обширный экспериментальный материал, касающийся СЛП, была написана Фистулем (см. [1] к гл. 11).

свойств эта область относится скорее к области сильного легирования. Поэтому можно надеяться, что описание структуры электронных состояний, предлагаемое в следующих главах, имеет некоторое отношение и к промежуточной области.

Ахиллесовой пятой теории СЛП является необходимость делать какие-либо предположения о корреляции в распределении примесей. Структура электронных состояний вблизи дна зоны проводимости, о которой пойдет речь в гл. 11, не очень чувствительна к этим предположениям. Однако для описания глубокого хвоста плотности состояний (гл. 12) и перехода металл — диэлектрик, происходящего под действием компенсации (гл. 13), эти предположения важны. Как и в первой части книги, мы начнем с пуассоновского распределения примесей, однако затем рассмотрим модели, позволяющие учесть корреляцию в их расположении. Примеси предполагаются неподвижными, что, несомненно, справедливо при низких температурах. При высоких температурах могут существовать подвижные возбуждения, состоящие и из электронов, и из примесей (флукуоны). Этот вопрос мы не рассматриваем, отсылая читателей к превосходному обзору Кривоглаза (см. [2] к гл. 11).

Наш подход к теории СЛП состоит в следующем. Мы рассматриваем различные соотношения между параметрами, предполагая те или иные сильные неравенства. Решая задачи в различных предельных случаях и сшивая их друг с другом, мы получаем полное представление о структуре электронных состояний и характере протекающих процессов. Однако в промежуточных областях эти представления являются скорее качественными, чем количественными, а именно промежуточные области и оказываются чаще всего экспериментально наблюдаемыми. Поэтому во второй части книги мы не будем стремиться к численному совпадению теории и эксперимента. Это — задача будущих теорий, которые, вероятно, должны быть основаны не на аналитических подходах, а на хорошей вычислительной технике. Наша цель состоит лишь в качественном объяснении экспериментальных фактов.

## СОСТОЯНИЯ ЭЛЕКТРОНОВ В СИЛЬНО ЛЕГИРОВАННЫХ ПОЛУПРОВОДНИКАХ

### § 35. Теория линейного экранирования

Экспериментальные данные указывают на то, что электроны в СЛП имеют сравнительно большую подвижность, т.е. они относительно свободны. На первый взгляд этот факт может показаться странным. Энергия взаимодействия электронов с примесями определяется средним расстоянием между примесями, и, следовательно, она порядка  $e^2 N^{1/2} / \kappa$ . Эта величина растет с ростом концентрации  $N$ . В невырожденной плазме, например, условие идеальности электронного газа имеет вид  $kT \gg e^2 N^{1/2}$ . Оно означает, что кинетическая энергия  $kT$  должна быть больше потенциальной и что при высоких концентрациях газ становится неидеальным.

Однако для вырожденного электронного газа ситуация совершенно иная. Дело в том, что кинетическая энергия большинства электронов порядка энергии Ферми. Если полупроводник не очень сильно компенсирован, то концентрация электронов близка к концентрации примесей  $N$  и энергия Ферми  $\mu$  порядка  $\hbar^2 N^{2/3} / m$ . Эта величина также возрастает с  $N$ , причем быстрее, чем энергия взаимодействия. Поэтому условие идеальности, имеющее вид  $\mu \gg e^2 N^{1/2} / \kappa$ , выполняется, если  $Na^3 \gg 1$ , где  $a = \hbar^2 \kappa / me^2$ . Итак, с увеличением концентрации степень идеальности электронного газа увеличивается. (Энергия взаимодействия электронов друг с другом также порядка  $e^2 N^{1/2} / \kappa$  и мала по сравнению с энергией Ферми.)

Этот факт и определяет подход к теории электронных состояний в СЛП. Ясно, что следует стартовать от свободных электронов, т.е. считать, что волновая функция



является плоской волной, а квазимпульс  $p$  представляет собой хорошее квантовое число.

При нулевой температуре должны быть заполнены все состояния с энергиями  $\epsilon < \mu$ . Как известно из квантовой механики, для того чтобы получить число состояний в единице объема, имеющих энергию в интервале от  $\epsilon$  до  $\epsilon + d\epsilon$ , необходимо найти объем в пространстве импульсов, который занимают эти состояния. Если спектр электронов изотропный и квадратичный,  $\epsilon = p^2/2m$ , то искомая область является сферическим слоем, а ее фазовый объем равен  $4\pi p^2 dp$ . Для определения числа состояний необходимо поделить этот объем на фазовый объем  $(2\pi\hbar)^3$ , соответствующий одному состоянию. Кроме того, следует учесть, что в каждом состоянии могут находиться два электрона с противоположно направленными спинами. В результате для плотности состояний получим

$$g_0(\epsilon) d\epsilon = \frac{8\pi p^2 dp}{(2\pi\hbar)^3} = \frac{8\pi p^2}{(2\pi\hbar)^3} \frac{dp}{d\epsilon} d\epsilon, \quad (35.1)$$

или

$$g_0(\epsilon) = \frac{(2m)^{3/2} \sqrt{\epsilon}}{2\pi^2 \hbar^3}. \quad (35.2)$$

Энергия Ферми в этом приближении определяется условием

$$\int_0^{\mu_0} g_0(\epsilon) d\epsilon = n$$

и равна

$$\mu_0 = (3\pi^2)^{2/3} \hbar^2 n^{2/3} / 2m. \quad (35.3)$$

В предыдущих рассуждениях распределение электронов считалось однородным в пространстве. Покажем теперь, каким образом можно описать неоднородные распределения. Неоднородное распределение электронов возникает в чистом полупроводнике вблизи границы или контакта с другим веществом. В легированном полупроводнике неоднородное распределение электронов в объеме является неизбежным следствием флуктуаций концентрации заряженных примесей. Это явление и будет предметом изучения в этой и в последующих главах.

Допустим, что в объеме СЛП возник электростатический потенциал  $\varphi(r)$ . Тогда квазиимпульс не является хорошим квантовым числом, и мы должны решить уравнение Шредингера, чтобы найти волновые функции и уровни энергии. Однако в наиболее важных для теории СЛП случаях эта процедура оказывается ненужной. Допустим, что потенциал  $\varphi(r)$  плавный в том смысле, что он мало меняется на длине волны электрона. (Поскольку большинство электронов имеет импульс порядка фермиевского, то под длиной волны здесь следует понимать величину  $\hbar/\sqrt{2m\mu}$ .) В этом случае электрон можно считать классическим, т. е. предполагать, что он «чувствует» только потенциал той точки, в которой находится. Пренебрегая соотношением неопределенности, напишем классическое выражение для его энергии:

$$\varepsilon = \frac{p^2}{2m} - e\varphi(r). \quad (35.4)$$

Сделанное приближение удобно изобразить графически (рис. 11.1, а, б). Энергия дна зоны проводимости совпадает теперь с потенциальной энергией электрона  $-e\varphi$  и дно зоны оказывается искривленным.

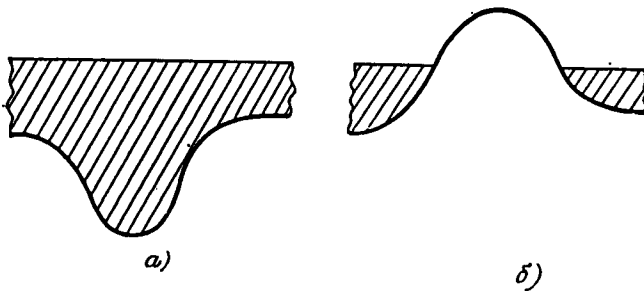


Рис. 11.1. Кривые изображают искривленное дно зоны проводимости. Горизонтальная прямая — уровень Ферми. Состояния, заполненные электронами, заштрихованы.

В термодинамическом равновесии максимальная энергия заполненных состояний  $\mu$  не может меняться от точки к точке. Зато меняются максимальная кинетическая энергия и концентрация электронов. В областях, где дно зоны понижено, электронов оказывается больше,

и, наоборот, там, где дно повышено, электронов меньше. В областях, где дно зоны выше уровня Ферми (рис. 11.1, б), совсем нет электронов, так как там нет классически разрешенных состояний, а экспоненциально малой вероятностью туннелирования мы пренебрегаем.

Легко выразить концентрацию электронов  $n(r)$  через  $\varphi(r)$ , воспользовавшись правилами квазиклассического квантования. Максимальный импульс электрона в точке  $r$  есть  $p_F = \sqrt{2m(\mu + e\varphi)}$ , а концентрация связана с ним соотношением  $n = 8\pi p_F^3 / 3 (2\pi\hbar)^3$ . Отсюда получаем

$$n(r) = \begin{cases} \frac{(2m)^{3/2}}{3\pi^2\hbar^3} (\mu + e\varphi(r))^{3/2} & \text{при } \mu + e\varphi > 0, \\ 0 & \text{при } \mu + e\varphi < 0. \end{cases} \quad (35.5)$$

Сгущения и разрежения электронной плотности создают свой собственный потенциал, который способствует распрямлению дна зоны проводимости, т. е. ослаблению внешнего потенциала. В этом и состоит электронное экранирование. В рамках сделанных предположений легко построить количественную теорию экранирования. Для этого достаточно написать уравнение Пуассона, в котором концентрация электронов связана с  $\varphi$  формулой (35.5):

$$\Delta\varphi = -\frac{4\pi e}{\kappa} [\xi(r) - n(r) + n]. \quad (35.6)$$

Здесь  $n$  — средняя концентрация электронов, а  $e\xi(r)$  — плотность заряда, не зависящая от электронной плотности. Уравнение (35.6) является квазиклассическим самосогласованным уравнением типа уравнения Томаса — Ферми. Оно описывает, например, изгиб зон на границе СЛП и другого вещества\*). Это уравнение нелинейно и в общем виде не решается. В некоторых случаях, однако, можно ограничиться линейным приближением. Допустим, что характерные значения энергии  $|e\varphi|$  малы по сравнению с  $\mu$ . Тогда (35.6) можно линеаризовать,

\*) По существу, уравнение такого же рода мы широко использовали в гл. 3 для изучения крупномасштабных флуктуаций потенциала. Однако там речь шла о слабо легированном полупроводнике с локализованными состояниями. Поэтому выражение для  $n(r)$  в (13.10) существенно отличалось от (35.5).

записав  $n(r) = n + \frac{dn}{d\mu} e\phi$ , где  $\frac{dn}{d\mu}$  определяется в приближении (35.3). В результате получаем линейное уравнение

$$\Delta\phi = \frac{\phi}{r_0^2} - \frac{4\pi e\xi(r)}{\kappa}, \quad (35.7)$$

где

$$r_0 = \left( \frac{4\pi e^2}{\kappa} \frac{dn}{d\mu} \right)^{-1/2} = \frac{a}{2} \left( \frac{\pi}{3} \right)^{1/6} \frac{1}{(na^3)^{1/6}}. \quad (35.8)$$

Длину  $r_0$  принято называть дебаевским радиусом экранирования вырожденного электронного газа. Уравнение (35.7) имеет следующее решение:

$$\phi = \int K(r-r') \xi(r') dr', \quad K(r) = \frac{e}{\kappa r} \exp\left(-\frac{r}{r_0}\right). \quad (35.9)$$

Для одного положительного заряда, помещенного в начале координат, получим

$$\xi(r) = \delta(r), \quad \phi = \frac{e}{\kappa r} \exp\left(-\frac{r}{r_0}\right). \quad (35.10)$$

Итак, кулоновский потенциал экранируется на расстоянии  $r_0$ . Следует иметь в виду, что учет квантовых эффектов приводит к дополнительным осциллирующим членам, которые малы по сравнению с (35.10) на малых расстояниях, но затухают гораздо слабее [3]. Это обстоятельство, однако, не является существенным в теории СЛП благодаря тому, что рассеяние на примесях создает значительную неопределенность в импульсе электронов, близких к уровню Ферми. Обсудим теперь условие применимости квазиклассической линейной теории для экранирования одного заряда. Легко убедиться, что неравенство  $na^3 \gg 1$  обеспечивает выполнение условия  $r_0 \gg \hbar/\sqrt{m\mu}$ , означающего, что длина волны электронов мала по сравнению с характерными размерами потенциальной ямы. Это же неравенство обеспечивает линейность экранирования. Необходимо, чтобы почти во всех точках ямы потенциальная энергия была мала по сравнению с энергией Ферми. Для этого нужно выполнение условия  $e^2/r_0\kappa \ll \mu$ , которое сводится к  $na^3 \gg 1$ . Наконец, заметим, что условие  $na^3 \gg 1$  эквивалентно

условию  $nr_0^3 \gg 1$ . Ясно, что это последнее условие необходимо, чтобы изложенная выше теория имела смысл. Действительно, в сфере с радиусом  $r_0$  должно быть много электронов. Только тогда малое изменение их плотности сможет привести к экранировке.

Отметим теперь основные выводы линейной теории экранирования.

1. Радиус экранирования  $r_0$  не зависит от величины экранируемого потенциала и определяется лишь концентрацией электронов и эффективной массой.

2. Экранирование каждого заряда системы происходит независимо. Например, суммарный потенциал, создаваемый всеми заряженными примесями, равен сумме потенциалов (35.10), т. е. каждая примесь экранируется сама по себе. Этот важный вывод следует из линейности уравнения (35.7).

Типичные примеры нелинейного экранирования мы находим в теории контактных явлений. В частности, с нелинейным экранированием связана резкая зависимость ширины слоя объемного заряда от приложенного напряжения. В гл. 13 мы столкнемся с нелинейным экранированием при описании сильно компенсированного полупроводника. Своеобразный пример нелинейного экранирования, осуществляемого электронами примесной зоны, рассматривался в § 13.

### § 36. Плотность состояний вблизи дна зоны проводимости

В 1962 г. на международной конференции по физике полупроводников в Экзетере независимо Бонч-Бруевичем, Кейном и Келдышем была высказана мысль о том, что для описания плотности состояний сильно легированного полупроводника можно пользоваться квазиклассическими представлениями. Соответствующие вычисления были сделаны в работах [4—8].

Строгое обоснование этой идеи мы отложим до следующего параграфа, а сначала посмотрим, к чему она приводит. Будем считать электроны классическими и вычислять плотность состояний по формуле (35.1), устанавливая связь между энергией и импульсом с помощью формулы (35.4). При этом получим локальную плотность

состояний  $g(\epsilon, r)$ , зависящую от координаты  $r$ :

$$g(\epsilon, r) = \frac{(2m)^{3/2}}{2\pi^2\hbar^3} \sqrt{\epsilon + e\varphi(r)}. \quad (36.1)$$

Здесь  $\varphi(r)$  — суммарный потенциал, создаваемый всеми примесями, с учетом электронного экранирования.

В областях, где  $\epsilon < -e\varphi$ , классически разрешенных состояний с энергией  $\epsilon$  нет и плотность состояний равна нулю. Для того чтобы найти полную плотность состояний  $g(\epsilon)$ , нужно просуммировать (36.1) по всем классически разрешенным областям и поделить на полный объем системы  $V_0$  (рис. 11.2):

$$g(\epsilon) = \frac{(2m)^{3/2}}{2\pi^2\hbar^3 V_0} \int_{\epsilon > -e\varphi} dr \sqrt{\epsilon + e\varphi(r)}. \quad (36.2)$$

В дальнейшем будет удобнее говорить о потенциальной энергии электрона  $V = -e\varphi$ . Интегрирование по координате в (36.2) можно заменить усреднением по всем

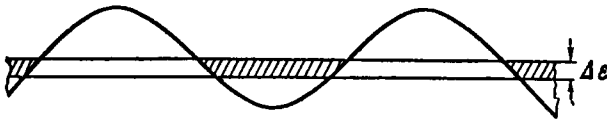


Рис. 11.2. Вычисление квазиклассической плотности состояний. Кривая изображает дно зоны проводимости. Классически разрешенные состояния, находящиеся в интервале  $\Delta\epsilon$ , заштрихованы.

значениям потенциальной энергии. Это соответствует тому, что мы изучаем потенциал в одной пространственной точке, меняя всевозможным образом конфигурации примесей. При этом получим

$$g(\epsilon) = \frac{(2m)^{3/2}}{2\pi^2\hbar^3} \int_{-\infty}^{\epsilon} \sqrt{\epsilon - V} F(V) dV, \quad (36.3)$$

где  $F(V)$  — функция распределения потенциальной энергии, которую нужно найти. Как уже говорилось, электронный газ в СЛП слабо неидеален. Следовательно, типичные значения потенциальной энергии электрона должны быть малыми по сравнению с энергией Ферми  $\mu$  и можно воспользоваться линейной теорией экранирова-

ния. Роль внешнего заряда в выражениях (35.7) и (35.9) должны играть неоднородности в распределении заряженных доноров. (Для определенности мы говорим СЛП  $n$ -типа и считаем компенсацию слабой.) Тогда

$$\xi(r) = N(r) - N, \quad (36.4)$$

где  $N$  — средняя концентрация доноров.

Будем считать, что корреляция в распределении примесей отсутствует и статистическим законом распределения примесей является «белый шум»:

$$\langle \xi(r) \xi(r') \rangle = N \delta(r - r'); \quad (36.5)$$

$\langle \dots \rangle$  означает среднее по конфигурациям. Предположим, что распределение примесей гауссово, т. е. более высокие корреляторы распадаются на произведение парных (это предположение будет обосновано ниже). Согласно (35.9) (36.5) получим

$$\langle V(r) V(r') \rangle = e^2 N \int K(r - r'') K(r' - r'') dr'', \quad (36.6)$$

а удвоенный средний квадрат потенциальной энергии выражается формулой

$$\gamma^2 = 2Ne^2 \int K^2(r) dr = 4\pi \frac{q^4}{x^2 r_0^2} Nr_0^3. \quad (36.7)$$

При линейном экранировании, согласно (35.9),  $V$  и связаны линейно. Поэтому более высокие корреляторы  $V(r)$  распадаются на произведения парных. Отсюда следует, что  $V(r)$ , так же как и  $\xi(r)$ , является гауссовой случайной функцией и

$$F(V) = (\sqrt{\pi} \gamma)^{-1} \exp(-V^2/\gamma^2). \quad (36.8)$$

Простая интерпретация формулы (36.8) состоит в следующем. Как показано в § 12, потенциал, который создается гауссовыми флуктуациями примесей, тем больше, чем больше масштаб флуктуаций. В нашем случае флуктуации масштаба, большего чем  $r_0$ , полностью экранируются электронами. Отсюда следует, что основной вклад в потенциал должны давать флуктуации с масштабом порядка  $r_0$ . Это видно и из (36.7), где основной вклад в интеграл дают длины порядка  $r_0$ . Оценивая потенциал, который создает флуктуация с размером

в приближении равномерно заряженной сферы, получим

$$V = e^2 \delta N^2 / \kappa r_0, \quad (36.9)$$

где  $\delta N^2$  — избыточное число заряженных доноров. Эти доноры мы считаем равномерно распределенными внутри шара с радиусом  $r_0$ . Среднее число доноров в таком шаре  $\bar{N} = N r_0^3$ . Как известно из статистики, среднеквадратичное значение  $\delta N^2$  равно  $(\bar{N})^{1/2}$ , т. е.  $(N r_0^3)^{1/2}$ . Подставляя это значение в (36.9), мы получим оценку для  $\gamma$ , с точностью до численных коэффициентов совпадающую с (36.7). Теперь легко оценить область применимости гауссовой статистики. Как известно, флуктуации являются гауссовыми, если избыточное число примесей  $\delta N^2$  мало по сравнению со средним, т. е.  $\delta N^2 \ll \bar{N}$ . Пользуясь (36.9), получим ограничение на потенциальную энергию электрона:

$$V \ll \bar{V}, \quad \bar{V} = \frac{e^2}{\kappa r_0} N r_0^3. \quad (36.10)$$

Заметим, что необходимое условие  $\gamma \ll \bar{V}$  следует из основного неравенства  $N a^3 \gg 1$ . В отсутствие компенсации  $n = N$  и  $\bar{V} \approx \mu_0$ . Тогда условие (36.10) эквивалентно условию линейного экранирования.

Подставив (36.8) в (36.3), получим

$$g(\varepsilon) = \frac{\sqrt{2} m^{3/2} \gamma^{1/2}}{\pi^2 \hbar^3} G_0\left(\frac{\varepsilon}{\gamma}\right), \quad (36.11)$$

где

$$G_0(X) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^X e^{-y^2} (X - y)^{1/2} dy. \quad (36.12)$$

График функции  $G_0(X)$ , полученный в [9], представлен на рис. 11.3.

При  $\varepsilon > 0$  и  $|\varepsilon| \gg \gamma$

$$g(\varepsilon) = g_0 - \frac{(2m)^{3/2} \gamma^2}{32\pi^2 \hbar^3 \varepsilon^{3/2}}, \quad (36.13)$$

откуда видно, что в этом случае величина  $g(\varepsilon)$  близка к плотности состояний свободных электронов  $g_0$ . При



$\varepsilon < 0$  и  $|\varepsilon| \gg \gamma$

$$g(\varepsilon) = \frac{m^{3/2} \gamma^{1/2}}{4\pi^2 \hbar^3} \left(\frac{\gamma}{\varepsilon}\right)^{3/2} \exp\left(-\frac{\varepsilon^2}{\gamma^2}\right), \quad (36.14)$$

т.е. плотность состояний экспоненциально убывает в глубь запрещенной зоны.

Легко убедиться, что при  $Na^3 \gg 1$  энергия Ферми  $\mu_0$  гораздо больше  $\gamma$ , т.е. уровень Ферми находится в области энергий, в которой плотность состояний близка к  $g_0$ . Это обстоятельство и позволяет считать экранирование флуктуаций дебаевским.

Измерение порога Мосса — Бурштейна, а также другие эксперименты свидетельствуют о том, что энергия Ферми в СЛП несколько меньше  $\mu_0$ . Это связано с неидеальностью электронного газа. Пусть  $\mu = \mu_0 - \mu_{ee} - \mu_{ei}$ ,

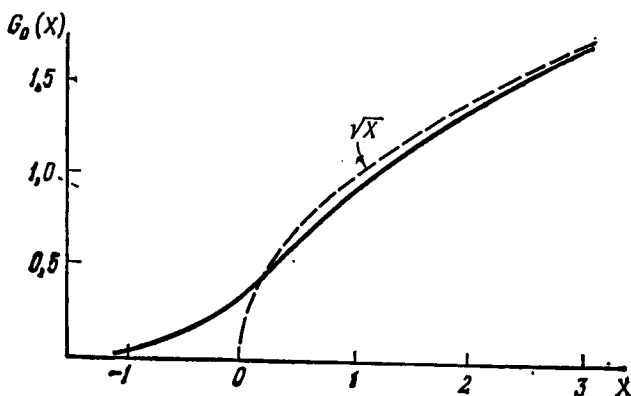


Рис. 11.3. Функция  $G_0(x)$ .

где  $\mu_{ee}$  и  $\mu_{ei}$  — поправки за счет взаимодействия электронов друг с другом и с примесями. Основная поправка связана с обменным взаимодействием электронов. Как известно (см., например, [10]),

$$\mu_{ee} = \frac{e^2}{\pi\kappa} \frac{\sqrt{2m\mu_0}}{\hbar}. \quad (36.15)$$

Поправка  $\mu_{ei}$  может быть получена с помощью (36.11):

$$\mu_{ei} = \frac{1}{8} \frac{\gamma^2}{\mu_0}. \quad (36.16)$$

Она описывает опускание уровня Ферми за счет хвоста плотности состояний.

Обобщение полученных результатов на случай компенсированного полупроводника не представляет особого труда [9], если по-прежнему считать, что применимо одноэлектронное приближение и примеси экранируются линейным образом. Помимо положительных доноров в этом случае имеются еще отрицательные акцепторы с концентрацией  $N_A$  и электроны с концентрацией  $n = N_D - N_A$ .

Очевидное обобщение состоит в том, что в среднеквадратичный потенциал  $\gamma$  следует подставлять вместо  $N$  суммарную концентрацию заряженных центров  $N_D + N_A$ . Так как  $N_D < N_D + N_A < 2N_D$ , то в оценках по порядку величины мы можем игнорировать это обстоятельство. Гораздо важнее, что в радиус экранирования  $r_0$  (35.8) и энергию Ферми  $\mu_0$  вместо  $N_D$  следует подставлять концентрацию электронов  $n$ , которая может быть гораздо меньше  $N_D$ . С уменьшением  $n$  радиус экранирования  $r_0$  и характерная величина потенциальной энергии  $\gamma$  возрастают, а энергия Ферми  $\mu_0$  уменьшается. Очевидно, что дебаевское приближение применимо, если

$$\alpha \equiv \frac{\gamma}{\mu_0} = \frac{4}{(3\pi)^{3/4}} \left(\frac{N}{n}\right)^{3/4} (Na^3)^{-1/4} \ll 1 \quad (N = N_D \approx N_A). \quad (36.17)$$

Это условие ограничивает применимость изложенной выше теории к сильно компенсированным полупроводникам. Случай  $\alpha > 1$  будет рассмотрен в гл. 13. Мы покажем там, что структура электронных состояний при  $\alpha > 1$  совершенно иная. Формула (36.11) легко обобщается также и на случай квантующего магнитного поля [9, 11].

Обратимся теперь к условиям применимости квазиклассического метода. Строгий вывод формулы (36.11), из которого видны сделанные приближения, приведен в следующем параграфе. Ниже мы ограничимся лишь качественными соображениями.

Как видно из (36.7), главный вклад в среднеквадратичный потенциал  $\gamma$  дают флуктуации масштаба  $r_0$ . Типичные потенциальные ямы имеют размер  $r_0$  и глубину  $\gamma$ . Энергия основного состояния в яме выше дна ямы на величину порядка  $E_s = \hbar^2/mr_0^2$ . Подавляющее большин-

ство состояний в таком потенциале может описываться квазиклассически, если  $\gamma \gg E_s$ . Это и есть основное условие применимости изложенной теории. Как показано в следующем параграфе, разложение, приводящее к (36.11), ведется по отношению

$$\left(\frac{E_s}{\gamma}\right)^{1/2} = \frac{1}{(Na^3)^{1/24}} \left(\frac{n}{N}\right)^{5/24}. \quad (36.18)$$

Этот параметр, конечно, очень плох. (Отношение  $n/N$ , согласно (36.17), не может быть очень малым). Тем не менее разложение по нему представляется нам единственно возможным аналитическим подходом к проблеме плотности состояний вблизи дна зоны, позволяющим в всяком случае понять физическую природу электронных состояний. В этом и состоит основное значение излагаемого метода.

В работе [12] показано, что мелкомасштабные флуктуации сильно ограничивают применимость (36.11) в области  $\epsilon < 0$ ,  $|\epsilon| \gg \gamma$ , где плотность состояний экспоненциально мала. Оказалось, что при  $\epsilon < 0$  область применимости ограничена условием

$$|\epsilon| < \min\{\epsilon_c, \mu_0\}, \quad \epsilon_c = \frac{\gamma^3 m a^3}{\hbar^2 (Na^3)^{2/5}}. \quad (36.19)$$

Так как  $\frac{\epsilon_c}{\gamma} = (Na^3)^{1/60} \left(\frac{N}{n}\right)^{1/12}$ , это означает фактически, что в области, где плотность состояний экспоненциально мала, квазиклассический подход неприменим и что для описания этой области нужны новые идеи и методы. Таким методом и является метод оптимальных флуктуаций, который изложен в следующей главе, где показано, что выражение для показателя экспоненты, которое фигурирует в (36.14), остается правильным за пределами области (36.19).

### § 37. Вывод квазиклассической формулы для плотности состояний

В случае СЛП строгий вывод формулы (36.11) и нахождение поправок к ней содержат некоторые нетривиальные моменты. Многие авторы [4, 13—15] получили квазиклассическое выражение (36.3), предполагая

примесный потенциал плавным. В идейном плане эти выводы не отличаются от вывода, приведенного в предыдущем параграфе, хотя, несомненно, являются более последовательными и строгими. Например, если считать потенциал гауссовым, то можно красиво и просто вывести (36.11) с помощью фейнмановских функциональных интегралов [13, 15]. Можно, однако, получить выражение (36.3) и не считая распределение потенциала гауссовым [4, 5]. Существенно только, что при выводах такого рода поправки к квазиклассическому выражению содержат среднее значение квадрата градиента потен-

циала  $\int (\nabla V)^2 dr$  [13] или квадрата более высоких производных [14]. Поэтому, строго говоря, эти выводы неприменимы в случае потенциала, который на малых расстояниях от центра ведет себя, как  $r^{-1}$ , так как для него поправочные члены обращаются в бесконечность.

Можно избежать этой расходимости, если не считать потенциал гауссовым и производить последовательное разложение плотности состояний по степеням постоянной Планка [16], подобное разложению статистической суммы ([4, § 33] к гл.3). В нулевом приближении по  $\hbar$  такое разложение действительно приводит к (36.11). Поправочный член оказывается порядка  $\hbar^2$ , и его отношение к нулевому члену — порядка  $a/r_0$ , где  $r_0$  определяется (35.8).

Этот результат легко понять. Рассматриваемый потенциал является суперпозицией экранированных кулоновских ям. Условие  $a/r_0 \ll 1$  означает, что каждая такая яма содержит много связанных состояний. При таком условии и справедливо разложение по  $\hbar$ . Однако при сильном легировании и не очень большой степени компенсации, когда  $\mu_0 \gg \gamma$  и можно считать экранирование линейным (а мы рассматриваем в этой главе только такой случай), обязательно выполняется условие  $r_0 \ll a$ , при котором одна яма вообще не обладает связанным состоянием. (Последнее неравенство легко получить, используя неравенства (36.17) и  $Na^3 \gg 1$ .) Таким образом, обычное квазиклассическое разложение в интересующем нас случае не приводит к цели.

Келдыш [17] предложил вывод (36.11), выполненный с помощью диаграммной техники. Впоследствии та-

кой же вывод независимо предложил Эфрос [12]. В работе [12] была построена техника, позволяющая находить поправки любого порядка к формуле (36.11). Основная идея этой техники состоит в том, что крупномасштабные флуктуации примесного потенциала учтены во всех порядках теории возмущений, но квазиклассическим образом, а мелкомасштабные флуктуации учтены квантовомеханически, но по теории возмущений.

В нулевом приближении этой модифицированной теории возмущений мы получаем формулу (36.11), а следующие приближения дают поправки к ней. Практические выводы состоят в том, что параметр разложения определяется формулой (36.18), а область применимости — условием (36.19). Отсюда, в частности, следует, что в случае СЛП формула (36.3) справедлива только там, где функция распределения гауссова. Ниже мы приведем вывод формулы (36.11), сделанный указанным методом.

Рассмотрим одноэлектронную функцию Грина

$$G = \left\langle \frac{1}{\varepsilon - H + i\delta} \right\rangle, \quad \delta \rightarrow +0, \quad (37.1)$$

где

$$H = \frac{p^2}{2m} + \sum_I e\varphi(r - r_I), \quad (37.2)$$

а  $\varphi(r)$  выражается (35.10). Плотность состояний  $g(\varepsilon)$  дается формулой

$$g(\varepsilon) = -\frac{2}{\pi} \text{Im Sp } G. \quad (37.3)$$

Простейшие диаграммы для  $G$  изображены на рис. 11.4, *a—г*. Каждой волнистой линии соответствует множитель  $\frac{4\pi e^2}{\kappa(q^2 + r_0^{-2})}$ , пучку волнистых линий соответствует множитель  $N$ , в каждом пучке выполняется закон сохранения импульса и по всем независимым импульсам ведется интегрирование. Свободная функция Грина равна

$$L_p = (\varepsilon - \varepsilon_p + i\delta)^{-1}, \quad \varepsilon_p = p^2/2m.$$

Диаграмма рис. 11.4, *a* соответствует сдвигу энергии, который точно компенсируется потенциалом электронов,

поэтому мы не будем ее рассматривать. Собственно энергетическая часть, соответствующая диаграмме рис. 11.4, б, имеет вид

$$\Sigma^{(1)} = \int \frac{d^3q}{(2\pi)^3} \frac{D(q)}{\varepsilon - \varepsilon_{p-\hbar q} + i\delta}, \quad (37.4)$$

где

$$D(q) = N \left[ \frac{4\pi e^2}{\kappa(q^2 + r_0^{-2})} \right]^2. \quad (37.5)$$

Выражения (37.4) достаточно для построения обычной теории возмущений, справедливой при больших энергиях. Интересно отметить, что такая теория возмущений

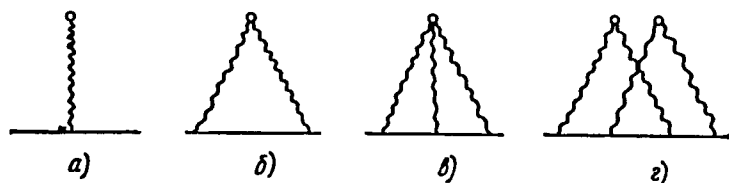


Рис. 11.4. Простейшие диаграммы для функции Грина.

применима при  $\varepsilon/\gamma \gg 1$  и что первая поправка к плотности состояний совпадает с поправкой в формуле (36.13).

Основная идея вывода состоит в том, что мы будем считать импульс передачи  $\hbar q$  в интегралах типа (37.4) малым по сравнению с  $p$ . Если положить в таком интеграле  $\varepsilon_{p-\hbar q} \approx \varepsilon_p$ , то он будет сходящимся и типичное  $q$  будет порядка  $r_0^{-1}$ , т. е. мы разлагаем по  $\hbar/r_0\rho$ . При малых  $\varepsilon$  в интеграле для плотности состояний окажутся существенными импульсы  $p \approx \sqrt{m\gamma}$ , так что параметром нашего разложения будут

$$\left( \frac{r_0 \sqrt{m\gamma}}{\hbar} \right)^{-1} = \left( \frac{E_s}{\gamma} \right)^{1/2}.$$

Такая процедура, однако, может быть проделана лишь в диаграммах, содержащих две волнистые линии в пучке. При большем числе линий в пучке (рис. 11.4, в) пренебрежение импульсом передачи приводит к расходимости. Действительно, каждой новой линии в пучке

соответствует интеграл  $\frac{e^2}{\kappa} \int \frac{L_{p-\hbar q} d^3 q}{(q^2 + r_0^{-2})}$ , линейно расходящийся при замене  $L_{p-\hbar q}$  на  $L_p$ . Поэтому прибавление новой линии соответствует умножению на величину  $G e^2 p / \kappa \hbar$ . При малых  $\epsilon$  существенны  $G \approx 1/\gamma$ ,  $p \approx \sqrt{m\gamma}$  так что

$$G \frac{e^2 p}{\kappa \hbar} \approx \frac{\hbar}{a \sqrt{m\gamma}} \ll \frac{\hbar}{\sqrt{m\gamma} r_0} \ll 1.$$

Поэтому в нескольких первых порядках разложения по  $\hbar/pr_0$  можно учитывать только диаграммы с двумя линиями в пучке. Фактически это следствие того, что к потенциалу  $\phi(r)$  при  $r_0/a \ll 1$  борновское приближение применимо при всех энергиях.

При учете только диаграмм с двумя линиями в пучке диаграммная техника оказывается такой же, как при упругом электрон-фононном взаимодействии с пропагатором  $D(q)$ , определяемым (37.5). Функция Грина удовлетворяет уравнению Дайсона

$$\epsilon - \frac{p^2}{2m} - G_p^{-1} = \int \Gamma(p, p - \hbar q, q) G_{p-\hbar q} D(q) \frac{d^3 q}{(2\pi)^3}, \quad (37.6)$$

где  $\Gamma(p, p - \hbar q, q)$  — вершинная часть. В соответствии со сказанным выше пренебрежем в правой части (37.6) импульсом передачи  $\hbar q$  в аргументах функций  $\Gamma$  и  $G$ . Тогда получим

$$\epsilon - \frac{p^2}{2m} - G_p^{-1} = G_p \Gamma(p, p, 0) \frac{\gamma^2}{2}, \quad (37.7)$$

где

$$\gamma^2 = 2 \int D(q) \frac{d^3 q}{(2\pi)^3}. \quad (37.8)$$

Разумеется, определения (37.8) и (36.7) совпадают. Воспользуемся, далее, тождеством Уорда [18]

$$\Gamma(p, p, 0) = \frac{dG_p^{-1}}{d\epsilon}. \quad (37.9)$$

Тогда получим для  $G_p$  замкнутое уравнение

$$\frac{dG}{dx} \frac{\gamma^2}{2} + Gx - 1 = 0 \quad (37.10)$$

(где  $x = \epsilon - \frac{p^2}{2m} + i\delta$ ), решение которого с граничным условием  $G(x) = 1/x$  при  $x \rightarrow \infty$  имеет вид

$$G = \frac{1}{\sqrt{\pi} \gamma} \int_{-\infty}^{\infty} \frac{\exp\left(-\frac{V^2}{\gamma^2}\right) dV}{\epsilon - \frac{p^2}{2m} - V + i\delta}. \quad (37.11)$$

Формулу (37.3) можно представить в виде

$$g(\epsilon) = -\frac{2}{\pi} \operatorname{Im} \int \frac{d^3 p}{(2\pi\hbar)^3} G_p. \quad (37.12)$$

Подставляя (37.11) в (37.12), найдем

$$g(\epsilon) = \frac{2}{\sqrt{\pi} \gamma} \int_{-\infty}^{\infty} \exp\left(-\frac{V^2}{\gamma^2}\right) dV \int \frac{d^3 p}{(2\pi\hbar)^3} \delta\left(\epsilon - \frac{p^2}{2m} - V\right). \quad (37.13)$$

Производя в (37.13) интегрирование по импульсу, получим (36.11).

Идея расчета более высоких приближений по  $\hbar/\sqrt{m\gamma} r_0$  состоит в следующем. Нужно представить пропагатор  $D(q)$  в виде

$$D(q) = D^0(q) + \tilde{D}(q), \quad (37.14)$$

причем

$$D^0(q) = \frac{1}{2} \gamma^2 (2\pi)^3 \delta(q). \quad (37.15)$$

Далее, нужно считать второй член (37.14) малым по сравнению с первым. Техника расчета поправок разработана в [12].



## ГЛАВА 12

# ГЛУБОКИЙ ХВОСТ ПЛОТНОСТИ СОСТОЯНИЙ И МЕЖЗОННОЕ ПОГЛОЩЕНИЕ СВЕТА

### § 38. Метод оптимальной флуктуации

В работах Гальперина и Лэкса [1, 2], Зиттарца и Ланжера [3] и И. М. Лифшица [4] был предложен изящный метод определения закона экспоненциального спада плотности состояний в области хвоста, который мы будем называть методом оптимальной флуктуации (см. подробный обзор Лифшица и др. [5]). Суть метода сводится к следующему.

Допустим, что характерный размер волновой функции много больше, чем среднее расстояние между примесями во флуктуации. Тогда квантовомеханическое усреднение сглаживает дискретность заряда отдельных примесей и флуктуацию можно характеризовать плавной функцией координат  $\xi(r)$ , представляющей отклонение примесной концентрации от среднего значения. Для нахождения плотности состояний  $g(\varepsilon)$  необходимо просуммировать вклады от всех конфигураций примесей, т. е. от всех потенциальных ям, которые создают электронный уровень с заданной энергией  $\varepsilon$ . В интересующей нас области энергий все эти вероятности экспоненциально малы и указанная сумма определяется вкладом наиболее вероятных конфигураций. Представим вероятность флуктуации  $\xi(r)$  в виде  $\exp(-\Omega\{\xi\})$ . Если среди всех флуктуаций, создающих уровень  $\varepsilon$ , найти самую вероятную флуктуацию  $\tilde{\xi}_\varepsilon$ , то можно утверждать, что плотность состояний спадает с ростом  $\varepsilon$  по закону  $\exp(-\Omega\{\tilde{\xi}_\varepsilon\})$ . При определении  $\tilde{\xi}_\varepsilon$  достаточно рассматривать лишь

такие флуктуации, в которых уровень с энергией  $\varepsilon$  является наименьшим, так как вероятность флуктуации, в которой этот уровень является возбужденным состоянием, безусловно, меньше.

Метод оптимальной флуктуации в том виде, в каком мы его изложили, не позволяет найти предэкспоненциальный множитель в выражении для плотности состояний, однако позволяет определить главный член в показателе экспоненты, т. е.  $\ln [g(\varepsilon)/g(0)]$ . Практически именно эта величина и измеряется в эксперименте.

Разумеется, задача о вычислении предэкспоненциального множителя оказывается гораздо сложнее. При этом, как и в обычном методе перевала, необходимо учитывать состояния, близкие к экстремальному. Кроме того, нужно учитывать трансляционную инвариантность, т. е. тот факт, что локализованное состояние может возникнуть в любой точке координатного пространства. Приближенный метод вычисления предэкспоненты был предложен Гальпериним и Лэксом [1]. Авторы предположили, что волновые функции всех состояний, имеющих заданную энергию и возникающих на флуктуациях, близких к экстремальной, совпадают с экстремальной волновой функцией. Естественно было сравнить полученный ими результат для одномерного «белого шума» с асимптотикой точного решения, существующего в этом случае. Обнаружилось различие предэкспоненциальных множителей в  $1/\sqrt{5}$  раз. В работе [2] был предложен способ отыскания поправок к сделанному приближению, позволяющий значительно уменьшить указанное расхождение.

Одновременно с первой работой Гальперина и Лэкса появилась работа Зиттарца и Ланжера [3], также посвященная отысканию плотности состояний в области, где она экспоненциально мала. В этой работе рассмотрен потенциал типа белого шума (в случаях одного, двух и трех измерений). Получен точный результат, отличающийся от результата Гальперина и Лэкса числом в предэкспоненциальном множителе. Само это число может быть найдено только в одномерном случае, и здесь результат Зиттарца и Ланжера совпадает с асимптотикой точного решения. В трехмерном случае для отыскания этого числа предложена весьма сложная процедура.

В работах Эвардса и др. (см. [13] к гл. 11, [26] к гл. 2, [6]) глубокий хвост плотности состояний исследуется с помощью фейнмановских интегралов по путям. Этот метод в принципе также позволяет получить предэкспоненциальный множитель.

Во всех цитированных выше работах, за исключением [4], исследовался случай гауссовой статистики. В работе Лифшица [4] этого ограничения не было, однако предполагалось, что потенциалы примесных центров перекрываются слабо (радиус действия порядка минимального расстояния между центрами). Это позволяло сделать важное математическое упрощение — считать, что потенциал локально связан с концентрацией  $\xi(r)$ . В СЛП, где радиус экранирования больше среднего расстояния между примесями, как мы увидим ниже, это приближение работает лишь в сравнительно узкой области значений параметров.

В работах Шкловского и Эфроса [7, 8], сделанных специально для полупроводников, подробно рассмотрена область, где гауссова статистика неприменима. Кроме того, в [7] составлена классификация возможных ситуаций, возникающих в полупроводниках при различных соотношениях между массами носителей и концентрациями примесей. Для этой цели использовался простой метод приближенного определения оптимальной флуктуации и плотности состояний, идея которого была высказана Кейном (см. [6] к гл. 11). В этом методе считается, что флуктуации имеют вид равномерно заряженных шаров. Шар характеризуется двумя параметрами — радиусом  $R$  и полным избыточным зарядом  $eZ$ . Один из этих параметров выбирается так, чтобы энергия основного состояния электрона в потенциале, создаваемом флуктуацией, равнялась заданному значению  $\varepsilon$ , а второй параметр варьируется, и находится оптимальная среди этих флуктуаций. Легко понять, что при таком расчете мы теряем лишь численные коэффициенты, зависящие от истинной формы оптимальной флуктуации. Простота метода позволяет легко проанализировать возможные случаи и составить их классификацию. Этот метод изложен в следующем параграфе, и с его помощью исследован спектр основных носителей в СЛП. В § 40 найдено точное распределение примесей в оптимальной флуктуации

и численные коэффициенты, опущенные в § 39. В § 41 рассмотрен спектр неосновных носителей.

Следует иметь в виду, что вся теория, изложенная в этой главе, основана на методе эффективной массы. Отсюда следует, что область энергий, которую мы рассматриваем, простирается в глубину запрещенной зоны не очень далеко. Ширина этой области должна быть значительно меньше ширины запрещенной зоны.

### § 39. Приближение равномерно заряженного шара. Спектр основных носителей

Флуктуационные уровни в запрещенной зоне могут возникнуть либо вследствие скопления притягивающих примесей, либо вследствие разрежения отталкивающих. Для начала примем во внимание только притягивающие примеси. Предположим, что флуктуации представляют собой однородные сферические скопления примесей радиуса  $R$ . Предположим также, что распределение примесей случайно. Тогда вероятность того, что  $k$  примесей попадет в объем, в котором в среднем находится  $\bar{k}$  примесей, определяется формулой Пуассона

$$W = \frac{(\bar{k})^k \exp(-\bar{k})}{k!}. \quad (39.1)$$

Считая, что  $\bar{k} \gg 1$ ,  $k \gg 1$ , и оставляя главный член логарифма вероятности, получим

$$\ln W = k \ln \frac{\bar{k}}{k} + k - \bar{k}. \quad (39.2)$$

В рассматриваемом скоплении в среднем находится  $\bar{k} = NR^3$  примесей\*). Представим  $k$  в виде

$$k = Z + NR^3, \quad (39.3)$$

где  $Z$  — избыточное число примесей во флуктуации. Подставляя  $k$  и  $\bar{k}$  в (39.2), получим вероятность флуктуации

\*) При изложении этого метода мы не будем писать численных коэффициентов, так как это было бы превышением точности. Выражение для  $\bar{k}$ , как и нижеследующие равенства, следует воспринимать как порядковые оценки.

как функцию  $R$  и  $Z$ :

$$\ln W(R, Z) = (Z + NR^3) \ln \frac{NR^3}{Z + NR^3} + Z. \quad (39.4)$$

Наша задача будет заключаться в том, чтобы с помощью уравнения Шредингера найти энергию основного состояния  $\epsilon(R, Z)$  в потенциале, создаваемом шаром с параметрами  $R$  и  $Z$ . Так как характерный размер волновой функции основного состояния оказывается гораздо больше среднего расстояния между примесями, заряд можно считать размазанным по рассматриваемому шару. Найдя  $\epsilon(R, Z)$ , выразим  $Z$  через  $R$  и  $\epsilon$  и, подставив в (39.4), получим  $W(R, \epsilon)$  — вероятность того, что однородное скопление радиуса  $R$  создает электронное состояние с энергией  $\epsilon$ .

Далее найдем  $R = r_{\max}$ , при котором  $\ln W(R, \epsilon)$  имеет абсолютный максимум при фиксированной энергии. Глубоко в запрещенной зоне, где плотность состояний  $g(\epsilon)$  очень мала, а число примесей в скоплении велико, этот максимум можно считать острым, так что

$$\ln \left[ \frac{g(\epsilon)}{g(0)} \right] = \ln W(r_{\max}, \epsilon). \quad (39.5)$$

Такова схема вычислений.

Введем вспомогательную функцию  $W(R, V)$ , определяющую вероятность того, что скопление примесей радиуса  $R$  создает яму, максимальная глубина которой есть  $V$ . Потенциальную энергию  $V$ , как и энергию электрона  $\epsilon$ , мы будем отсчитывать вниз от невозмущенного дна зоны проводимости. Тогда, если  $R \ll r_0$ , где  $r_0$  — радиус экранирования (35.8), то

$$V = \left( \frac{e^2}{\kappa R} \right) Z \quad (R \ll r_0). \quad (39.6)$$

(Мы учитываем только флуктуационный потенциал, создаваемый избыточными примесями. Средний потенциал доноров компенсируется электронами.) Если  $R \gg r_0$ , то вклад в потенциал дают не все примеси скопления, а лишь те, которые находятся на расстоянии, меньшем чем радиус экранирования  $r_0$ . Концентрация избыточных примесей  $Z/R^3$ , а их число в радиусе действия равно

$(Z/R^3)r_0^3$ . Поэтому вместо (39.6) получим

$$V = \frac{e^2}{\kappa r_0} \frac{r_0^3}{R^3} Z \quad (R \gg r_0). \quad (39.7)$$

Выражая  $Z$  через  $R$  и  $V$  с помощью (39.6) и (39.7) и подставляя в (39.4), получим  $\ln W(R, V)$ .

Вспомогательная функция  $W(R, V)$  имеет некоторое отношение к искомой функции  $W(R, \varepsilon)$ . Действительно, потенциальная яма, создаваемая равномерно заряженным шаром, имеет электронный уровень  $\varepsilon$ , близкий к потенциальной энергии в центре шара  $V$ , если  $V \gg \hbar^2/mR^2$ . Поэтому при  $R \gg r_q \equiv \hbar/\sqrt{mV}$  мы можем получить  $\ln W(R, \varepsilon)$  из  $\ln W(R, V)$ , просто заменяя  $V$  на  $\varepsilon$ . При  $R \leq r_q$ , однако, в (39.4) следует подставлять функцию  $Z(R, \varepsilon)$ , полученную квантовомеханическим путем.

В этом и следующем параграфах мы будем изучать спектр основных носителей в сильно легированном и слабо компенсированном полупроводнике (для определенности  $n$ -типа). Как было показано в гл. 11, для  $\varepsilon < \varepsilon_c$  (см. (36.19)) плотность состояний выражается квазиклассической формулой (36.11). Так как, строго говоря,  $\varepsilon_c > \gamma$ , то область применимости метода оптимальной флуктуации ( $|\ln[g(\varepsilon)/g(0)]| = \varepsilon^2/\gamma^2 \gg 1$ ) перекрывается с областью применимости формулы (36.11). Поэтому при не очень больших энергиях мы должны снова получить эту формулу.

Введем характерную длину  $r_2 = (V\kappa/Ne^2)^{1/2}$ . Отношение  $r_2/r_0 = (V/\bar{V})^{1/2}$ , где  $\bar{V} = (e^2/\kappa r_0)NR_0^3$ . Согласно (36.10)  $\bar{V}$  есть средняя потенциальная энергия, создаваемая донорами. Так как  $\bar{V} > \gamma$ , то отношение  $r_2/r_0$  в области, где плотность состояний экспоненциально мала, может быть произвольным. График функции  $\ln W(R, V)$  при  $V < \bar{V}$  и  $V > \bar{V}$  схематически изображен на рис. 12.1.

Рассмотрим сначала случай  $V < \bar{V}$ . Легко убедиться, что область  $R \gg r_2$  на рис. 12.1, *a* соответствует гауссовой статистике. Действительно, условие применимости гауссовой статистики есть  $Z \ll NR^3$ . Если  $R \ll r_0$ , то, воспользовавшись (39.6), получим

$$\frac{Z}{NR^3} = \frac{V}{(e^2/\kappa R)NR^3} = \left(\frac{r_2}{R}\right)^2. \quad (39.8)$$

Итак, при  $R \gg r_2$

$$\ln W = -\frac{Z^2}{NR^3}. \quad (39.9)$$

Подставляя (39.6) и (39.7) в (39.9), получим

$$\ln W = -\frac{V^2}{[(e^2/\kappa R)(NR^3)^{1/2}]^2} \quad (r_2 \ll R \ll r_0) \quad (39.10)$$

и

$$\ln W = -\frac{V^2 R^3 \kappa^2}{e^4 N r_0^4} \quad (R \gg r_0). \quad (39.11)$$

Таким образом, при  $R \gg r_2$  функция  $\ln W(R, V)$  растет с  $R$  вплоть до  $R=r_0$ , а затем убывает (см. рис. 12.1, а).

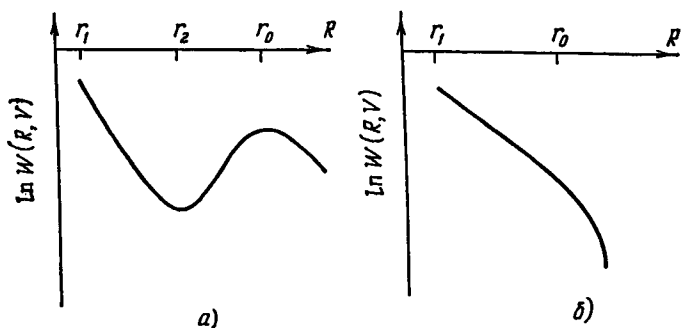


Рис. 12.1. Логарифм вероятности создания потенциальной ямы глубиной  $V$  скоплением радиуса  $R$  как функция  $R$ : а)  $V < \bar{V}$ ; б)  $V > \bar{V}$ .

Флуктуации с  $R > r_0$  оказываются менее вероятными из-за экранирования. Таким образом, показано, что функция  $\ln W(R, V)$  имеет относительный максимум в точке  $R = r_0$ . Из неравенства  $r_q/r_0 = [(\gamma/V)(E_s/\gamma)]^{1/2} \ll 1$  следует, что этот максимум находится в области  $R \gg r_q$ , т. е. там, где энергия основного состояния  $\epsilon$  совпадает с дном ямы  $V$ . Тогда, в соответствии с (39.5), подставим  $R = r_0$  в (39.10) или (39.11) и получим вклад в плотность состояний от только что найденного максимума:

$$\ln \frac{g(\epsilon)}{g(0)} = -\frac{\epsilon^2}{\gamma^2}. \quad (39.12)$$

Эта формула совпадает с результатом квазиклассической теории (36.11), однако, как мы увидим, она имеет гораздо более широкую область применимости.

Как видно из рис. 12.1, максимум при  $R = r_0$  существует и формула (39.12) имеет место только при  $\varepsilon \ll V$ . Подставляя в (39.8)  $R = r_0$ , убеждаемся в том, что при  $V = V$  флуктуации в яме размера  $r_0$  перестают быть гауссовыми. Легко показать, что значения  $V$ , большие чем  $V$ , вообще не могут быть получены с помощью гауссовых флуктуаций (на кривой рис. 12.1, б совсем нет гауссова участка). Однако еще при  $\varepsilon < V$  флуктуации, лежащие левее минимума на рис. 12.1, а, могут оказаться более вероятными, чем флуктуации с  $R = r_0$ . Для этого необходимо (но недостаточно), чтобы при  $V = \varepsilon$  выполнялось неравенство  $r_q < r_2$ . Тогда  $\ln W(R, \varepsilon)$  возрастает с уменьшением  $R$  вплоть до  $R = \hbar/\sqrt{m\varepsilon}$ .

При  $R \ll r_2$ , учитывая (39.8), из (39.4) получим

$$\ln W = -Z \ln \frac{Z}{NR^3}. \quad (39.13)$$

Подставляя (39.6) в (39.13), найдем

$$\ln W(R, V) = -\frac{VR\kappa}{e^2} \ln \frac{V\kappa}{e^2 NR^2}. \quad (39.14)$$

Функция (39.14) возрастает с уменьшением  $R$ , что и изображено на рис. 12.1, а. При этом  $Z$  уменьшается, обращаясь в единицу при  $R = r_1 = e^2/V\kappa$ . Однако, если  $V \gg E_0$  (где  $E_0 = me^4/\hbar^2\kappa^2$ ), яма окажется мелкой (не будет иметь связанного состояния) при гораздо больших значениях  $R$ , чем  $r_1$ . Действительно,  $r_1/r_q = (E_0/V)^{1/2} < 1$ .

Легко сообразить, как выглядит функция  $\varepsilon(R, Z)$ , если  $R \ll \hbar/\sqrt{m\varepsilon}$ . В этом случае флуктуация представляет собой почти точечное ядро атома, наименьший терм которого реализует энергию  $\varepsilon$ . С поправкой на неточечность ядра имеем

$$\varepsilon = Z^2 E_0 \left( 1 - \frac{R^2}{(a/Z)^2} \right). \quad (39.15)$$

Таким образом, при  $R \ll \hbar/\sqrt{m\varepsilon}$  энергия основного состояния  $\varepsilon$  отнюдь не совпадает с максимальной глубиной  $V$  (39.6), стремящейся к бесконечности при  $R \rightarrow 0$ , а стремится к постоянной величине  $Z^2 E_0$ .



Обращая (39.15), получим

$$Z = \sqrt{\frac{\epsilon}{E_0}} \left( 1 + \frac{R^2}{a^2} \frac{\epsilon}{E_0} \right), \quad R \ll \frac{\hbar}{\sqrt{m\epsilon}} \approx a \sqrt{\frac{E_0}{\epsilon}}. \quad (39.16)$$

Подставляя (39.16) в (39.13), видим, что при  $R \rightarrow 0$  функция  $\ln W(R, \epsilon) \rightarrow -\infty$ . Так как при  $R \gg \hbar/\sqrt{m\epsilon}$  для  $\ln W(R, \epsilon)$  применима формула (39.14), в которой можно положить  $V = \epsilon$ , это означает, что  $\ln W(R, \epsilon)$  имеет второй максимум (рис. 12.2). Он определяется конкуренцией комбинаторного логарифма в (39.13), который показывает, что более вероятны «рыхлые» ядра, и поправки на неточность ядра в (39.16), свидетельствующей о том, что точечные ядра могут обеспечить заданную энергию при меньшем  $Z$ . Так как поправка на неточность входит в (39.13) степенным образом, она влияет сильнее и оптимальное ядро оказывается точечным ( $r_{\max} \ll \hbar/\sqrt{m\epsilon}$ ). Поэтому для нахождения  $r_{\max}$  достаточно подставить (39.16) в (39.13) и найти максимум по  $R$ . В результате для оптимальной флуктуации имеем

$$r_{\max} \approx \frac{\hbar}{\sqrt{m\epsilon}} \frac{1}{\left\{ \ln \left[ \left( \frac{\epsilon}{E_0} \right)^2 \frac{1}{Na^3} \right] \right\}^{1/2}}. \quad (39.17)$$

При выводе предполагалось, что аргумент логарифма в (39.17) много больше единицы. При энергиях, для которых плотность состояний определяется рассматриваемым максимумом ( $\epsilon > E_1$ ; см. ниже), это условие выполняется.

Таким образом, размер оптимального скопления мал по сравнению с его боровским радиусом, т. е. это скопление представляет собой почти точечное ядро атома.

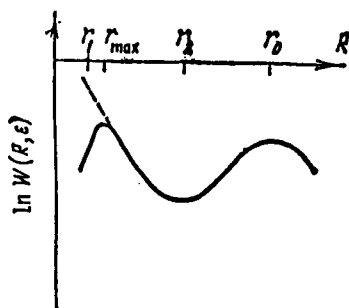


Рис. 12.2. Логарифм вероятности создания электронного состояния с энергией  $\epsilon$  скоплением радиуса  $R$  как функция  $R$  ( $E_s < \epsilon < \bar{V}$ ). Пунктир показывает, как бы шла эта кривая без учета квантовых эффектов.

Подставляя  $Z = \sqrt{\varepsilon/E_0}$  и  $r_{\max}$  из (39.17) в (39.13), получим [7, 8]

$$\ln \frac{g(\varepsilon)}{g(0)} = - \sqrt{\frac{\varepsilon}{E_0}} \ln \left[ \left( \frac{\varepsilon}{E_0} \right)^2 \frac{D(\varepsilon)}{Na^3} \right], \quad (39.18)$$

где

$$D(\varepsilon) \approx \left\{ \ln \left[ \left( \frac{\varepsilon}{E_0} \right)^2 \frac{1}{Na^3} \right] \right\}^{3/2}. \quad (39.19)$$

Из сравнения (39.18) и (39.12) видно, что (39.18) имеет место при  $\varepsilon > E_1$ , а (39.12) — при  $\varepsilon < E_1$ , где

$$E_1 = \frac{\gamma^{4/3}}{E_0^{1/3}} \left\{ \ln \left[ \left( \frac{\gamma}{E_0} \right)^{8/3} \frac{1}{Na^3} \right] \right\}^{2/3}. \quad (39.20)$$

Легко убедиться в том, что в рассматриваемом случае  $E_1 < \bar{V}$ . Если  $\varepsilon > \bar{V}$ , то, как уже говорилось, максимум при  $R = r_0$  исчезает, а рассуждения, относящиеся к максимуму при  $R = r_{\max}$  (39.17), совершенно не меняются. Таким образом, формула (39.18) справедлива при всех энергиях  $\varepsilon > E_1$ , лежащих в пределах метода эффективной массы. Обсудим теперь предположения, сделанные при выводе этой формулы.

При выводе (39.18) мы рассматривали  $Z$  как непрерывную переменную. Это верно, если ширина максимума функции  $\ln W(Z, \varepsilon)$  по переменной  $Z$  велика по сравнению с единицей. Легко убедиться в том, что это имеет место, если

$$\left( \frac{\varepsilon}{E_0} \right)^{1/4} \gg \ln \left[ \left( \frac{\varepsilon}{E_0} \right)^2 \frac{1}{Na^3} \right]. \quad (39.21)$$

Формула (39.18) получена в приближении равномерно заряженного шара. При определении размера оптимальной флуктуации мы не учитывали численных множителей, так как в действительности они зависят от распределения плотности по флуктуации. Однако утверждение, что  $r_{\max} < a/Z$ , не зависит от сделанного приближения. Поэтому предлогарифмический множитель в (39.18) не содержит никаких дополнительных чисел, что и подтверждается теорией, развитой в § 40; в которой определяется истинное распределение заряда по флуктуации и точно находится функция  $D(\varepsilon)$ . К этому следует добавить, что в рамках метода эффективной массы множитель перед логарифмом не зависит от вида спек-

ра. Действительно, вид этого множителя является следствием того, что  $\epsilon \propto Z^2$ . Но это выполняется всегда, когда гамильтониан есть квадратичная форма компонент импульса.

Все рассуждения этого параграфа основывались на одноэлектронном приближении и предположении, что примеси экранируются линейным образом. Однако, по существу, линейное экранирование требовалось только при выводе (39.12). Так как в слабо компенсированном полупроводнике  $E_1 < \mu_0$ , где  $\mu_0$  — энергия Ферми, предположение о линейном экранировании не накладывает ограничений на область применимости (39.12). Что же касается формулы (39.18), то учет электрон-электронного взаимодействия не влияет на нее и при  $\epsilon > \mu_0$ . Рассмотренное нами образование представляет собой многоэлектронный атом, наинизший терм которого реализует требуемую энергию  $\epsilon$ . Но, как известно, при большом заряде ядра внешние электроны слабо меняют энергию этого терма (см. [8] к гл. 1).

Таким образом, (39.18) является асимптотическим выражением для плотности состояний при больших энергиях. Важно, однако, что мы считали примеси распределенными случайно. Если центры созданы с помощью облучения при низких температурах, это предположение может быть правильным. Если же образец подвергался высокотемпературной обработке, в процессе которой примеси могли свободно перемещаться, в их распределении возникает корреляция, связанная со взаимодействием. Коэффициент диффузии примесей очень резко убывает с температурой. Так как образцы остывают достаточно быстро, можно считать, что распределение примесей является мгновенной фотографией плазмы, существовавшей при некоторой температуре  $T_0$ , соответствующей прекращению диффузии. Эта температура близка обычно к температуре плавления.

Учет корреляции оказывается важным при вычислении асимптотики  $g(\epsilon)$  при больших  $\epsilon$ . Действительно, мы считали, что вероятность образования ядра из  $Z$  примесей, имеющего радиус  $R$ , есть  $\exp[-Z \ln(Z/NR^3)]$ . Однако при температуре  $T_0$  эти примеси отталкивались друг от друга. Для того чтобы их собрать, необходимо совершить работу  $Z^2 e^2 / \kappa R$ . Поэтому вероятность такой

флуктуации есть

$$\exp\left(-Z \ln \frac{Z}{NR^3}\right) \exp\left(-\frac{Z^2 e^2}{\kappa R k T_0}\right). \quad (39.22)$$

При достаточно большом  $Z$  второй множитель обязательно станет важнее первого. Найдем размер оптимальной флуктуации в этом случае. Если  $R \gg \hbar/\sqrt{m\varepsilon}$ , то яма классическая и  $\varepsilon \approx Ze^2/\kappa R$ . (Имеются в виду значения  $R$ , малые по сравнению с радиусом экранирования.) Отсюда  $Z \approx \kappa R \varepsilon/e^2$ , а вероятность такой ямы  $\exp\left(-\frac{\kappa R \varepsilon^2}{e^2 k T_0}\right)$  убывает с увеличением  $R$ . Если  $R \ll \hbar/\sqrt{m\varepsilon}$ , яма представляет собой точечное ядро:  $Z = \sqrt{\varepsilon/E_0}$ , а вероятность равна  $\exp\left(-\frac{\varepsilon e^2}{E_0 \kappa R k T_0}\right)$  и убывает с уменьшением  $R$ . Таким образом, оптимальная флуктуация имеет размер  $R \approx \hbar/\sqrt{m\varepsilon}$ , равный по порядку величины радиусу первой орбиты. При этом  $Z \approx \sqrt{\varepsilon/E_0}$ . Подставляя указанные значения  $R$  и  $Z$  во второй множитель (39.22), получим [8, 9, 10]

$$\ln \frac{g(\varepsilon)}{g(0)} = -C \frac{\varepsilon^{3/2}}{E_0^{1/2} k T_0}, \quad (39.23)$$

где  $C$  — численный коэффициент. Это и есть асимптотика плотности состояний с учетом корреляции примесей. В работе [8] получено безразмерное уравнение, определяющее форму оптимальной флуктуации и величину  $C$ .

Итак, мы рассмотрели плотность состояний основных носителей в сильно легированном и слабо компенсированном полупроводнике и доказали, что в отсутствие корреляции она описывается формулами (39.12) и (39.18), а при наличии корреляции — формулами (39.12) и (39.23).

#### § 40. Точное распределение примесей в оптимальных флуктуациях

а) Вывод основных уравнений. В предыдущем параграфе мы произвели качественный анализ, в котором предполагалось, что флуктуации представляют собой однородно заряженные шары, отличающиеся друг от

друга только радиусом и плотностью заряда. Переходим теперь к изложению строгой теории, позволяющей найти истинную форму оптимальных флуктуаций. При этом рассматривается только некоррелированное распределение примесей. Обобщение метода на случай коррелированного распределения содержится в работе [8].

Введем функцию  $\xi(r) = N(r) - N$ , где  $N(r)$  — концентрация примесей в точке  $r$ , усредненная по объему, большему чем  $N^{-1}$ . Найдем вероятность флуктуации  $\xi(r)$ , считая, что корреляция в распределении примесей отсутствует. Для этого рассмотрим элемент объема  $\Delta V$ , достаточно малый, чтобы считать внутри него концентрацию  $N(r)$  постоянной, но содержащий большое число примесей  $\mathcal{N} = N(r)\Delta V \gg 1$ . Для подсчета числа состояний допустим, что примеси могут находиться только в узлах решетки, причем концентрация этих узлов  $\rho \gg \gg N(r)$ . Найдем число способов, которым можно разместить  $\mathcal{N}$  примесей по узлам решетки, находящимся в объеме  $\Delta V$ . Число узлов  $\mathcal{P} = \rho \Delta V$ , а искомое число способов  $\Gamma_{\Delta V}$  есть число сочетаний из  $\mathcal{P}$  по  $\mathcal{N}$ :

$$\Gamma_{\Delta V} = \frac{\mathcal{P}!}{\mathcal{N}!(\mathcal{P} - \mathcal{N})!}. \quad (40.1)$$

Для энтропии объема  $\Delta V$  получим

$$S_{\Delta V} = \ln \Gamma_{\Delta V} = \mathcal{N} \ln \frac{\mathcal{P}e}{\mathcal{N}}. \quad (40.2)$$

Введем теперь плотность энтропии

$$\sigma(r) = \frac{S_{\Delta V}}{\Delta V} = N(r) \ln \frac{\rho e}{N(r)}. \quad (40.3)$$

Полная энтропия кристалла, имеющего объем  $V_0$ ,

$$S = \int_{V_0} d^3r \sigma(r). \quad (40.4)$$

Найдем энтропию такого распределения примесей, когда в некотором объеме  $V$ , составляющем ничтожную долю  $V_0$ , имеет место флуктуация  $\xi(r)$ , вообще говоря сравнимая с  $N(r)$ , а в остальной части объема  $\xi(r) \ll N$ . Из

закона сохранения числа частиц имеем

$$\int_{V_0-V} \xi(\mathbf{r}) d^3r = - \int_V \xi(\mathbf{r}) d^3r. \quad (40.5)$$

Энтропия такой флуктуации

$$S = \int_V (N + \xi) \ln \frac{pe}{N + \xi} d^3r + (V_0 - V) N \ln \frac{pe}{N} + \\ + \int_{V_0-V} \left( \frac{d\sigma}{dN(\mathbf{r})} \right)_{N(\mathbf{r})=N} \xi d^3r. \quad (40.6)$$

Преобразуя последний член с помощью (40.5), найдем, что отличие  $S$  от энтропии равномерного распределения примесей  $S_0 = NV_0 \ln (pe/N)$  равно

$$S - S_0 = \int_V \left[ (N + \xi) \ln \frac{N}{N + \xi} + \xi \right] d^3r. \quad (40.7)$$

Если  $\xi(\mathbf{r})$  убывает достаточно быстро, так что интеграл (40.7) сходится, интегрирование можно распространить до бесконечности. Тогда вероятность того, что в некоторой части объема возникла флуктуация  $\xi(\mathbf{r})$ , есть  $\exp(-\Omega\{\xi\})$ , где

$$\Omega\{\xi\} = - \int \left[ (N + \xi) \ln \frac{N}{N + \xi} + \xi \right] d^3r. \quad (40.8)$$

Введем теперь функционалы  $V_r\{\xi\}$  и  $\lambda_0\{\xi\}$ , первый из которых есть потенциальная энергия пробного электрона в точке  $\mathbf{r}$ , созданная флуктуацией  $\xi$ , а второй — энергия основного состояния электрона, определяемая из уравнения Шредингера

$$-\frac{\hbar^2}{2m} \Delta\Psi + V_r\{\xi\} \Psi = \lambda_0\Psi. \quad (40.9)$$

В приближении линейного экранирования (см. (35.9))

$$-V_r\{\xi\} = e \int \xi(\mathbf{r}') K(\mathbf{r} - \mathbf{r}') d^3r', \quad (40.10)$$

$$K(\mathbf{r}) = \left( \frac{e}{\kappa r} \right) \exp\left(-\frac{r}{r_0}\right).$$

Оптимальная флуктуация  $\xi(r)$  должна осуществлять минимум  $\Omega\{\xi\}$  при условии

$$\lambda_0\{\xi\} = -\varepsilon. \quad (40.11)$$

Соответствующее уравнение имеет вид

$$\delta(\Omega\{\xi\} + \beta\lambda_0\{\xi\}) = 0, \quad (40.12)$$

где  $\beta$  — множитель Лагранжа, который следует находить из (40.11).

Варьируя (40.8) и учитывая, что

$$\delta\lambda_0\{\xi\} = - \int |\Psi(r')|^2 d^3r' e \int K(r' - r'') \delta\xi(r'') d^3r'', \quad (40.13)$$

получим окончательно уравнение [4]

$$\ln \frac{N + \xi}{N} - \beta \int |\Psi(r')|^2 e K(r - r') d^3r' = 0, \quad (40.14)$$

которое вместе с (40.9) определяет оптимальную флуктуацию  $\xi$ . Главный член логарифма плотности состояний выражается формулой

$$\ln \frac{g(\varepsilon)}{g(0)} = -\Omega(\xi). \quad (40.15)$$

Теперь мы приведем строгий вывод формул (39.12), (39.18) и определим численные множители, зависящие от формы флуктуации.

**б) Вывод формулы (39.12).** Оптимальная флуктуация, соответствующая (39.12), описывается гауссовой статистикой. В этом случае ( $\xi \ll N$ ) (40.8) имеет вид

$$\Omega\{\xi\} = \frac{1}{2N} \int \xi^2 d^3r, \quad (40.16)$$

а вместо (40.14) получим

$$\xi = N\beta \int |\Psi(r')|^2 e K(r - r') d^3r'. \quad (40.17)$$

Предположим, что характерный размер волновой функции  $\lambda \ll r_0$ , тогда

$$\xi = N\beta e K(r) \quad (r \gg \lambda), \quad (40.18)$$

а потенциальная энергия оптимальной ямы, согласно (40.10), равна

$$V_r \{\xi\} = -N\beta e^2 \int K(r') K(r-r') d^3r', \quad (40.19)$$

или

$$V_r \{\xi\} = -\frac{e^4 N \beta r_0}{\kappa^2} 2\pi \exp\left(-\frac{r}{r_0}\right). \quad (40.20)$$

Потенциальная энергия спадает на расстоянии порядка  $r_0$ , и при  $\epsilon \gg E_s = \hbar^2/mr_0^2$  энергия основного состояния близка ко дну ямы. Поэтому можно определить  $\beta$  из условия

$$V_r \{\xi\} |_{r=0} = -\epsilon. \quad (40.21)$$

Тогда

$$\beta = \frac{\epsilon \kappa^2}{2\pi e^4 N r_0}. \quad (40.22)$$

Подставляя (40.22) в (40.18) и (40.16) в (40.15), получим формулу (39.12), причем численный коэффициент в  $\gamma$  оказывается таким же, как в (36.7).

Найдем теперь длину волны электрона в основном состоянии. Для этого заметим, что согласно (40.17) при  $r \ll \lambda$

$$\xi \approx \frac{N\beta e^2}{\kappa \lambda} \quad (r \ll \lambda) \quad (40.23)$$

и не зависит от  $r$ . Таким образом, внутренняя часть флуктуации ( $r < \lambda$ ) представляет собой равномерно заряженный шар. Для того чтобы решение было самосогласованным, радиус шара должен быть таким, чтобы в потенциале, созданном шаром, был электронный уровень с длиной волны порядка радиуса шара. Шар радиуса  $\lambda$  с избыточной плотностью (40.23) создает потенциальную яму глубиной  $N\beta (e^2/\kappa)^2 \lambda$ . Радиус  $\lambda$  можно найти, приравняв эту величину к  $\hbar^2/m\lambda^2$ . В результате получим

$$\lambda \approx r_0 (E_s/\epsilon)^{1/3}. \quad (40.24)$$

При  $\epsilon > E_s$  предположенное при выводе (40.18) неравенство  $\lambda \ll r_0$  выполняется.

В работе Гальперина и Лэкса [1] уравнения (40.9), (40.17) решались численно. Интересно сравнить приведенное выше аналитическое решение с результатами этих расчетов. На рис. 12.3 представлены оптимальный



потенциал и волновая функция, полученные при  $\nu = \epsilon/E_s \gg 1$ .

На графиках отложены безразмерный потенциал  $U(r) = V_0 \kappa^2 / 2\pi e^4 N \beta r_0$  и безразмерная волновая функция

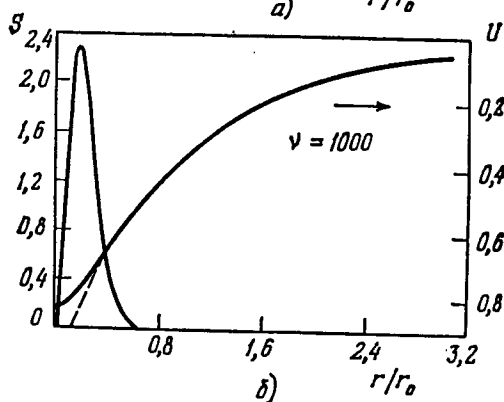
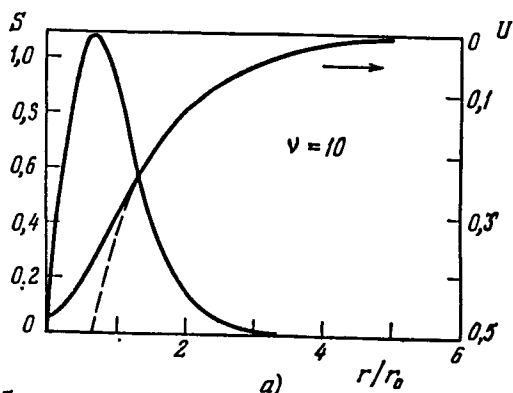


Рис. 12.3. Зависимости безразмерного потенциала  $U(r)$  и волновой функции  $S(r)$  от  $r/r_0$ , полученные в [1] с помощью ЭВМ. Пунктиром показан потенциал, вычисленный по формуле (40.20), в области, где он отличается от  $U(r)$ : а)  $\nu = 10$ ; б)  $\nu = 10^3$ .

$S(r) = \sqrt{4\pi} \Psi(r) r r_0^{1/2}$ . В соответствии с (40.20) в тех местах, где  $S(r)$  мала,  $U(r)$  точно совпадает с графиком функции  $\exp(-r/r_0)$ , который нанесен на рис. 12.3 пунктиром. Так как мы не получили точного вида волновой функции, то можем лишь приближенно сравнить (40.24) с рис. 12.3. Для этого допустим, что длина волны  $\lambda$  пропорциональна координате максимума  $S(r)$ . Отношение

полученных таким образом длин волн при  $\nu = 10$  и  $\nu = 10^3$  равно  $\lambda_{10}/\lambda_{1000} = 4,4$ . Согласно (40.24) это отношение равно  $(100)^{1/2} = 4,6$ . Таким образом, аналитическое и численное решения хорошо согласуются друг с другом.

в) Асимптотика плотности состояний в отсутствие корреляции примесей. Перепишем (40.14) в виде

$$\xi = N [\exp \chi(r) - 1], \quad (40.25)$$

где

$$\chi(r) = \frac{\beta e^2}{\kappa} \int \frac{|\Psi(r')|^2}{|r - r'|} \exp\left(-\frac{|r - r'|}{r_0}\right) d^3 r'. \quad (40.26)$$

Покажем теперь, что в области больших энергий существует сферически симметричное решение уравнений (40.9), (40.25). Это решение обладает тем свойством, что функция  $\xi$  экспоненциально спадает на расстоянии, малом по сравнению с длиной волны электрона в основном состоянии  $\lambda$ . (По порядку величины это расстояние совпадает с  $r_{\max}$  (39.17).) Кроме того, оказывается, что  $\lambda \ll r_0$ . Предположив все это заранее, следует считать, что флуктуация образует почти точечное ядро иона с зарядом

$$Z = \int \xi(r) d^3 r, \quad (40.27)$$

а волновая функция основного состояния имеет вид

$$\Psi(r) = \frac{1}{\sqrt{\pi \lambda^{3/2}}} \exp\left(-\frac{r}{\lambda}\right), \quad (40.28)$$

причем  $\lambda = a/Z$ . Поскольку основное состояние должно иметь энергию  $\epsilon$ , то

$$Z = \sqrt{\frac{\epsilon}{E_0}}, \quad \lambda = a \sqrt{\frac{E_0}{\epsilon}}. \quad (40.29)$$

Докажем, что такое решение удовлетворяет уравнениям (40.9), (40.25). Для этого подставим (40.28) в (40.26) и найдем явный вид  $\chi(r)$  и  $\xi(r)$ . Для определения  $\beta$  подставим  $\xi$  в (40.27) и воспользуемся (40.29).

Подставляя (40.28) в (40.26) и разлагая по  $r/\lambda$ , получим

$$\chi(r) = t - \frac{r^2}{\lambda^2}, \quad (40.30)$$

где

$$t = \frac{\beta e^2}{\kappa \lambda}, \quad x^2 = \frac{3}{2} \frac{\lambda^2}{t}. \quad (40.31)$$

Видно, что сделанное выше предположение о резком спаде  $\xi(r)$  на малых расстояниях эквивалентно предположению  $t \gg 1$ . Если действительно  $t \gg 1$ , то при  $r \leq \lambda$  показатель экспоненты в (40.25) велик и

$$\xi(r) = N e^t \exp\left(-\frac{r^2}{x^2}\right). \quad (40.32)$$

Подставив (40.32) и (40.29) в (40.27), получим трансцендентное уравнение для определения  $t$ :

$$t = \ln \left[ \left( \frac{e}{E_0} \right)^2 \frac{t^{3/2}}{N a^3} \left( \frac{2}{3\pi} \right)^{3/2} \right]. \quad (40.33)$$

Отсюда при достаточно больших энергиях

$$t \approx \ln \left[ \left( \frac{e}{E_0} \right)^2 \frac{1}{N a^3} \right] \gg 1, \quad (40.34)$$

что и оправдывает сделанное предположение. Таким образом, предсказанное решение действительно существует.

Для вычисления  $\Omega\{\xi\}$  учтем, что  $\xi \gg N$  и (40.8) можно представить в виде

$$\Omega\{\xi\} = \int \xi \left[ \ln \frac{\xi}{N} - 1 \right] d^3r. \quad (40.35)$$

Подставляя (40.32) в (40.35), легко получить

$$\Omega\{\xi\} = Z(t-1) - \frac{3}{2}Z. \quad (40.36)$$

В этом выражении имеются члены порядка  $Zt$  и  $Z$ . Поэтому сюда нельзя подставлять  $Z = \sqrt{e/E_0}$ , а следует учесть поправки к этому выражению порядка  $t^{-1}$ . Энергия основного состояния с поправкой на неточность ядра имеет вид [11]

$$e = Z^2 E_0 \left( 1 - \frac{4}{3} \frac{\langle r^2 \rangle}{\lambda^2} \right). \quad (40.37)$$

Отсюда

$$Z = \sqrt{\frac{e}{E_0}} \left( 1 + \frac{2}{3} \frac{\langle r^2 \rangle}{\lambda^2} \right). \quad (40.38)$$

Здесь

$$\langle r^2 \rangle = \frac{1}{Z} \int \xi(r) r^2 d^3r. \quad (40.39)$$

Подставляя (40.32) в (40.39) и (40.38), получим

$$Z = \sqrt{\frac{\varepsilon}{E_0}} \left( 1 + \frac{3}{2t} \right). \quad (40.40)$$

Поправки к уравнению (40.33), полученные с помощью (40.38), порядка  $t^{-1}$ , и их можно не учитывать. Подставляя (40.40) в (40.36), найдем

$$\Omega \{ \xi \} = \left( \frac{\varepsilon}{E_0} \right)^{1/2} (t - 1). \quad (40.41)$$

Подстановка (40.41) в (40.15) приводит к результату (39.18), причем

$$D(\varepsilon) = \left( \frac{2}{3\pi} \right)^{3/2} t^{3/2} \exp(-1), \quad (40.42)$$

а  $t$  определяется уравнением (40.33).

Заметим, однако, что экспоненциальный спад  $\xi(r)$ , описываемый (40.32), происходит только при  $r < \lambda$ . На больших расстояниях он оказывается более плавным. При  $r \gg \lambda$  из (40.26) получим

$$\chi(r) = \frac{\beta e^2}{\kappa r} \exp\left(-\frac{r}{r_0}\right). \quad (40.43)$$

Подставляя  $\beta e^2/\kappa$  из (40.31), убеждаемся, что при  $r \gg \lambda$  величина  $\chi(r) \ll 1$  и, следовательно,

$$\xi(r) = Nt \left( \frac{\lambda}{r} \right) \exp\left(-\frac{r}{r_0}\right) \quad (r \gg \lambda). \quad (40.44)$$

Потенциал, создаваемый этой частью заряда в центре флуктуации, приводит к классическому сдвигу уровня на величину

$$\frac{e^2}{\kappa} \int \frac{\xi(r') d^3r'}{r'} \approx \frac{e^2}{\kappa} t N \lambda r_0. \quad (40.45)$$

(Сдвиг можно считать классическим, так как основной вклад в интеграл дает область, где  $r \approx r_0 \gg \lambda$ .) Изложенная выше теория справедлива, если правая часть (40.45) мала по сравнению с  $\varepsilon$ . Для этого необходимо условие  $\varepsilon \geq E_1$ , где  $E_1$  определяется (39.20). Кроме того,

разумеется, необходимо еще условие  $\lambda < r_0$ , т. е.  $\epsilon > E_s$ . Этих двух условий также достаточно для того, чтобы выражение, стоящее под знаком логарифма, было велико. Таким образом, область применимости формулы (39.18) в предыдущем параграфе была получена правильно.

### § 41. Спектр неосновных носителей

Как будет показано в следующем параграфе, в вырожденном полупроводнике частотная зависимость коэффициента межзонного поглощения света на частотах ниже пороговой определяется именно спектром неосновных носителей. Задача о вычислении плотности состояний имеет в этом случае две важные особенности.

1) Основные примеси являются для неосновных носителей отталкивающими центрами (например, заряженный донор для дырки).

2) Обычно массы электрона и дырки резко отличаются друг от друга. Это обстоятельство привносит в задачу значительное разнообразие (см. [8] к гл. 11). Как мы видели, при выводе формулы (39.12) было существенно условие  $\gamma > E_s \equiv \hbar^2/mr_0^2$ , означающее, что типичная

потенциальная яма с размером  $r_0$  содержит квантовые уровни. Если речь идет о неосновных носителях, то в выражение для  $E_s$  следует подставлять их массу, а в выражение для  $r_0$  — массу основных носителей, осуществляющих экранирование. При этом возможны различные ситуации, которые мы сейчас и проанализируем.

а) Классический случай ( $\gamma \gg E_s$ ). Рассмотрим спектр дырок в электронном полупроводнике, предполагая, что масса дырки  $m_h$  больше массы электрона  $m_e$ . Предположим сначала, что компенсация вообще отсутствует, т. е. имеются только донорные примеси, причем для электронов выполняется условие сильного легирования  $Na^3 \gg 1$ , где  $a = \chi \hbar^2/m_e e^2$  — боровский радиус электрона. Условие  $\gamma \gg E_s$  теперь выполняется легче, поскольку оно имеет вид  $\hbar^2/m_h r_0^2 \ll \gamma$ , или  $(m_h/m_e)(Nr_0^3)^{1/6} \gg 1$ .

Так как по отношению к дыркам доноры являются отталкивающими примесями, глубокие дырочные состояния создаются не скоплениями, а разрежениями доноров. В гауссовой статистике скопления и разрежения

равновероятны. Легко поэтому убедиться в том, что вывод, приводящий к формуле (39.12), остается без изменений. Согласно (36.10) гауссова статистика применима при  $\epsilon \ll V$ . При больших энергиях симметрия скоплений и разрежений нарушается.

В отсутствие акцепторов спектр дырок не простирается сколь угодно глубоко в запрещенную зону, а имеет точку окончания. Эта точка лежит выше вершины валентной зоны кристалла без примесей на величину, близкую к  $\mu$ , где  $\mu$  — уровень Ферми электронов, отсчитанный от дна зоны проводимости. Уровни дырок, близкие к точке окончания, возникают в таких областях пространства, где вследствие флуктуации доноры почти полностью отсутствуют. Фактически речь идет о флуктуациях, представляющих макроскопическое вкрапление чистого полупроводника. При нулевой температуре уровень Ферми в такой области близок к дну зоны проводимости. Однако при термодинамическом равновесии уровень Ферми должен быть одинаковым во всей системе. Поэтому в области флуктуации возникает локальное повышение дна зоны проводимости и вершины валентной зоны (рис. 12.4) на величину, близкую к  $\mu$ .

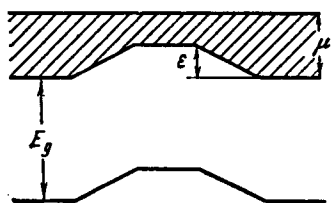


Рис. 12.4. Изгиб зон в области, обедненной донорами. Заполненные состояния зоны проводимости заштрихованы.

Этот перепад представляет собой просто контактную разность потенциалов на границе чистого и легированного полупроводников. Очевидно, что никакая флуктуация концентрации доноров не может поднять дно валентной зоны на величину, большую чем  $\mu$ .

Для того чтобы найти вид плотности состояний вблизи точки окончания, необходимо оценить минимальный размер области, в которой должны отсутствовать доноры. Изменение потенциала на контакте сильно и слабо легированных полупроводников происходит плавно на длине, которую обычно называют шириной области объемного заряда. В случае, когда концентрация примесей изменяется на контакте резко, эта длина оказывается порядка радиуса экранирования в слабо легированной

области. Для того чтобы контактная разность потенциалов достигла своего макроскопического значения, необходимо, чтобы размер флуктуации превышал эту длину, которая и является минимальным размером флуктуации.

Допустим, что вследствие флуктуации образовалась область с концентрацией доноров  $\bar{N}$  значительно меньшей, чем средняя концентрация  $N$ . Контактная разность потенциалов, создающая повышение дна зоны проводимости и верха валентной зоны на величину  $\epsilon$ , определяется соотношением  $\mu - \epsilon \approx \hbar^2 \bar{N}^{2/3} / m_e$  (см. рис. 12.4). Такое повышение действительно имеет место, если размер области, в которой произошла флуктуация, больше радиуса экранирования  $r_s$  в этой области. Согласно § 35  $r_s = r_0(N/\bar{N})^{1/6}$ . Вероятность такой флуктуации можно оценить с помощью (39.1), положив  $\bar{k} = Nr_s^3$ ,  $k \ll \bar{k}$ . Она равна  $\exp(-Nr_s^3)$ . Выражая  $r_s$  через  $(\mu - \epsilon)$ , получим для плотности состояний дырок

$$\ln \frac{g(\epsilon)}{g(0)} \approx -Nr_0^3 \left( \frac{\mu}{\mu - \epsilon} \right)^{3/4}. \quad (41.1)$$

Энергия  $\epsilon$  отсчитана вверх от вершины валентной зоны чистого кристалла. Заметим, что при выводе (41.1) предполагалась идеальность электронного газа во флуктуации. Для этого необходимо условие  $\bar{N}a^3 \gg 1$  или  $\mu - \epsilon \gg E_0$  (во все эти формулы входит эффективная масса электрона). Так как  $E_0 \ll \mu$ , формула (41.1) описывает спад  $g(\epsilon)$  в довольно широком интервале.

Если допустить, что в полупроводнике имеются компенсирующие акцепторы, то плотность состояний будет отличной от нуля сколь угодно глубоко. При достаточно большом значении  $\epsilon$  оказывается применимой формула (39.18), в которую вместо  $N$ ,  $a$  и  $E_0$  следует подставлять

$$N_A, a_h = \frac{\hbar^2 \kappa}{m_h e^2}, E_h = \frac{\hbar^2}{2m_h a_h^2} \quad (\text{подробнее см. [7]}).$$

**б) Квантовый случай ( $\gamma \ll E_s$ ).** В типичных полупроводниках масса тяжелых дырок значительно больше массы электрона. Поэтому условие квантового случая  $\gamma \ll E_s$  легко реализуется, когда в полупроводнике  $p$ -типа изучается спектр электронов. Будем по-прежнему считать, что для основных носителей выполняется условие сильного легирования  $N_A a_h^3 \gg 1$ . Так как это

эквивалентно условию  $N_A r_0^3 \gg 1$ , то из неравенства  $\gamma \ll E_s$  следует, что  $r_0 \ll a$ , где  $a$  — боровский радиус электрона. (Напомним, что экранирование определяется дырками, а  $E_s = \hbar^2/m_e r_0^2$ ). Это означает, что отдельная донорная примесь не имеет связанного состояния.

В области гауссовых флуктуаций ( $|Z| \ll N_A R^3$ ) можно пользоваться формулами (39.6), (39.7), изменив в них знак, т. е. считая, что отрицательным значениям  $Z$  (областям с пониженным числом акцепторов) соответствуют положительные  $V$  и  $\epsilon$ , которые теперь отсчитываются вниз от дна зоны проводимости.

Как уже говорилось, график функции  $\ln W(R, \epsilon)$  совпадает с кривыми рис. 12.1 при  $R \gg r_q = \hbar/\sqrt{m_e \epsilon}$ . Пусть сначала  $\epsilon < E_s$ , т. е.  $r_q > r_0$ . Тогда при  $R > r_q$  функция  $\ln W(R, \epsilon)$  убывает с ростом  $R$  (рис. 12.5).

Флуктуации, которые имеют радиус  $R$ , удовлетворяющий условию  $r_0 \ll R \ll \hbar/\sqrt{m_e \epsilon}$ , могут создавать уровень  $\epsilon$ , только если глубина ямы  $V \approx (\hbar^2/m_e R^2) \gg \epsilon$

(узкая яма с мелким уровнем). Подставляя в это условие (39.7) и выражая  $Z$  через  $R$ , получим

$$|Z| = \frac{aR}{r_0^2}, \quad (41.2)$$

$$\ln W(R, \epsilon) =$$

$$= -\frac{Z^2}{N_A R^3} = -\frac{a^2}{N_A R r_0^4}. \quad (41.3)$$

Отсюда видно, что при  $R \ll \hbar/\sqrt{m_e \epsilon}$  функция  $\ln W(R, \epsilon)$  убывает с уменьшением  $R$  (см. рис. 12.5). Легко показать, что она продолжает убывать и при  $R < r_0$ . Таким образом, оптимальная флуктуация имеет радиус  $R = \hbar/\sqrt{m_e \epsilon}$ . Подставляя это выражение в (41.3), получим

$$\ln \frac{g(\epsilon)}{g(0)} = -\left(\frac{\epsilon}{E_0}\right)^{1/2} \frac{E_s}{\bar{V}} C, \quad \bar{V} = \frac{e^2}{\kappa r_0} N_A r_0^3. \quad (41.4)$$

Формула (41.4) хорошо известна в теории неупорядоченных систем. Она соответствует гауссову потенциалу с

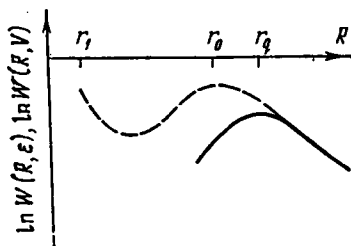


Рис. 12.5. Сплошная кривая —  $\ln W(R, \epsilon)$ , штриховая кривая —  $\ln W(R, V)$ .



нулевым радиусом корреляции (белый шум). На выводе этой формулы оттачивались самые разнообразные методы [1—7]. Разумеется, найден и численный коэффициент  $C$  в этой формуле; согласно [1],  $C = 3/4\pi 2^{3/2}$ . В СЛП закону (41.4) должна подчиняться плотность состояний неосновных носителей в квантовом случае.

Таким образом, в отличие от классического случая, где экспоненциальный спад  $g(\varepsilon)$  начинается при  $\varepsilon = \gamma$  и описывается (39.12), в квантовом случае спад начинается при  $\varepsilon = E_0(\nabla/E_s)^2$  и описывается (41.4).

Дальнейший ход  $g(\varepsilon)$  зависит от соотношения  $E_s$  и  $\nabla$ . Если  $E_s \ll \nabla$ , то при  $\varepsilon = E_s$  длины  $r_q$  и  $r_0$  сравнимы, (41.4) переходит в (39.12) и дальнейший ход  $g(\varepsilon)$  такой же, как в классическом случае. Если же  $E_s \gg \nabla$ , то гауссова статистика нарушается при  $r_q > r_0$  и (39.12) вообще не имеет места. Если притягивающие центры отсутствуют, спектр электронов обрывается при  $\varepsilon \approx \nabla$ . При наличии компенсации плотность состояний отлична от нуля и при больших энергиях и определяется флуктуациями доноров. Асимптотика по энергии по-прежнему описывается формулой (39.18). Переход от (41.4) к (39.18) происходит довольно сложным образом. Он подробно изучен в работе [7].

## § 42. Теория межзонного поглощения света

а) О связи между коэффициентом межзонного поглощения света (КМПС) и плотностью состояний. Важным способом исследования флуктуационных уровней являются эксперименты по межзонному поглощению света вблизи порога фундаментального поглощения. Типичные экспериментальные данные, полученные на  $n$ -GaAs, представлены на рис. 12.6.

Видно, что КМПС убывает на 4 порядка при уменьшении энергии кванта на величину, ничтожную по сравнению с шириной запрещенной зоны ( $E_g = 1,51$  эв при  $T = 77^\circ\text{K}$ ). С увеличением концентрации электронов кривые сдвигаются в коротковолновую область и становятся более пологими. Сдвиг кривых связан с заполнением зоны проводимости (эффект Мосса — Бурштейна). В СЛП  $n$ -типа при низких температурах переход происходит на уровень Ферми, так как все лежащие ниже

состояния заполнены (рис. 12.7). Пороговая энергия оказывается при этом больше ширины запрещенной зоны  $E_g$  на величину энергии Ферми  $\mu$ , если учитывать переходы с вершины валентной зоны, происходящие с изменением

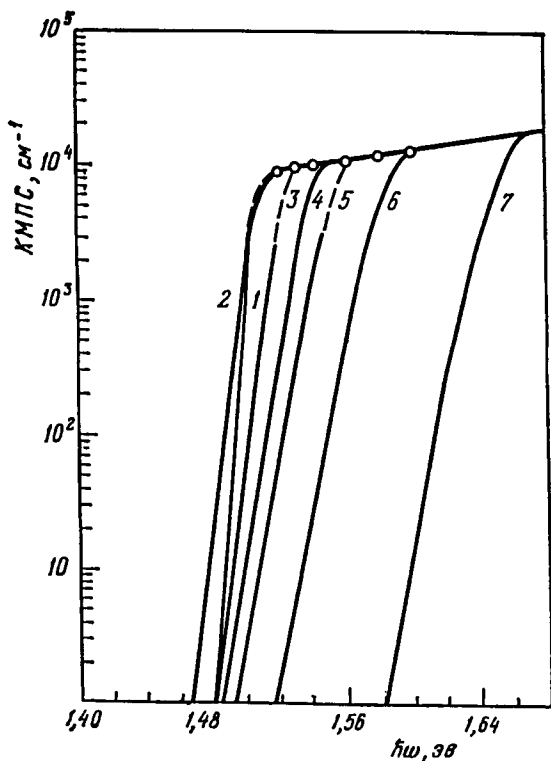


Рис. 12.6. КМПС как функция энергии кванта [12] при  $T = 77^\circ\text{K}$  в образцах  $n\text{-GaAs}$  с концентрациями электронов (в  $\text{см}^{-3}$ ): 1 —  $2 \cdot 10^{16}$ ; 2 —  $2,2 \cdot 10^{17}$ ; 3 —  $5,3 \cdot 10^{17}$ ; 4 —  $1,2 \cdot 10^{18}$ ; 5 —  $1,62 \cdot 10^{18}$ ; 6 —  $3,15 \cdot 10^{18}$ ; 7 —  $6,5 \cdot 10^{18}$ .

импульса (непрямые переходы), или на величину  $\mu(1 + m_e/m_h)$ , если учитывать только прямые переходы (см. рис. 12.7). Непрямые переходы оказываются возможными из-за наличия примесей, причем надо думать, что в СЛП их вероятность слабо отличается от вероятности прямых переходов. Поэтому в дальнейшем мы будем понимать под пороговой энергией минимальную энергию непрямого перехода.

При конечной температуре КМПС может быть отличным от нуля на частотах ниже пороговой за счет дырок в ферми-заполнении. Так как вероятность существования дырки с энергией, меньшей  $\mu$  на величину  $\epsilon$ , пропорциональна  $\exp(-\epsilon/kT)$ , КМПС убывает с частотой  $\omega$ , как  $\exp\left(-\frac{\mu + E_g - \hbar\omega}{kT}\right)^*$ .

При понижении температуры этот механизм поглощения оказывается маловероятным\*\*). Единственным механизмом поглощения, существующим при нулевой температуре, являются переходы с флуктуационных уровней, лежащих выше верха валентной зоны. (Для определенности мы будем по-прежнему говорить о полупроводнике  $n$ -типа.)

При поглощении кванта света с энергией  $\hbar\omega < E_g + \mu$  должна образоваться дырка, находящаяся выше верха невозмущенной валентной зоны на величину, не меньшую чем  $\epsilon_h = E_g + \mu - \hbar\omega$ . Схема таких переходов при наличии флуктуационного потенциала, изгибающего границы обеих зон, изображена на рис. 12.8. Число флуктуационных уровней дырок экспоненциально убывает с ростом  $\epsilon_h$ . Вследствие этого КМПС экспоненциально убывает с уменьшением частоты.

Однако КМПС пропорционален не только вероятности существования глубокого дырочного уровня, но и вероятности межзонного перехода. Последняя также может быть экспоненциально малой, поскольку дырочный

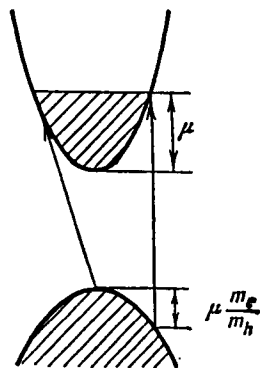


Рис. 12.7. Переходы с пороговой энергией кванта. Вертикальная стрелка — прямой переход, наклонная стрелка — непрямой переход.

\*) При учете только прямых переходов КМПС пропорционален

$$\exp \left\{ \left[ -\mu - \frac{m_h}{m_e + m_h} (E_g - \hbar\omega) \right] (kT)^{-1} \right\}.$$

\*\*) Аналогичным образом, при низких температурах вымораживается механизм поглощения, представляющий переходы с одновременным поглощением фононов.

уровень создается скоплением отрицательно заряженных акцепторов и волновая функция электрона в этой области может быть мала. Как было показано в § 39, в области гауссовых флуктуаций потенциал слабо искажает электронные состояния вблизи уровня Ферми (мы пользовались этим обстоятельством, когда считали экранирование линейным). В этом случае вероятность перехода на уровень Ферми не мала и показатель экспоненты КМПС получается из показателя экспоненты плотности состояний дырок путем замены  $\epsilon$  на  $\epsilon_h = E_g + \mu - \hbar\omega$ .

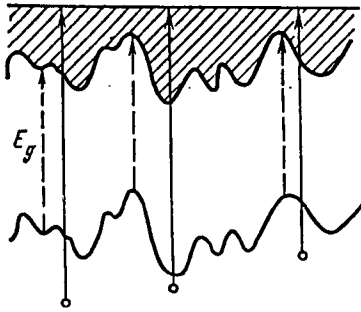


Рис. 12.8. Схема межзонных переходов в вырожденном полупроводнике при  $T = 0^\circ\text{K}$ . Извилистые линии — границы зон, искривленные потенциалом флуктуаций; занятые состояния зоны проводимости заштрихованы.

энергию  $\epsilon$ . При  $\epsilon > \mu$  обеднение волновой функции электрона с энергией  $\mu$  может быть значительным. Однако это обеднение происходит в сравнительно малой области пространства, и можно показать [8], что оно не влияет на частотную зависимость КМПС.

Итак, в сильно легированном и слабо компенсированном полупроводнике с вырожденными носителями тока частотная зависимость КМПС воспроизводит плотность состояний неосновных носителей (например, дырок в рассмотренном выше случае полупроводника  $n$ -типа).

б) Поглощение света за счет гауссовых флуктуаций в отсутствие вырождения носителей. Рассмотрим сильно легированный полупроводник, у которого уровень Ферми лежит глубоко в запрещенной зоне. Это может быть следствием высокой температуры или сильной компенсации. Сюда же относится ситуация (существующая, например, в германии), когда при поглощении кванта происходит переход электрона в незаполненную зону с экстремумом в точке  $\Gamma$ , а электроны образуют ферми-

электронные состояния вблизи уровня Ферми (мы пользовались этим обстоятельством, когда считали экранирование линейным). В этом случае вероятность перехода на уровень Ферми не мала и показатель экспоненты КМПС получается из показателя экспоненты плотности состояний дырок путем замены  $\epsilon$  на  $\epsilon_h = E_g + \mu - \hbar\omega$ .

В области негауссовых флуктуаций основную роль в плотности состояний дырок играют скопления акцепторов, образующие ядра многозарядных ионов, основное состояние которых имеет

заполнение в боковом экстремуме и уровень Ферми проходит гораздо ниже дна незаполненной зоны, в которую переходит электрон.

Пусть, далее, в образце имеются заряженные доноры и акцепторы с суммарной концентрацией  $N_t = N_D + N_A$ . Их потенциал будем считать кулоновским вплоть до некоторого расстояния  $r_0$ , а затем быстро спадающим. Этот спад может быть вызван как электронным экранированием, так и корреляцией в расположении примесей. Для нас будет существенно только то, что в гауссовых флуктуациях, имеющих размер  $R \ll r_0$ , корреляцию примесей можно не учитывать и что длина  $r_0$  велика, так что флуктуационный потенциал можно считать, по терминологии § 41, классическим и для электронов и для дырок, т. е.

$$\frac{\hbar^2}{m_e r_0^2} \ll \gamma, \quad \frac{\hbar^2}{m_h r_0^2} \ll \gamma, \quad \gamma = \frac{e^2}{\kappa r_0} (N_t r_0^3)^{1/2}. \quad (42.1)$$

Тогда плотность состояний электронов и дырок спадает в глубь запрещенной зоны по закону (39.12), а оптимальные флуктуации представляют собой ямы размера  $r_0$ , имеющие много уровней. Легко понять, что в этом случае частотная зависимость коэффициента поглощения не повторяет плотности состояний. Это хорошо видно из рис. 12.9.

Допустим, что мы изучаем переход, в результате которого образовались электрон и дырка с энергиями  $\epsilon_e$  и  $\epsilon_h$  (мы отсчитываем их вверх от дна зоны проводимости и вершины валентной зоны). При этом поглощается квант с дефицитом  $\Delta = \epsilon_h - \epsilon_e$ . (Дефицитом кванта  $\Delta$  мы называем энергию, на которую этот квант меньше ширины запрещенной зоны, т. е.  $\Delta = E_g - \hbar\omega$ .)

Как видно из рис. 12.9, если  $\Delta > 0$ , то области, в которых находятся электрон и дырка, пространственно

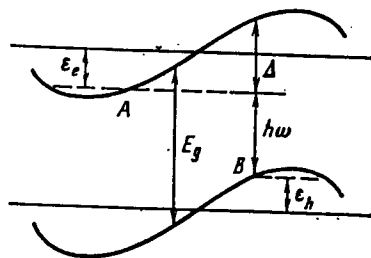


Рис. 12.9. Схема поглощения кванта с дефицитом  $\Delta$  в невырожденном полупроводнике. Сплошные горизонтальные прямые — границы невозмущенных зон; изогнутые линии — границы зон, искривленные потенциалом примесей.

разделены. Действительно, область правее точки  $A$  является классически недоступной для электрона, а область левее точки  $B$  классически недоступна для дырки. Коэффициент поглощения содержит интеграл перекрытия волновых функций электрона и дырки. Устремим массы электрона и дырки к бесконечности (при фиксированном  $r_0$ ). При этом вероятность туннелирования обратится в нуль и поглощение кванта с  $\Delta > 0$  окажется невозможным. В то же время плотность состояний, определяемая только вероятностью создания потенциальной ямы, выражается классической формулой (39.12), которая вообще не содержит массы. Таким образом, поглощение света в этом случае требует специального рассмотрения. По существу, это есть эффект Келдыша — Франца, но не во внешнем электрическом поле, а в случайном поле заряженных примесей.

Мы получим решение этой задачи методом оптимальной флуктуации [13], т. е. найдем конфигурацию примесей, дающую максимальный вклад в поглощение света с заданным дефицитом. В области больших дефицитов, где коэффициент поглощения  $K(\Delta)$  экспоненциально мал, максимум, соответствующий оптимальной флуктуации, оказывается очень резким. Поэтому главный член  $\ln K(\Delta)$  определяется вкладом оптимальной флуктуации. Разумеется, такой метод не позволяет определить предэкспоненциальный член в  $K(\Delta)$ .

Мы приведем лишь упрощенный вывод, аналогичный приближению равномерно заряженного шара (§ 39), позволяющий определить показатель экспоненты в  $K(\Delta)$  с точностью до численного множителя. Этот множитель найден с помощью теории, построенной в работах [13, 14].

Допустим, что, вследствие флуктуации концентрации примесей в объеме с линейным размером  $R$ , возникло однородное электрическое поле  $\mathcal{E}$ . Будем считать, что  $R$  и  $\mathcal{E}$  связаны соотношением  $e\mathcal{E}R = \Delta$ . Это необходимо для того, чтобы в поле такой флуктуации мог поглотиться квант с дефицитом  $\Delta$ . (Поглощение кванта произойдет и при  $R > \Delta/e\mathcal{E}$ , однако, разумеется, такая флуктуация менее вероятна.) Избыточное число примесей  $Z$ , необходимое для создания требуемой флуктуации,

определяется из условия  $\mathcal{E} = Zv/\kappa R^2$ , т. е.

$$Z = \frac{\mathcal{E}\kappa R^2}{v} = \frac{\kappa R\Delta}{v^2}.$$

Мы не пишем здесь численных множителей, зависящих от формы флуктуации. Вклад в коэффициент поглощения от такой флуктуации пропорционален величине

$$\exp\left(-\frac{Z^2}{N_t R^3}\right) \exp\left(-\frac{R\sqrt{m_e\Delta}}{\hbar}\right). \quad (42.2)$$

Первый множитель представляет собой вероятность появления  $Z$  избыточных примесей в объеме  $R^3$ . Вторым множителем есть вероятность туннелирования электрона на расстояние  $R$  (мы предполагаем для простоты, что масса дырки больше, чем масса электрона, и что дырка не туннелирует). Выразив  $Z$  через  $R$  и  $\Delta$ , перепишем (42.2) в виде

$$\exp\left(-\frac{\kappa^2\Delta^2}{e^4 N_t R} - \frac{R\sqrt{m_e\Delta}}{\hbar}\right), \quad (42.3)$$

откуда ясно, что при заданном  $\Delta$  вероятность туннелирования растет с уменьшением  $R$ , в то время как вероятность образования флуктуации падает.

Найдем размер оптимальной флуктуации  $\tilde{R}$  из условия максимума показателя экспоненты в (42.3):

$$\tilde{R} = a \left(\frac{\Delta}{E_0}\right)^{3/4} \frac{1}{(N_t a^3)^{1/2}}. \quad (42.4)$$

Подставляя (42.4) в (42.3), получим вклад в  $K(\Delta)$  от оптимальной флуктуации. В соответствии со сказанным выше, окончательно имеем

$$\ln \frac{K(\Delta)}{K(0)} = -\beta \left(\frac{\Delta}{E_0}\right)^{5/4} \frac{1}{(N_t a^3)^{1/2}}. \quad (42.5)$$

Здесь  $\beta$  — численный множитель, который не может быть получен из столь упрощенного вывода. В работе [13] показано, что если  $m_h \gg m_e$ , то  $\beta = 2/5\sqrt{\pi}$ . В дальнейшем Меркулов и Перель [14] показали, что при сравнимых массах электрона и дырки в (42.5) следует подставлять приведенную массу  $m_h m_e / (m_e + m_h)$ , а коэффициент  $\beta$  не меняется. Из (42.5) видно, что закон спада

коэффициента поглощения существенно отличается от закона спада плотности состояний (39.12).

Исследуем теперь область применимости (42.5). В проделанном выше выводе было очень существенно, что потенциал примесей является кулоновским, по крайней мере, на расстоянии порядка  $\bar{R}$ . Поэтому результат справедлив, только если  $\bar{R} < r_0$ , или, с учетом (42.4),

$$\Delta < \Delta_1, \quad \Delta_1 = E_0 \left( \frac{r_0}{a} \right)^{4/3} (N_1 a^3)^{2/3}. \quad (42.6)$$

С другой стороны, для применимости (42.5) необходимо выполнение условия  $|\ln [K(\Delta)/K(0)]| \gg 1$ , или

$$\Delta \gg \Delta_2, \quad \Delta_2 = E_0 (N_1 a^3)^{2/5}. \quad (42.7)$$

Из первого неравенства (42.1) следует, что  $\Delta_1 \gg \Delta_2$ , т. е. область применимости (42.5) существует, и тем шире, чем больше  $r_0$ .

Рассмотренная нами флуктуация, как видно из рис. 12.9, представляет потенциальную яму для электрона и расположенную рядом с ней потенциальную яму для дырки. В нашем выводе предполагалось, что энергия электрона опущена, а энергия дырки поднята на величину  $\Delta$ , т. е. потенциальные ямы имеют уровень с энергией порядка их глубины. Для этого достаточно выполнения неравенства  $(\hbar^2/m_e \bar{R}^2) \ll \Delta$ , которое вытекает из (42.7) и не накладывает дополнительных ограничений.

Выясним теперь, что произойдет при  $\Delta > \Delta_1$ . Если  $\bar{R} > r_0$ , то при  $R = r_0$  второй член в (42.3) мал по сравнению с первым, т. е. туннелирование несущественно. Как мы видели в § 39, вероятность создания потенциальной ямы глубиной  $\Delta$  и размером  $R$  при  $R > r_0$  убывает с ростом  $R$  (потенциал далеких примесей не достигает центра). Поэтому, как и в теории плотности состояний, размер оптимального скопления в этом случае есть  $r_0$ . Подставляя в (42.3)  $R = r_0$ , получим при  $\Delta > \Delta_1$

$$\ln \frac{K(\Delta)}{K(0)} = - \frac{\Delta^2}{\gamma^2}, \quad (42.8)$$

что воспроизводит плотность состояний. При некотором соотношении между параметрами возможно, что еще при  $\Delta < \Delta_1$  негауссовы флуктуации типа атома будут давать больший вклад в КМПС, чем гауссовы флуктуа-



ции, рассмотренные выше. В этом случае (42.8) вообще не имеет области применимости.

В заключение отметим, что с помощью изложенного выше метода оптимальной флуктуации Меркулов и Перель [14] построили теорию электропоглощения в невырожденных СЛП.

в) Обсуждение экспериментальных результатов. Обычно в полупроводниках экспериментально наблюдаемый КМПС на частотах ниже пороговой хорошо описывается формулой

$$\ln \frac{K(\omega)}{K(\omega_n)} = - \frac{\hbar(\omega_n - \omega)}{\Delta_0}, \quad (42.9)$$

где  $\hbar\omega_n$  — пороговая энергия, а  $\Delta_0$  — характерная энергия, не зависящая от частоты. Первая трудность, с которой мы сталкиваемся, состоит в том, что ни в одном из рассмотренных выше случаев поглощения за счет флуктуационных уровней не возникала линейная зависимость  $\ln K(\omega) \propto (\omega - \omega_n)$ . Более того, мы не знаем способа, не вызывающего серьезных возражений, который бы позволил получить такую зависимость при низких температурах. Выражаясь точнее, такая зависимость может быть получена лишь как интерполяционная, т. е. связывающая две области с разными законами для КМПС, однако мы не знаем такой непротиворечивой комбинации параметров, при которой линейная зависимость являлась бы самостоятельным физическим законом, справедливым в широкой области частот. Нам представляется, однако, что для объяснения эксперимента этого и не требуется, так как измерения всегда проводятся в довольно узком интервале частот, в котором зависимости  $\Delta^{5/4}$ ,  $\sqrt{\Delta} \ln \Delta$  трудно отличить от линейной.

К сожалению, до сих пор хорошего количественного объяснения хвостов КМПС не существует. Нам представляется, что в вырожденных образцах  $n$ -GaAs, изученных наиболее полно (см. [15]), эти хвосты связаны с флуктуациями концентрации акцепторов и с дырочными уровнями, возникающими над вершиной валентной зоны (см. [7, 16]).

Однако этот вывод нельзя считать окончательным. Предположение о том, что за хвост КМПС ответственны неосновные примеси, трудно проверить количественно,

так как сложно определить степень компенсации в тех случаях, когда она мала.

На важную роль неосновных примесей в формировании хвоста плотности состояний указывают также эксперименты по излучательной рекомбинации в эпитаксиальном GaAs, выполненные Алферовым и др. [17].

Иной подход к описанию эксперимента предлагается в работах Доу и Редфилда [18]. Идея этих работ состоит в том, что хвост КМПС нужно понимать как следствие эффекта Келдыша — Франца в случайном электрическом поле. Однако результаты отличаются от полученных в разделе б) по следующим причинам. 1) Авторы учитывали взаимодействие электрона с дыркой, возникающей при межзонном переходе. Численные расчеты, сделанные ими, показали, что КМПС в однородном электрическом поле  $\mathcal{E}$  зависит от дефицита кванта по закону

$$K(\Delta) \propto \exp(-\text{const } \Delta/\mathcal{E}), \quad (42.10)$$

что существенно отличается от асимптотики Келдыша — Франца

$$K_g(\Delta) \propto \exp\left[-\frac{4\sqrt{2m}\Delta^{3/2}}{3\hbar e\mathcal{E}}\right], \quad (42.11)$$

в которой не учитывается взаимодействие электрона с дыркой. 2) Функцию распределения полей они считали гауссовой. Достоинство этих работ состоит в том, что благодаря формуле (42.10) удается объяснить линейный ход (42.9), однако это же обстоятельство и вызывает основное возражение.

Меркулов и Перель [19] получили аналитическое решение задачи о КМПС в однородном поле с учетом электронно-дырочного взаимодействия. Они показали, что линейная зависимость (42.10) имеет место в сравнительно узком интервале значений  $\Delta$  вблизи  $\Delta \approx 2E_{ex}$ , где  $E_{ex}$  — энергия связи экситона. При больших  $\Delta$  имеет место (42.11), что следует также из простых физических соображений. В связи с этим нам кажется, что предлагаемое Доу и Редфилдом объяснение не может быть универсальным.

## ГЛАВА 13

# ТЕОРИЯ СИЛЬНО ЛЕГИРОВАННЫХ И СИЛЬНО КОМПЕНСИРОВАННЫХ ПОЛУПРОВОДНИКОВ (СЛКП)

### § 43. Некоррелированное распределение примесей

а) Качественное обсуждение. Для определенности будем по-прежнему говорить о полупроводнике  $n$ -типа. В этом случае каждый акцептор забирает электрон от донора и заряжается отрицательно. Оставшиеся в зоне проводимости электроны имеют концентрацию  $n = N_D - N_A$  и находятся в случайном потенциале заряженных доноров и акцепторов.

Как указывалось в § 36, квазиклассическая теория, основанная на линейном экранировании, применима к СЛКП при условии, что выполнено неравенство (36.17). Оно означает, что характерная величина случайного потенциала мала по сравнению с энергией Ферми. В двумерном случае потенциальный рельеф можно представить в виде холмистого дна электронного океана с очень редкими горами, вершины которых поднимаются выше уровня океана (см. рис. 12.8).

Допустим теперь, что мы увеличиваем концентрацию акцепторов  $N_A$ , приближая ее к  $N_D$ . Энергия Ферми  $\mu_0$  уменьшается, как  $n^{2/3}$ , а потенциальный рельеф растет, так как электронное экранирование ослабляется ( $r_0$  в (36.7) возрастает). Легко понять, что двумерным аналогом картины, возникающей при больших степенях компенсации, является пересыхающий водоем, в самых глубоких местах которого остались изолированные друг от друга лужи. Как мы покажем, в трехмерном случае элект-

троны собираются в металлические капли с плотностью, которая остается конечной при  $n \rightarrow 0$ . Подобно лужам в пересыхающем водоеме, эти капли изолированы друг от друга. Поэтому статическая проводимость всего образца

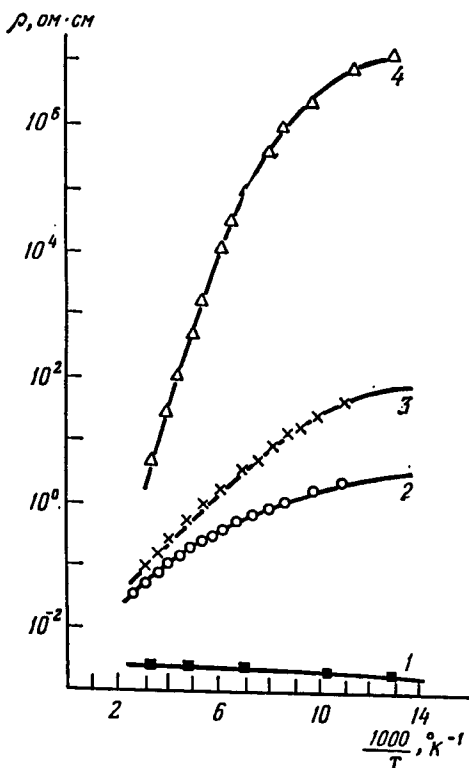


Рис. 13.1. Удельное сопротивление  $n$ -Ge с концентрацией As  $N_D = 8 \cdot 10^{18} \text{ см}^{-3}$  как функция обратной температуры: 1 — исходный образец, 2—4 — компенсированные образцы [1].

в широком интервале температур оказывается активационной. На рис. 13.1 приведены экспериментальные кривые [1], из которых видно, что компенсация действительно приводит к срыву металлической проводимости.

Теория, изложенная в этой главе, позволяет найти зависимость критической степени компенсации, при которой происходит срыв, от концентрации основных примесей и зависимость энергии активации неметаллической фазы от степени компенсации и концентрации основных

примесей [2, 3]. В этом параграфе будем считать, что корреляция в расположении примесей отсутствует.

Ниже мы увидим, что переход от металлической проводимости к активационной является квазиклассическим вариантом перехода Андерсона (см. гл. 2). Критическая концентрация электронов, соответствующая переходу, имеет вид

$$n_c = \frac{\beta N^{2/3}}{a}, \quad (43.1)$$

причем численный коэффициент  $\beta$  в настоящее время неизвестен. Порядковую оценку величины  $n_c$  можно получить, превратив неравенство (36.17) в равенство.

Значительно более сложным является вопрос о структуре потенциального рельефа и электронных состояний при надкритической компенсации, когда  $\alpha \gg 1$  (см. (36.17)). С этим вопросом мы уже сталкивались в § 13, где речь шла о сильно компенсированном, но слабо легированном полупроводнике. Основная трудность состояла в том, что крупномасштабные флуктуации концентрации примесей создают очень большой потенциал. Так, типичная флуктуация в объеме с линейными размерами порядка  $R$  создает понижение (или повышение) энергии электрона на величину

$$\gamma(R) = \frac{e^2 (N_t R^3)^{1/2}}{\kappa R} \quad (N_t = N_A + N_D). \quad (43.2)$$

Эта величина обращается в бесконечность при  $R \rightarrow \infty$ , вследствие чего возникает необходимость учитывать электронное экранирование, даже если концентрация электронов очень мала. Эти рассуждения, конечно, полностью переносятся на случай СЛКП. Уже при  $n \approx n_c$  потенциальная энергия электронов оказывается сравнимой с энергией Ферми. Поэтому ясно, что при  $n \ll n_c$  горбы потенциального рельефа окажутся столь большими, что электронная плотность будет сильно неоднородной (см. рис. 13.2). Следовательно, линейная теория экранирования, изложенная в § 35, оказывается совершенно неприменимой. Однако основные идеи нелинейной теории экранирования, построенной в § 13, являются плодотворными и в случае СЛКП. Простые рассуждения, приведенные в § 13, показывают, что флуктуации, имею-

щие размер, больший чем  $r_s$ , экранируются электронами и что типичное значение флуктуации потенциальной энергии электрона равно  $\gamma(r_s)$ , где, согласно (13.8), (13.9),

$$r_s = \frac{N_t^{1/3}}{n^{2/3}}, \quad \gamma(r_s) = \frac{e^2 N_t^{2/3}}{\kappa n^{1/3}}. \quad (43.3)$$

В рассуждениях, приводящих к (43.3), фигурировала электронная плотность, усредненная по большим масштабам. Поэтому результат не зависит от вида локальной плотности состояний, т. е. от того, как распределены электронные состояния в более мелких масштабах. Таким образом, выражения (43.3) окажутся оправданными, когда мы покажем, что в объеме  $r_s^3$  содержится много электронных капель.

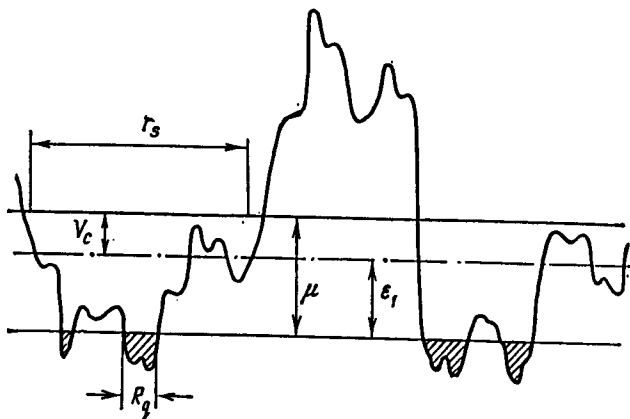


Рис. 13.2. Энергетическая схема компенсированного полупроводника; извилистая линия изображает искривление дна зоны проводимости, верхняя сплошная линия — энергия дна зоны проводимости в отсутствие примесного потенциала, нижняя сплошная — уровень Ферми, штрих-пунктирная — уровень протекания. Заштрихованы области, занятые электронами (электронные капли).

Перейдем теперь к изучению мелкомасштабной структуры электронных состояний. Она оказывается иной, чем в случае слабого легирования. Внутри ямы размера  $r_s$  существуют ямы и горбы мелкомасштабных флуктуаций. Разумеется, электроны будут проваливаться в эти ямы (рис. 13.2). Поэтому электронная плотность должна

содержать гармоники всех размеров, меньших  $r_s$ . Допустим, что примесный потенциал не содержит флуктуаций масштаба, меньшего чем  $L$ , причем  $L \ll r_s$ . Тогда электроны образуют капли размера  $L$ , находящиеся в самых глубоких потенциальных ямах мелкого масштаба, которые в свою очередь находятся в самых глубоких ямах более крупного масштаба (см. рис. 13.2).

Характерная глубина потенциальной ямы масштаба  $L$  определяется ее примесным зарядом. Избыточное число доноров внутри такой ямы порядка  $(N_t L^3)^{1/2}$ , и, следовательно,  $\gamma(L) = e^2 \chi^{-1} L^{-1} (N_t L^3)^{1/2}$ . Число электронов в капле не может превышать избыточного числа доноров, так как иначе яма превратится в потенциальный горб. Таким образом, число электронов в капле равно  $(N_t L^3)^{1/2}$ , а концентрация электронов в капле  $\bar{n}(L) = (N_t L^3)^{1/2} / L^3$ . С уменьшением  $L$  электронные капли дробятся, а концентрация электронов в них возрастает.

Теперь следует вспомнить, что число электронов в яме лимитируется не только ее примесным зарядом, но и принципом Паули. В яме глубиной  $\gamma(L)$  число состояний  $\mathcal{N}_q(L) = L^3 (m\gamma(L))^{3/2} / \hbar^3$ , при малых  $L$  имеем  $\mathcal{N}_q(L) < (N_t L^3)^{1/2}$  и квантовое ограничение становится важнее зарядового. Введем длину  $R_q$  (*quantum*), определив ее равенством  $\mathcal{N}_q(R_q) = (N_t R_q^3)^{1/2}$ , т. е.

$$R_q = \frac{a}{(N_t a^3)^{1/9}}. \quad (43.4)$$

Для типичных ям, имеющих размер, больший  $R_q$ , существенно зарядовое ограничение, а для меньших ям — квантовое ограничение. Для того чтобы все сказанное выше имело смысл, необходимо условие  $R_q \ll r_s$ , которое выполняется, если  $\alpha \gg 1$  (см. (36.17)).

Длину  $R_q$  можно иначе определить как такое значение  $L$ , при котором энергия Ферми электронов с концентрацией  $\bar{n}(L)$  сравнивается с глубиной ямы  $\gamma(L)$ . Легко понять, что флуктуации примесного потенциала, имеющие масштаб  $R \ll R_q$ , не приводят к дроблению капель. Действительно, создаваемые этими флуктуациями горбы и ямы имеют масштаб  $\gamma(R)$  меньший, чем энергия Ферми электронов в капле, которая равна  $\gamma(R_q)$ . Поэтому такие горбы не образуют островов в электронном море

(см. рис. 13.2). Иными словами, можно сказать, что дробления капли размера  $R_q$  не происходит, так как в мелких каплях, располагающихся на территории старой капли, не хватает квантовых состояний для размещения всех электронов старой капли.

Итак, мы доказали, что электроны разбиваются на капли, имеющие размер  $R_q$ . Легко убедиться в том, что этот размер гораздо больше среднего расстояния между примесями и что ямы, имеющие размер  $R_q$  и глубину  $\gamma(R_q)$ , содержат много уровней ( $\hbar^2/mR_q^2 \ll \gamma(R_q)$ ). Концентрация электронов в капле

$$\tilde{n} = \frac{(N_t R_q^3)^{1/2}}{R_q^3} = \frac{N_t}{(N_t a^3)^{1/3}} \quad (43.5)$$

не зависит от средней концентрации электронов  $n$ . По отношению к взаимодействию друг с другом электроны в капле образуют слабо неидеальный ферми-газ, так как  $\tilde{n} a^3 \gg 1$ . Кинетическая энергия электронов порядка

$$\gamma(R_q) = E_0 (N_t a^3)^{4/9} \quad (43.6)$$

и мала по сравнению с  $\gamma(r_s)$ .

В предыдущих рассуждениях считалось, что температура равна нулю. На самом деле эти рассуждения справедливы, пока  $kT \ll \gamma(R_q)$ . Это неравенство и является условием фермиевского вырождения электронного газа. При  $kT \gg \gamma(r_s)$  имеем идеальный бoльцмановский газ, и описанная выше структура случайного потенциала полностью разрушается. В интервале  $\gamma(R_q) \ll kT \ll \gamma(r_s)$  крупномасштабная структура потенциала сохраняется, но размер капель зависит от температуры. Термодинамика электронного газа в этой области была изучена Гуляевым и Плесским [4].

**б) Основные уравнения и асимптотические свойства потенциала.** Из приведенного выше качественного анализа структуры потенциала и электронных состояний должно быть ясно, что основой математического описания такой системы должно быть самосогласованное уравнение типа Томаса — Ферми (35.6), где концентрация электронов связана с потенциалом  $\phi(r)$  соотношением (35.5), а избыточная плотность примесного заряда  $e\xi(r)$



равна

$$e\xi(r) = e[N_D(r) - N_D - N_A(r) + N_A]. \quad (43.7)$$

Из качественных соображений следует, что наиболее существенный вклад в потенциал электронов дают гауссовы флуктуации, и потому  $\xi(r)$  следует считать гауссовой случайной функцией. Ее корреляторы разбиваются на произведения парных корреляторов, причем последние имеют вид

$$\langle \xi(r) \xi(r') \rangle = (N_D + N_A) \delta(r - r'), \quad (43.8)$$

так как мы не учитываем здесь корреляции, связанной со взаимодействием примесей.

Введем функцию распределения потенциальной энергии  $F(V)$ , где  $V = -e\phi$ . Тогда уровень Ферми  $\mu$  определяется уравнением

$$n = \frac{(2m)^{3/2}}{3\pi^2 \hbar^3} \int_{-\infty}^{\mu} (\mu - V)^{3/2} F(V) dV. \quad (43.9)$$

В § 36 сформулированная система уравнений была решена в предположении, что  $\alpha \ll 1$ . Это условие позволяло линеаризовать уравнение (35.6). Как ясно из качественного анализа задачи, при сильной компенсации эта линеаризация недопустима. Поэтому можно лишь переписать систему уравнений в безразмерных переменных и сформулировать математическую задачу, не содержащую буквенных параметров. Единицы, в которых выражаются величины, входящие в безразмерные уравнения, и определяют зависимость этих величин от параметров задачи. Кроме того, можно найти асимптотические свойства функции  $F(V)$ , воспользовавшись методом оптимальной флуктуации, сформулированным в гл. 12.

Введем безразмерные переменные:

$$\begin{aligned} \chi &= \frac{2^{2/3} V}{\gamma(r_s)}, & x &= \frac{2^{1/3} r}{r_s}, \\ \eta &= \frac{-2^{2/3} \mu}{\gamma(r_s)}, & f(x) &= \frac{\xi}{n}, \end{aligned} \quad (43.10)$$

где  $r_s$  и  $\gamma(r_s)$  определены формулами (43.3). В этих переменных получим (см. (35.5), (35.6))

$$\Delta_x \chi = 4\pi [f(x) + 1 - u(x)], \quad (43.11)$$

где

$$u(x) = \begin{cases} \frac{3^{1/2} \alpha^2}{\pi^{1/2} 4 \sqrt{2}} (-\eta - \chi)^{3/2}, & \chi < -\eta, \\ 0, & \chi > -\eta. \end{cases} \quad (43.12)$$

Функция  $f(x)$  есть случайная гауссова функция с коррелятором

$$\langle f(x) f(x') \rangle = 2\delta(x - x'). \quad (43.13)$$

Более высокие корреляторы расцепляются на произведения парных.

Из уравнений (43.11), (43.12) следует, что при  $\alpha \gg 1$  в областях, где  $u(x) > 0$ , потенциальная энергия  $\chi(x)$  близка к  $-\eta$ . Из (43.11) следует, что с точностью до членов, содержащих степени  $\alpha^{-1}$ , в этих областях должно выполняться условие

$$u(x) = f(x) + 1, \quad (43.14)$$

т. е. в областях, где электроны присутствуют, они полностью нейтрализуют заряд примесей.

В результате в нулевом приближении по  $\alpha^{-1}$  мы получили любопытную математическую задачу, состоящую в нахождении областей, в которых электроны должны полностью нейтрализовать примесный заряд. Эту задачу можно сформулировать следующим образом.

В некоторой области (или во всем пространстве) дана непрерывная плотность заряда  $\rho(x)$ , создающая потенциал  $\chi_0$  по закону

$$\chi_0(x) = - \int \frac{\rho(x') d^3x'}{|x - x'|}. \quad (43.15)$$

Требуется найти области  $O$ , обладающие следующими свойствами:

- 1)  $\rho(x) > 0$  везде в  $O$ .
- 2) Потенциал  $\chi(x)$ , создаваемый всеми зарядами, кроме зарядов, находящихся в  $O$ , равен постоянной величине  $-\eta$  во всех точках  $O$  (электрическое поле в  $O$  равно нулю).
- 3)  $\chi(x) \geq -\eta$  везде вне  $O$ .

В интересующем нас случае  $\rho(x)$  есть полный примесный заряд ( $\rho = f + 1$ ), а химический потенциал  $\eta$

следует определять из условия нейтральности. Это условие гласит, что суммарный нескомпенсированный заряд, т. е. суммарный заряд, находящийся вне  $O$ , равен нулю.

Сформулированная выше задача является новой задачей теории потенциала. В случае потенциала тяготения, например, она может звучать так: внутри тела произвольной формы и с произвольным распределением плотности вырезать полость такую, что по удалении массы, находящейся внутри полости, оставшиеся снаружи массы создают внутри полости постоянный заданный потенциал (сила тяготения равна нулю во всех точках полости).

Для этой задачи можно доказать теорему единственности решения (см. [16] к гл. 11). Кроме того, можно написать ограниченный снизу функционал такой, что искомая функция  $u(x)$  осуществляет его минимум [2]. На основании этого функционала можно доказать некоторые теоремы, имеющие отношение к вопросу о существовании решения (см. [16] к гл. 11).

Система уравнений (43.11), (43.13), (43.14), определяющая химический потенциал  $\eta$  в нулевом приближении по  $\alpha^{-1}$ , не содержит параметров. Поэтому характерный масштаб электростатического потенциала и глубина, на которую опускается уровень Ферми по отношению к среднему значению потенциала, имеют порядок величины  $\gamma(r_s)$ , как и объяснялось в разделе а). Согласно (43.10)

$$\mu = -\eta \frac{e^2 N_D^{2/3}}{\kappa n^{1/3}}. \quad (43.16)$$

Метод оптимальной флуктуации, изложенный в гл. 12, позволяет найти асимптотический вид функции распределения  $F(V)$  при больших положительных и отрицательных значениях  $V$ . Воспользуемся для этого уравнениями (43.11) — (43.13). Вероятность того, что случайная гауссова функция, определенная (43.13), имеет заданное значение, пропорциональна  $\exp(-\Omega\{f\})$ , где

$$\Omega\{f\} = \frac{1}{4} \int f^2 dx. \quad (43.17)$$

Этот результат можно также получить из (40.16), заменив в этой формуле  $N$  на  $N_t$  и подставив в нее (43.10).

Метод оптимальной флуктуации состоит в том, чтобы искать функцию  $\bar{f}$ , для которой  $\Omega\{f\}$  имеет минимум при дополнительном условии, что функционал  $\chi\{f\}$ , представляющий решение уравнения Пуассона (43.11) в некоторой точке пространства, равен  $\chi$ . Тогда, если  $\Omega\{f\} \gg 1$ ,

$$\ln\left(\frac{F(\chi)}{F(0)}\right) = -\Omega\{\bar{f}\}. \quad (43.18)$$

Поместим начало координат в точку, где  $\chi\{f\} = \chi \gg 1$ . Учтем, что ввиду сферической симметрии задачи экстремальная функция не зависит от углов. В области, где  $\chi > -\eta$ , электроны отсутствуют, и последний член (43.11) можно выкинуть. Электроны появляются лишь на расстоянии  $x_0 \gg 1$  от центра флуктуации, на котором потенциал уменьшается почти до нуля. Они быстро (на расстоянии порядка единицы) нейтрализуют флуктуацию, так что область, где есть электроны, дает малый вклад в потенциал центра флуктуации  $\chi\{f\}$ . Таким образом, в сферических координатах имеем

$$\chi\{f\} = - \int_{|x| < x_0} \frac{(f+1)d^3x}{x} = -4\pi \int_0^{x_0} (f+1)x dx. \quad (43.19)$$

На расстоянии  $x_0$  потенциал должен упасть до нуля. При сферической симметрии это означает, что суммарный заряд области  $|x| < x_0$  должен быть равным нулю, т. е.

$$\int_0^{x_0} (f+1)x^2 dx = 0. \quad (43.20)$$

Экстремальная функция  $\bar{f}$  находится из уравнения

$$\delta(\Omega\{f\} + t\chi\{f\}) = 0, \quad (43.21)$$

где  $t$  — множитель Лагранжа, который следует определить из условия  $\chi\{f\} = \chi$ . Варьируя (43.19), получим

$$\delta\chi\{f\} = -4\pi(1+f(x_0))x_0\delta x_0 - 4\pi \int_0^{x_0} \delta f x dx. \quad (43.22)$$

Вариацию  $\delta x_0$  найдем, варьируя условие (43.20):

$$[f(x_0) + 1] x_0^2 \delta x_0 = - \int_0^{x_0} \delta f x^2 dx. \quad (43.23)$$

Наконец, варьируя  $\Omega\{f\}$  и подставляя полученные формулы в (43.21), найдем

$$\bar{f} = \frac{2t}{x} \left[ 1 - \frac{x}{x_0} \right]. \quad (43.24)$$

Из условия (43.20) найдем, что  $t = -x_0$ , а из условия  $\chi\{f\} = \chi$  получим  $x_0 = \sqrt{\chi/2\pi}$ . Подставляя (43.24) в (43.18), имеем при  $V \gg \gamma(r_3)$

$$\ln \left( \frac{F(V)}{F(0)} \right) = - \frac{\sqrt{2}}{3\sqrt{\pi}} \chi^{3/2} = - \frac{\sqrt{2}}{3\sqrt{\pi}} \left( \frac{V \kappa n^{1/3}}{e^2 N_D^{2/3}} \right)^{3/2}. \quad (43.25)$$

Перейдем теперь к изучению функции  $F(V)$  при больших положительных значениях  $\epsilon = \mu - V$ , т. е. в области потенциальных энергий, лежащих значительно глубже энергии Ферми. В этом случае мы ограничимся приближением равномерно заряженного шара (см. § 39), поскольку численный коэффициент в  $\Omega\{f\}$  в настоящее время все равно не известен. Согласно (39.10) вероятность того, что потенциальная энергия электрона опустилась на величину  $\epsilon$  за счет гауссовой флуктуации примесного заряда, случившейся в шаре радиуса  $R$ , равна

$$\exp \left( - \frac{\epsilon^2}{[(e^2/\kappa R)(N_D R^3)^{1/2}]^2} \right) \quad (43.26)$$

при условии, что электроны не экранируют примесный заряд. Эта вероятность тем больше, чем больше  $R$ , и потому оптимальная флуктуация определяется радиусом электронного экранирования. Если дно ямы лежит на величину  $\epsilon$  ниже уровня Ферми, то концентрация электронов в яме  $\tilde{n} \approx (\sqrt{m\epsilon}/\hbar)^3$ . Тогда, согласно (35.8), радиус экранирования порядка  $\tilde{r}_s = a^{1/2} \tilde{n}^{-1/6}$ . Флуктуации примесной концентрации, имеющие размер  $R > \tilde{r}_s$ , нейтрализуются электронами. Поэтому в (43.26) следует

подставить  $R = \tilde{r}_s$ . В результате получим

$$\ln \left( \frac{F(\varepsilon)}{F(0)} \right) = - \left( \frac{\varepsilon}{\gamma(R_q)} \right)^{9/4} \lambda. \quad (43.27)$$

В работе [3] сформулирован рецепт определения численного коэффициента  $\lambda$ .

#### § 44. Коррелированное распределение примесей

Тот факт, что при достаточно большой степени компенсации электронные состояния оказываются локализованными и происходит переход от металлической проводимости к активационной, не вызывает сомнений и в случае, когда распределение примесей не является пуассоновским. Однако критическая концентрация  $n_c$ , а также масштабы длин и энергий, характеризующие распределение потенциала в СЛКП, могут существенно зависеть от корреляции в распределении примесей.

В работе Келдыша и Прошко (см. [7] к гл. 11) было указано, что корреляция в расположении примесей приводит к своеобразному неэлектронному экранированию. Основная идея этой работы состоит в том, что распределение примесей в компенсированном образце, полученном вытягиванием из расплава, следует считать мгновенной фотографией распределения, имевшего место при температуре замораживания диффузии примесей  $T_0$ . Коэффициент диффузии примесей спадает с температурой экспоненциально с большой энергией активации. Поэтому величина  $T_0$  определена довольно четко. В высокотемпературной плазме заряженные доноры и акцепторы экранируют друг друга. Корреляция в их распределении приводит к тому, что флуктуация заряда больших объемов оказывается гораздо меньше, чем та, которая соответствует пуассоновскому распределению заряженных центров. Поэтому результаты предыдущего параграфа должны быть пересмотрены, если введенная там длина  $r_s$  больше, чем радиус примесного экранирования  $r_i$ .

Если предположить, что примесная плазма при температуре  $T_0$  является слабо неидеальной, то легко вычислить распределение флуктуирующего потенциала. Распределение потенциалов, создаваемых донорами и

акцепторами в идеальной плазме, является гауссовым:

$$F(V) = (\sqrt{\pi} \gamma_i)^{-1} \exp\left(-\frac{V^2}{\gamma_i^2}\right), \quad \gamma_i^2 = 16\pi \frac{e^4 N_D r_i}{x^2}, \quad (44.1)$$

где  $r_i = (\kappa k T_0 / 8\pi N_D e^2)^{1/2}$  — дебаевский радиус экранирования плазмы. (Оценку для  $\gamma_i$  легко получить, подставив  $R = r_i$  в (43.2).) Согласно идее Келдыша и Прошко, в таком потенциале и находятся электроны при низких температурах. Воспользовавшись теорией, развитой в предыдущем параграфе, можно убедиться в том, что если  $r_i < R_q, r_s$ , то собственным потенциалом электронов можно пренебрегать. В этом случае заполнение ям с размером  $r_i$  определяется принципом Паули, а заряд всех электронов, которые может вместить яма, мал по сравнению с примесным зарядом ямы. Если  $r_i$  удовлетворяет условию  $\hbar^2 / m r_i^2 \ll \gamma_i$ , то типичные ямы содержат много уровней и уровень Ферми определяется уравнением (43.9), куда следует подставить (44.1). Если  $\hbar^2 n^{2/3} / m \ll \gamma_i$ , то электроны разбиваются на изолированные капли. В этом случае из (43.9) для  $\mu$  легко получить уравнение

$$\eta^2 = \ln \left[ \frac{(m \gamma_i)^{3/2}}{8\pi^2 \hbar^3 n} \eta^{5/2} \right], \quad (44.2)$$

где  $\eta = -\mu/\gamma$ . При очень малых  $n$

$$\eta \approx \left\{ \ln \left[ \frac{(m \gamma_i)^{3/2}}{8\pi^2 \hbar^3 n} \right] \right\}^{1/2}. \quad (44.3)$$

Легко найти концентрацию электронов в капле  $\bar{n}$ . Для этого следует выяснить, какие значения  $\mu - V$  существенны в интеграле (43.9). Им соответствуют значения концентрации  $n \approx [m(\mu - V)]^{3/2} / \hbar^3$ . В результате с точностью до логарифмических множителей получим

$$\bar{n} = N_D (N_D a^3)^{-1/4} \left( \frac{r_i}{a} \right)^{3/4}. \quad (44.4)$$

Как и при случайном распределении примесей,  $\bar{n}$  не зависит от  $n$ . При  $r_i = R_q$  формулы (44.4) и (43.5) совпадают.

Если  $R_q \ll r_i < r_s$ , то максимальный размер флуктуаций потенциала по-прежнему равен  $r_i$ , а амплитуда типичных флуктуаций порядка  $\gamma_i$ . Однако при исследовании структуры капле необходимо учитывать потенциал электронов. Этот случай рассмотрен в [2]. Размер капле и концентрация электронов в капле оказываются такими же, как и без корреляции примесей. Опускание уровня Ферми описывается (44.3), однако выражение под знаком логарифма несколько иное. Наконец, если  $r_i > r_s$ , корреляция примесей несущественна и применима теория, развитая в § 43.

Из сказанного выше следует, что если существует корреляция в распределении примесей, то она обязательно окажется существенной при достаточно больших степенях компенсации, когда длина  $r_s$  сравняется с  $r_i$ . При дальнейшем увеличении степени компенсации масштаб типичных флуктуаций остается равным  $r_i$  и более не растет. Поэтому дальнейшее углубление уровня Ферми пропорционально не  $n^{-1/3}$ , а  $|\ln n|^{1/2}$ . Однако в полупроводниках с узкой запрещенной зоной это не так. Дело в том, что при температуре  $T_0$  в полупроводнике может быть большое количество собственных электронов и дырок, которые также принимают участие в экранировании. При понижении температуры, после того, как диффузия примесей заморозится, эти носители рекомбинируют, и примесный заряд оказывается частично разэкранированным. Это явление рассмотрено Гальперн и Эфросом [5]. Основной результат состоит в следующем. Пусть концентрация собственных носителей при температуре  $T_0$  равна  $p$ , и пусть  $p \ll N_D$ . Тогда фурье-компонента коррелятора плотности  $\langle \xi(r) \xi(r') \rangle$  имеет вид

$$\langle (\xi_q)^2 \rangle = \frac{4N_D (q^2 + Q_p^2)}{V_0 (q^2 + r_i^{-2})}, \quad (44.5)$$

где  $Q_p^2 = 8\pi p e^2 / \kappa k T_0$ ,  $V_0$  — нормировочный объем. Средний квадрат потенциальной энергии имеет вид

$$\gamma_i^2 = 2 \langle V^2 \rangle = V_0 \int \frac{dq}{(2\pi)^3} \langle \xi_q^2 \rangle \left( \frac{4\pi e^2}{\kappa q^2} \right)^2. \quad (44.6)$$

Если  $p = 0$ , то  $\langle \xi_q^2 \rangle \rightarrow 0$  при  $q \rightarrow 0$ , что и соответствует примесному экранированию. Подставляя (44.5) в (44.6)



при  $p = 0$ , получим (44.1). Однако при  $p \neq 0$  интеграл (44.6) расходится при малых значениях  $q$ . Причина расходимости в том, что  $\langle \xi_q^2 \rangle = 4p/V_0$  при  $q \rightarrow 0$ . На больших расстояниях коррелятор имеет вид (43.8), где вместо  $N_D + N_A$  стоит  $2p$ . Таким образом, дело обстоит так, как будто бы есть примеси с концентрацией  $2p$ , которые совершенно не коррелированы. Как известно из § 43, при  $p \gg n$  эти примеси создают флуктуации потенциала с размером порядка  $r_s$  и амплитудой порядка  $\gamma(r_s)$ , где

$$r_s = \frac{p^{1/3}}{n^{2/3}}, \quad \gamma(r_s) = \frac{e^2 p^{2/3}}{\kappa n^{1/3}}. \quad (44.7)$$

При  $p \gg n^{1/2}(N_{Dr_i})^{3/4}$  имеем  $\gamma(r_s) \gg \gamma_i$ , и мы возвращаемся к ситуации, описанной в § 43. Таким образом, при учете собственных носителей мы получаем, что корреляция не уничтожает крупномасштабные флуктуации, а лишь уменьшает долю примесей, которая принимает в них участие. Разумеется, если  $p > N_D$ , то не происходит даже этого уменьшения. В этом случае корреляцию можно вообще считать отсутствующей.

Оценить величину  $p$  не очень просто, поскольку ширина запрещенной зоны при высоких температурах не известна достаточно хорошо. Однако, по-видимому, можно сказать, что, например, в антимониде индия при  $N_D < 10^{17} \text{ см}^{-3}$  собственные носители полностью уничтожают примесную корреляцию.

## § 45. Кинетические свойства СЛКП

В первую очередь мы остановимся на статической проводимости СЛКП. Результаты предыдущих параграфов состоят в том, что при сильной компенсации электроны образуют изолированные друг от друга капли. Внутри каждой капли концентрация электронов велика и проводимость имеет металлический характер. Однако капли занимают малую долю всего объема кристалла и, следовательно, протекание по ним невозможно. Капли изолированы друг от друга потенциальными барьерами с высотой  $\gamma(r_s)$  и толщиной  $r_s$ . Формула туннельной

прозрачности этого рельефа содержит

$$\exp\left(-\sqrt{\gamma(r_s)} m \frac{r_s}{\hbar}\right), \quad (45.1)$$

и при  $\alpha \gg 1$  прозрачность очень мала. Поэтому при не очень низких температурах проводимость осуществляется путем заброса электронов на уровень протекания  $V_c$ , который, как ясно из сказанного, лежит выше уровня Ферми на величину порядка  $\gamma(r_s)$  (см. рис. 13.2). Таким образом, для энергии активации  $\epsilon_1$  имеем

$$\epsilon_1 = |V_c - \mu| = \phi_1 \frac{e^2 N^{2/3}}{\kappa n^{1/3}}. \quad (45.2)$$

Численный коэффициент  $\phi_1$  неизвестен, хотя в принципе его можно вычислить с помощью безразмерных уравнений, приведенных в § 43. Можно вычислить также и коэффициент  $\beta$  в формуле (43.1) для критической концентрации электронов, при которой происходит переход от металлической проводимости к активационной. Заметим, что при  $\alpha \approx 1$  потенциальный рельеф остается квазиклассическим. Таким образом, рассматриваемый переход является квазиклассическим вариантом перехода Андерсона. Он происходит, когда уровень протекания пересекает уровень Ферми. На возможность такого перехода впервые указал Займан (см. [54] к гл. 5).

Приведенные выше формулы относились к случаю, когда корреляция в распределении примесей отсутствует. Если  $\sqrt{m\gamma_i} r_i/\hbar \gg 1$ , то качественная картина явления остается такой же и при корреляции. Однако выражения (43.1) и (45.2) меняются. Если  $r_i < R_q$ ,  $r_s$ , то, как говорилось в § 44, задача оказывается одноэлектронной и можно найти даже численные коэффициенты в формулах. Как указывалось в § 17, для гауссовой случайной функции критическая доля пространства, соответствующая возникновению протекания, равна 0,17. Отсюда легко получить (см. (17.2)), что уровень протекания  $V_c = -0,67\gamma_i$ . Для определения  $\mu$  следует воспользоваться (43.9), подставив туда (44.1). При сильной компенсации, согласно (44.3), получим

$$\epsilon_1 \approx \gamma_i \left\{ \ln \left[ \frac{(m\gamma_i)^{3/2}}{8\pi^2 \hbar^3 n} \right] \right\}^{1/2}. \quad (45.3)$$

Из (43.9) получим также, что  $\mu = V_c$  при

$$n_c = \frac{\beta (2m\gamma_l)^{3/2}}{3\pi^3 \hbar^3}. \quad (45.4)$$

где  $\beta = \int_{0,67}^{\infty} (x - 0,67)^{3/2} \exp(-x^2) dx = 0,11$ .

При температурах  $T < T_s$ , где  $T_s$  определяется условием

$$\frac{\sqrt{m\gamma(r_s)} r_s}{\hbar} = \frac{\gamma(r_s)}{kT_s} \quad (45.5)$$

или аналогичным условием с  $\gamma_l$  и  $r_l$ , вероятность активации на уровень протекания становится меньше, чем вероятность туннелирования в окрестности уровня Ферми. В этой области доминирует прыжковая проводимость по каплям. Ее температурная зависимость без учета кулоновской щели изучалась в работе [6]. При достаточно низкой температуре наличие кулоновской щели ведет к закону (33.18), причем в (33.19) следует заменить радиус состояния  $a$  на величину  $\hbar/\sqrt{m|\mu|}$ , учитывающую туннелирование под барьер с высотой  $\mu$ .

Переход от металлической проводимости к активационной под влиянием компенсации наблюдался в  $n$ -Ge (Сасаки и Яманучи [7], Гаджиев и др. [1], Забродский и др. [17] к гл. 9), в GaAs (Редфилд и Крендл [8]), в CdTe (Агринская и Крымова [9], Дидковский и Хиврич [10]), в  $n$ -InSb (Гершензон и др. [11], Яременко [12]).

К сожалению, эти данные не позволяют проанализировать зависимость  $n_c(N)$ , так как почти все эксперименты выполнены в области, где параметр сильного легирования  $Na^3$  был равен единице или незначительно превышал ее. В работах [9, 11] указывается, что зависимость (45.2) выполняется хорошо. Тем не менее ситуацию в целом нельзя считать удовлетворительной. Дело в том, что область с линейной зависимостью  $\ln \rho$  от  $T^{-1}$ , как правило, оказывается в экспериментах очень узкой, что трудно объяснить в рамках изложенной выше теории, хотя попытки в этом направлении предпринимались [6]. Узость линейного участка может свидетельствовать о такой корреляции в расположении примесей, при ко-

торой крупномасштабные квазиклассические флуктуации отсутствуют. Кроме того, при очень низких температурах температурная зависимость сопротивления иногда становится слабой [11, 12], что также трудно объяснить.

Переход от металлической проводимости к активационной в СЛП происходит не только под действием компенсации, но и под действием магнитного поля. Это связано с тем, что магнитное поле, сжимая волновые функции электронов, способствует их локализации. Указанное явление называется магнитным вымораживанием электронов (см. также гл. 7). Возникает вопрос, как происходит вымораживание в СЛКП. В частности, интересно знать зависимость критической концентрации электронов  $n_c$ , при которой происходит срыв металлической проводимости, от магнитного поля и энергии активации  $e_1(n, H)$ . Эта задача решена Шкловским и Эфросом [13] для модели некоррелированных примесей и Эфросом и Янчевым [14] для коррелированного распределения.

В случае некоррелированных примесей возникает интересная теоретическая задача о квантовом нелинейном экранировании. До сих пор мы делили экранирование на линейное и нелинейное. Однако есть еще один способ классификации. Экранирование, которое мы описывали в § 35 на основе квазиклассического уравнения, можно назвать классическим. Оно отличается тем, что радиус экранирования велик по сравнению с длиной волны электрона. В квантующем магнитном поле длина волны поперек поля мала, а вдоль поля велика. Это приводит к интересным особенностям электронного экранирования. В частности, экранирование точечного заряда оказывается сильно анизотропным. Такое экранирование естественно назвать квантовым. Оно описано в обзоре Хоринга [15] (см. также [13]). Квантовое экранирование в свою очередь может быть линейным и нелинейным. В работе [13] нелинейная теория экранирования, изложенная в § 43, обобщена на квантовый случай.

Сильная неоднородность электронной плотности, возникающая в СЛКП, проявляется не только в статической проводимости. Как известно, образцы  $n$ -InSb при гелиевых температурах широко используются в качестве приемного элемента для субмиллиметровых и миллимет-

ровых волн. Оказалось [16], что чувствительность приемников и величина фотоотклика существенно возрастают при компенсации образцов. В связи с этим появилось большое количество работ [17—20], в которых делаются попытки объяснить свойства компенсированного  $n$ -InSb на основе изложенных выше представлений, хотя обсуждаются главным образом образцы с  $N_D \approx 10^{16} \text{ см}^{-3}$ , т. е. с  $N_{Da^3} \approx 1$ , и потому речь может идти о качественном понимании.

Наличие высоких потенциальных барьеров в СЛКП позволяет также объяснить явление остаточной (или замороженной) фотопроводимости. Качественное объяснение этого явления было дано Рывкиным и Шлимаком [21]. Оно состоит в том, что родившиеся электроны и дырки быстро отделяются друг от друга и находятся в разных областях пространства, разделенных потенциальными барьерами. Это сильно затрудняет рекомбинацию, приводя к долговременной релаксации фототока. Качественная теория этого явления была построена Ткачем [22] и Шиком [23]. Наконец, теория люминесценции СЛКП, основанная на представлениях о крупномасштабных флуктуациях, была построена Леванюком и Осиповым [24].

#### § 46. Полностью компенсированный полупроводник

Теория, изложенная в § 43, предполагает, что опускание уровня Ферми вследствие компенсации меньше, чем половина ширины запрещенной зоны  $E_g$ . Возникает интересный вопрос о том, как должны быть устроены электронные состояния, если компенсация столь точная, что это условие не выполняется. Ниже мы разберем качественно этот случай, а затем объясним, почему он может вызвать интерес.

Рассмотрим полностью компенсированный полупроводник, содержащий  $N$  доноров и  $N$  акцепторов в единице объема. Если бы доноры и акцепторы были равномерно распределены в пространстве, электроны и дырки при нулевой температуре полностью рекомбинировали бы друг с другом. Фактически это был бы собственный полупроводник, в котором уровень Ферми лежит посередине запрещенной зоны.

Учтем теперь флуктуации концентрации примесей, считая, что корреляция в их расположении отсутствует. При этом мы снова сталкиваемся с вопросом, каким образом экранируются крупномасштабные флуктуации. В полностью компенсированном полупроводнике свободные носители отсутствуют при низких температурах и в зоне проводимости, и в валентной зоне. Поэтому в рассматриваемой модели ничто не мешает случайным колебаниям потенциальной энергии достигать значений порядка  $E_g$ . Легко, однако, видеть, что колебаний потенциала больших, чем  $E_g/2$ , быть не может. Действительно, если в некоторой области избыток положительного заряда примесей опустит дно зоны проводимости ниже уровня Ферми, то в этой области возникнут электроны в количестве, достаточном для того, чтобы не позволить зоне проводимости опуститься существенно глубже (рис. 13.3).

Аналогичным образом, избыток отрицательного заряда не может поднять дно зоны проводимости выше, чем  $E_g/2$ , так как при этом вершина валентной зоны окажется выше энергии Ферми и в ней возникнут дырки.

Определим наименьший размер  $R_g$  типичной флуктуации, изгибающей зоны на величину порядка  $E_g$ . Используя (43.2), получим для  $R_g$

$$\gamma(R_g) = E_g, \quad R_g = \frac{E_g^2 \kappa^2}{N e^4}. \quad (46.1)$$

Число электронных состояний  $\mathcal{N}_q(R_g)$  в яме с размером  $R_g$  и глубиной  $E_g$  есть \*)

$$\mathcal{N}_q(R_g) = \frac{(m_e E_g)^{3/2} R_g^3}{\hbar^3}. \quad (46.2)$$

Избыточное число примесей в такой яме  $Z(R_g) = (N R_g^3)^{1/2}$ . В типичных полупроводниках при концентрациях примесей, малых по сравнению с концентрацией основных атомов, хорошо выполняется неравенство ( $a_e = \hbar^2 \kappa / m_e e^2$ )

$$\frac{Z(R_g)}{\mathcal{N}_q(R_g)} = \left( \frac{e^2 N^{1/3}}{\kappa E_g} \right)^{1/2} (N a_e^3)^{1/2} \ll 1. \quad (46.3)$$

\*) Здесь, как и во всех качественных рассуждениях этой главы, мы не пишем численных множителей.

Это означает, что электроны, необходимые для компенсации заряда ямы размера  $R_g$ , могут разместиться в самых низших ее состояниях, другими словами, дно такой ямы может опуститься ниже уровня Ферми лишь на глубину, малую по сравнению с глубиной ямы, как и показано на рис. 13.3. Из сказанного ясно, что флуктуации

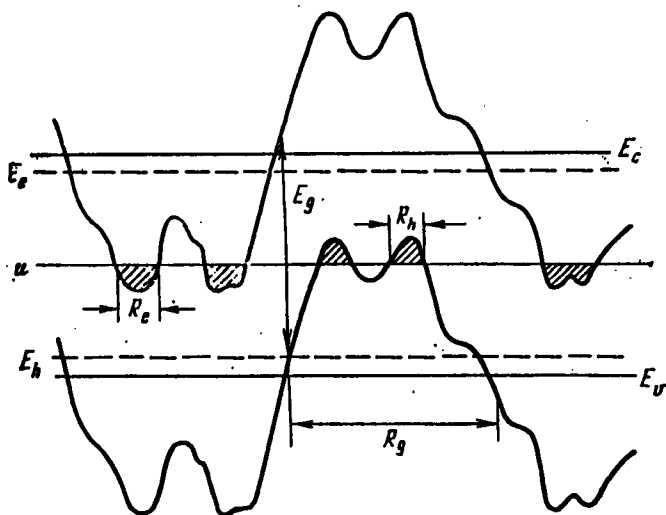


Рис. 13.3. Энергетическая схема полностью компенсированного полупроводника. Верхняя и нижняя прямые — невозмущенное положение дна зоны проводимости и вершины валентной зоны, средняя — уровень Ферми. Извилистые линии — границы зон, модулированные электростатическим потенциалом зарядечных примесей, штриховые — уровни протекания для электронов и дырок. Области, занятые носителями (капли), заштрихованы.  $R_e$ ,  $R_h$  — размеры электронных и дырочных капель.

с размером, большим  $R_g$ , нейтрализуются электронами и дырками, т. е.  $R_g$  играет роль радиуса экранирования потенциала примесей.

Таким образом, крупномасштабные флуктуации превращают собственный полупроводник в систему, похожую на полуметалл. При  $T = 0$  имеется довольно большое число электронов и дырок, расположенных, однако, в изолированных друг от друга областях пространства. По этой причине статическая проводимость такой системы не имеет металлического характера и должна осу-

ществляться путем теплового заброса носителей на уровне протекания.

В предыдущих рассуждениях предполагалось, что концентрация доноров  $N_D$  точно равна концентрации акцепторов  $N_A$ . Легко сообразить, что эти рассуждения справедливы, если концентрация несобственных электронов  $n = N_D - N_A$  мала по сравнению со средней концентрацией электронно-дырочных пар  $n_p$ , автоматически возникающих за счет изгиба зон. Очевидно, что концентрация  $n_p$  близка к средней концентрации избыточного примесного заряда в объеме  $R_g^3$ , т. е.

$$n_p = \frac{(NR_g^3)^{1/2}}{R_g^3} = \frac{N^2 e^6}{E_g^3 \kappa^3}. \quad (46.4)$$

Если  $n \gg n_p$ , то, согласно (43.3),  $r_s \ll R_g$ , а  $\gamma(r_s) \ll E_g$ , т. е. уровень Ферми близок к зоне проводимости. Таким образом, по мере увеличения степени компенсации растет радиус экранирования, увеличивается амплитуда крупномасштабных флуктуаций и уровень Ферми погружается в глубь запрещенной зоны. При приближении уровня Ферми к середине запрещенной зоны экранирование начинает осуществляться собственными носителями и дальнейшее опускание уровня Ферми прекращается.

Согласно (46.4) ситуацию, при которой  $n < n_p$ , легче всего осуществить в узкозонных полупроводниках. Кроме того, согласно сказанному в § 44, в этом случае легче всего реализовать некоррелированное распределение примесей, необходимое для применимости изложенной выше теории.

Фрицше [25] предложил картину изгиба зон, изображенную на рис. 13.3, для объяснения экспериментальных данных, относящихся к аморфным полупроводникам. Согласно идеям Фрицше, такая картина в состоянии объяснить противоречие между большой плотностью состояний на уровне Ферми, наблюдаемой в экспериментах по СВЧ-поглощению и эффекту поля, и малой плотностью состояний в запрещенной зоне, наблюдаемой в экспериментах по межзонному оптическому поглощению. В связи с этим Шкловский и Эфрос [3] и развили изложенную выше теорию, предложив схему полностью компенсированного полупроводника в качестве модели аморф-



ного полупроводника. Как видно из рис. 13.3, плотность электронных состояний в запрещенной зоне в рассматриваемой модели очень велика, однако благодаря плавности потенциала «локальная» ширина запрещенной зоны всюду близка к  $E_g$  и коэффициент межзонного поглощения света должен убывать с дефицитом кванта очень резко (он должен описываться формулой (42.5)). Из предположения о плавном изгибе зон также следует (см. [62] к гл. 5), что оптическая ширина запрещенной зоны должна быть больше, чем удвоенная энергия активации статической электропроводности (электрическая ширина запрещенной зоны).

Рыкин и Шлимак [21] также подчеркивали аналогию между свойствами СЛКП и аморфного полупроводника и связывали ее с наличием крупномасштабного потенциала в обеих системах. Они высказывали идею о моделировании аморфного полупроводника путем создания разными способами компенсирующих примесей в кристаллическом легированном полупроводнике.

## ПРИЛОЖЕНИЯ

### 1. Вычисление показателя экспоненты энергетического интеграла перекрытия $I_{ij}$ в магнитном поле

Показатели экспонент слагаемых (15.5) одинаковы. Поэтому остановимся лишь на вычислении показателя экспоненты интеграла  $\int F_i(\mathbf{r}) F_j(\mathbf{r}) d\mathbf{r}$ . Начнем со случая сильного поля. В этом случае, поместив донор  $i$  в начало координат, получаем

$$\begin{aligned}
 I_{ij} \sim \int_{-\infty}^{\infty} dx dy dz \exp \left\{ -\frac{x^2 + y^2}{4\lambda^2} - \frac{|z|}{a_H} - \right. \\
 \left. - \frac{(x - x_j)^2 + (y - y_j)^2}{4\lambda^2} - \frac{|z - z_j|}{a_H} + \frac{i}{2\lambda^2} (y_j x - x_j y) \right\} \infty \\
 \infty \exp \left\{ -\frac{|z_{ij}|}{a_H} + \frac{x_{ij}^2 + y_{ij}^2}{4\lambda^2} \right\}. \quad (\text{П1.1})
 \end{aligned}$$

В слабом магнитном поле декартову систему координат удобнее выбрать таким образом, чтобы донор  $i$  находился в начале координат, а донор  $j$  — на оси  $x$  на расстоянии  $x_0 \equiv r_{ij}$  от донора  $i$ . Пусть вектор  $\mathbf{H}$  лежит в плоскости  $(x, z)$ . Обозначим  $\vartheta$ ,  $\vartheta_1$  и  $\vartheta_2$  углы, составляемые вектором  $\mathbf{H}$  с осью  $x$  и векторами  $\mathbf{r}$  и  $\mathbf{r} - \mathbf{r}_j$  соответственно. Тогда

$$\begin{aligned}
 I_{ij} \sim \int d\mathbf{r} \exp \left\{ -\frac{r}{a} - \frac{r^2 a \sin^2 \vartheta_1}{24\lambda^4} - \frac{|r - r_j|}{a} - \right. \\
 \left. - \frac{|r - r_j|^3 a \sin^2 \vartheta_2}{24\lambda^4} - \frac{i x_0 y}{2\lambda^2} \sin \vartheta \right\}, \quad (\text{П1.2})
 \end{aligned}$$

Подынтегральное выражение (П1.2) имеет острый максимум вблизи точки  $(x_0/2, 0, 0)$ . Поэтому можно разложить величины  $r$  и  $|r - r_I|$  вблизи этой точки и пренебречь отличием углов  $\theta_1$  и  $\pi - \theta_2$  от  $\theta$ . В результате получаем

$$\begin{aligned}
 I_{IJ} \sim \exp \left\{ -\frac{x_0}{a} - \frac{x_0^3 a \sin^2 \theta}{96\lambda^4} \right\} \times \\
 \times \int_{-\infty}^{\infty} dx \exp \left\{ -\frac{x_0 a \sin^2 \theta}{8\lambda^4} \left(x - \frac{x_0}{2}\right)^2 \right\} \int_{-\infty}^{\infty} dz \exp \left\{ -\frac{2z^2}{x_0 a} \right\} \times \\
 \times \int_{-\infty}^{\infty} dy \exp \left\{ -\frac{2y^2}{x_0 a} - \frac{ix_0 y}{2\lambda^2} \sin \theta \right\} \infty \\
 \infty \exp \left\{ -\frac{x_0}{a} - \frac{x_0^3 a \sin^2 \theta}{24\lambda^4} \right\}. \quad (\text{П1.3})
 \end{aligned}$$

Возвращаясь в систему координат, в которой ось  $z$  направлена по вектору  $H$ , окончательно имеем

$$I_{IJ} \infty \exp \left\{ -\frac{r_{IJ}}{a} - \frac{r_{IJ} a (x_{IJ}^2 + y_{IJ}^2)}{24\lambda^4} \right\}. \quad (\text{П1.4})$$

## 2. Энергия активации $\epsilon_3$ в случае анизотропных волновых функций

**А.** Сдвиг порога протекания при малом изменении поверхности  $Q_\xi$ . Рассмотрим сначала вспомогательную задачу, с частными случаями которой мы уже сталкивались в § 25. Допустим, что нам известен порог протекания  $\xi_c^0$  в задаче охватывающих поверхностей (см. § 18) с критерием связности  $\xi_{IJ}^0(r_{IJ}) \leq \xi$ . Предположим, что в условии связности внесено малое возмущение:  $\xi_{IJ}^0(r_{IJ}) + \Delta\xi(r_{IJ}) \leq \xi$ . Требуется найти сдвиг порога протекания  $\Delta\xi_c = \xi_c - \xi_c^0$ , связанный с этим возмущением. При произвольной функции  $\xi_{IJ}^0$  эта задача не имеет точного решения. Ниже будет изложен приближенный метод, основанный на том, что критическое число связей на узел  $B_c$  слабо зависит от формы поверхности  $Q_\xi$ , которая определяется уравнением  $\xi_{IJ}(r_{IJ}) = \xi$  (см. § 18).

Отсюда следует приближенная инвариантность объема  $V_{\xi_c}$ , ограниченного поверхностью  $Q_{\xi_c}$ . Запишем уравнения поверхностей  $Q_{\xi}$  для невозмущенной и возмущенной задач соответственно:

$$r = r_0(\theta, \varphi, \xi), \quad (\text{П2.1})$$

$$r = r_0(\theta, \varphi, \xi) + \Delta r(\theta, \varphi, \xi); \quad (\text{П2.2})$$

здесь  $r, \theta, \varphi$  — сферические координаты точек поверхности. При фиксированном значении  $\xi$  объем  $V_{\xi}^0$ , ограниченный поверхностью (П2.1), отличается от объема  $V_{\xi}$ , ограниченного поверхностью (П2.2). Следует найти  $\Delta \xi_c$  из условия, что

$$V_{\xi_c^0}^0 = V_{\xi_c}. \quad (\text{П2.3})$$

Из (П2.1) и (П2.2) имеем

$$V_{\xi_c^0}^0 - V_{\xi_c^0}^0 = \int r_0^2(\theta, \varphi, \xi_c^0) \Delta r(\theta, \varphi, \xi_c^0) d\Omega. \quad (\text{П2.4})$$

Подставив  $V_{\xi_c}$  в виде  $V_{\xi_c^0}^0 + \left. \frac{dV_{\xi}}{d\xi} \right|_{\xi=\xi_c^0} \Delta \xi_c$  и используя (П2.4), получаем из (П2.3)

$$\begin{aligned} \Delta \xi_c &= \left( \left. \frac{dV_{\xi}}{d\xi} \right|_{\xi=\xi_c^0} \right)^{-1} \int r_0^2(\theta, \varphi, \xi_c^0) \Delta r(\theta, \varphi, \xi_c^0) d\Omega = \\ &= \left( \left. \frac{dV_{\xi}}{d\xi} \right|_{\xi=\xi_c^0} \right)^{-1} \int r_0^2(\theta, \varphi, \xi_c^0) \left. \frac{dr}{d\xi} \right|_{\xi_c^0} \Delta \xi [r_0(\theta, \varphi, \xi_c^0), \theta, \varphi] d\Omega. \end{aligned} \quad (\text{П2.5})$$

Используя идеи теории возмущений (см. § 28), можно показать, что в частных случаях, когда невозмущенная поверхность  $Q_{\xi}$  является сферой или эллипсоидом, результат (П2.5) оказывается точным в первом порядке по  $\Delta \xi$  при произвольном возмущении. Подчеркнем еще раз, что в общем случае вывод (П2.5) основывался на приближенной инвариантности  $V_c$ .

**Б. Энергия активации  $\varepsilon_3$ .** Функция распределения величин  $\varepsilon_{ij}$ , вообще говоря, зависит от  $r_{ij}$ . Обозначим ее  $f_r(\varepsilon)$ . Как показано в § 28, энергия активации  $\varepsilon_3$  определяется величинами  $\varepsilon_{ij}$  тех сопротивлений, для которых  $\xi_{ij}^0 = \xi_c^0$ . В изотропном случае эти сопротивления

соединяют узлы с  $r_{ij} = r_c$  и, следовательно,  $\Delta\xi$  является функционалом  $f_{r_c}(\varepsilon)$ . В анизотропном случае уравнение поверхности  $\xi_{ij}^0 = \xi_c^0$  можно записать в виде  $r = r_0(\vartheta, \varphi, \xi_c^0)$ . В результате на энергию активации  $\varepsilon_3$  влияют величины  $\varepsilon_{ij}$ , принадлежащие сопротивлениям с различными  $r_{ij}$ . Поэтому в линейном приближении по  $\varepsilon_{ij}$  наиболее общее выражение для поправки  $\Delta\xi_c$  имеет вид

$$\begin{aligned} \Delta\xi_c &= \frac{1}{kT} \int d\Omega d\varepsilon dr \varepsilon f_r(\varepsilon) F(\vartheta, \varphi) \delta\{r - r_0(\vartheta, \varphi, \xi_c^0)\} = \\ &= \frac{1}{kT} \int d\Omega F(\vartheta, \varphi) \varepsilon \{r(\vartheta, \varphi, \xi_c^0)\}, \quad (\text{П2.6}) \end{aligned}$$

где  $F(\vartheta, \varphi)$  — неизвестная функция, а  $\varepsilon(r)$  — среднее значение  $\varepsilon_{ij}$  при  $r_{ij} = r$ :

$$\varepsilon(r) = \int \varepsilon d\varepsilon f_r(\varepsilon). \quad (\text{П2.7})$$

Согласно (П2.6) для всех функций  $f_r(\varepsilon)$ , имеющих одинаковое значение  $\varepsilon(r)$ , сдвиг  $\Delta\xi_c$  одинаков. В частности, для вычисления  $\Delta\xi_c$  можно заменить исходную функцию  $f_r(\varepsilon)$  функцией  $\delta(\varepsilon - \varepsilon(r))$ . При этом возникает рассмотренная в пункте А задача о сдвиге порога протекания за счет малого возмущения поверхности  $Q_\xi$ . Решением этой задачи и определяется  $\varepsilon_3$ .

Воспользовавшись приближенным рецептом (П2.5), найдем  $F(\vartheta, \varphi)$ . Подставляя  $\Delta\xi = \varepsilon(r)/kT$  в (П2.5) и сравнивая результат с (П2.6), найдем, что

$$F(\vartheta, \varphi) = r_0^2(\vartheta, \varphi, \xi_c^0) \left( \frac{dr_0}{d\xi} / \frac{dV_\xi}{d\xi} \right) \Big|_{\xi = \xi_c^0}. \quad (\text{П2.8})$$

В частном случае сферической поверхности  $Q_\xi$  формула (П2.8) является точной, а  $F(\vartheta, \varphi) = 1/4\pi$ .

### 3. Оценка коэффициентов $t_1$ и $n_{c1}$ в формулах (30.2) и (30.8)

Рассмотрим задачу об определении  $\xi_c$  для прыжковой проводимости с переменной длиной прыжка в четырехмерном пространстве безразмерных переменных  $z, \Delta$  (см. (29.11)). Вокруг каждого узла построим четырехмер-

ную гиперповерхность  $Q_i$ , внутри которой должен находиться узел  $j$ , чтобы выполнялось условие связности, и вычислим объем  $V_i$ , заключенный внутри этой поверхности. Тогда  $\xi_c$  определится из условия

$$\frac{n(\xi_c)}{2} \bar{V}_i = B_c^{(4)}, \quad (\text{ПЗ.1})$$

где  $\bar{V}_i$  — среднее значение  $V_i$ ;  $n(\xi)$  — концентрация узлов в трехмерном  $s$ -пространстве, даваемая формулой (29.15);  $n(\xi)/2$  — число узлов в единице объема четырехмерного пространства, ограниченного гиперплоскостями  $\Delta = \pm 1$ ;  $B_c^{(4)}$  — критическое число связей на узел. При  $H = 0$  уравнение поверхности  $Q_i$  имеет вид  $s_{ij} + \Delta_{ij} = 1$ , так что

$$V_i = \frac{4\pi}{3} \int_{\Delta_{ij} < 1} (1 - \Delta_{ij})^3 d\Delta_{ij},$$

$$\bar{V}_i = \bar{V}_i^0 = \frac{4\pi}{3} \iint_{\Delta_{ij} < 1} (1 - \Delta_{ij})^3 d\Delta_{ij} \frac{d\Delta_{ij}}{2} = \frac{\pi}{5}. \quad (\text{ПЗ.2})$$

Пусть  $\xi_c^0$  — значение  $\xi_c$  при  $H = 0$ . Тогда, согласно § 29,  $n(\xi_c^0) = n_c = 5,3$  и  $B_c^{(4)} = 1/2 n_c \bar{V}_i^0 = 1,66$ . В слабом магнитном поле уравнение поверхности  $Q_i$  имеет вид

$$s_{ij} + \Delta_{ij} + \beta s_{ij}^3 \sin^2 \theta = 1, \quad \text{где } \beta = \frac{a^4}{96\lambda^4} \left(\frac{T_0}{T}\right)^{1/2} \ll 1. \quad (\text{ПЗ.3})$$

Последнее слагаемое в левой части происходит из добавки (30.1) после замены  $r_{ij}$  на  $s_{ij} \frac{a\xi_c^0}{2}$  в соответствии с (29.11) и (29.10). Вычисляя объем  $\bar{V}_i$  в линейном по  $\beta$  приближении, получаем

$$\bar{V}_i = \frac{1}{3} \iint_{\Delta_{ij} < 1} \frac{d\Delta_{ij}}{2} d\Delta_{ij} \int d\Omega [1 - \Delta_{ij} - \beta(1 - \Delta_{ij})^3 \sin^2 \theta]^3 =$$

$$= \bar{V}_i^0 - \frac{4\pi}{21} \beta \equiv \bar{V}_i^0 - \Delta \bar{V}_i. \quad (\text{ПЗ.4})$$

Будем считать, что магнитное поле практически не изменяет  $B_c^{(4)}$ . Запишем  $\xi_c$  в виде  $\xi_c = \xi_c^0 + \Delta \xi_c$ , где  $\Delta \xi_c \ll \xi_c^0$ , и, используя (29.15), линеаризуем произведение  $n(\xi_c) \bar{V}_i/2$

по  $\Delta \xi_{sc}$  и  $\Delta \bar{V}_i$  (см. аналогичную процедуру в § 25 и Приложении 2А). Тогда из (П3.1) получаем

$$\Delta \xi_{sc} = \frac{\xi_{sc}^0}{4} \frac{\Delta \bar{V}_i}{\bar{V}_i^{(0)}} = \frac{5}{21} \beta_{sc}^0 = \frac{5}{2016} \frac{a^4}{\lambda^1} \left( \frac{T_0}{T} \right)^{3/4}. \quad (\text{П3.5})$$

Таким образом,  $t_1 = 5/2016$ .

В сильном магнитном поле удобнее использовать четырехмерное пространство переменных (30.4). Согласно (30.6) в этом случае уравнение поверхности  $Q_i$  имеет вид  $\tau_{ij}^2 + \eta_{ij}^2 + |\xi_{ij}| + \Delta_{ij} = 1$  и, как нетрудно показать,  $\bar{V}_i = 1/\pi$ . Тогда, считая, что критическое число связей на узел остается таким же, как при  $H = 0$ , получаем

$$\frac{n_{c1}}{2} \bar{V}_i = B_c^{(4)}, \quad n_{c1} = \frac{8B_c^{(4)}}{\pi} \approx 4,2. \quad (\text{П3.6})$$

## ЛИТЕРАТУРА

### К предисловию

1. *Pollak M.* Hopping conduction in vitreous or amorphous semiconductors. — Труды 6-й Межд. конф. по аморфным и жидким полупроводникам. — Л.: Наука, 1976, с. 79.
2. *Böttger H., Bryksin V. V.* Hopping conductivity in ordered and disordered solids. — *Phys. Stat. Sol.*, 1976, v. 78, p. 9.

### К главе 1

1. *Kohn W.* Shallow impurity states in Silicon and Germanium. — *Solid St. Phys.* — N. Y., 1957, v. 5
2. *Bassani F., Iadonisi G., Preziosi B.* Electronic impurity levels in semiconductors. — *Reports on Prog. Phys.*, 1974, v. 37, p. 1099.
3. *Бир Г. Л., Пикус Г. Е.* Симметрия и деформационные эффекты в полупроводниках. — М.: Наука, 1972.
4. *Kohn W., Luttinger J. M.* Theory of donor states in silicon. — *Phys. Rev.*, 1955, v. 98, p. 915. См. перевод в сб.: «Проблемы физики полупроводников». — М.: ИЛ, 1957, с. 551.
5. *Luttinger J. M., Kohn W.* Motion of electrons and holes in perturbed periodic fields. — *Phys. Rev.*, 1955, v. 97, p. 869. См. перевод в сб.: «Проблемы физики полупроводников». — М.: ИЛ, 1957, с. 515.
6. *Luttinger J. M.* Quantum theory of cyclotron resonances in semiconductors. — *Phys. Rev.*, 1956, v. 102, p. 1030.
7. *Гельмонт Б. Л., Дьяконов М. И.* Примесные состояния в полупроводнике с нулевой запрещенной зоной. — *ЖЭТФ*, 1972, т. 62, с. 713.
8. *Ландау Л. Д., Лифшиц Е. М.* Квантовая механика. — М.: Наука, 1974.
9. *Гельмонт Б. Л., Дьяконов М. И.* Акцепторные уровни в полупроводнике со структурой алмаза. — *ФТП*, 1971, т. 5, с. 2191.
10. *Balderness A., Lipari N. O.* Spherical model of shallow acceptor states in semiconductors. — *Phys. Rev.*, 1973, v. B8, p. 2697.

### К главе 2

1. *Пайерлс Р.* Квантовая теория твердых тел. — М.: ИЛ, 1956, с. 98.
2. *Lieb E., Wu F. Y.* Absence of Mott transition in an exact solution of the short-range, one band model in one dimension. — *Phys. Rev. Lett.*, 1968, v. 20, p. 1445.



3. *Mott N. F.* Metal-Insulator Transitions. — London: Taylor and Francis, 1974, p. 124.
4. *Холмский Д. И.*, Электронные корреляции в узких зонах. — ФММ, 1970, т. 29, с. 31.
5. *Anderson P. W.* Absence of diffusion in certain random lattice. — Phys. Rev., 1958, v. 109, p. 1492.
6. *Mott N. F., Twose W. D.* The theory of impurity conduction. — Adv. Phys., 1961, v. 10, p. 107.
7. *Borland R. E.* The nature of electronic states in disordered one-dimensional systems. — Proc. Roy. Soc., 1963, v. A274, p. 529.
8. *Halperin B. I.* Properties of a particle in a one-dimensional random potential. — Adv. Chem. Phys., 1967, v. 13, p. 123.
9. *Березинский В. Л.* Кинетика квантовой частицы в одномерном случайном потенциале. — ЖЭТФ, 1973, т. 65, с. 1251.
10. *Ishii K.* Localization of eigenstates and transport phenomena in the one-dimensional disordered system. — Prog. Theor. Phys. Suppl., 1973, v. 53, p. 77.
11. *Papatriantafyllou C., Economou E. N., Eggarter T. P.* Eigenfunctions in one-dimensional disordered systems, I; II. — Phys. Rev., 1976, v. B13, p. 910; *ibid.*, 1976, v. B13, p. 920.
12. *Thouless D. J.* Electrons in disordered systems and the theory of localization. — Phys. Rep., 1974, v. 13C, p. 94.
13. *Edwards J. T., Thouless D. J.* Numerical studies of localization in disordered systems. — J. Phys., 1972, v. C5, p. 807.
14. *Licciardello D. C., Thouless D. J.* Conductivity and mobility edges for two-dimensional disordered systems. — J. Phys., 1975, v. C8, p. 4157.
15. *Weaire D., Srivastava V.* The critical strength of disorder for the Anderson transition in three dimensions. — Communications on Physics, 1977, v. 2, p. 73. New numerical results for Anderson localization. — Proc. 7th Int. Conf. on Amorphous and Liquid Semicond./Ed. W. E. Spear. — Edinburgh, 1977, p. 286.
16. *Мотт Н., Дэвис Э.* Электронные процессы в некристаллических веществах. — М.: Мир, 1974, с. 47.
17. *Mott N. F.* Conduction in amorphous materials. — Electronics and Power, 1973, v. 19, p. 321.
18. *Stern F.*, Evidence for a mobility edge in inversion layers. — Phys. Rev., 1974, v. B9, p. 2762.
19. *Mott N. F., Peper M., Pollit S., Wallis R. H., Adkins C. J.* The Anderson transition. — Proc. Roy. Soc., 1975, v. A345, p. 169.
20. *Mott N. F.* Electrons in disordered structures. — Adv. Phys., 1967, v. 16, p. 49.
21. *Mott N. F.* Conduction in non-crystalline systems. — Phil. Mag., 1974, v. 29, p. 613.
22. *Cohen M. H.*, Review of the theory of amorphous semiconductors. — J. Non-Cryst. Solids, 1970, v. 4, p. 391.  
*Cohen M. H., Jortner J.* Inhomogeneous transport regime in disordered materials. — Phys. Rev. Lett., 1973, v. 30, p. 699.
23. *Ioffe A. F., Regel A. R.* Non-crystalline, amorphous and liquid electronic semiconductors. — Prog. Semicond., 1960, v. 4, p. 237.
24. *Licciardello D. C., Thouless D. J.* Constancy of minimum metallic conductivity. — Phys. Rev. Lett., 1975, v. 35, p. 1475.

25. *Anderson P. W.* The size of localized states near the mobility edge. — Proc. Nat. Acad. Sci. USA, 1972, v. 69, p. 1097.
26. *Abram R. A., Edwards S. F.* The nature of the electronic states of a disordered system. — J. Phys., 1972, v. C5, p. 1183.
27. *Freed K. F.* Selfconsistent theory of electron localization. — J. Phys. 1971, v. C4, p. L331.
28. *Садовский М. В.,* Локализация электронов в неупорядоченных системах. — ЖЭТФ, 1976, т. 70, с. 1936.
29. *De Gennes P. G.* Exponents for the excluded volume problem as derived by the Wilson method. — Phys. Lett., 1972, v. 38A, p. 339.
30. *Abram R. A.* Localized states in inverted silicon-silicon dioxide interfaces. — J. Phys., 1973, v. C6, p. L379.
31. *Last B. J., Thouless D. J.* Evidence for power law localization in disordered systems. — J. Phys., 1974, v. C7, p. 699.
32. *Licciardello D. C., Economou E. N.* Study of localization in Anderson's model for random lattice. — Phys. Rev., 1975, v. B11, p. 3697.
33. *Иванов М. А., Погорелов Ю. Г.,* Электронные состояния в кристалле с примесными уровнями вблизи края зоны. — ЖЭТФ, 1977, т. 72, с. 2198.
34. *Lloyd P.* Exactly solvable model of electronic states in a three-dimensional disordered Hamiltonian: non existing of localized states. — J. Phys., 1969, v. C2, p. 1717.
35. *Abou-Chacra R., Anderson P. W., Thouless D. J.* A self-consistent theory of localization. — J. Phys., 1973, v. C6, p. 1734.
36. *Abou-Chacra R., Thouless D. J.* Self-consistent theory of localization. — J. Phys., 1974, v. C7, p. 65.
37. *Лифшиц И. М.,* О структуре энергетического спектра и квантовых состояниях неупорядоченных конденсированных систем. — УФН, 1964, т. 83, с. 617.
38. *Knor K. E., Smith P. V.* Localization of electron states in two-dimensional disordered potential arrays. — J. Phys., 1971, v. C4, p. 2029.
39. *Kikuchi M.* Numerical studies of localization in structurally disordered system. — J. Phys. Soc. Japan, 1974, v. 37, p. 904.

### К главе 3

1. *Efros A. L., Shklovskii B. I., Yanchev J. Y.* Impurity conductivity in low compensated semiconductors. — Phys. Stat. Sol. (b), 1972, v. 50, p. 45.  
*Шкловский Б. И., Эфрос А. Л., Янчев И. Я.* Энергия активации прыжковой проводимости. — Письма в ЖЭТФ, 1971, т. 14, с. 348.
2. *Mott N. F.* On the transition to metallic conduction in semiconductors. — Can. Journ. Phys., 1956, v. 34, p. 1356.
3. *Miller A., Abrahams E.* Impurity conduction at low concentrations. — Phys. Rev., 1960, v. 120, p. 745.
4. *Ландау Л. Д., Лифшиц Е. М.* Статистическая физика. — М.: Наука, 1964.
5. *Ивченко Е. А., Такунов Л. В.* Поляризация излучения донорно-акцепторных пар при оптическом...

6. *Абакумов В. Н., Перель В. И., Ясиевич И. Н.*, Захват носителей на притягивающие центры при фотовозбуждении. — ЖЭТФ, 1977, т. 72, с. 674.
7. *Шкловский Б. И., Эфрос А. Л.* Примесная зона и проводимость компенсированных полупроводников. — ЖЭТФ, 1971, т. 60, с. 867.
8. *Stern F.* Low temperature limit of screening length in semiconductors. — Phys. Rev., 1974, v. B9, p. 4597.
9. *Арендарчук В. В., Гершензон Е. М., Литвак-Горская Л. Б., Рабинович Р. И.* Резонансное поглощение локализованными на донорных парах электронами в компенсированном  $n$ -InSb. — Письма в ЖЭТФ, 1973, т. 17, с. 237. Особенности энергетических состояний локализованных электронов в сильно компенсированном  $n$ -InSb. — ЖЭТФ, 1973, т. 65, с. 2387.
10. *Кальфа А. А., Коган Ш. М.* Ширина линий примесного оптического поглощения в сильно компенсированных полупроводниках. — ФТП, 1972, т. 6, с. 2175.

## К главе 4

1. *Ансельм А. И.* Введение в теорию полупроводников. — Л.: Физматгиз, 1962.
2. *Блэкмор Дж.* Статистика электронов в полупроводниках. — М.: Мир, 1964.
3. *Gudden B., Schottky W.* Probleme der Ionen und Elektronenleitung in nichtmetallischen festen Körpern. — Z. Techn. Phys., 1935, Bd. 16, s. 323.
4. *Bush G., Labhart H.* Über den Mechanismus der elektrischen Leitfähigkeit des Silicium-Carbids. — Helv. phys. acta, 1946, v. 14, p. 463.
5. *Hung C. S., Gliessman J. R.* Resistivity and Hall effect of germanium at low temperatures. — Phys. Rev., 1954, v. 96, p. 1226.
6. *Fritzsche H., Cuevas M.* Impurity conduction in transmutation-doped  $p$ -type germanium. — Phys. Rev., 1960, v. 119, p. 1238.
7. *Mott N., Twose W. D.* The theory of impurity conduction. — Adv. Phys., 1961, v. 10, p. 707. См. перевод. — УФН, 1963, т. 79, с. 691.
8. *Доброго В. П., Ермолаев О. П.* Прыжковая проводимость германия  $n$ -типа, облученного  $\gamma$ -квантами  $^{60}\text{Co}$ . — ФТП, 1975, т. 9, с. 926.
9. *Davis E. A., Compton W. D.* Compensation dependence of impurity conduction in antimony-doped germanium. — Phys. Rev., 1965, v. 140A, p. 2183.
10. *Yamanouchi C.* Hall coefficient and resistivity in the intermediate impurity conduction of  $n$ -type germanium. — J. Phys. Soc. Japan, 1965, v. 20, p. 1029.
11. *Nishimura H.* Impurity conduction in the intermediate concentration range. — Phys. Rev., 1965, v. 138A, p. 815.
12. *Гершензон Е. М., Гольцман Г. Н., Мельников А. П.* Об энергии связи носителя заряда с нейтральным примесным центром.

14. *Miller A., Abrahams E.* Impurity conduction at low concentrations. — *Phys. Rev.*, 1960, v. 120, p. 745.
15. *Kirkpatrick S.* Percolation and conduction. I. Transport theory of percolation processes. — *Rev. Mod. Phys.*, 1973, v. 45, p. 574.
16. *Ambegaokar V., Halperin B. I., Langer J. S.* Hopping conductivity in disordered systems. — *Phys. Rev.*, 1971, v. B4, p. 2612.
17. *Pollak M.* A percolation treatment of d. c. hopping conduction. — *J. Non-Cryst. Solids*, 1972, v. 11, p. 1.
18. *Шкловский Б. И., Эфрос А. Л.* Примесная зона и проводимость компенсированных полупроводников. — *ЖЭТФ*, 1971, т. 60, с. 867.

### К главе 5

1. *Broadbent S. R., Hammersley J. M.* Percolation processes. I. Crystals and mazes. — *Proc. Camb. Phil. Soc.*, 1957, v. 53, p. 629.
2. *Frisch H. L., Hammersley J. M., Welsh D. J. A.* Monte-Carlo estimates of percolation probabilities for various lattices. — *Phys. Rev.* 1962, v. 126, p. 949.
3. *Frisch H. L., Hammersley J. M.* Percolation processes and related topics. — *J. Soc. Ind. Appl. Math.*, 1963, v. 11, p. 894.
4. *Kirkpatrick S.* The nature of percolation «channels». — *Sol. St. Commun.*, 1973, v. 12, p. 1279.
5. *Shante V. K. S., Kirkpatrick S.* An introduction to percolation theory. — *Adv. Phys.*, 1971, v. 20, p. 325.
6. *Kikuchi R.* Concept of the long-range order in percolation problems. — *J. Chem. Phys.*, 1970, v. 53, p. 2713.
7. *Elliott R. J., Heap B. R., Morgan D. J., Rushbrooke G. S.* Equivalence of the critical concentrations in the Ising and Heisenberg models of ferromagnetism. — *Phys. Rev. Lett.*, 1960, v. 5, p. 366.
8. *Kumar D., Bhargava R.* Variation of critical temperature with concentration for dilute Heisenberg model. — *J. Phys.*, 1975, v. C8, p. L478.
9. *Harris A. B.* Upper bounds for the transition temperatures. — *J. Phys.*, 1974, v. C7, p. 3082.
10. *Kumar D., Harris A. B.* Spin waves and long range order in the dilute Heisenberg ferromagnet at zero temperature. — *Phys. Rev.*, 1973, v. B8, p. 2166.
11. *Shender E. F., Shklovski B. I.* The Curie temperature of dilute ferromagnetic alloys near the percolation threshold. — *Phys. Lett.*, 1975, v. 55A, p. 77.
12. *Шендер Е. Ф.*, Термодинамика неупорядоченных гейзенберговских ферромагнетиков вблизи пороговой концентрации (порога протекания). — *ЖЭТФ*, 1976, т. 70, с. 2251.
13. *Левинштейн М. Е., Шур М. С., Шкловский Б. И., Эфрос А. Л.* О связи между критическими индексами теории протекания. — *ЖЭТФ*, 1975, т. 69, с. 386.
14. *Скал А. С., Шкловский Б. И., Эфрос А. Л.* Энергия активации прыжковой проводимости. — *ФТТ*, 1975, т. 17, с. 506.
15. *Sykes M. F., Essam J. W.* Some exact critical percolation probabilities for bond and site problems in two dimensions. — *Phys. Rev. Lett.*, 1963, v. 10, p. 3.

16. *Sykes M. F., Essam J. W.* Exact critical percolation probabilities for site and bond problems in two dimensions. — *J. Math. Phys.*, 1964, v. 5, p. 1117.
17. *Frisch H. L., Gordon S. B., Vyssotsky V. A., Hammersley J. M.* Monte-Carlo solution of bond percolation processes in various crystal lattices. — *Bell Syst. Tech. J.*, 1962, v. 41, p. 909.
18. *Vyssotsky V. A., Gordon S. B., Frisch H. L., Hammersley J. M.* Critical percolation probabilities (bond problem). — *Phys. Rev.*, 1961, v. 123, p. 1566.
19. *Frisch H. L., Sonneblick E., Vyssotsky V. A., Hammersley J. M.* Critical percolation probabilities (site problem). — *Phys. Rev.*, 1961, v. 124, p. 1021.
20. *Sykes M. F., Essam J. W.* Critical percolation probabilities by series methods. — *Phys. Rev.*, 1964, v. 133A, p. 310.
21. *Dean P.* A new Monte-Carlo method for percolation problems on a lattice. — *Proc. Camb. Phil. Soc.*, 1963, v. 59, p. 397.
22. *Dean P., Bird N. F.* Monte-Carlo estimates of critical percolation probabilities. — *Proc. Camb. Phil. Soc.*, 1967, v. 63, p. 477.
23. *Sykes M. F., Gaunt D. S., Glen M.* Percolation processes in two dimensions, II. Critical concentrations and mean size index. — *J. Phys.*, 1976, v. A9, p. 97.
24. *Neal D. G.* Estimates of critical percolation probabilities for set of two-dimensional lattices. — *Proc. Camb. Phil. Soc.*, 1972, v. 71, p. 97.
25. *Kurkijärvi J., Padmore T. C.* Percolation in two-dimensional lattices. — *J. Phys.*, 1975, v. A8, p. 683.
26. *Шейнман А. В.* Прыжковая проводимость полупроводниковых пленок и критические величины теории протекания. — *ФТП*, 1975, т. 9, с. 2146.
27. *Dunn A. G., Essam J. W., Ritchie D. S.* Series expansion study of the pair connectedness in bond percolation models. — *J. Phys.*, 1975, v. C8, p. 4219.
28. *Kirkpatrick S.* Percolation phenomena in higher dimensions: Approach to the mean-field limit. — *Phys. Rev. Lett.*, 1976, v. 36, p. 69.
29. *Sykes M. F., Gaunt D. S., Essam J. W.* The percolation probability for the site problem on the face-centred lattice. — *J. Phys.*, 1976, v. A9, p. L43.
30. *Sykes M. F., Glen M., Gaunt D. S.* The percolation probability for the site problem on the triangular lattice. — *J. Phys.*, 1974, v. A7, p. L105.
31. *Sykes M. F., Gaunt D. S., Glen M.* Percolation processes in two dimensions. IV. Percolation probability. — *J. Phys.*, 1976, v. A9, p. 725.
32. *Harris A. B., Lubensky T. C., Holcomb W. K., Dasgupta C.* Renormalization-group approach to percolation problems. — *Phys. Rev. Lett.*, 1975, v. 35, p. 327, 1397.
33. *Blease J., Essam J. W., Plöze C. M.* Critical behaviour of the percolation probability on a three dimensional lattice. — *Phys. Lett.*, 1976, v. 57A, p. 199.
34. *Leath P. L.* Cluster size and boundary distribution near percolation threshold. — *Phys. Rev.*, 1976, v. B14, p. 5046.

35. *Kurkijärvi J.* Conductivity in random systems. II. Finite-size-system percolation. — *Phys. Rev.*, 1974, v. B9, p. 770.
36. *Watson B. P., Leath P. L.* Conductivity in the two dimensional problem. — *Phys. Rev.*, 1974, v. B9, p. 4893.
37. *Levinshtein M. E.* Critical indexes of conductivity in two dimensional percolation problems. — *J. Phys.*, 1977, v. C10, p. 1895.
38. *Straley J. P.* Critical exponents for the conductivity of random resistor lattices. — *Phys. Rev.*, 1977, v. B15, p. 5733.
39. *Adler D., Flora L. P., Santuria S. D.* Electrical conductivity in disordered systems. — *Sol. St. Commun.*, 1973, v. 12, p. 9.
40. *Onizuka K.* Computer experiment on a 3D site percolation model of porous materials — its connectivity and conductivity. — *J. Phys. Soc. Japan*, 1975, v. 39, p. 527.
41. *Webman I., Jortner J., Cohen M. H.* Numerical simulation of electrical conductivity in microscopically inhomogenous materials. — *Phys. Rev.*, 1975, v. B11, p. 2885.
42. *Fisher M. E.* Critical probabilities for cluster size and percolation problems. — *J. Math. Phys.*, 1961, v. 2, p. 620.
43. *Fisher M. E., Essam J. W.* Some cluster size and percolation problems. — *J. Math. Phys.*, 1961, v. 2, p. 609.
44. *Stinchcombe R. B.* The branching model for percolation and electrical conductivity. — *J. Phys.*, 1973, v. C6, p. L1.
45. *Stinchcombe R. B.* Conductivity and spinwave stiffness in disordered systems — an exactly soluble model. — *J. Phys.*, 1974, v. C7, p. 179.
46. *Straley J. P.* Critical phenomena in resistor networks. — *J. Phys.*, 1976, v. C9, p. 783.
47. *Essam J. W.* Percolation and cluster size. In: *Phase transition and critical phenomena*./Ed. by C. Domb, M. Green. 1972, v. 2, p. 192.
48. *Domb C.* Fluctuation phenomena and stochastic processes. — *Nature*, 1959, v. 184, p. 509.
49. *Domb C., Sykes M. F.* Cluster size in random mixtures and percolation processes. — *Phys. Rev.*, 1960, v. 122, p. 77.
50. *Фишер М.* Природа критического состояния. — М.: Мир, 1968.
51. *Скал А. С., Шкловский Б. И., Эфрос А. Л.* Вычисление уровня протекания в двухмерном случайном потенциале. — ФТТ, 1973, т. 15, с. 1423.
52. *Скал А. С., Шкловский Б. И., Эфрос А. Л.* Уровень протекания в трехмерном случайном потенциале. — Письма в ЖЭТФ, 1973, т. 17, с. 522.
53. *Скал А. С., Шкловский Б. И.* Топология бесконечного кластера в теории протекания и теория прыжковой проводимости. — ФТП, 1974, т. 8, с. 1586.
54. *Ziman J. M.* The localization of electrons in ordered and disordered systems. — *J. Phys.*, 1968, v. C1, p. 1532.
55. *Scher H., Zallen R.* Critical density in percolation processes. — *J. Chem. Phys.*, 1970, v. 53, p. 3759.
56. *Eggarter T. P.* Semiclassical theory of electron transport properties in a disordered material. — *Phys. Rev.*, 1972, v. A5, p. 2496.
57. *Eggarter T. P., Cohen M. H.* Simple model for density of states and mobility of an electron in a gas of hard-core scatterer. — *Phys. Rev. Lett.*, 1970, v. 25, p. 807.

58. Шкловский Б. И., Эфрос А. Л. Примесная зона и проводимость компенсированных полупроводников. — ЖЭТФ, 1971, т. 60, с. 867.
59. Шкловский Б. И., Эфрос А. Л. Переход от металлической проводимости к активационной в компенсированных полупроводниках. — ЖЭТФ, 1971, т. 61, с. 816.
60. Zallen R., Scher H. Percolation on a continuum and the localization—delocalization transition in amorphous semiconductors. — Phys. Rev., 1971, v. B4, p. 4471.
61. Дыхне А. М. Проводимость двумерной двухфазной системы. — ЖЭТФ, 1970, т. 59, с. 110.
62. Шкловский Б. И. Оптическая и электрическая запрещенные зоны аморфного полупроводника. — Письма в ЖЭТФ, 1971, т. 14, с. 397.
63. Duckers L. J., Ross R. G. Percolation with non-random site occupation. — Phys. Lett., 1974, v. 49A, p. 361.
64. Odagaki T., Ogita N., Matsuda H. Percolation approach to the metal-insulator transition in super-critical fluid metals. — J. Phys. Soc. Japan, 1975, v. 39, p. 618.
65. Müller-Krumbhaar H. Percolation in a lattice system with particle interaction. — Phys. Lett., 1974, v. 50A, p. 27.
66. Pike G. E., Seager C. H. Percolation and conductivity: a computer study, I. — Phys. Rev., 1974, v. B10, p. 1421.
67. Dalton N. W., Domb C., Sykes M. F. Dependence of critical concentration of dilute ferromagnet on the range of interaction. — Proc. Phys. Soc., 1964, v. 83, p. 496.
68. Domb C., Dalton N. W. Crystal statistics with long-range forces, I. The equivalent neighbour model. — Proc. Phys. Soc., 1966, v. 89, p. 859.
69. Gilbert E. N. Random plane networks. — J. S. I. A. M., 1961, v. 9, p. 533.
70. Holcomb D. F., Rehr J. J. Percolation in heavily doped semiconductors. — Phys. Rev., 1969, v. 183, p. 773.
71. Roberts F. D. K. A Monte-Carlo solution of a two-dimensional unstructured cluster problem. — Biometrika, 1967, v. 54, p. 625.
72. Roberts F. D. K., Storey S. H. A three-dimensional cluster problem. — Biometrika, 1968, v. 55, p. 253.
73. Holcomb D. F., Iwasawa M., Roberts F. D. K. Clustering of randomly placed spheres. — Biometrika, 1972, v. 59, p. 207.
74. Ottavi H., Gayda J. P., Percolation in a continuous two dimensional medium. — J. de Physique, 1974, v. 35, p. 631.
75. Gayda J. P., Ottavi H. Cluster of atoms coupled by long range interactions. — J. de Physique, 1974, v. 35, p. 393.
76. Fremlin D. H. The clustering problem: some Monte-Carlo results. — J. de Physique, 1976, v. 37, p. 813.
77. Скал А. С., Шкловский Б. И. Концентрационная зависимость прыжковой проводимости. — ФТП, 1973, т. 7, с. 1589.
78. Fitzpatrick J. P., Malt R. B., Spaepen F. Percolation theory and the conductivity of random close packed mixtures of hard spheres. — Phys. Lett., 1974, v. 47A, p. 207.
79. Kasteleyn P. W., Fortuin C. M. Phase transitions in lattice systems with random local properties. — J. Phys. Soc. Japan Suppl., 1969, v. 26, p. 11.

80. *Essam J. W., Gwilym R. M.* The scaling law for percolation process. — *J. Phys.*, 1971, v. C4, p. L228.
81. *Dunn A. G., Essam J. W., Loveluck J. M.* Scaling theory for the pair-connectedness in percolation models. — *J. Phys.*, 1975, v. C8, p. 743.
82. *Стенли Г.* Фазовые переходы и критические явления. — М.: Мир, 1973.
83. *Gaunt D. S., Sykes M. F.* Percolation processes in two dimensions. V. The exponent  $\delta$  and scaling theory. — *J. Phys.*, 1976, v. A9, p. 1109.
84. *Toulouse G.* Perspectives from the theory of phase transitions. — *Nuovo Cimento*, 1974, v. 23B, p. 234.
85. *Young A. P., Stinchcombe R. B.* A renormalization group theory for percolation problems. — *J. Phys.*, 1975, v. C8, p. L535.
86. *Гинзбург С. Л.* Микроскопическая теория подобия в задаче протекания. — *ЖЭТФ*, 1976, т. 71, с. 1129.
87. *Dasgupta C.* Renormalization group calculation of the crystal exponents for percolation. — *Phys. Rev.*, 1976, v. B14, p. 1221.
88. *Last B. J., Thouless D. J.* Percolation theory and electrical conductivity. — *Phys. Rev. Lett.*, 1971, v. 27, p. 1719.
89. *Kirkpatrick S.* Percolation and conduction. I. Transport theory of percolation processes. — *Rev. Mod. Phys.*, 1973, v. 45, p. 574.
90. *Lightsley P. A.* Percolation view on transport properties of  $\text{Na}_2\text{WO}_3$ . — *Phys. Rev.*, 1973, v. B8, p. 3586.
91. *Kirkpatrick S.* Percolation aspects of metal-insulator transitions in inhomogenous materials. — *Proc. 2nd Int. Conf. Prop. of Liquid Metals.* — Токио, 1972, p. 351.
92. *Abeles B., Pinch H. L., Gittleman J. I.* Percolation conductivity in  $\text{W}-\text{Al}_2\text{O}_3$  granular metal films. — *Phys. Rev. Lett.*, 1975, v. 35, p. 249.
93. *Liang N. T., Shan Y., Wang S.* Electronic conductivity and percolation theory in aggregated films. — *Phys. Rev. Lett.*, 1976, v. 37, p. 526.
94. *De Gennes P. G.* On a relation between percolation theory and the elasticity of gels. — *J. de Physique*, 1976, v. 37, p. L1.
95. *Stauffer D.* Gelation in concentrated critically branched polymer solutions. Percolation scaling theory of intramolecular bond cycles. — *J. Chem. Soc. Faraday Transactions II*, 1976, v. 72, p. 1354.
96. *Harris A. B., Kirkpatrick S.* Low frequency response functions of random magnetic systems. — *Phys. Rev.*, 1977, v. B16, p. 542.
97. *Lubensky T. C.* Scaling theory of phase transitions in diluted systems near the percolation threshold. — *Phys. Rev.*, 1977, v. B15, p. 311.
98. *Левинштейн М. Е., Шур М. С., Эфрос А. Л.* Гальваномагнитные явления в неупорядоченных системах. Теория и моделирование. — *ЖЭТФ*, 1975, т. 69, с. 2203.
99. *Watson B. P., Stinchcombe R. B.* Renormalization group approach for percolation conductivity. — *J. Phys.*, 1976, v. C9, p. 3221.
100. *Kirkpatrick S.* Classical transport in disordered media: scaling and effective-medium theories. — *Phys. Rev. Lett.*, 1971, v. 27, p. 1722.



100. *Kirkpatrick S.* Percolation and conduction. — Technical report of ISSP, ser. B, 1973, No. 15, p. 18.
101. *Seager C. H., Pike G. E.* Percolation and conductivity: a computer study II. — *Phys. Rev.*, 1974, v. B10, p. 1435.
102. *Шкловский Б. И., Эфрос А. Л.* Предэкспоненциальный множитель прыжковой проводимости. — Письма в ЖТФ, 1975, т. 1, с. 174.

## К главе 6

1. *Емельяненко О. В., Лагунова Т. С., Наследов* и др. Проводимость по примесям в *n-GaAs*. — ФТП, 1973, т. 7, с. 1919.
2. *Asai Sh., Toyabe T., Hirao M.* Impurity conduction and non-ohmic properties of high purity *n*-type gallium arsenide. — *Proc. X Int. Conf. Phys. of Semicond.* — Boston, 1970, p. 578.
3. *Oliver D.* Electrical properties of *n*-type gallium arsenide. — *Phys. Rev.*, 1962, v. 127, p. 1045.
4. *Eddols D. V.* Electrical properties of *n*-type epitaxial gallium arsenide. — *Phys. Stat. Sol.*, 1966, v. 17, p. 67.
5. *Kahlert H., Geiger J., Landwehr G., Schlachetzki A., Salow H.* Hopping conduction and non-ohmic properties of lightly doped *n*-type gallium arsenide. — *Proc. XII Int. Conf. Phys. of Semicond.* — Stuttgart, 1974, p. 269.
6. *Lemoine D., Pelletier C., Rolland S., Granger R.* Hopping conduction in epitaxial *n-GaAs* layers. — *Phys. Lett.*, 1976, v. 56A, p. 497.
7. *Емельяненко О. В., Масагутов К. Г., Наследов Д. Н., Тимченко И. Н.*, Прыжковая проводимость по примесям в *n-InP*. — ФТП, 1975, т. 9, с. 503.
8. *Fritzsche H.* Electrical properties of germanium semiconductors at low temperatures. — *Phys. Rev.*, 1955, v. 99, p. 406.
9. *Rollin B. V., Russel J. P.* Impurity conduction in copper doped germanium. — *Proc. Phys. Soc.*, 1963, v. 81, p. 571.
10. *Ray R. K., Fan H. Y.* Impurity conduction in silicon. — *Phys. Rev.*, 1961, v. 121, p. 768.
11. *Скал А. С., Шкловский Б. И.* Концентрационная зависимость прыжковой проводимости. — ФТП, 1973, т. 7, с. 1589.
12. *Шкловский Б. И., Нгуен Ван Лиен*, Прыжковое магнитосопротивление *n*-германия. — ФТП, 1978, т. 12, с. 1346.
13. *Fritzsche H.* Resistivity and Hall coefficient of antimony-doped germanium at low temperatures. — *J. Phys. Chem. Sol.*, 1958, v. 6, p. 69.
14. *Шкловский Б. И., Шлимак И. С.* Прыжковая проводимость германия. — ФТП, 1972, т. 6, с. 129.
15. *Fritzsche H.* Effect of stress on the donor wave functions in germanium. — *Phys. Rev.*, 1962, v. 125, p. 1560.
16. *Fritzsche H.* Effect of uniaxial compression on impurity conduction in *n*-type germanium, *Phys. Rev.*, 1962, v. 125, p. 1552.
17. *Шкловский Б. И.* Анизотропия прыжковой проводимости. — ФТП, 1977, т. 11, с. 2142.
18. *Бир Г. Л., Пикус Г. Е.* Симметрия и деформационные эффекты в полупроводниках. — М.: Наука, 1972, § 30.
19. *Pollak F. H.* Effect of uniaxial compression on impurity conduction in *p*-germanium. — *Phys. Rev.*, 1965, v. 138A, p. 619.

## К главе 7

1. *Sladek R. J.* Magnetically induced impurity banding. — *J. Phys. Chem. Sol.*, 1958, v. 5, p. 157.
2. *Гершензон Е. М., Ильин В. А., Литвак-Горская Л. Б.* Влияние магнитного поля на прыжковую проводимость в *n*-InSb. — ФТП, 1974, т. 8, с. 295.
3. *Гершензон Е. М., Куриленко И. Н., Литвак-Горская Л. Б.* Проводимость компенсированного *p*-InSb в области промежуточной концентрации примесей. — ФТП, 1974, т. 8, с. 1057.
4. *Chroboczek J. A., Sladek R. J.* Magnetoresistance of *p*-type germanium in the phonon-assisted hopping conduction range at high magnetic fields. — *Phys. Rev.*, 1966, v. 151, p. 595.
5. *Lee W. W., Sladek R. J.* Low temperature magnetoresistance of *n*-type germanium doped with small and intermediate concentrations of phosphorus. — *Phys. Rev.*, 1967, v. 158, p. 788.
6. *Lee W. W., Sladek R. J.* Hop-conduction magnetoresistance in *p*-type germanium. — *Phys. Rev.*, 1967, v. 158, p. 794.
7. *Гаджиев А. Р., Шлимак И. С.* Влияние магнитного поля на прыжковую проводимость *p*-германия. — ФТП, 1972, т. 6, с. 1582.
8. *Chroboczek J., Klokoski A., Koralko K.* The d-c. magnetoresistance of *p*-type germanium in the hop conduction range. — *J. Phys.*, 1974, v. C7, p. 3042.
9. *Емельяненко О. В., Наследов Д. Н., Недеогло Д. Д., Сиукаев Н. В.* Эффект Холла и магнетосопротивление в чистых кристаллах *n*-InP при низких температурах. — ФТП, 1971, т. 5, с. 334.
10. *Емельяненко О. В., Масагутов К. Г., Наследов Д. Н., Тимченко И. Н.* Прыжковая проводимость по примесям в *n*-InP. — ФТП, 1975, т. 9, с. 503.
11. *Halbo L., Sladek R. J.* Magnetoresistance of undoped *n*-type GaAs at low temperatures. — *Phys. Rev.*, 1968, v. 173, p. 794.
12. *Емельяненко О. В., Лагунова Т. С., Наследов Д. Н., Недеогло Д. Д., Тимченко И. Н.* Проводимость по примесям в *n*-GaAs. — ФТП, 1973, т. 7, с. 1919.
13. *Kahlert H., Landwehr G., Schlachetzki A., Salow H.* Impurity conduction and magneto-resistance in lightly doped *n*-type GaAs. — *Z. Physik*, 1976, v. 276B, p. 1.
14. *Lemoine D., Pelletier C., Rolland S., Ganger G.* Hopping conduction in epitaxial *n*-GaAs layer. — *Phys. Lett.*, 1976, v. 56A, p. 493.
15. *Mikoshiba N.* Strong-field magnetoresistance of impurity conduction in *n*-type germanium. — *Phys. Rev.* 1962, v. 127, p. 1961.
16. *Шкловский Б. И.* Прыжковая проводимость в сильном магнитном поле. — ЖЭТФ, 1971, т. 61, с. 2033.
17. *Шкловский Б. И.* К теории экспоненциального магнетосопротивления полупроводников — ФТП, 1973, т. 8, с. 416.
18. *Elliott R. J., Loudon R. J.* Theory of the absorption edge in semiconductors in a high magnetic field. — *J. Phys. Chem. Sol.*, 1960, v. 15, p. 196.
19. *Hasegawa H., Howard R. E.* Optical absorption spectrum of hydrogenic atom in strong magnetic field. — *J. Phys. Chem. Sol.*, 1961, v. 21, p. 179.

20. *Yafet Y., Keyes R. W., Adams E. N.* Hydrogen atom in a strong magnetic field. — *J. Phys. Chem. Sol.*, 1956, v. 1, p. 137.
21. *Holstein T.* Hall effect in impurity conduction. — *Phys. Rev.*, 1961, v. 124, p. 1329.
22. *Шкловский Б. И.* Прыжковая проводимость слабо легированных полупроводников. — *ФТП*, 1972, т. 6, с. 1197.
23. *Chroboczek J., Fritzsche H., Jiang C. L., Pollak M.* Piezomagneto-resistance studies of *p*-type germanium in the hopping conduction region. — *Proc. XIII Int. Conf. Phys. Sem. — Rome*, 1976, p. 1113.
24. *Larsen D. M.*, Shallow donor levels of InSb in a magnetic field. — *J. Phys. Chem. Soc.*, 1968, v. 29, p. 271.
25. *Гершензон Е. М., Ильин В. А., Куриленко И. Н., Литвак-Горская Л. Б.* Примесная проводимость в *n*-InSb. — *ФТП*, 1972, т. 6, с. 1687.
26. *Ferre D., Dubois H., Biskupski G.* Structure of the impurity band and magnetic-field-induced metal — non-metal transition in *n*-type InSb. — *Phys. Stat. Sol. (b)*, 1975, v. 70, p. 81.
27. *Walton A. K., Dutt J. C.* Freeze-out and hopping conduction in InSb subjected to intense magnetic fields. — *J. Phys.*, 1977, v. C10, p. L29.
28. *Клинггер М. И., Воронюк П. И.* Гальваномагнитные явления при сильных магнитных полях в полупроводниках. — *ЖЭТФ*, 1957, т. 33, с. 77.

## К главе 8

1. *Шкловский Б. И., Шлимак И. С.* Прыжковая проводимость германия. — *ФТП*, 1972, т. 6, с. 129.
2. *Fritzsche H.* Electrical properties of germanium semiconductors at low temperatures. — *Phys. Rev.*, 1955, v. 99, p. 406.
3. *Fritzsche H.* Resistivity and Hall coefficient of antimony doped germanium at low temperatures. — *J. Phys. Chem. Sol.*, 1958, v. 6, p. 69.
4. *Chroboczek J. A., Prohofsky E. W., Sladek R. J.* Magnetically induced spin-reversal transitions in impurity hop conduction in *n*-type germanium. — *Phys. Rev.*, 1968, v. 169, p. 593.
5. *Scher H., Lax M.* Stochastic transport in a disordered solid. II. Impurity conduction. — *Phys. Rev.*, 1973, v. B7, p. 4502.
6. *Pollak M., Geballe T. H.* Low-frequency conductivity due to hopping processes in silicon. — *Phys. Rev.*, 1961, v. 122, p. 1742.
7. *Pollak M., Knotek M. L.* Evidence for correlated hopping in impurity concentrations. — *Sol. St. Commun.*, 1977, v. 21, p. 183.
8. *Knotek M. L.* High field magnetoresistance of hopping transport in the disordered impurity system of transmutation doped Ge. — *Phys. Rev.*, 1977, v. B16, p. 2629.
9. *Шкловский Б. И., Янчев И. Я.* Определение степени компенсации полупроводников по эффекту насыщения прыжковой проводимости. — *ФТП*, 1972, т. 6, с. 1616.
10. *Shlimak I. S., Emtzev V. V.* Dependence of the activation energy of conductivity on the compensation degree in germanium. — *Phys. Stat. Sol. (b)*, 1971, v. 47, p. 325.

11. Cuevas M. Mobility of electrons in compensated semiconductors. I. Experiment. — Phys. Rev., 1967, v. 164, p. 1021.
12. Гершензон Е. М., Ивлева В. С., Куриленко И. Н., Литвак-Горская Л. Б. Влияние компенсации на электропроводность  $p$ -InSb. — ФТП, 1972, т. 6, с. 1982.
13. Гельмонт Б. Л., Гаджиев А. Р., Шкловский Б. И., Шлимак И. С., Эфрос А. Л. Прыжковая проводимость твердых растворов германия с кремнием. — ФТП, 1974, т. 8, с. 2377.
14. Шлимак И. С., Эфрос А. Л., Янчев И. Я. Исследование роли флуктуаций состава в твердом растворе Ge—Si. — ФТП, 1977, т. 11, с. 257.

## К главе 9

1. Mott N. F. Conduction in glasses containing transition metal ions. — J. Non-Cryst. Solids, 1968, v. 1, p. 1.
2. Ambegaokar V., Halperin B. I., Langer J. S. Hopping conductivity in disordered systems. — Phys. Rev., 1971, v. B4, p. 2612.
3. Pollak M. A percolation treatment of d.c. hopping conduction. — J. Non-Cryst. Solids, 1972, v. 11, p. 1.
4. Seager C. H., Pike G. E. Percolation and conductivity: A computer study. II. — Phys. Rev., 1974, v. B10, p. 1435.
5. Jones R., Schaich W. Theory of impurity band hopping conduction. — J. Phys., 1972, v. C5, p. 43.
6. Скал А. С., Шкловский Б. И. О формуле Мотта для низкотемпературной прыжковой проводимости. — ФТП, 1976, т. 16, с. 1820.
7. Brenig W., Döhler G. H., Heyszenau H. Hopping conductivity in highly anisotropic systems. — Phil. Mag., 1973, v. 27, p. 1093.
8. Kurkijärvi J. Hopping conductivity in one dimension. — Phys. Rev., 1973, v. B8, p. 922.
9. Shante V. K. S., Varma C. M., Bloch A. N. Hopping conductivity in «one-dimensional» disordered compounds. — Phys. Rev., 1973, v. B8, p. 4885.
10. Hamilton E. M. Variable range hopping in a non-uniform density of states. — Phil. Mag., 1972, v. 26, p. 1043.
11. Hill R. M. Variable range hopping. — Phys. Stat. Sol. (a), 1976, v. 34, p. 601.
12. Hill R. M. On the observation of variable range hopping. — Phys. Stat. Sol. (a), 1976, v. 35, p. K29.
13. Забродский А. Г., Прыжковая проводимость и ход плотности локализованных состояний в окрестности уровня Ферми. — ФТП, 1977, т. 11, с. 595.
14. Шлимак И. С., Никулин Е. И. Проводимость легированного германия при сверхнизких температурах. — Письма в ЖЭТФ, 1972, т. 15, с. 30.
15. Sadasiv G. Magnetoresistance in germanium in the impurity conduction range. — Phys. Rev., 1962, v. 128, p. 1131.
16. Allen F. R., Adkins C. J. Electrical conduction in heavily doped germanium. — Phil. Mag., 1972, v. 26, p. 1027.
17. Забродский А. Г., Ионов А. Н., Корчажкина Р. Л., Шлимак И. С. Проводимость сильно легированного и компенсированного германия. — ФТП, 1973, т. 7, с. 1914.

18. Емельяненко О. В., Наследов Д. Н., Никулин Е. И., Тимченко И. Н. Проводимость по примесям в  $n$ -GaAs при сверхнизких температурах. — ФТП, 1972, т. 6, с. 2283.
19. Redfield D. Observation of  $\log \sigma \propto T^{-1/2}$  in three-dimensional energy-band tails. — Phys. Rev. Lett., 1973, v. 30, p. 1319.
20. Redfield D. Transport of electrons in energy band tails. — Adv. Phys. 1975, v. 24, p. 463.
21. Suprapto B. B., Butcher P. N. DC hopping conduction by magnetically frozen electrons. — J. Phys., 1975, v. C8, p. L517.
22. Гершензон Е. М., Ильин В. А., Литвак-Горская Л. Б. Влияние магнитного поля на прыжковую проводимость в  $n$ -InSb. — ФТП, 1974, т. 8, с. 295.
23. Pollak M., Hauser J. J. Note on the anisotropy of the conductivity in thin amorphous films. — Phys. Rev. Lett., 1973, v. 31, p. 1304.
24. Hauser J. J., Staudinger A. Electrical and structural properties of amorphous germanium. — Phys. Rev., 1973, v. B8, p. 607.
25. Hauser J. J. Electrical properties and anisotropy in amorphous Si and  $\text{Si}_{0.5}\text{Ge}_{0.5}$  alloy. — Phys. Rev., 1973, v. B8, p. 3817.
26. Hauser J. J. Electrical structural and optical properties of amorphous carbon. — J. Non-Cryst. Solids, 1977, v. 23, p. 21.
27. Hauser J. J. Hopping conductivity in amorphous carbon films. — Sol. St. Commun., 1975, v. 17, p. 1577.
28. Knotek M. L., Pollak M., Donovan T. M., Kurtzman H. Thickness dependence of hopping transport in amorphous — Ge films. — Phys. Rev. Lett., 1973, v. 30, p. 853.
29. Brenig W., Döhler G. H., Heyszenau H.  $T^{-1/2}$ -hopping in organic charge transfer materials. — Phys. Lett., 1972, v. 39A, p. 175.
30. Shante V. K. S. Variable range hopping conduction in thin films. — Phys. Lett., 1973, v. 43A, p. 249.
31. Knotek M. L. Temperature and thickness dependence of low temperature transport in amorphous silicon thin films: a comparison to amorphous germanium. — Sol. St. Commun., 1975, v. 17, p. 1431.
32. Shklovskii B. I. Hopping conduction in semiconductor films. — Phys. Lett., 1975, v. 51A, p. 289.
33. Шейнман А. В. Прыжковая проводимость полупроводниковых пленок и критические величины теории протекания. — ФТП, 1975, т. 9, с. 2146.
34. Shklovskii B. I. Thickness dependence of hopping conduction in amorphous films. Experiment versus theory. — Phys. Stat. Sol. (b), 1977, v. 83, K11.
35. Efros A. L., Shklovskii B. I. Hopping conduction and critical indexes of the percolation theory. — Труды 6-й Межд. конф. по аморфным и жидким полупроводникам. — Л.: Наука, 1976, с. 69.
36. Kirkpatrick S. Hopping conduction: Experiment versus theory. — Proc. 5th Int. Conf. Amorphous and Liquid Semiconductors. — London: Taylor and Fransis, 1974, p. 183.
37. Butcher P. N. The exponent and prefactor in the d.c. hopping conductivity formula. — Труды 6-й Межд. конф. по аморфным и жидким полупроводникам. — Л.: Наука, 1976, с. 89.

## К главе 10

1. *Pollak M.* Effect of carrier-carrier interactions on some transport properties in disordered semiconductors. — *Disc. Faraday Soc.*, 1970, v. 50, p. 13.
2. *Knotek M. L., Pollak M.* Correlation effects in hopping conduction: Hopping as multielectron transition. — *J. Non-Cryst. Solids*, 1972, v. 8—10, p. 505; Correlation effects in hopping conduction: A treatment in terms of multielectron transitions. — *Phys. Rev.*, 1974, v. B9, p. 664.
3. *Srinivasan G.* Statistical mechanics of charged traps in an amorphous semiconductor. — *Phys. Rev.*, 1971, v. B4, p. 2581.
4. *Kurosawa T., Sugimoto H.* Effects of coulomb correlation on hopping conduction in disordered systems. — *Prog. Theor. Phys. Suppl.*, 1975, No. 57, p. 217.
5. *Efros A. L., Shklovskii B. I.* Coulomb gap and low-temperature conductivity of disordered systems. — *J. Phys.*, 1975, v. C8, p. L49.
6. *Гельмонт Б. Л., Эфрос А. Л.* Кулоновская щель в неупорядоченных системах. Численный эксперимент в одномерной модели. — *Письма в ЖЭТФ*, 1977, т. 25, с. 77.
7. *Baranovskii S. D., Efros A. L., Gelmont B. L., Shklovskii B. I.* Coulomb gap in disordered systems. Computer simulation. — *Solid St. Commun.*, 1978, v. 27, p. 1.
8. *Efros A. L.* Coulomb gap in disordered systems. — *J. Phys.*, 1976, v. C9, p. 2021.
9. *Mott N. F.* The effect of electron interaction on variable-range hopping. — *Phil. Mag.*, 1976, v. 34, p. 643.
10. *Wolf E. L.* Nonsuperconducting electron tunneling spectroscopy. — *Solid. St. Phys.*, 1975, v. 30, p. 1.
11. *Дубров В. Е., Левинштейн М. Е., Шур М. С.* Аномалия диэлектрической проницаемости при переходе металл — диэлектрик. — *ЖЭТФ*, 1976, т. 70, с. 2014.
12. *Efros A. L., Shklovskii B. I.* Critical behaviour of conductivity and dielectric constant near the metal — non-metal transition threshold. — *Phys. Stat. Sol. (b)*, 1976, v. 76, p. 475.
13. *Castner T. G., Lee N. K., Cieloszyk G. S., Sallinger G. L.* Dielectric anomaly and the metal-insulator transition in *n*-type silicon. — *Phys. Rev. Lett.*, 1975, v. 34, p. 1627.
14. *R. Kikuchi.* Cation diffusion and conductivity in solid electrolytes. In: *Fast ion transport in solids. Solid state batteries and devices.* — Amsterdam: North Holland, 1973, p. 248.
15. *Richards P. M.* Theory of one-dimensional hopping conductivity and diffusion. — *Phys. Rev.*, 1977, v. B16, p. 1393.

## К главе 11

1. *Фистуль В. И.* Сильно легированные полупроводники. — М.: Наука, 1967.
2. *Кривоглаз М. А.* Флуктуонные состояния электронов. — *УФН*, 1973, т. 111, с. 617.
3. *Займан Дж.* Принципы теории твердого тела. — М.: Мир, 1966.
4. *Бонч-Бруевич В. Л.* К теории сильно легированных полупроводников. — *ФТТ*, 1962, т. 4, с. 2660.

5. *Бонч-Бруевич В. Л.* Вопросы электронной теории сильно легированных полупроводников. — В сб.: Физика твердого тела. — М.: ВИНТИ, 1965, с. 174.
6. *Kane E. O., Thomas* — Fermi approach to impure semiconductor band structure. — *Phys. Rev.*, 1963, v. 131, p. 79.
7. *Келдыш Л. В., Прошко Г. П.* Инфракрасное поглощение в сильно легированном германии. — ФТТ, 1963, т. 5, с. 3378.
8. *Багаев В. С., Березашвили Ю. Н., Вул Б. М., Заварицкая Э. И., Келдыш Л. В., Шотов А. П.* Об энергетическом спектре сильно легированного арсенида галлия. — ФТТ, 1964, т. 6, с. 1399.
9. *Dyakonov M. I., Efros A. L., Mitchell D. L.* Magnetic freeze-out of electrons in extrinsic semiconductors. — *Phys. Rev.*, 1969, v. 180, p. 813.
10. *Абрикосов А. А., Горьков Л. П., Дзялошинский И. Е.* Методы квантовой теории поля в статистической физике. — М.: Физматгиз, 1962, с. 255.
11. *Цицишвили Е. Г.* Сильно легированный полупроводник в магнитном поле. — ФТТ, 1966, т. 8, с. 1193.
12. *Эфрос А. Л.* Теория электронных состояний в сильно легированных полупроводниках. — ЖЭТФ, 1970, т. 59, с. 880.
13. *Edwards S. F., Gulyaev Y. V.* The density of states of a highly impure semiconductor. — *Proc. Phys. Soc.*, 1964, v. 83, p. 495.
14. *Андреев И. В.* Электрон в случайном поле. — ЖЭТФ, 1965, т. 48, с. 1437.
15. *Bezak V.* Path-integral theory of electron gas in a random potential. — *Proc. Roy. Soc. Lond.*, 1970, v. A315, p. 339.
16. *Эфрос А. Л.* Электронные свойства сильно легированных полупроводников. Докт. дисс. — Л.: ФТИ, 1971.
17. *Келдыш Л. В.* Полупроводники в сильном электрическом поле. Докт. дисс. — М.: ФИАН, 1965.
18. *Лифшиц Е. М., Питаевский Л. П.* Релятивистская квантовая теория, ч. 2. — М.: Наука, 1971, с. 35.

## К главе 12

1. *Halperin B. I., Lax M.* Impurity-band tails in the high-density limit. I. — *Phys. Rev.*, 1966, v. 148, p. 722.
2. *Halperin B. I., Lax M.* — Impurity-band tails in the high-density limit. II. — *Phys. Rev.*, 1967, v. 153, p. 802.
3. *Zittartz J., Langer J. S.* Theory of bound states in a random potential. — *Phys. Rev.*, 1966, v. 148, p. 741.
4. *Лифшиц И. М.* Теория флуктуационных уровней в неупорядоченных системах. — ЖЭТФ, 1967, т. 53, с. 743.
5. *Лифшиц И. М., Гредескул С. А., Пастур Л. А.* Флуктуационные уровни в неупорядоченных системах. — ФНТ, 1976, т. 2, с. 1093.
6. *Edwards S. F.* The localization of electrons in disordered systems. — *J. Non-Cryst. Solids*, 1970, v. 4, p. 417.
7. *Шкловский Б. И., Эфрос А. Л.* Хвосты плотности состояний в сильно легированных полупроводниках. — ФТП, 1970, т. 4, с. 305.
8. *Шкловский Б. И., Эфрос А. Л.* Глубокие хвосты плотности состояний и поглощение света в полупроводниках. — ЖЭТФ, 1970, т. 58, с. 657.

9. *Рогачев А. А., Саблина Н. И.* Рекомбинационное излучение сильно легированного германия. — ФТТ, 1966, т. 8, с. 866.
10. *Koipov Z. G., Yanchev I. Y.* Density-of-states tail in heavily doped closely compensated semiconductors. — J. Phys., 1978, v. C11, p. L253.
11. *Бете Г., Солпитер Э.* Квантовая механика атомов с одним и двумя электронами. — М.: Физматгиз, 1960, с. 163.
12. *Hwang C. J.* Optical properties on *n*-type GaAs. — J. Appl. Phys., 1969, v. 40, p. 3731.
13. *Шкловский Б. И., Эфрос А. Л.* Межзонное поглощение света в сильно легированных полупроводниках. — ЖЭТФ, 1970, т. 59, с. 1343.
14. *Меркулов И. А., Перель В. И.* Электропоглощение в полупроводниках с крупномасштабными флуктуациями концентрации примесей. — ФТП, 1973, т. 7, с. 1197.
15. *Панков Ж.* Оптические процессы в полупроводниках. — М.: Мир, 1973, с. 53.
16. *Эфрос А. Л.* Плотность состояний и межзонное поглощение света в полупроводниках. — УФН, 1973, т. 111, с. 451.
17. *Алферов Ж. И., Андреев В. М., Гарбузов Д. З., Трукан М. К.* Излучательная рекомбинация в эпитаксиальном компенсированном арсениде галлия. — ФТП, 1972, т. 6, с. 2015.
18. *Dow J. D., Redfield D.* Electroabsorption in semiconductors. — Phys. Rev., 1970, v. B1, p. 3358.  
*Dow J. D., Redfield D.* Toward a unified theory of Urbach's rule and exponential absorption edges. — Phys. Rev., 1972, v. B5, p. 594.
19. *Merkulov I. A., Perel V. I.* Effects of electron-hole interaction on electroabsorption in semiconductors. — Phys. Lett., 1973, v. 45A, p. 83.  
*Меркулов И. А., Перель В. И.* Влияние экситонного эффекта на электропоглощение в полупроводниках. — ЖЭТФ, 1974, т. 66, с. 2314.

## К главе 13

1. *Гаджиев А. Р., Рывкин С. М., Шлимак И. С.* Компенсированный быстрыми нейтронами *n*-Ge как модель аморфного полупроводника. — Письма в ЖЭТФ, 1972, т. 15, с. 605.
2. *Шкловский Б. И., Эфрос А. Л.* Переход от металлической проводимости к активационной в компенсированных полупроводниках. — ЖЭТФ, 1971, т. 61, с. 816.
3. *Шкловский Б. И., Эфрос А. Л.* Полностью компенсированный кристаллический полупроводник как модель аморфного полупроводника. — ЖЭТФ, 1972, т. 62, с. 1156.
4. *Гуляев Ю. В., Плесский В. П.* Об электронных свойствах невырожденных сильно легированных компенсированных полупроводников. — ЖЭТФ, 1976, т. 71, с. 1475.
5. *Гальперн Ю. С., Эфрос А. Л.* Электронные свойства компенсированных полупроводников с коррелированным распределением примесей. — ФТП, 1972, т. 6, с. 1081.
6. *Шкловский Б. И.* Прыжковая проводимость сильно легированных полупроводников. — ФТП, 1973, т. 7, с. 112.



7. *Sasaki W., Yamanouchi C.* Quantitative study of the effect of compensation on impurity conduction in heavily doped *n*-type germanium. — *J. Non-Cryst. Solids*, 1970, v. 4, p. 183.
8. *Redfield D., Crandall R.* Energy dependence of conduction in band tails. — *Proc. X Int. Conf. on the Physics of Semiconductors.* — Cambridge, 1970, p. 574.
9. *Агринская Н. В., Крымова Е. Д.* Хвосты плотности состояний в компенсированных кристаллах CdTe. — *ФТП*, 1972, т. 6, с. 1783.
10. *Didkovskii A. P., Khivrich V. I.* Electrical properties of semiconducting CdTe. — *Phys. Stat. Sol. (a)*, 1975, v. 32, p. 621.
11. *Гершензон Е. М., Ильин В. А., Куриленко И. Н., Литвак-Горская Л. Б.* Электропроводность сильно легированного и компенсированного *n*-InSb. — *ФТП*, 1975, т. 9, с. 1324.  
*Банная В. Ф., Гершензон Е. М., Литвак-Горская Л. Б.* Электропроводность *n*-InSb при температуре 0,3—20 °K. — *ФТП*, 1968, т. 2, с. 978.
12. *Яременко Н. Г.* Проводимость сильно компенсированного *n*-InSb при низких температурах. — *ФТП*, 1975, т. 9, с. 840.
13. *Шкловский Б. И., Эфрос А. Л.* Локализация электронов в магнитном поле. — *ЖЭТФ*, 1973, т. 64, с. 2222.
14. *Эфрос А. Л., Янчев И. Я.* Магнитное вымораживание электронов в компенсированных полупроводниках. — *ФТП*, 1973, т. 7, с. 2099.
15. *Horring N. J.* Quantum theory of static shielding of an impurity charge by an electron gas plasma in a magnetic field. — *Ann. Phys.*, 1969, v. 54, p. 405.
16. *Выставкин А. Н., Губанков В. Н., Листвин В. Н., Мигулин В. В.* Исследование влияния примесей и магнитного поля на детектирующие свойства *n*-InSb в миллиметровом и субмиллиметровом диапазонах. — *ФТП*, 1967, т. 1, с. 844.
17. *Гальперн Ю. С., Эфрос А. Л.* Внутризонное поглощение электромагнитных волн в легированных полупроводниках. — *ФТП*, 1969, т. 11, с. 2301.
18. *Потапов В. Т., Трифонов В. И., Чусов И. И., Яременко Н. Г.* Эффект Фарадея в сильно компенсированном *n*-InSb. — *ФТП*, 1972, т. 6, с. 1227.
19. *Гуляев Ю. В., Листвин В. Н., Потапов В. Т., Чусов И. И., Яременко Н. Г.* О природе внутризонной фотопроводимости в компенсированном *n*-InSb. — *ФТП*, 1975, т. 9, с. 1471.
20. *Шкловский Б. И., Шур М. С., Эфрос А. Л.* S-образная вольт-амперная характеристика в компенсированном полупроводнике. — *ФТП*, 1971, т. 5, с. 1938.
21. *Ryvkín S. M., Shlimak J. S.* A doped highly compensated crystalline semiconductor as a model of amorphous semiconductors. — *Phys. Stat. Sol. (a)*, 1973, v. 16, p. 515.
22. *Ткач Ю. Я.* Фотопроводимость аморфного полупроводника в модели «искривленных» зон. — *ФТП*, 1975, т. 9, p. 1071.
23. *Шик А. Я.* Фотопроводимость случайно-неоднородных полупроводников. — *ЖЭТФ*, 1975, т. 68, с. 1859; *ФТП*, 1975, т. 9, с. 2129.
24. *Леванюк А. П., Осипов В. В.* — *ФТП*, 1973, т. 7, с. 1069.
25. *Fritzsche H.* — *J. Non-Cryst. Solids*, 1971, v. 6, p. 49.

## ОГЛАВЛЕНИЕ

Предисловие . . . . .	3
<b>ЧАСТЬ I. СЛАБО ЛЕГИРОВАННЫЕ ПОЛУПРОВОДНИКИ</b>	
<b>Глава 1. Структура одиночных примесных состояний . . . . .</b>	<b>7</b>
§ 1. Мелкие примеси . . . . .	7
§ 2. Примесные уровни вблизи невырожденной зоны . . . . .	10
§ 3. Примесные уровни вблизи точки вырождения зон . . . . .	20
§ 4. Асимптотика волновых функций примесных уровней . . . . .	30
<b>Глава 2. Локализация электронных состояний . . . . .</b>	<b>38</b>
§ 5. Узкие зоны и переход Мотта . . . . .	39
§ 6. Переход Андерсона . . . . .	45
§ 7. Примеры переходов Андерсона. Электропроводность и волновые функции вблизи перехода . . . . .	52
§ 8. Некоторые аспекты теории перехода Андерсона . . . . .	58
§ 9. Локализация в модели Лифшица . . . . .	61
<b>Глава 3. Структура примесной зоны слабо легированных полупроводников . . . . .</b>	<b>69</b>
§ 10. Общие замечания . . . . .	69
§ 11. Примесная зона в случае малой степени компенсации . . . . .	74
§ 12. Крупномасштабный потенциал при малой степени компенсации . . . . .	83
§ 13. Примесная зона при большой степени компенсации . . . . .	88
<b>Глава 4. Общее представление о прыжковой электропроводности слабо легированных полупроводников . . . . .</b>	<b>100</b>
§ 14. Основные экспериментальные факты . . . . .	100
§ 15. Модель сетки сопротивлений Миллера и Абрахамса . . . . .	111
<b>Глава 5. Теория протекания . . . . .</b>	<b>126</b>
§ 16. Решеточные задачи . . . . .	126
§ 17. Континуальные задачи . . . . .	145
§ 18. Задачи на случайных узлах . . . . .	150
§ 19. Теория критических индексов . . . . .	159
§ 20. Электропроводность случайных сетей из проводящих и непроводящих элементов и топология бесконечного кластера . . . . .	166
§ 21. Теория протекания и электропроводность сильно неоднородных сред . . . . .	174

Глава 6. Зависимость прыжковой проводимости от концентрации примесей и деформации кристалла . . . . .	184
§ 22. Удельное сопротивление $\rho_z$ для полупроводников с изотропными волновыми функциями примесных состояний	184
§ 23. Удельное сопротивление $\rho_z$ для полупроводников с анизотропными примесными состояниями . . . . .	193
Глава 7. Прыжковая проводимость в магнитном поле . . . . .	207
§ 24. Сопротивление элемента $R_{ij}$ в магнитном поле . . . . .	207
§ 25. Вычисление магнитосопротивления и обсуждение экспериментальных данных . . . . .	218
Глава 8. Энергия активации прыжковой проводимости . . . . .	237
§ 26. Энергия активации $\epsilon_3$ при слабой компенсации . . . . .	237
§ 27. Энергии активации $\epsilon_1$ и $\epsilon_3$ при сильной компенсации	245
§ 28. Метод малых возмущений в теории протекания и общая теория энергии активации $\epsilon_3$ . . . . .	252
Глава 9. Прыжковая проводимость с переменной длиной прыжка . . . . .	263
§ 29. Закон Мотта . . . . .	263
§ 30. Магнитосопротивление в области действия закона Мотта . . . . .	273
§ 31. Зависимость прыжковой проводимости аморфных пленок от толщины . . . . .	276
§ 32. Предэкспоненциальный множитель прыжковой проводимости . . . . .	284
Глава 10. Влияние корреляционных эффектов на плотность состояний и прыжковую проводимость . . . . .	288
§ 33. Кулоновская щель в плотности состояний . . . . .	299
§ 34. Роль многочастичных корреляций в прыжковой проводимости . . . . .	306
 ЧАСТЬ II. СИЛЬНО ЛЕГИРОВАННЫЕ ПОЛУПРОВОДНИКИ	
Введение . . . . .	314
Глава 11. Состояния электронов в сильно легированных полупроводниках . . . . .	316
§ 35. Теория линейного экранирования . . . . .	316
§ 36. Плотность состояний вблизи дна зоны проводимости	321
§ 37. Вывод квазиклассической формулы для плотности состояний . . . . .	327
Глава 12. Глубокий хвост плотности состояний и межзонное поглощение света . . . . .	333
§ 38. Метод оптимальной флуктуации . . . . .	333
§ 39. Приближение равномерно заряженного шара. Спектр основных носителей . . . . .	326
§ 40. Точное распределение примесей в оптимальных флуктуациях . . . . .	344

§ 41. Спектр неосновных носителей . . . . .	353
§ 42. Теория межзонного поглощения света . . . . .	357
Глава 13. Теория сильно легированных и сильно компенсированных полупроводников (СЛКП) . . . . .	367
§ 43. Некоррелированное распределение примесей . . . . .	367
§ 44. Коррелированное распределение примесей . . . . .	378
§ 45. Кинетические свойства СЛКП . . . . .	381
§ 46. Полностью компенсированный полупроводник . . . . .	385
Приложения . . . . .	390
1. Вычисление показателя экспоненты энергетического интеграла перекрытия $I_{ij}$ в магнитном поле . . . . .	390
2. Энергия активации $e_a$ в случае анизотропных волновых функций . . . . .	391
3. Оценка коэффициентов $t_i$ и $\nu_{ci}$ в формулах (30.2) и (30.8) . . . . .	393
Литература . . . . .	396

*Борис Иванович Шкловский, Алексей Львович Эфрос*

Электронные свойства легированных полупроводников

(Серия: «Физика полупроводников и полупроводниковых приборов»)

М., 1979 г., 416 стр. с илл.

Редактор *Г. М. Карасева*

Техн. редактор *И. Ш. Аксельрод.*

Корректоры *Т. С. Плетнева, Т. А. Панькова*

ИБ № 2279

Сдано в набор 19.09.78. Подписано к печати 31.01.79. Т-05310. Бумага 84×108<sup>1</sup>/<sub>16</sub>.  
Тип. № 1. Литературная гарнитура. Высокая печать. Услови. печ. л. 21,84.  
Уч.-изд. л. 21,41. Тираж 4500 экз. Заказ № 1272. Цена книги 2 р. 40 к.

Издательство «Наука»

Главная редакция физико-математической литературы  
117071, Москва, В-71, Ленинский проспект, 15

Ордена Трудового Красного Знамени Ленинградская типография № 2 имени  
Евгении Соколовой «Союзполиграфпрома» при Государственном комитете  
СССР по делам издательств, полиграфии и книжной торговли. 198052,  
Ленинград, Л-52, Измайловский проспект, 29.